A Fourier Analysis of Extremal Events

Yuwei Zhao

This thesis has been submitted to
the PhD School of The Faculty of Science,
University of Copenhagen

Supervisor Professor Thomas Mikosch, University of Copenhagen

Date Submitted November 30th, 2013.

Department of Mathematical Sciences

Faculty of Science

University of Copenhagen

Author	Yuwei Zhao Universitetsparken 5, 2100 Copenhagen, Denmark email: zhao@math.ku.dk

Date Submitted	November 30th, 2013
Supervisor	Professor Thomas Mikosch, University of Copenhagen
Assessment Committee	Professor Michael Sørensen, University of Copenhagen Professor Claudia Klüppelberg, Technische Universität München Professor Richard A. Davis, Columbia University

ISBN 978-87-7078-982-0

Contents

Abstract **iv**

Summary **v**

Sammenfatning **vi**

1 Introduction **2**
1.1 Basic motivation . . . 2
1.1.1 Some facts from time series analysis . . . 4
1.1.2 Regular Variation . . . 6
1.2 The extremogram and its estimation . . . 8
1.2.1 Definition and basic properties . . . 8
1.2.2 The extremogram of some regularly varying time series models . . . 9
1.2.3 Estimation of the extremogram . . . 10
1.3 Extremal periodogram and smoothed extremal periodogram . . . 12
1.4 Integrated extremal periodogram . . . 15
1.4.1 Integrated extremal periodogram . . . 15
1.4.2 The bootstrapped integrated extremal periodogram . . . 18
1.5 Directions of future research . . . 19
1.6 Acknowledgments . . . 20

2 A Fourier analysis of extremal events **23**
2.1 Introduction . . . 23
2.2 Preliminaries . . . 25
2.2.1 Regular variation . . . 25
2.2.2 The mixing and dependence conditions (M), (M1) and (M2) . . . 26
2.2.3 The periodogram of extreme events . . . 27
2.3 Examples . . . 29
2.3.1 IID sequence . . . 29
2.3.2 Stochastic volatility model . . . 29
2.3.3 ARMA process . . . 30
2.3.4 Max-moving averages . . . 32
2.4 Basic properties of the periodogram . . . 33

2.4.1 The periodogram ordinates at distinct frequencies are asymptotically uncorrelated . . . 34
2.4.2 Central limit theorem . . . 37
2.5 Smoothing the periodogram . . . 42
2.6 A discussion of related results and possible extensions . . . 52
2.7 Appendix . . . 53
2.7.1 Some trigonometric sum formulas . . . 53
2.7.2 The spectral density f_A of an ARMA(1,1) process . . . 56

3 Measures of serial extremal dependence **63**
3.1 Introduction . . . 63
3.1.1 The extremal index as reciprocal of the expected extremal cluster size . . 64
3.1.2 The extremogram: an asymptotic correlogram for extreme events . . . 66
3.2 Estimation of the extremogram . . . 79
3.2.1 Asymptotic theory . . . 79
3.2.2 Cross-extremogram for bivariate time series . . . 84
3.3 An example: Equity indices . . . 84
3.4 A Fourier analysis of extreme events . . . 88

4 Integrated extremal periodogram **96**
4.1 Introduction . . . 96
4.1.1 Regularly varying sequences . . . 96
4.1.2 The extremogram . . . 97
4.1.3 The sample extremogram . . . 97
4.1.4 Spectral density and periodogram . . . 98
4.1.5 The integrated periodogram . . . 98
4.2 Preliminaries . . . 100
4.2.1 Some moment calculations . . . 100
4.2.2 Mixing conditions . . . 100
4.2.3 Central limit theory for the sample extremogram . . . 102
4.2.4 Mean square consistency of the integrated periodogram . . . 105
4.3 Functional central limit theorem for the integrated periodogram . . . 105
4.4 The bootstrapped integrated periodogram . . . 109
4.4.1 Stationary bootstrap . . . 109
4.4.2 The bootstrapped sample extremogram . . . 110
4.4.3 The bootstrapped integrated periodogram . . . 111
4.5 Proof of Lemma 4.2.9 . . . 112
4.6 Proof of Theorem 4.3.1 . . . 115
4.7 Proof of Theorem 4.3.3 . . . 122
4.8 Proof of Theorem 4.4.2 . . . 123

Abstract

The extremogram is an asymptotic correlogram for extreme events constructed from a regularly varying strictly stationary sequence. Correspondingly, the spectral density generated from the extremogram is introduced as a frequency domain analog of the extremogram. Its empirical estimator is the *extremal periodogram*. The extremal periodogram shares numerous asymptotic properties with the periodogram of a linear process in classical time series analysis: the asymptotic distribution of the periodogram ordinates at the Fourier frequencies have a similar form and smoothed versions of the periodogram are consistent estimators of the spectral density. By proving a functional central limit theorem, the *integrated extremal periodogram* can be used for constructing asymptotic tests for the hypothesis that the data come from a strictly stationary sequence with a given extremogram or extremal spectral density. A numerical method, the *stationary bootstrap*, can be applied to the estimation of the integrated extremal periodogram.

Summary

This thesis consists of three papers which are contained in Chapter 2–4.

Chapter 2 is based on the paper

> [19] Mikosch, T. and Zhao, Y. (2013) A Fourier analysis of extreme events. Bernoulli, to appear.

It yields the basic asymptotic theory for the *extremal periodogram*. These results include the proof of the asymptotic independence of the extremal periodogram at distinct frequencies and the consistency of the *smoothed extremal periodogram*. This chapter illustrates that there are numerous similarities between the extremal periodogram and the periodogram of a stationary sequence.

Chapter 3 is based on the paper

> [10] Davis, R.A., Mikosch, T. and Zhao, Y. (2013) Measures of serial extremal dependence and their estimation. Stoch. Proc. Appl., 123, 2575–2602.

It is a review of the recent developments on measuring extremal dependence in a time series. The paper starts with a critique of the *extremal index* as a measure of the extremal cluster size, then various regularly varying time series models and their extremograms and extremal periodograms are discussed. In this framework, max-stable processes with Fréchet marginals get some special attention.

Chapter 4 is based on the paper

> [20] Mikosch, T. and Zhao, Y. (2013) The integrated periodogram of a dependent extremal event sequence. Working paper.

It is devoted to the asymptotic properties of the *integrated extremal periodogram* and its applications for constructing goodness-of-fit tests for a time series model based on its extremes. The main results are functional central limit theorems for the weighted integrated periodogram with Gaussian limits and their stationary bootstrap analogs.

Sammenfatning

Denne afhandling indeholder en introduktion til emnet i Kapitel 1 og tre videnskabelige artikler, som findes i Kapitlerne 2–4.

Kapitel 2 er baseret på artiklen

> [19] Mikosch, T. and Zhao, Y. (2013) A Fourier analysis of extreme events. Bernoulli, to appear.

Her er givet en fundamental asymptotisk teori for det ekstremale periodogram. Blandt resultaterne er en bevis af den asymptotiske uafhængighed for det ekstremale periodogram for forskellige frekvenser og bevisen af konsistensen for det udglattede ekstremale periodogram. I kapitlen også sammenlignes egenskaberne af det ekstremale periodogram og det sadvanlige periodogram af en stationær tidsrcekke.

Kapitel 3 er baseret på artiklen

> [10] Davis, R.A., Mikosch, T. and Zhao, Y. (2013) Measures of serial extremal dependence and their estimation. Stoch. Proc. Appl., 123, 2575–2602.

Artiklen er en review af moderne metoder for målingen af ekstremale afhængigheder i en tidsrække. I begyndelsen undersøges den ekstremale indeks, som er et standardmål af den ekstremale cluster størrelse. Bagefter undersøges regulær varierende tidsrækker, deres extremogram og ekstremale periodogram. Der er fokus på egenskaberne af max-stabile processer med en Fréchet marginalfordeling.

Kapitel 4 er baseret på artiklen

> [20] Mikosch, T. and Zhao, Y. (2013) The integrated periodogram of a dependent extremal event sequence. Working paper.

Her undersøges de asymptotiske egenskaber af det integrerede ekstremale periodogram og anvendelser til goodness-of-fit tests for tidsrækkemodeller med ekstremale hændelser. Hovedresultaterne er funktionale centrale grænseværdisætninger for det udglættede integrerede periodogram med gaussiske grænseprocesser. Der også bevises funktionale centrale grænseværdisætninger for det ekstremale periodogram ved anvendelsen af den stationære bootstrap metode.

Chapter 1

Introduction

1.1 Basic motivation

When the financial crisis started in 2007 the general public learned from the media about the inadequate use of probabilistic and statistical models in finance. In contrast to the original goals of applying these models — increasing the profit and reducing the risk — it seemed that these models fell short in achieving their objectives, in particular, they underestimated the risk. To avoid similar mistakes in the future one needs to study the pitfalls of these models and to test their reliability under stress situations. In this thesis, we will propose some statistical tools for measuring serial extremal dependence in a (financial) time series and study its advantages and limitations.

The hypothesis of (approximate) normality of the data has been the basis for building stochastic models due to traditional preferences in probability theory and statistics which are at its best under the (approximate) normality assumption. For example, this assumption is imposed on some standard models in finance. The Nobel Prize winning model of Black-Scholes-Merton which is used for pricing European-style options assumes that the log-price of the stock price is a geometric Brownian motion. By the year 2007, the geometric Brownian motion was widely accepted as a feasible model for speculative prices despite the fact that real-life log-returns exhibit significant deviations from the normality hypothesis.

Data in areas as diverse as insurance, finance, telecommunication and the earth sciences typically fluctuate strongly in certain periods of time, and these fluctuations tend to occur in clusters. This oberservation disagrees with the assumption of Gaussianity. The topics of this thesis are closely related to time series analysis. Also in this area, the assumption of Gaussianity is a classical paradigm. In this thesis, we will deal with time series models which are highly non-Gaussian: they exhibit heavy tails in a sense which will be made precise later; rougly speaking, we will deal with time series models whose marginal distributions have infinite moments of a certain order.

This thesis is located at the boundary beetwen time series analysis and extreme value theory. In the latter field one is particularly interested in events which happen far away from the median of the underlying probability distribution and manifest themselves as natural or man made catastrophes, big losses of an investment portfolio, huge files transferred via the Internet, etc.

Classical extreme value theory deals with the extremes of iid sequences of univariate data $(X_t)_{t\in\mathbb{Z}}$. The Fisher-Tippett theorem (Theorem 3.2.3 in Embrechts et al. [11]) gives the basis for the asymptotic theory of the sequence of the maxima

$$M_n = \max_{i=1,\ldots,n} X_i\,, \quad n \geq 1\,.$$

This theorem states that the only possible non-degenerate limit laws H for the suitably normalized and centered maxima M_n, i.e.

$$c_n^{-1}(M_n - d_n) \xrightarrow{d} Y \sim H\,, \tag{1.1.1}$$

for appropriate $c_n > 0$ and $d_n \in \mathbb{R}$, $n \geq 1$, are given by the following extreme value distributions (also called max-stable distributions):

- Fréchet distribution:

$$\Phi_\alpha(x) = \begin{cases} 0\,, & x \leq 0 \\ \exp\{-x^{-\alpha}\}\,, & x > 0 \end{cases} \qquad \alpha > 0\,.$$

- Weibull distribution:

$$\Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^{\alpha}\}\,, & x \leq 0 \\ 1\,, & x > 0 \end{cases} \qquad \alpha > 0\,.$$

- Gumbel distribution:

$$\Lambda(x) = \exp\{-e^{-x}\}\,, \quad x \in \mathbb{R}\,.$$

In this thesis, we will be concerned mainly with the Fréchet distribution Φ_α and its maximum domain of attraction, i.e. those distributions F of the X_i's such that (1.1.1) holds for $H = \Phi_\alpha$ (we write $F \in \mathrm{MDA}(\Phi_\alpha)$). It is easy to see that the right tail of Φ_α is of power law type:

$$\overline{\Phi}_\alpha(x) \sim x^{-\alpha}\,, \quad x \to \infty\,,$$

and then, necessarily, the tail $\overline{F} = 1 - F$ for $F \in \mathrm{MDA}(\Phi_\alpha)$ is of similar type.

If one wants to judge how well a model fits the time series of the data there are several ways of doing this. Classical time series analysis is mostly concerned with fitting the covariance structure of the data to a suitable model and to draw conclusions from the theoretical properties of the fitted model about the data; see Brockwell and Davis [4]. When dealing with the serial dependence of extremes, covariances are less informative: extremes happen in the tail of the distribution and covariances do not contain a lot of information about the tails. The basic theoretical tool of this thesis is the extremogram introduced in Davis and Mikosch [8] as an alternative to the covariance function of the data: it is an asymptotic covariance function derived from the sequence of the indicator functions of the extreme events in a time series. As such, one can still use the notions of time series analysis in the extreme value context. In particular, we will borrow tools and basic methodology from the Fourier analysis of time series and build a corresponding theory for the extremal events in a time series.

In what follows, we will explain the essence of our findings.

1.1.1 Some facts from time series analysis

Classical time series analysis is mostly concerned with the second order (or covariance) structure of a univariate time series $(X_t)_{t\in\mathbb{Z}}$; see Brockwell and Davis [4] as a major reference. For this reason, it is common to assume second order or covariance stationarity of the data, i.e.

$$\gamma_X(h) = \operatorname{cov}(X_t, X_{t+h}) = \operatorname{cov}(X_0, X_h)\,, \quad t \in \mathbb{Z}\,, \quad h \in \mathbb{Z}\,,$$

and we refer to γ_X and $\rho_X(h) = \gamma_X(h)/\gamma_X(0)$ as the autocovariance and autocorrelation functions of the time series (ACF and ACVF), respectively. In the context of extreme value theory this kind of stationarity is less meaningful (the case of stationary Gaussian time series being an exception) and therefore we will assume strict stationarity of the underlying time series whenever we study the extremes of this series. Under this assumption, we will denote a generic element of (X_t) by X.

The classical estimators of γ_X and ρ_X are their sample counterparts, given by

$$\widetilde{\gamma}_X(h) = \frac{1}{n}\sum_{t=1}^{n}(X_t - \overline{X}_n)(X_{t+h} - \overline{X}_n)\,, \quad \text{and} \quad \widetilde{\rho}_X(h) = \widetilde{\gamma}_X(h)/\widetilde{\gamma}_X(0)\,, \quad h \in \mathbb{Z}\,,$$

where $\overline{X}_n$ is the sample mean of $X_1, \ldots, X_n$. The sample autocorrelation function (sample ACF) is the main tool in the time domain to judge about the dependence structure of the data. The sample ACF plot is widely used as an exploratory tool and it is also a major building block for many parameter estimation techniques in time series analysis. The time domain approach is intuitive and simple to apply. Therefore its tools and estimators are contained in all major software packages.

Early on, the time domain approach to a time series was supplemented by the frequency domain approach. Based on fundamental theory (e.g. Herglotz's Theorem 4.3.1 in Brockwell and Davis [4]), the Fourier series based on the ACVF γ_X of the real-valued stationary time series (X_t) given by

$$f_X(\lambda) = \sum_{h=-\infty}^{\infty} \gamma_X(h)\, e^{-ih\lambda} = \gamma_X(0) + 2\sum_{h=1}^{\infty}\gamma_X(h)\cos(h\lambda)\,, \quad \lambda \in [0,\pi]\,, \tag{1.1.2}$$

is the quantity of major interest. In view of the spectral representation of a stationary time series (Theorem 4.8.2 in Brockwell and Davis [4]) the spectral density (or, more generally, spectral distribution) completely characterizes the second order properties of a stationary time series and, therefore, the spectral or frequency domain of time series analysis is just another language which expresses the time domain theory in the world of Fourier analysis. Although less elementary (the quantities of Fourier analysis often do not have a "straightforward" interpretation) Fourier analysis is often very powerful.

This comment also applies to the estimation in the frequency domain. Estimation is frequently based on the sample version of the spectral density, the periodogram, given by

$$I_{n,X}(\lambda) = \Big| n^{-1/2}\sum_{t=1}^{n} X_t e^{-it\lambda}\Big|^2 = \sum_{|h|<n} e^{-ih\lambda}\widetilde{\gamma}_X(h)\,, \quad \lambda \in [0,\pi]\,,$$

which is obtained by replacing the ACF by its sample version in (1.1.2). Calculations based on the periodogram are often executed at the Fourier frequencies $2\pi j/n \in (0,\pi)$. The Fast

Fourier Transform (FFT) is a fast and efficient way of calculating the periodogram at the Fourier frequencies; see Section 10.7 of Brockwell and Davis [4].

For the sake of comparison with the periodogram for extremal events, we recall some of the classical results in the frequency domain of time series analysis; see Chapter 10 in Brockwell and Davis [4]. We will assume that (X_t) is a linear process with iid white noise with variance σ^2. Most of the results have to be formulated in terms of Fourier frequencies closest to a fixed frequency $\lambda \in (0,\pi)$, but for the ease of presentation we will neglect this fact.

For any frequency $\lambda \in (0,\pi)$,

$$EI_{n,X}(\lambda) \to f_X(\lambda)\,, \quad n \to \infty\,,$$

and the vector of the periodogram ordinates at the distinct frequencies $0 < \lambda_1 < \cdots < \lambda_h < \pi$, $h \geq 1$, satisfies the limit relation

$$\big(I_{n,X}(\lambda_i)\big)_{i=1,\ldots,h} \stackrel{d}{\to} \sigma^2 \big(f_X(\lambda_i)\, E_i\big)_{i=1,\ldots,h}\,, \quad n \to \infty\,, \tag{1.1.3}$$

where (E_i) is an iid standard exponential sequence. The proof of this fact is based on the representation

$$\begin{aligned} I_{n,X}(\lambda) &= \Big|n^{-1/2}\sum_{t=1}^n X_t e^{-it\lambda}\Big|^2 = \Big(n^{-1/2}\sum_{t=1}^n X_t\cos(t\lambda)\Big)^2 + \Big(n^{-1/2}\sum_{t=1}^n X_t\sin(t\lambda)\Big)^2 \\ &= \alpha^2_{n,X}(\lambda) + \beta^2_{n,X}(\lambda)\,, \end{aligned} \tag{1.1.4}$$

and an application of a multivariate central limit theorem to $\big(\alpha_{n,X}(\lambda_i), \beta_{n,X}(\lambda_i)\big)_{i=1,\ldots,h}$. The representation of the limiting vector is then crucial for smoothing the periodogram which leads to consistent estimation. For the periodogram of extremal events, similar ideas and methods of proof apply; see Section 1.3 on page 12. The *integrated periodogram*

$$J_{n,X}(x) = \int_0^x I_{n,X}(\lambda) g(\lambda)\, d\lambda\,, \quad x \in [0,\pi]\,,$$

where g is a smooth function, shows a close relationship with the empirical distribution of an iid sequence by numerous results. Therefore it is sometimes referred to as the *the empirical spectral distribution function*; see for example Dahlhaus [5] and Dahlhaus and Polonik [6]. The functional central limit theorem for the integrated periodogram can be taken as the basis for constructing goodness-of-fit tests for iid sequences and linear processes. For example, for an iid sequence (X_t) with mean zero and variance σ^2 the following limit result holds in $\mathbb{C}[0,\pi]$, the space of continuous functions on $[0,\pi]$:

$$\sqrt{n}(J_{n,X}(\cdot) - \sigma^2\cdot) \stackrel{d}{\to} G\,,$$

where the limit process G is a Brownian bridge on $[0,\pi]$. Then, similarly to the Kolmogorov-Smirnov test, one can consider the supremum functional acting on the converging process to obtain

$$\sqrt{n} \sup_{\lambda\in[0,\pi]} \big|J_{n,X}(\lambda) - \sigma^2\lambda\big| \stackrel{d}{\to} \sup_{\lambda\in[0,\pi]} |G(\lambda)|\,.$$

Thus, under the null hypothesis that (X_t) is iid the left-hand expression can be taken as a goodness-of-fit test statistic and its limit distribution can be used to construct an asymptotic test for this hypothesis. The corresponding test is called *Grenander-Rosenblatt test*; see Priestley [22] for a general reference on the integrated periodogram. Similar results also hold for the integrated periodogram for extremal events; see Section 1.4 on page 15.

1.1.2 Regular Variation

The notion of regular variation is basic in extreme value theory and limit theory for partial sums of iid random variables. In multivariate extreme value theory, regular variation with index $\alpha > 0$ of the d-dimensional iid random vectors X_t, $t \in \mathbb{R}$, with values in $(0,\infty)^d$ is necessary and sufficient for the fact that the suitably normalized sequence of component-wise maxima $(a_n^{-1} \max_{t\le n} X_t^{(i)})_{i=1,\dots,d}$, $t = 1,2,\dots$, converges in distribution to a d-dimensional extreme value distribution H on $(0,\infty)^d$ whose marginal distributions are Fréchet Φ_α-distributed; see Resnick [24] for a general theory of multivariate extremes for iid sequences. Similarly, for a general $\mathbb{R}^d$-valued iid sequence (X_t), the sequence of suitably normalized and centered partial sums $a_n^{-1}(X_1 + \cdots + X_n - b_n)$ converges in distribution to an infinite variance α-stable limit if and only if the distribution of X_0 is regularly varying with index α. The index α is then necessarily in the range $\alpha \in (0,2)$. We refer to Rvačeva [26] and Resnick [25] for proofs of this fact.

Various definitions of a d-dimensional regularly varying vector X exist; we refer to Resnick [23, 24, 25]. We start with a definition in terms of spherical coordinates. We say that X is regularly varying with index $\alpha > 0$ and spectral measure $P(\Theta \in \cdot)$ on the Borel σ-field of the unit sphere $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : |x| = 1\}$ if [1] the following weak limits exist for every fixed $t > 0$:

$$\frac{P(|X| > tx\,, X/|X| \in \cdot)}{P(|X| > x)} \stackrel{w}{\to} t^{-\alpha}\, P(\Theta \in \cdot)\,, \quad x \to \infty\,. \tag{1.1.5}$$

Relation (1.1.5) can be written in an equivalent form as a pair of conditions:

1. The norm $|X|$ is regularly varying in the classical sense, i.e. $P(|X| > tx)/P(|X| > x) \to t^{-\alpha}$, $t > 0$, or, equivalently, $P(|X| > x) = x^{-\alpha} L(x)$, $x > 0$, for a slowly varying function L; cf. Bingham et al. [2] for an encyclopedia on regularly varying functions.

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

$$P(X/|X| \in \cdot \mid |X| > x) \stackrel{w}{\to} P(\Theta \in \cdot)\,, \quad x \to \infty\,. \tag{1.1.6}$$

In any of these limit relations, it is possible to replace the converging parameter x by a sequence (a_n) such that $P(|X| > a_n) \sim n^{-1}$. Then (1.1.5) and (1.1.6), respectively, read as

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

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

The convergence relation (1.1.5) can be understood as convergence on the particular sets $\{\mathbf{x} \in \mathbb{R}^d : |\mathbf{x}| > t, \mathbf{x}/|\mathbf{x}| \in S\}$ for Borel sets $S \subset \mathbb{S}^{d-1}$ with a smooth boundary. This convergence can be extended to the Borel σ-field on $\overline{\mathbb{R}}_0^d = \overline{\mathbb{R}}^d \setminus \{\mathbf{0}\}$, $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$:

$$\mu_x(\cdot) = \frac{P(x^{-1}X \in \cdot)}{P(|X| > x)} \xrightarrow{v} \mu(\cdot)\,, \quad x \to \infty\,. \tag{1.1.7}$$

Here $\xrightarrow{v}$ refers to vague convergence of measures on the Borel σ-field on $\overline{\mathbb{R}}_0^d$, i.e. $\int_{\overline{\mathbb{R}}_0^d} f\, d\mu_x \to \int_{\overline{\mathbb{R}}_0^d} f\, d\mu$ as $x \to \infty$ for any continuous and compactly supported f on $\overline{\mathbb{R}}_0^d$; see Kallenberg [15], Resnick [24]. This means in particular, that the support of f is bounded away from zero. In view of (1.1.5), $\mu(\{\mathbf{x} \in \overline{\mathbb{R}}_0^d : |\mathbf{x}| > t, \mathbf{x}/|\mathbf{x}| \in S\}) = t^{-\alpha} P(\Theta \in S)$, and therefore μ is a Radon measure (i.e. finite on sets bounded away from zero) satisfying $\mu(tA) = t^{-\alpha}\mu(A)$, $t > 0$. In particular, μ does not charge points containing infinite components. Again, the parameter x in (1.1.7) can be replaced by a sequence (a_n) satisfying $P(|X| > a_n) \sim n^{-1}$ and then we get

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

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

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

where N is a Poisson random measure with mean measure μ and state space $\overline{\mathbb{R}}_0^d$; see Resnick [23, 24]. Since point process convergence is basic to extreme value theory, the notion of multivariate regular variation is very natural in the context of extreme value theory for multivariate observations with heavy-tailed components; see also the recent monograph by Resnick [25] who stresses the importance of the notion of regular variation as relevant for many applications in finance, insurance and telecommunications.

Regularly varying stationary sequences. In the context of this thesis, the notion of a regularly varying strictly stationary sequence will be relevant. This notion was coined by Davis and Hsing [7]. It simply means that the finite-dimensional distributions of an $\mathbb{R}^d$-valued strictly stationary sequence (X_t) are regularly varying for some index $\alpha > 0$. In the context of time series analysis, it is convenient to use the alternative definition

$$\frac{P(x^{-1}(X_1, \ldots, X_h) \in \cdot)}{P(|X_0| > x)} \xrightarrow{v} \mu_h(\cdot)\,, \tag{1.1.8}$$

with non-null limit measures μ_h on $\overline{\mathbb{R}}_0^{dh}$, i.e. the converging measures are all normalized by $P(|X_0| > x)$. This normalization is natural since it does not depend on h.

1.2 The extremogram and its estimation

1.2.1 Definition and basic properties

To facilitate the application of the time domain and frequency domain approaches to the extremes in a strictly stationary $\mathbb{R}^d$-valued sequence, the extremogram

$$\gamma_{AB}(h) = \lim_{n\to\infty} n \operatorname{cov}\big(I_{\{a_n^{-1}X_0\in A\}}, I_{\{a_n^{-1}X_h\in B\}}\big)\,, \quad h \geq 0\,, \tag{1.2.1}$$

was introduced by Davis and Mikosch [8]. Here (a_n) is a suitably chosen normalization sequence and A, B are two fixed sets bounded away from zero. The quantity $\gamma_{AB}(h)$ measures the influence of the extremal event $\{X_0 \in a_nA\}$ at time zero on the extremal event $\{X_h \in a_nB\}$, h lags apart.

For fixed n, $(I_{\{a_n^{-1}X_t\in A\}})$ and $(I_{\{a_n^{-1}X_t\in B\}})$ constitute strictly stationary sequences and the limit sequence $(\gamma_{AB}(h))$ inherits the property of covariance function from $\operatorname{cov}\big(I_{\{a_n^{-1}X_0\in A\}}, I_{\{a_n^{-1}X_h\in B\}}\big)$. Moreover, for any dimension d and suitable sets A, B, the limiting sequence

$$\begin{pmatrix} \gamma_{AA}(h) & \gamma_{AB}(h) \\ \gamma_{BA}(h) & \gamma_{BB}(h) \end{pmatrix}, \quad h \geq 0\,,$$

is a matrix covariance function. In an asymptotic sense, one can use the notions of classical time series analysis for the sequences of indicator functions $(I_{\{a_n^{-1}X_t\in A\}})$ and $(I_{\{a_n^{-1}X_t\in B\}})$. Of course, there are several crucial differences to classical time series analysis. Firstly, the value of $\gamma_{AB}(h)$ cannot be negative and the strictly stationary sequences of indicator functions $(I_{\{a_n^{-1}X_t\in A\}})$ and $(I_{\{a_n^{-1}X_t\in B\}})$ depend on n, i.e., we are dealing with an array of strictly stationary sequences. Last but not least, the notion of autocovariance or cross-covariance function is only defined in an asymptotic sense.

A motivating example is the so-called *(upper) tail dependence coefficient* of the vector (X_0, X_h) given as the limit:

$$\rho(h) = \lim_{x\to\infty} P(X_h > x \mid X_0 > x)\,. \tag{1.2.2}$$

Here we assume that X has infinite right endpoint. These pairwise tail dependence coefficients have attracted some attention in the literature on quantitative risk management; see for example McNeil et al. [17]. To avoid ambiguity, we assume that (a_n) satisfies the relation $nP(|X_n| > a_n) \sim 1$. With this choice of (a_n), $\gamma_{AB}(h) = \lim_{n\to\infty} nP(a_n^{-1}X_0 \in A, a_n^{-1}X_h \in B)$, which coincides with $\rho(h)$ if $d = 1$, $X \geq 0$ a.s. and $A = B = (1,\infty)$. Indeed,

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

The limit in (1.2.1) does not exist in general, but if the sequence (X_t) is strictly stationary and regularly varying, there exist non-null limiting measures μ_{h+1} on $\overline{\mathbb{R}}_0^{d(h+1)}$ such that

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

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

1.2.2 The extremogram of some regularly varying time series models

Numerous real-valued time series models constitute regularly varying sequences; we give some examples and the corresponding extremograms ρ in (1.2.2); these examples are taken from Davis and Mikosch [8], Davis et al. [9, 10].

IID regularly varying sequence

The simplest example is an iid regularly varying sequence (Z_t) with index α when Z is regularly varying with the same index; the limit measures μ_h are concentrated on the axes and $\rho(h) = 0$, $h \geq 1$.

Regularly varying linear process

Building upon this iid regularly varying sequence (Z_t), we can define a regularly varying (causal) linear process

$$X_t = \sum_{t=0}^{\infty} \psi_j Z_{t-j}\,, \quad t \in \mathbb{Z}\,. \tag{1.2.3}$$

In the history of extreme value theory and in time series analysis a lot of attention was paid to this model. The process (X_t) inherits regular variation under mild conditions on the deterministic sequence (ψ_i) which are close to those dictated by the 3-series theorem, ensuring the a.s. convergence of the series in (1.2.3); see Mikosch and Samorodnitsky [18] for the case of the tails of the marginals. The regular variation of the finite-dimensional distributions of (X_t) follows since regular variation is preserved under affine transformations of regularly varying vectors. The class (1.2.3) includes causal ARMA processes which are relevant for applications. We refer to Chapter 7 of Embrechts et al. [11] for various applications of regularly varying linear processes. Under the tail balance condition $P(Z > x) \sim p\,P(|Z| > x)$, $P(Z \leq -x) \sim q\,P(|Z| > x)$, as $x \to \infty$, for some $p, q \geq 0$ with $p + q = 1$,

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

Max-stable processes with Fréchet marginals

The class of max-stable processes has recently attracted some attention since it is a flexible class for modeling heavy tails and spatio-temporal dependence. Since the finite-dimensional distributions of max-stable processes are explicitly given it is often simple to verify properties (such as regular variation) and to calculate certain quantities (e.g. mixing coefficients, extremogram). We will focus on stationary ergodic max-stable processes (X_t) with Fréchet marginals given by

$$P(X_t \leq x) = \exp\Big\{ - x^{-\alpha} \int_E f_t^\alpha(y)\,\nu(dy)\Big\}\,, \quad x > 0\,,$$

where the non-negative functions $f_t \in L^\alpha(E,\mathcal{E},\nu)$ and $(E,\mathcal{E},\nu)$ is a σ-finite measure space. The corresponding tail dependence coefficient is given by

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

GARCH and stochastic volatility (SV) processes

They are well-known time series models for financial returns which are widely used in practice. Both processes are defined via the relation

$$X_t = \sigma_t Z_t\,, \quad t \in \mathbb{Z}\,, \tag{1.2.4}$$

where (σ_t) is a strictly stationary sequence of non-negative random variables and (Z_t) is an iid sequence. Depending on the dependence structure of the process (σ_t) one can achieve very different tail and dependence behavior of the return sequence (X_t).

For the SV model, (σ_t) and (Z_t) are independent. If we further assume that Z is regularly varying with index $\alpha > 0$ and $E(\sigma^{\alpha+\varepsilon}) < \infty$ for some $\varepsilon > 0$ then (X_t) is regularly varying with index α. This is a simple consequence of Breiman's lemma [3] about the tail of products of independent random variables. In this case, $\rho(h) = 0$, $h \geq 1$, as in the iid case. Thus the extremes in a regularly varying SV model are asymptotically independent at all lags, indicating that the SV model does not exhibit extremal clustering through time.

This is in contrast to the GARCH model. For the ease of presentation, we focus on the GARCH$(1,1)$ process. Assume that the iid noise sequence (Z_t) has mean zero and unit variance and the volatility sequence (σ_t) is given by the equations

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2 = \alpha_0 + \sigma_{t-1}^2 C_{t-1}\,, \quad C_t = \alpha_1 Z_t^2 + \beta_1\,, \quad t \in \mathbb{Z}\,.$$

Here the parameters $\alpha_0, \alpha_1, \beta_1 > 0$ are chosen such that (X_t) is strictly stationary and a unique positive solution α to the equation $EC^{\kappa/2} = 1$, $\kappa > 0$, exists. Then the sequences (σ_t) and (X_t) are regularly varying with index α (under additional regularity conditions); see Kesten [13], Goldie [12]. Moreover,

$$\rho(h) = E(\min(1, C_0 \cdots C_{h-1}))^{\alpha/2}\,, \quad h \geq 1\,,$$

and $0 < \rho(h) \to 0$ as $h \to \infty$ at an exponential rate.

The examples of the GARCH and SV models show that their extremograms ρ could be used for distinguishing between these two processes, i.e. it might be possible to dicriminate between these processes solely based on their extremal behavior. This is similar to the ACVF and ACF in classical time series analysis where the second order structure of a time series is taken as a means to judge the goodness of fit of a stationary time series.

1.2.3 Estimation of the extremogram

For the sake of simplicity, we focus on the estimation of the extremogram in the case $A = B$; we write $\gamma_A = \gamma_{AA}$ and $\rho_A = \rho_{AA}$. For the estimation of the extremogram γ_{AB} we refer to Davis et al. [8, 9]. Recall the fact that

$$\gamma_A(h) = \lim_{n\to\infty} n\,P(a_n^{-1} X_0 \in A, a_n^{-1} X_h \in A)$$

Natural estimators are obtained by replacing the probabilities in the limiting relation by their empirical counterparts. We refer to such an estimator as the *sample extremogram*:

$$\widetilde{\gamma}_A(h) = \frac{m_n}{n} \sum_{t=1}^{n-h} I_{\{a_m^{-1} X_i \in A\}} I_{\{a_m^{-1} X_{i+h} \in A\}}\,, \quad h \geq 0\,. \tag{1.2.5}$$

Here $m = m_n \to \infty$ and $m/n \to 0$. We can also define an estimator of the standardized extremogram $\rho_A(h) = \gamma_A(h)/\gamma_A(0)$ by

$$\widetilde{\rho}_A(h) = \widetilde{\gamma}_A(h)/\widetilde{\gamma}_A(0)\,, \quad h \geq 0\,.$$

For the derivation of the consistency and asymptotic normality of the estimators, we introduce mixing and anti-clustering conditions:

(M) The sequence (X_t) is *strongly mixing with rate function* (ξ_t) given by

$$\xi_t = \sup\left\{|P(A \cap B) - P(A)P(B)| : t \geq 0\,, A \in \mathcal{F}_{-\infty}^0\,, B \in \mathcal{F}_t^\infty\right\}, \quad t \geq 1\,,$$

where $\mathcal{F}_{-\infty}^0$ is the σ-algebra generated by $\{\ldots, X_{-1}, X_0\}$ and $\mathcal{F}_t^\infty$ is the σ-algebra generated by $\{X_t, X_{t+1}, \ldots\}$. Moreover, there exist sequences $m = m_n \to \infty$ and $r_n \to \infty$ such that $m_n/n \to 0$, $r_n/m_n \to 0$ and

$$\lim_{n\to\infty} m_n \sum_{h=r_n}^{\infty} \xi_h = 0\,, \tag{1.2.6}$$

and for all $\epsilon > 0$, an *anti-clustering condition* holds:

$$\lim_{k\to\infty} \limsup_{n\to\infty} \sum_{h=k}^{r_n} P(|X_h| > \epsilon\, a_m \mid |X_0| > \epsilon\, a_m) = 0\,. \tag{1.2.7}$$

(M1) The sequences (m_n), (r_n), $k_n = [n/m_n]$ from (M) satisfy the growth conditions $k_n \xi_{r_n} \to 0$, and $m_n = o(n^{1/3})$.

Condition (1.2.6) is easily satisfied if the mixing rate (ξ_h) is geometric, i.e. exponentially decaying to zero. Under mild conditions, the popular classes of ARMA, max-stable, GARCH and stochastic volatility processes are strongly mixing with geometric rate; see Davis et al. [8, 9, 10] , Mikosch and Zhao [19] for discussions of these examples.

Condition (1.2.7) is similar in spirit to condition (2.8) used in Davis and Hsing [7] for establishing convergence of a sequence of point processes to a limiting cluster point process. It is much weaker than the anti-clustering condition $D'(\epsilon a_n)$ of Leadbetter which is well known in the extreme value literature; see Leadbetter et al. [16] or Embrechts et al. [11].

The quantities m_n and r_n have some straightforward interpretation as size in a large-small block scheme: the sample $X_1, \ldots, X_n$ consists of roughly k_n large disjoint blocks of size m_n. After chopping off the first r_n elements in each large block one aims at ensuring the asymptotic independence of the resulting large blocks.

By applying this method, one can prove the following pre-asymptotic central limit theorem; see Section 3 in Davis and Mikosch [8], cf. Section 2 in Mikosch and Zhao [20].

Lemma 1.2.1. *Assume that (X_t) is an $\mathbb{R}^d$-valued strictly stationary regularly varying sequence with index $\alpha > 0$. Let $A \subset \overline{\mathbb{R}}_0^d$ be bounded away from zero and $\mu_1(\partial A) = 0$. If the mixing conditions* (M), (M1) *hold and $\sum_{l=1}^{\infty} \gamma_A(l) < \infty$ then for $h \geq 0$,*

$$\widetilde{\gamma}_A(h) \stackrel{P}{\to} \gamma_A(h)\,, \tag{1.2.8}$$

$$(n/m)^{1/2}\big(\widetilde{\gamma}_A(i) - E\widetilde{\gamma}_A(i)\big)_{i=0,\ldots,h} \stackrel{d}{\to} (Z_i)_{i=0,\ldots,h}\,, \tag{1.2.9}$$

where $(Z_i)_{i=0,\ldots,h}$ is mean zero Gaussian with covariance matrix $\Sigma_h = (\sigma_{ij})_{i,j=0,\ldots,h}$ whose entries are given by

$$\sigma_{ij} = \gamma_A(i,j) + \sum_{l=1}^{\infty} \big[\gamma_A(i,l,l+j) + \gamma_A(j,l,l+i)\big]\,, \quad i,j = 0,\ldots,h\,,$$

and for $u, s, t \geq 0$,

$$\gamma_A(u,s,t) = \lim_{n\to\infty} n\,P(a_n^{-1}X_0 \in A\,, a_n^{-1}X_u \in A, a_n^{-1}X_s \in A, a_n^{-1}X_t \in A)\,,$$

with the convention that $\gamma_A(u,t) = \gamma_A(u,u,t)$.

Unlike the classical central limit theorem, the centering part $E\widetilde{\gamma}_A(h)$ in (1.2.9), called the *pre-asymptotic extremogram*, in general cannot be replaced by its limit $\gamma_A(h)$. It is difficult to show that

$$(n/m)^{-1/2}|E\widetilde{\gamma}_A(h) - \gamma_A(h)| \to 0\,, \quad n \to \infty\,,$$

even for "nice" models such as GARCH(1,1). For this well-studied model, the information about the tail behavior is not sufficiently known. Meanwhile, the pre-asymptotic extremogram has a very concrete interpretation in contrast to its less intuitive limit − the extremogram.

The covariance matrix Σ_h in (1.2.9) is in general unfamiliar and very hard to calculate. Davis et al. [9] applied the stationary bootstrap introduced by Politis and Romano [21] for the estimation of Σ_h. The algorithm will be explained in Section 1.4 below.

1.3 Extremal periodogram and smoothed extremal periodogram

In this section we present the results of Chapter 2. Observing that the extremogram γ_A is an ACVF of a stationary process, in analogy with the spectral density (1.1.2) of a stationary process (see (1.1.2)) we define the following spectral density

$$f_A(\lambda) = \sum_{h=-\infty}^{\infty} \gamma_A(h)e^{-ih\lambda} = \gamma_A(0) + 2\sum_{h=1}^{\infty} \gamma_A(h)\cos(h\lambda)\,, \quad \lambda \in \Pi = [0,\pi]\,. \tag{1.3.1}$$

In classical time series analysis the periodogram is a natural estimator of the spectral density; see (1.1.4). In analogy, we define the *extremal periodogram* of a strictly stationary regularly varying $\mathbb{R}^d$-valued sequence (X_t) by

$$\begin{aligned} I_{nA}(\lambda) &= \frac{m_n}{n}\Big|\sum_{t=1}^{n} I_{\{a_m^{-1}X_t\in A\}}e^{-it\lambda}\Big|^2 \\ &= \frac{m_n}{n}\Big(\sum_{t=1}^{n} I_{\{a_m^{-1}X_t\in A\}}\cos(\lambda t)\Big)^2+\frac{m_n}{n}\Big(\sum_{t=1}^{n} I_{\{a_m^{-1}X_t\in A\}}\sin(\lambda t)\Big)^2 \\ &= \alpha_{nA}^2(\lambda)+\beta_{nA}^2(\lambda)\,,\quad \lambda\in\Pi\,, \end{aligned} \tag{1.3.2}$$

where (a_m) is chosen such that $mP(|X|>a_m)\sim 1$, A is bounded away from zero and the sequence (m_n) satisfies $m_n\to\infty$ with $m_n/n\to 0$ as $n\to\infty$. As mentioned above, the sequence $(I_{\{a_m^{-1}X_t\in A\}})$ is strictly stationary for fixed n and therefore we expect that the extremal periodogram shares the asymptotic properties of the periodogram of a stationary process.

Applying the large-small block technique, one can prove joint central limit theory for $\alpha_{nA}(\lambda)$ and $\beta_{nA}(\lambda)$; see Theorem 4.4 in Mikosch and Zhao [19] for details; cf. Theorem 2.4.4 in Chapter 2 of this thesis.

Theorem 1.3.1. *Consider a strictly stationary $\mathbb{R}^d$-valued sequence (X_t) which is regularly varying with index $\alpha>0$. Let $A\subset\overline{\mathbb{R}}_0^d$ be bounded away from zero and $\mu_1(\partial A)=0$. Assume that the mixing and anti-clustering conditions* (M), (M1) *hold and $\sum_{h\ge 1}\gamma_A(h)<\infty$. Consider any fixed frequencies $0<\lambda_1<\cdots<\lambda_N<\pi$ for some $N\ge 1$. Then the following central limit theorem holds:*

$$\big(\alpha_{nA}(\lambda_i),\beta_{nA}(\lambda_i)\big)_{i=1,\ldots,N}\stackrel{d}{\to}\big((\alpha(\lambda_i),\beta(\lambda_i)\big)_{i=1,\ldots,N}\,,\quad n\to\infty\,, \tag{1.3.3}$$

where the limiting vector has $N(\mathbf{0},\Sigma_N)$ distribution with

$$\Sigma_N=\operatorname{diag}\big(f_A(\lambda_1),f_A(\lambda_1),\ldots,f_A(\lambda_N),f_A(\lambda_N)\big)\,.$$

The limit relation (1.3.3) *remains valid if the frequencies λ_i, $i=1,\ldots,N$, are replaced by distinct Fourier frequencies $\omega_i(n)=2\pi i_n/n\to\lambda_i\in(0,\pi)$ for integers $i_n\ge 1$ as $n\to\infty$. The limits λ_i do not have to be distinct.*

As a corollary of Theorem 1.3.1 and as a consequence of the continuous mapping theorem, we can derive the asymptotic distribution of the extremal periodogram ordinates. For any fixed frequencies $0<\lambda_1<\cdots<\lambda_N<\pi$ and $N\ge 1$,

$$\big(I_{nA}(\lambda_i)\big)_{i=1,\ldots,N}\quad\stackrel{d}{\to}\quad\big(f_A(\lambda_i)E_i\big)_{i=1,\ldots,N}\,,\quad n\to\infty\,, \tag{1.3.4}$$

where (E_i) is a sequence of iid standard exponential random variables. Similarly, for any distinct Fourier frequencies $\omega_i(n)\to\lambda_i\in(0,\pi)$ as $n\to\infty$, $i=1,\ldots,N$, where the limits λ_i do not have to be distinct, the following relation hold:

$$\big(I_{nA}(\omega_i(n))\big)_{i=1,\ldots,N}\quad\stackrel{d}{\to}\quad\big(f_A(\lambda_i)E_i\big)_{i=1,\ldots,N}\,,\quad n\to\infty\,.$$

As (1.3.4) shows, the raw periodogram is not a consistent estimator of the spectral density. An intuitive idea is to smooth the raw periodogram by replacing the periodogram ordinates at

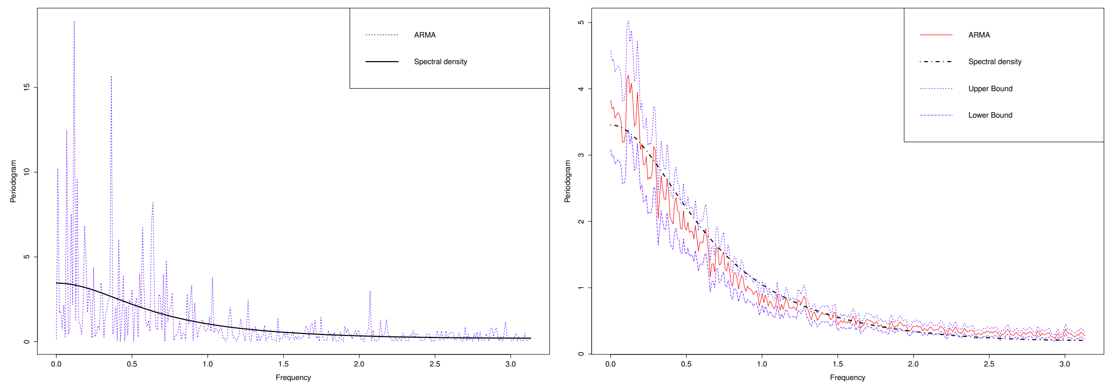

Figure 1.1: Left: The normalized raw periodogram $I_{nA}(\omega_i(n))/\widetilde{\gamma}_A(0)$ of an ARMA(1,1) process $X_t = 0.8\,X_{t-1} + 0.1\,Z_{t-1} + Z_t$ with iid t-distributed noise (Z_t) of degree 3 and the corresponding theoretical spectral density f_A with $A = (1, \infty)$ (solid line). Right: The smoothed periodogram with Daniell window, $s_n = 50$.

Fourier frequencies by a weighted sum of periodgram ordinates at nearby Fourier frequencies. The smoothed periodgram at a frequency λ is given by

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

where λ_0 is the closest Fourier frequency to the frequency λ and $\lambda_j = \lambda_0 + 2j\pi/n$. Consistency of this estimator means that the relations $E\widetilde{f}_{nA}(\lambda) = f_A(\lambda)$ and $\mathrm{var}(\widetilde{f}_{nA}(\lambda)) \to 0$ hold. For this purpose, we assume some conditions on the sequence (s_n) and the weights $(w_n(j))$. In view of the asymptotic properties of the periodogram ordinates, we need $s_n \to 0$ and $s_n/n \to 0$ as $n \to \infty$. These conditions imply that the number of the periodogram ordinates taken into consideration grows with n while the Fourier frequencies λ_j, $|j| \le s_n$, converge to the frequency λ. To ensure that $E\widetilde{f}_{nA}(\lambda) = f_A(\lambda)$, the condition $\sum_{|j| \le s_n} w_n(j) = 1$ is required. We observe that

$$\mathrm{var}(\widetilde{f}_{nA}(\lambda)) = \sum_{|j| \le s_n} w_n^2(j) c_{jj} + \sum_{-s_n \le j_1 \ne j_2 \le s_n} w_n(j_1) w_n(j_2) c_{j_1 j_2}\,,$$

where $c_{j_1 j_2} = \mathrm{cov}(I_{nA}(\lambda_{j_1}), I_{nA}(\lambda_{j_2}))$. Under (M), (M1) and other mild conditions, it is shown in Mikosch and Zhao [19] (cf. Theorem 2.5.1 in Chapter 2) that $c_{j_1 j_2} \to 0$ for $j_1 \ne j_2$ and $c_{jj} = \mathrm{var}(I_{nA}(\lambda_j)) \to f_A^2(\lambda)$ uniformly for $j, j_1, j_2 \in [-s_n, s_n]$, $j_1 \ne j_2$. Therefore, with the condition $\sum_{|j| \le s_n} w_n^2(j) \to 0$ as $n \to \infty$, it is enough to show that $\mathrm{var}(\widetilde{f}_{nA}(\lambda)) \to 0$.

Consistency of the extremal periodogram can also be achieved by reducing the number of trigonometric functions in its definition. Using this approach, consistency was proved in Davis and Mikosch [8]. However, it is more convenient to work with the complete periodogram for various reasons. First, one does not have to choose the number of trigonometric functions involved (this number is a theoretical quantity only) and, second, the complete periodogram can be calculated by using all standard software for the frequency domain.

Of course, the estimator $\widetilde{f}_{nA}(\lambda)$ is determined by the choice of the sequence (s_n) and, more importantly, the choice of the weights $(w_n(j))$. The simplest choice is $w_n(j) = (2s_n + 1)^{-1}$ for $j \in [-s_n, s_n]$, which is called *Daniel window.* As a consequence we also have

$$\widetilde{f}_{nA}(\lambda)/\widetilde{\gamma}_A(0) \stackrel{P}{\to} f_A(\lambda)/\gamma_A(0)\,.$$

The normalized smoothed extremal periodogram $\widetilde{f}_{nA}(\lambda)/\widetilde{\gamma}_A(0)$ gives a satisfactory approximation to the spectral density. This fact is indicated in Figure 1.1.

Comparing the extremal periodogram with the periodogram of a linear process (see (1.1.4)), we find several similarities: the asymptotic distribution of the periodogram ordinates at the Fourier frequencies have a similar form and the smoothed version of the periodogram is a consistent estimator of the spectral density. Unlike the periodogram of a linear process, the proof of the asymptotic properties of the extremal periodogram is mainly based on regular variation, the anti-clustering condition and conditions on mixing rates, such as (1.2.6) and (M1). These conditions are satisfied by numerous classes of processes, including the regularly varying processes introduced in Section 1.2.2. This fact implies wide application of the extremal periodogram.

1.4 Integrated extremal periodogram

1.4.1 Integrated extremal periodogram

In this section we present the main results of Chapter 4. We mentioned in Section 1.1.1 that the integrated periodogram of a stationary sequence has properties similar to to the empirical distribution of an iid sequence and therefore it can be used as the basis for testing the goodness of fit of the underlying second order structure of the stationary process. We will go a similar way for the extremal periodogram which bears some similarities with the periodogram of a stationary sequence. We will study the *integrated extremal periodogram*

$$J_{nA}(g) = \int_\Pi I_{nA}(\lambda) g(\lambda)\, d\lambda = c_0(g)\widetilde{\gamma}_A(0) + 2\sum_{h=1}^{n-1} c_h(g)\widetilde{\gamma}_A(h)\,, \tag{1.4.1}$$

where g is non-negative and square integrable with respect to Lebesgue measure on Π (we write $\gamma \in L^2_+(\Pi)$) with corresponding Fourier coefficients

$$c_h(g) = \int_\Pi \cos(h\lambda) g(\lambda)\, d\lambda\,, \quad h \in \mathbb{Z}\,.$$

The integrated extremal periodogram $J_{nA}(g)$ is a natural estimator of

$$J_A(g) = \int_\Pi f_A(\lambda) g(\lambda)\, d\lambda\,,$$

where f_A is the spectral density defined in (1.3.1). This is confirmed by the following consistency results; see Lemma 4.2.9 in Chapter 4.

Lemma 1.4.1. *Consider an $\mathbb{R}^d$-valued strictly stationary regularly varying sequence (X_t) with index $\alpha > 0$. Assume that the set $A \subset \overline{\mathbb{R}}^d_0$ is bounded away from zero and $\mu_1(\partial A) = 0$. If $\sum_{l=1}^\infty \gamma_A(l) < \infty$ and* (M) *holds then the following asymptotic relations hold for $g \in L^2_+(\Pi)$.*

1. *$EJ_{nA} \to J_A(g)$ as $n \to \infty$.*
2. *If in addition, $m \log^2 n/n = O(1)$ as $n \to \infty$, and there exists a constant $c > 0$ such that*
$$|c_h(g)| \le c/h\,, \quad h \in \mathbb{Z}\,, \tag{1.4.2}$$
then $E(J_{nA}(g) - J_A(g))^2 \to 0$ as $n \to \infty$.

Condition (1.4.2) holds under mild smoothness conditions on g, e.g. if g is Lipschitz or has bounded variation on Π; see Theorem 4.7 on p. 46 and Theorem 4.12 on p. 47 in Zygmund [27].

We are particularly interested in functions of the form $h = gI_{[0,\cdot]}$, where the function g is smooth. Abusing notation, we write for such functions
$$J_{nA}(\lambda) = \int_0^\lambda I_{nA}(x)\, g(x)\, dx\,, \quad \lambda \in \Pi\,,$$
and
$$\psi_h(\lambda) = \int_0^\lambda \cos(h\,x)\, g(x)\, dx\,, \quad \lambda \in \Pi\,,$$
suppressing the dependence on g in the notation.

The following result yields a functional central limit theorem for this integrated periodogram in the space $\mathbb{C}(\Pi)$ of continuous functions on Π; see Theorem 4.3.1 in Chapter 4. It is one of the main results in Chapter 4 of this thesis. This results is another confirmation of the parallel worlds of extremogram and ACVF of a second order stationary process.

Theorem 1.4.2. *Assume that (X_t) is an $\mathbb{R}^d$-valued strictly stationary regularly varying sequence with index $\alpha > 0$ and the Borel set A is bounded away from zero with $\mu_1(\partial A) = 0$. Let g be non-negative β-Hölder continuous function with $\beta \in (3/4, 1]$. If the conditions* (M), (M1) *and $\sum_{l=1}^\infty \gamma_A(l) < \infty$ hold then in $\mathbb{C}(\Pi)$,*
$$(n/m)^{0.5}(J_{nA} - EJ_{nA}) \xrightarrow{d} G\,, \quad n \to \infty\,, \tag{1.4.3}$$
and the limit process is given by the infinite series
$$G = \psi_0 Z_0 + 2\sum_{h=1}^\infty \psi_h Z_h\,,$$
which converges in distribution in $\mathbb{C}(\Pi)$, and (Z_h) is a mean zero Gaussian sequence such that $(Z_0, \dots, Z_h)$ has covariance matrix (Σ_h), $h \ge 0$, given in Lemma 1.2.1.

This functional central limit theorem is pre-asymptotic, like the central limit theorem for the sample extremogram $\widetilde{\gamma}_A$, i.e. in general one cannot center J_{nA} by $\int_0^\cdot f_A(x)\, g(x)\, dx$. This result differs from classical theory for the periodogram of a stationary sequence (see Section 1.1.1): the rate of convergence $(n/m)^{0.5}$ is significantly smaller than the classical $\sqrt{n}$-rate. Moreover, the limit process G has a rather unfamiliar covariance structure which is hardly tractable. In the case of an iid sequence (X_t), $Z_h = 0$ for $h \ge 1$ and (1.4.3) collapses into
$$(n/m)^{0.5}(J_{nA} - EJ_{nA}) \xrightarrow{d} \psi_0 \sqrt{\gamma_A(0)} N$$
for a standard normal random variable N. However, for an iid sequence one can prove the following functional limit theory; see Theorem 4.3.3 in Chapter 4.

Theorem 1.4.3. *Assume that (X_t) is an $\mathbb{R}^d$-valued iid regularly varying sequence with index $\alpha > 0$ for some $\eta \geq 0$ and the Borel set $A \subset \overline{\mathbb{R}}_0^d$ is bounded away from zero, $\mu_1(\partial A) = 0$ and $\mu_1(A) > 0$. Also assume that the limits in (4.2.11) exist. Let g be a non-negative β-Hölder continuous function with $\beta \in (3/4, 1]$. Then the relation*

$$\sqrt{n}\big(J_{nA} - \psi_0 \widetilde{\gamma}_A(0)\big) \quad \stackrel{d}{\to} \quad \overline{G}\,,$$

holds in $\mathbb{C}(\Pi)$, where the limit process is given by the a.s. converging infinite series

$$\overline{G} = 2\sum_{h=1}^{\infty} \psi_{\eta+h}\, Z_h\,,$$

and (Z_h) is a mean zero Gaussian sequence such that $(Z_1, \ldots, Z_h)$ has covariance matrix $\overline{\Sigma}_h$, $h \geq 1$, given in Lemma 4.2.5.

In the case of an iid sequence and $g = 1$, the limiting process is a Brownian bridge and the convergence rate is much faster than in the case of a dependent sequence.

The proof of Theorem 1.4.2 requires two ideas. Lemma 1.2.1 immediately yields that for every fixed $k \geq 1$,

$$\psi_0\left(\widetilde{\gamma}_A(0) - E\widetilde{\gamma}_A(0)\right) + 2\sum_{h=1}^{k} \psi_h\left(\widetilde{\gamma}_A(h) - E\widetilde{\gamma}_A(h)\right) \stackrel{d}{\to} \psi_0 Z_0 + 2\sum_{h=1}^{k} \psi_h Z_h\,,$$

Thus it remains to show that for any $\varepsilon > 0$,

$$\lim_{k\to\infty}\limsup_{n\to\infty} P\big((n/m)^{0.5}\sup_{\lambda\in\Pi}\big|\sum_{h=k+1}^{n-1}\psi_h(\lambda)(\widetilde{\gamma}_A(h) - E\widetilde{\gamma}_A(h))\big| > \varepsilon\big) = 0\,.$$

The proof of the latter relation borrows ideas from Theorem 3.2 in Klüppelberg and Mikosch [14]: the time interval $[k+1, n-1]$ and frequency interval Π are both divided into disjoint subintervals of small size. On these subintervals, one can control the covariances of the increments of the process under consideration and, finally, apply a maximal inequality for stochastic processes provided in Billingsley [1].

An application of the continuous mapping theorem to Theorem 1.4.2 immediately yields the following limit relations.

- *Grenander-Rosenblatt test:*

$$(n/m)^{0.5}\sup_{x\in\Pi}|J_{nA}(x) - EJ_{nA}(x)| \quad \stackrel{d}{\to} \quad \sup_{x\in\Pi}|G(x)|\,. \tag{1.4.4}$$

- ω^2- *or Cramér-von Mises test:*

$$(n/m)\int_{x\in\Pi}\Big(J_{nA}(x) - EJ_{nA}(x)\Big)^2\,dx \stackrel{d}{\to} \int_{x\in\Pi} G^2(x)\,dx\,. \tag{1.4.5}$$

They are the analogs of the limit relations for the integrated periodogram in classical time series analysis; see Section 1.1.1 above. As in classical theory, these limit results can be used to construct asymptotic tests for the hypothesis that the data come from a strictly stationary sequence (X_t) with a given extremogram or extremal spectral density. However, with the exception of an iid sequence (X_t), the asymptotic distribution of the limiting random variables of the functionals in (1.4.4) and (1.4.5) is untractable and therefore one needs to come up with confidence bands in a different way. This is the content of the following subsection.

1.4.2 The bootstrapped integrated extremal periodogram

With a few exceptions, the limit processes G in Theorem 1.4.2 have an unfamiliar dependence structure and then it is impossible to give confidence bands for the test statistics mentioned in the previous section. One faces a similar problem when dealing with the sample extremograms whose asymptotic covariance matrix is a complicated function of the measures μ_h in (1.1.2). Davis et al. [9] proposed to apply the stationary bootstrap for constructing confidence bands for the sample extremogram. The stationary bootstrap can also be used for the integrated periodogram, as will be illustrated below.

The *stationary bootstrap* was introduced by Politis and Romano [21] as an alternative block bootstrap method. First, we describe this procedure for a strictly stationary sequence (Y_t). Given a sample $Y_1, \dots, Y_n$, consider the bootstrapped sequence

$$Y_{K_1}, \dots, Y_{K_1+L_1-1}, \dots, Y_{K_N}, \dots, Y_{K_N+L_N-1}, \dots , \tag{1.4.6}$$

where (Y_i), (K_i), (L_i) are independent sequences, (K_i) is an iid sequence of random variables uniformly distributed on $\{1, \dots, n\}$, (L_i) is an iid sequence of geometrically distributed random variables with distribution $P(L_1 = i) = \theta(1-\theta)^{i-1}$, $i = 1, 2, \dots$, for some $\theta = \theta_n \in (0,1)$ such that $\theta_n \to 0$ as $n \to \infty$ and $N = N_n = \inf\{i \geq 1 : \sum_{j=1}^{i} L_j \geq n\}$. If any element Y_t in (1.4.6) has an index $t > n$, we replace it by $Y_{t \mod n}$. As a matter of fact, $(Y_t)_{t\geq 1}$ constitutes a strictly stationary sequence. The stationary bootstrap sample is now chosen as the block of the first n elements in (1.4.6). In what follows, we write $(Y_{t^*})_{t\geq 1}$ for the bootstrap sequence (1.4.6), indicating that this sequence is nothing but the original Y-sequence sampled at the random indices $(K_1, \dots, K_1 + L_1 - 1, K_2, \dots, K_2 + L_2 - 1, \dots)$ with the convention that indices larger than n are taken modulo n.

In what follows, the probability measure generated by the bootstrap procedure is denoted by P^*, i.e. $P^*(\cdot) = P(\cdot \mid (X_t))$. The corresponding expected value, variance and covariance are denoted by E^*, var^* and cov^*.

We will apply the stationary bootstrap directly to (I_t). Write

$$\overline{I}_n = n^{-1} \sum_{t=1}^{n} I_t \quad \text{and} \quad \widehat{I}_t = I_t - \overline{I}_n \,, \quad t \in \mathbb{Z}\,.$$

and define the corresponding bootstrap sample extremogram

$$\widehat{\gamma}^*_A(h) = \frac{m}{n} \sum_{t=1}^{n-h} \widehat{I}_{t^*} \widehat{I}_{(t+h)^*} \,, \quad h = 0, \dots, n-1\,,$$

and the bootstrap periodogram

$$I^*_{nA}(\lambda) = \frac{m}{n}\Big|\sum_{t=1} \widehat{I}_{t^*}\,\mathrm{e}^{-it\,\lambda}\Big|^2\,, \quad \lambda \in \Pi\,.$$

In the definition of J_{nA} in (1.4.1), we simply replace (I_t) by $(\widehat{I}_{t^*})$, resulting in its bootstrap version

$$J^*_{nA}(\lambda) \quad = \quad \int_0^\lambda I^*_{nA}(x)\,g(x)\,dx = \psi_0\,\widehat{\gamma}^*_A(0) + 2\sum_{h=1}^{n-1}\psi_h\,\widehat{\gamma}^*_A(h)\,, \quad \lambda \in \Pi\,.$$

Now we can formulate a bootstrap analog of Theorem 1.4.2 which shows the consistency of the stationary bootstrap procedure. This theorem is another main result of Chapter 4 ; see Theorem 4.4.2.

Theorem 1.4.4. *Assume the conditions of Theorem 1.4.2 and the following conditions:*

1. $\sum_{h=1}^\infty h\xi_h < \infty$.
2. *The growth conditions* $\theta = \theta_n \to 0$ *and* $n\theta^2/m \to \infty$.
3. *The set* A *is bounded away from zero and a continuity set with respect to* μ_1.

Then

$$d_{P^*}\Big((n/m)^{1/2}\big(J^*_{nA} - E^*J^*_{nA}\big), G\Big) \overset{P}{\to} 0\,, \quad n \to \infty\,,$$

where the Gaussian process G *is defined in Theorem 1.4.2 and* d_{P^*} *is any metric which describes weak convergence in* $\mathbb{C}(\Pi)$ *relative to the probability measure* P^*.

1.5 Directions of future research

Various interesting questions about the extremal periodogram and the integrated extremal periodogram are open. They might become directions of future research. An important question is about the choice of the threshold a_n in the definition of the extremogram (1.2.1), which affects the value of the extremogram, the extremal periodogram and the integrated extremal periodogram. In the data examples of this thesis, the quantity a_n is taken as a high empirical quantile of the absolute values of the sample, considered as a constant. It is more realistic to assume that the threshold a_n is random. This might lead to a different limit theory for the extremogram and correspondingly for the extremal periodgram.

In classical time series analysis, the integrated periodogram can be used for constructing an estimator of the parameters of ARMA process, which is called *Whittle estimator.* So far, a similar estimator based on the integrated extremal periodogram is not available. Therefore, it is also meaningful to develop such estimators for certain classes of heavy-tailed processes. Among these classes are the max-stable processes with Fréchet marginals. The parameters of such processes are typically estimated through a pairwise likelihood estimation procedure. Whittle estimation based on the spectral density of max-stable processes might be an elegant alternative to pairwise likelihood estimation.

In Chapter 4, we already studied the integrated periodogram indexed by a special classes of functions. As in classical time series analysis (e.g. Dahlhaus [5]) the integrated periodogram indexed by classes of functions can be interpreted as an empirical spectral distribution indexed by functions. Whittle estimation is closely related to this topic as well. It may be of interest to prove functional central limit theory for large classes of index functions. The goal will be to cover large classes of interesting statistical functionals as described in Dahlhaus [5].

1.6 Acknowledgments

First of all, I would like to thank my supervisor Prof. Thomas Mikosch, who gave me the chance to work on this project. I am grateful for his patience and confidence in me. Numerous hours of discussions with him were vital for completing this thesis. The last three years' study under his supervision includes a lot of unforgettable experiences, especially the dinners in his apartment. Because of such a great mentor, I have completed the work which I could never imagine when I left China in 2011.

In addition, I would like to thank Prof. Richard A. Davis, especially for his great hospitality during my visit to the Department of Statistics, Columbia University. I had a wonderful time in New York and it has always been a great pleasure to work and discuss with him.

During my work on this project, I am grateful for fruitful discussions with Prof. Gennady Samorodnitsky, Prof. Claudia Klüppelberg, Prof. Paul Embrechts, Prof. Johan Segers, Prof. Robert Stelzer, Prof. Zakhar Kabluchko and Prof. Vicky Fasen. Their comments and advices have been very helpful for the progress of this project.

My research is supported by the Danish Research Council (FNU) 10-084172 "Heavy tail phenomena: Modeling and estimation", to which I owe many thanks. I am also grateful for the excellent working conditions provided by the Department of Mathematical Sciences, University of Copenhagen. I would like to thank the Department of Statistics, Columbia University and RiskLab, ETH Zurich for their support during my visits.

Furthermore, I owe many thanks to my colleagues and friends at the Department of Mathematical Sciences, University of Copenhagen, especially the colleagues from the group for Mathematical and Statistical Methods in Insurance and Economics (MSMIE), for their company and many helpful discussions.

Finally, I would like to express my sincere gratitude to my parents Taiqi Zhao and Shen'e Chen for their love and unconditional support ever since my birth.

Bibliography

[1] Billingsley, P. (1999) *Convergence of Probability Measures.* 2nd Edition. Wiley, New York.

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

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

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

[5] Dahlhaus, R. (1988) Empirical spectral processes and their applications to time series analysis. *Stoch. Proc. Appl.* **30**, 69–83.

[6] Dahlhaus, R. and Polonik, W. (2002) Empirical spectral processes and nonparametric maximum likelihood estimation for time series. In: *Dehling, H.G. Mikosch, T. and Sørensen, M.* (Eds.) (2002) *Empirical Process Techniques for Dependent Data.* Birkhäuser, Boston, pp. 275–298,

[7] 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.

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

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

[10] Davis, R.A., Mikosch, T. and Zhao, Y. (2013) Measures of serial extremal dependence and their estimation. *Stoch. Proc. Appl.*, **123**, 2575–2602.

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

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

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

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

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

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

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

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

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

[20] Mikosch, T. and Zhao, Y. (2013) The integrated periodogram of a dependent extremal event sequence. *Working paper.*

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

[22] Priestley, M.B. (1981) *Spectral Analysis and Time Series.* Academic Press, London, New York.

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

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

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

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

[27] Zygmund, A. (2002) *Trigonometric Series. Vol. I, II.* 3rd Edition Cambridge University Press, Cambridge (UK).

Chapter 2

A Fourier analysis of extreme events

Abstract

The *extremogram* is an asymptotic correlogram for extreme events constructed from a regularly varying stationary sequence. In this paper, we define a frequency domain analog of the correlogram: a periodogram generated from a suitable sequence of indicator functions of rare events. We derive basic properties of the periodogram such as the asymptotic independence at the Fourier frequencies and use this property to show that weighted versions of the periodogram are consistent estimators of a spectral density derived from the extremogram.

2.1 Introduction

In this paper we study an analog of the periodogram for extremal events. In classical time series analysis, the periodogram is a method of moments estimator of the spectral density of a second order stationary time series (X_t); see for example the standard monographs Brillinger [8], Brockwell and Davis [9], Grenander and Rosenblatt [23], Hannan [26], Priestley [42]. The notions of spectral density and periodogram are the respective frequency domain analogs of the autocorrelation function and the sample autocorrelation function in the time domain. In the context of extremal events these notions are not meaningful since second order characteristics are not suited for describing the occurrence of rare events.

However, Davis and Mikosch [15] introduced a time domain analog of the autocorrelation function, the *extremogram* for rare events. For an $\mathbb{R}^d$-valued strictly stationary time series (X_t) and a Borel set A bounded away from zero *the extremogram at lag* $h \geq 0$ is given as the limit

$$\rho_A(h) = \lim_{x \to \infty} P(x^{-1} X_h \in A \mid x^{-1} X_0 \in A)\,. \tag{2.1.1}$$

This definition requires that the support of X (here and in what follows, X denotes a generic element of any stationary sequence (X_t)) is unbounded and, more importantly, that the limit on the right-hand side exists. In general, these limits do not exist. A sufficient condition for their existence is *regular variation* of all pairs (X_0, X_h) or, more generally, *regular variation of the finite-dimensional distributions* of the process (X_t). A precise definition of regular variation will be given in Section 2.2.1. Since A is assumed to be bounded away from zero the probabilities $P(x^{-1}X \in A)$ converge to zero as $x \to \infty$. Then the following calculation is straightforward for A:

$$\begin{aligned}\lim_{x\to\infty} \operatorname{corr}(I_{\{x^{-1}X_0\in A\}}, I_{\{x^{-1}X_h\in A\}}) &= \lim_{x\to\infty} \frac{P(x^{-1}X_0 \in A\,, x^{-1}X_h \in A) - [P(x^{-1}X \in A)]^2}{P(x^{-1}X \in A)(1 - P(x^{-1}X \in A))} \\ &= \lim_{x\to\infty} P(x^{-1}X_h \in A \mid x^{-1}X_0 \in A) = \rho_A(h)\,.\end{aligned}$$

For fixed x, $(I_{\{x^{-1}X_t\in A\}})_{t\in\mathbb{Z}}$ constitutes a strictly stationary sequence. The limit sequence $(\rho_A(h))$ inherits the property of correlation function from $(\operatorname{corr}(I_{\{x^{-1}X_0\in A\}}, I_{\{x^{-1}X_h\in A\}}))$. Therefore, in an asymptotic sense, one can use the notions of classical time series analysis (such as the autocorrelation function) for the sequences of indicator functions $(I_{\{x^{-1}X_t\in A\}})_{t\in\mathbb{Z}}$. Of course, there are several crucial differences to classical time series analysis.

- The notion of autocorrelation function is only defined in an asymptotic sense.
- The strictly stationary sequence of indicator functions $(I_{\{x^{-1}X_t\in A\}})_{t\in\mathbb{Z}}$ depends on the threshold x, i.e., we are dealing with an array of strictly stationary processes.
- By definition, the values $\rho_A(h)$ cannot be negative.

Davis and Mikosch [15, 16] introduced the extremogram and calculated the extremogram for various standard regularly varying time series models such as the GARCH model, stochastic volatility and linear processes with regularly varying noise, and infinite variance stable processes; see also Section 2.3. They studied the basic asymptotic properties of the extremogram (consistency, asymptotic normality) and also introduced a frequency domain analog of the correlation function ρ_A given as the Fourier series

$$f_A(\lambda) = \sum_{h\in\mathbb{Z}} \rho_A(h)\, \mathrm{e}^{-i\,h\,\lambda}\,, \quad \lambda \in [0, \pi]\,. \tag{2.1.2}$$

A natural estimator of $f_A(\lambda)$ is found by replacing the correlations $\rho_A(h)$ by sample analogs. The convergence in the mean square sense of such an analog of the classical periodogram estimator towards the spectral density $f_A(\lambda)$ at a fixed frequency λ was shown in [15]. However, the periodogram of $(I_{\{x^{-1}X_t\in A\}})_{t\in\mathbb{Z}}$ used in [15] had to be truncated to achieve consistency; the truncation level depended on some mixing rate which is unknown for real-life data. In this paper, we overcome this inconvenience. In addition, we study the periodogram ordinates of the indicator functions at finitely many frequencies. We show that the limiting vector of the periodogram ordinates at distinct fixed or Fourier frequencies converges in distribution to a vector of independent exponential random variables. This property parallels the asymptotic theory for the periodogram of a second order stationary sequence; see e.g. Brockwell and Davis [9], Chapter 10.

In classical time series analysis, the asymptotic independence of the periodogram at distinct frequencies is the theoretical basis for consistent estimation of the spectral density via weighted averages or kernel based methods. We show that weighted average estimators of the periodogram evaluated at Fourier frequencies in the neighborhood of a fixed non-zero frequency are consistent estimators of the limiting spectral density.

The paper is organized as follows. In Section 2.2 we introduce basic notions and conditions used throughout this paper. In Section 2.2.1 we define regular variation of a strictly stationary sequence. In Section 2.2.2 we consider those mixing conditions which are relevant for the results of this paper. The periodogram of extreme events is introduced in Section 2.2.3. In Section 2.3 we discuss some regularly varying strictly stationary sequences. Among them are linear, stochastic volatility and max-moving average processes with regularly varying noise. We give expressions for the extremogram and, if possible, for the corresponding spectral density. In Section 2.4 we give the main results of this paper. We start in Section 2.4.1 by showing that the periodogram ordinates of extreme events are asymptotically uncorrelated at distinct fixed or Fourier frequencies in the interval $(0, \pi)$. Next, in Section 2.4.2 we show that the periodogram ordinates at distinct fixed or Fourier frequencies converge to independent exponential random variables. This property is exploited in Section 2.5 to show that weighted averages of periodogram ordinates evaluated at Fourier frequencies in a small neighborhood of a fixed frequency yield consistent estimates of the underlying spectral density at the given frequency. In Section 2.6 we give a short discussion of work related to the extremogram or the spectral analysis of sequences of indicator functions. The proofs depend on various calculations involving formulas for sums of trigonometric functions. Some of these formulas and related calculations are given in the Appendix.

2.2 Preliminaries

2.2.1 Regular variation

It was mentioned in Section 2.1 that one needs conditions to ensure that the limits $\rho_A(h)$ in (2.1.1) exist. A sufficient condition for this to hold is *regular variation* of the strictly stationary sequence (X_t). Regular variation is a convenient tool for modeling multivariate heavy-tail phenomena and serial extremal dependence in a time series; see Resnick's monographs [44, 45], Resnick [43], Basrak and Segers [5, 4], Davis and Hsing [11], Embrechts et al. [20], Jakubowski [30, 31], Bartkiewicz et al. [2], and the references therein. Regular variation is particularly useful for modeling extremes in financial time series; see Basrak et al. [3], Mikosch and Stărică [39], Davis and Mikosch [12, 13, 14] ; cf. Andersen et al. [1] and the references therein. See also the examples in Section 2.3.

A random vector X with values in $\mathbb{R}^d$ for some $d \geq 1$ is *regularly varying* if there exists a non-null Radon measure μ on the Borel σ-field of $\overline{\mathbb{R}}_0^d = \overline{\mathbb{R}}^d \setminus \{\mathbf{0}\}$, where $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$, such that

$$\frac{P(x^{-1}X \in \cdot)}{P(|X| > x)} \xrightarrow{v} \mu(\cdot)\,, \quad x \to \infty\,. \tag{2.2.1}$$

Here $\xrightarrow{v}$ denotes *vague convergence* on the Borel σ-field of $\overline{\mathbb{R}}_0^d$; for definitions see Kallenberg [33], Resnick [44, 43]. In this context, bounded sets are those which are bounded away from zero and the Radon measure μ charges finite mass to these sets. Then, necessarily, there exists an $\alpha \geq 0$ such that $\mu(tA) = t^{-\alpha}\mu(A)$, $t > 0$, for all A in the Borel σ-field of $\overline{\mathbb{R}}_0^d$. We refer to *regular* *variation of X with limiting measure μ and index α.* A multivariate t-distributed random vector is regularly varying and the index α is the degree of freedom. Other well known multivariate regularly varying distributions are the multivariate F- and Fréchet distributions; see Resnick [44], Chapter 5, in particular Section 5.4.2.

We will often use an equivalent sequential version of (2.2.1): there exists (a_n) such that $a_n \to \infty$ as $n \to \infty$ and

$$n\, P(a_n^{-1} X \in \cdot) \stackrel{v}{\to} \mu(\cdot)\,, \quad n \to \infty\,. \tag{2.2.2}$$

A possible choice of (a_n) is given by the $(1-1/n)$-quantile of $|X|$.

Now, a strictly stationary d-dimensional sequence (X_t) is *regularly varying* if the lagged vectors $Y_h = \text{vec}(X_0,\dots,X_h)$, $h \ge 0$, are regularly varying with index α. Of course, the limiting non-null Radon measures μ_h in (2.2.1) now depend on the lag h and the normalization in (2.2.2) would also change with h. In the context of this paper it is convenient to choose the normalizations of the rare event probabilities independently of h. In particular, we will use the following relations for $h \ge 0$,

$$\begin{aligned} \frac{P(x^{-1}Y_h \in \cdot)}{P(|X_0| > x)} &\stackrel{v}{\to} \mu_h(\cdot)\,, \quad x \to \infty\,. \\ n\, P(a_n^{-1} Y_h \in \cdot) &\stackrel{v}{\to} \mu_h(\cdot)\,, \quad n \to \infty\,, \end{aligned}$$

where (a_n) satisfies $nP(|X_0| > a_n) \to 1$, as $n \to \infty$. These relations are equivalent to the definitions (2.2.1) and (2.2.2) of regular variation of Y_h.

Now we are in the position to verify that the limits $\rho_A(h)$ in (2.1.1) exist for any Borel set $A \subset \overline{\mathbb{R}}_0^d$ bounded away from zero. Write $\widetilde{A} = A \times \overline{\mathbb{R}}_0^{dh}$ and $\widetilde{B} = A \times \overline{\mathbb{R}}_0^{d(h-1)} \times A$. These sets are bounded away from zero in $\overline{\mathbb{R}}_0^{d(h+1)}$. If these sets are continuity sets with respect to μ_h we obtain from the sequential definition of regular variation of Y_h for $h \ge 0$,

$$\begin{aligned} \rho_A(h) &= \lim_{n\to\infty} P(a_n^{-1} X_h \in A \mid a_n^{-1} X_0 \in A) \\ &= \lim_{n\to\infty} \frac{n\, P(a_n^{-1} Y_h \in \widetilde{B})}{n\, P(a_n^{-1} Y_h \in \widetilde{A})} = \frac{\mu_h(\widetilde{B})}{\mu_h(\widetilde{A})}\,. \end{aligned}$$

2.2.2 The mixing and dependence conditions (M), (M1) and (M2)

The results in Davis and Mikosch [15, 16] were proved under the following mixing/dependence condition on the sequence (X_t).

(M) The sequence (X_t) is strongly mixing with rate function (ξ_t). There exist $m = m_n \to \infty$ and $r_n \to \infty$ such that $m_n/n \to 0$ and $r_n/m_n \to 0$ and

$$\lim_{n\to\infty} m_n \sum_{h=r_n}^{\infty} \xi_h = 0\,, \tag{2.2.3}$$

and for all $\epsilon > 0$,

$$\lim_{k\to\infty} \limsup_{n\to\infty} m_n \sum_{h=k}^{r_n} P(|X_h| > \epsilon\, a_m\,, |X_0| > \epsilon\, a_m) = 0\,. \tag{2.2.4}$$

Condition (2.2.4) is similar in spirit to condition (2.8) used in Davis and Hsing [11] for establishing convergence of a sequence of point processes to a limiting cluster point process. It is much weaker than the anti-clustering condition $D'(\epsilon a_n)$ of Leadbetter which is well known in the extreme value literature; see Leadbetter et al. [34] or Embrechts et al. [20]. Since we choose (a_n) such that $n\,P(|X|>a_n)\to 1$ as $n\to\infty$, (2.2.4) is equivalent to

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

In addition, we also need the following technical condition, using the same notation as in (M).

(M1) The sequences (m_n), (r_n), $k_n=[n/m_n]$ from (M) also satisfy the growth conditions $k_n\xi_{r_n}\to 0$, and $m_n=o(n^{1/3})$.

Remark 2.2.1. Some of the examples in Section 2.3 are strongly mixing with geometric rate, i.e. there exists $a\in(0,1)$ such that $\xi_h\le a^h$ for sufficiently large h. Then (2.2.3) is satisfied if $m_n a^{r_n}=o(1)$. If $m_n=n^\gamma$ for some $\gamma\in(0,1)$ then (2.2.3) is satisfied for $r_n=c\log n$ if c is chosen sufficiently large and (M1) trivially holds as well. If $\xi_h\le h^{-s}$ for some $s>1$ and sufficiently large h then (2.2.3) is satisfied if $m_n r_n^{-s+1}=o(1)$. Thus, if $m_n=n^\gamma$ for some $\gamma\in(0,1)$ and $r_n=n^\delta$ for some $\delta\in(\gamma/(s-1),\gamma)$, some $s>2$, then (2.2.3) holds. Condition (M1) is satisfied if $(1+s)^{-1}<\gamma<1/3$ and $\delta\in((1-\gamma)/s,\gamma)$. Thus (2.2.3) and (M1) are always satisfied if s can be chosen arbitrarily large.

For our main result on the smoothed periodogram (see Theorem 2.5.1) we finally need the condition:

(M2) The sequences (m_n), (r_n) from (M) also satisfy the growth conditions

$$m_n^2 n\sum_{h=r_n+1}^{n}\xi_h\to 0,\quad m_n r_n^3/n\to 0\,.$$

Remark 2.2.2. Condition (M2) is stronger than (2.2.3). If (X_t) is strongly mixing with geometric or polynomial rate, a similar argument as in Remark 2.2.1 shows that (M2) holds for suitable choices of (r_n) and (m_n).

2.2.3 The periodogram of extreme events

In this section we recall some of the results from Davis and Mikosch [15] concerning the estimation of the spectral density f_A defined in (2.1.2). Write

$$I_t=I_{\{X_t/a_m\in A\}}\,,\quad \widetilde{I}_t=I_t-p_0\,,\quad p_0=EI_t=P(a_m^{-1}X\in A)\,,\quad t=1,\ldots,n\,,$$

for some sequence $m=m_n\to\infty$ such that $m_n/n\to 0$ as in condition (M) above. We suppress the dependence of I_t on A and a_m. We introduce the estimators

$$I_{nA}(\lambda)=\frac{m_n}{n}\Big|\sum_{t=1}^{n}\widetilde{I}_t\,\mathrm{e}^{-i\,t\,\lambda}\Big|^2,\quad \lambda\in[0,\pi]\,,\quad\text{and}\quad \widehat{P}_m(A)=\frac{m_n}{n}\sum_{t=1}^{n}I_t\,. \tag{2.2.5}$$

It follows from Theorem 3.1 in [15] that

$$\widehat{P}_m(A) = \frac{m_n}{n} \sum_{t=1}^{n} I_t \stackrel{L^2}{\to} \mu_0(A) = \lim_{n\to\infty} m_n \, P(a_m^{-1} X \in A)\,, \tag{2.2.6}$$

provided A is a continuity set with respect to the limiting measure μ_0. The conditions $m_n \to \infty$ and $m_n/n \to 0$ cannot be avoided since we need that $E\widehat{P}_m(A) = m_n P(a_m^{-1} X \in A) \to \mu_0(A)$ and then we also get $\mathrm{var}(\widehat{P}_m(A)) = O(m_n/n)$.

Davis and Mikosch [15], Theorem 5.1, also proved that the *lag-window estimator* or *truncated periodogram*

$$\widehat{f}_{nA}(\lambda) = \widetilde{\gamma}_n(0) + 2 \sum_{h=1}^{r_n} \cos(\lambda h) \widetilde{\gamma}_n(h) \tag{2.2.7}$$

with $\widetilde{\gamma}_n(0) = (m/n) \sum_{t=1}^{n} I_t$ and $\widetilde{\gamma}_n(h) = (m/n) \sum_{t=1}^{n-h} \widetilde{I}_t \widetilde{I}_{t+h}$, $h > 0$, for fixed $\lambda \in (0, \pi)$, satisfies the relations

$$E\widehat{f}_{nA}(\lambda) \to \mu_0(A) f_A(\lambda) \quad \text{and} \quad E\big(\widehat{f}_{nA}(\lambda) - \mu_0(A) f_A(\lambda)\big)^2 \to 0\,. \tag{2.2.8}$$

under condition (M), if A is a μ_0-continuity set and the sets $A \times \overline{\mathbb{R}}_0^{k-1} \times A$ are continuity sets with respect to μ_k, $k \ge 1$, and $m_n r_n^2 = O(n)$. If we combine (2.2.6) and (2.2.8) we have for fixed $\lambda \in (0, \pi)$,

$$\frac{\widehat{f}_{nA}(\lambda)}{\widehat{P}_m(A)} \stackrel{P}{\to} f_A(\lambda)\,. \tag{2.2.9}$$

A natural self-normalized estimator of the spectral density $f_A(\lambda)$ in (2.1.2) is the following analog of the periodogram

$$\widetilde{I}_{nA}(\lambda) = \frac{I_{nA}(\lambda)}{\widehat{P}_m(A)} = \frac{\Big| \sum_{t=1}^{n} \widetilde{I}_t \, \mathrm{e}^{-it\lambda} \Big|^2}{\sum_{t=1}^{n} I_t}\,, \quad \lambda \in [0, \pi]\,,$$

In contrast to $\widehat{f}_{nA}(\lambda)$ one does not need to know the quantities m_n and r_n which appear in the definition of $\widehat{f}_{nA}(\lambda)$ and are hard to determine for practical estimation purposes. We call $\widetilde{I}_{nA}(\lambda)$ the *standardized periodogram*. However, we know from theory for the classical periodogram of the stationary process (X_t), given by

$$J_{n,X}(\lambda) = n^{-1} \Big| \sum_{t=1}^{n} X_t \, \mathrm{e}^{-it\lambda} \Big|^2\,, \quad \lambda \in [0, \pi]\,,$$

that $J_{n,X}(\lambda)$ is *not* a consistent estimator of the spectral density $f_X(\lambda)$ of the process (X_t) even in the case when (X_t) is iid and has finite variance; see e.g. Proposition 10.3.2 in Brockwell and Davis [9]. To achieve consistent estimation of $f_X(\lambda)$ one needs to truncate the periodogram, similarly to $\widehat{f}_{nA}(\lambda)$, or to apply smoothing techniques to neighboring periodogram ordinates. A similar observation applies to the periodogram for extremal events, $I_{n,A}(\lambda)$; see Section 2.4.

2.3 Examples

In this section we collect some examples of regularly varying stationary time series models, give their extremograms (2.1.1) and, if possible, give an explicit expression of the corresponding spectral density (2.1.2). However, in general, the extremogram is too complicated and one cannot calculate the Fourier series (2.1.2). Some of the examples below are taken from Davis and Mikosch [15].

2.3.1 IID sequence

Consider an iid real-valued sequence (Z_t) such that

$$P(Z > x) \sim p\, x^{-\alpha}\, L(x) \quad \text{and} \quad P(Z \le -x) \sim q\, x^{-\alpha}\, L(x)\,, \quad x \to \infty\,, \tag{2.3.1}$$

where $\alpha > 0$, $p, q \ge 0$, $p+q = 1$ and L is a slowly varying function. It is well known (e.g. Resnick [43, 44]) that (Z_t) is regularly varying with index α. The limiting measures μ_h are concentrated on the axes:

$$\mu_h(dx_0, \dots, dx_h) = \sum_{i=0}^{h} \lambda_\alpha(dx_i) \prod_{i \ne j} \varepsilon_0\, dx_j\,,$$

where ε_y denotes Dirac measure at y, $\lambda_\alpha(x, \infty] = px^{-\alpha}$, $\lambda_\alpha[-\infty, -x] = qx^{-\alpha}$, $x > 0$. Then for any A bounded away from zero,

$$\rho_A(h) = 0\,, \quad h \ge 1\,, \quad \text{and} \quad f_A \equiv 1\,.$$

The conditions (M), (M1) and (M2) are trivially satisfied in this case.

2.3.2 Stochastic volatility model

Let (σ_t) be a strictly stationary sequence of non-negative random variables with $E\sigma^{\alpha+\delta} < \infty$ for some $\delta > 0$, independent of the iid regularly varying sequence (Z_t) with index $\alpha > 0$, satisfying the tail balance condition (2.3.1). The process

$$X_t = \sigma_t\, Z_t\,, \quad t \in \mathbb{Z}\,,$$

is a *stochastic volatility process*. It is a regularly varying sequence with index α and limiting measures concentrated on the axes. The extremogram and the spectral density coincide with these quantities in the iid case; see Davis and Mikosch [12]. As discussed in Davis and Mikosch [14], the process (X_t) inherits the strong mixing property and the same rate function from the volatility process (σ_t). In particular, if (σ_t) is strongly mixing with *geometric rate*, (X_t) is also strongly mixing with geometric rate, and then the conditions (2.2.3), (M1) and (M2) are satisfied; see Remarks 2.2.1 and 2.2.2. Condition (2.2.4) also holds if $E\sigma^{4\alpha} < \infty$; see Davis and Mikosch [15].

The situation of a vanishing ρ_A is rather incomplete information about tail dependence. Hill [28] proposed to use an alternative lag-wise dependence measure of the form $\lim_{x\to\infty} P(X_h > x, X_0 > x)/[P(X_0 > x)]^2 - 1$ which in general does not vanish. This measure is in agreement with the asymptotic tail independence conditions of Ledford and Tawn [35].

The mentioned literature [12, 14] focuses on stochastic volatility processes with iid regularly varying noise (Z_t) with index α and stochastic volatility satisfying the moment condition $E\sigma^{\alpha+\delta} < \infty$ for some $\delta > 0$. Mikosch and Rezapur [37] consider regularly varying stochastic volatility processes with index α when the sequence (σ_t) is regularly varying with index α, $E|Z|^{\alpha+\delta} < \infty$ for some $\delta > 0$ and they give examples with $\rho_A \neq 0$ and $f_A \not\equiv 1$ for A bounded away from zero. The aforementioned comments about mixing also apply in this setting.

2.3.3 ARMA process

Consider the linear process

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}\,, \quad t \in \mathbb{Z}\,, \tag{2.3.2}$$

where (Z_t) is an iid one-dimensional regularly varying sequence with index $\alpha > 0$ and tail balance condition (2.3.1). We choose the coefficients from the ARMA equation $\psi(z) = 1 + \sum_{i=1}^{\infty} \psi_i z^i = \theta(z)/\phi(z)$, $z \in \mathbb{C}$, where

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_r z^r \quad \text{and} \quad \theta(z) = 1 + \theta_1 z + \cdots + \theta_s z^s\,,$$

for integers $r, s \geq 0$, and the coefficients θ_i, ϕ_i are chosen such that $\phi(z)$ and $\theta(z)$ have no common zeros and $\phi(z) \neq 0$ for $|z| \leq 1$. It is well known that X is regularly varying with index α; see e.g. Appendix A3.3 in Embrechts et al. [20] or Mikosch and Samorodnitsky [38]. The proofs in the latter references use the fact that $X_t^{(s)} = \sum_{j=0}^{s} \psi_j Z_{t-j}$, $s \geq 1$, is regularly varying as a simple consequence of the fact that linear combinations of iid regularly varying random variables are regularly varying; see Feller [22], p. 278; cf. Lemma 1.3.1 in [20]. Moreover,

$$\lim_{s\to\infty} \limsup_{n\to\infty} nP(a_n^{-1}|X_t - X_t^{(s)}| > \varepsilon) = 0\,, \quad \varepsilon > 0\,. \tag{2.3.3}$$

Then it follows from Lemma 3.6 in Jessen and Mikosch [32] that X_t is regularly varying.

The vector $(X_0^{(s)}, \ldots, X_h^{(s)})$ is also regularly varying with index α. This fact follows from an application of a multivariate version of Breiman's lemma [7] (see Basrak et al. [3]) or the fact that linear operations preserve regular variation; see Lemma 4.6 in [32]. Since (2.3.3) holds a straightforward multivariate extension of Lemma 3.6 in [32] yields that $(X_0, \ldots, X_h)$ is regularly varying for every $h \geq 0$.

The same arguments leading to the asymptotic tail behavior of X_t (see e.g. Appendix A3.3 in Embrechts et al. [20], Mikosch and Samorodnitsky [38]) yield for $A = (1, \infty)$,

$$\rho_A(h) = \frac{\sum_{i=0}^{\infty} \Big[p\,(\min(\psi_i^+, \psi_{i+h}^+))^\alpha + q\,(\min(\psi_i^-, \psi_{i+h}^-))^\alpha\Big]}{\sum_{i=0}^{\infty} \Big[p\,(\psi_i^+)^\alpha + q(\psi_i^-)^\alpha\Big]}\,, \quad h \geq 1\,. \tag{2.3.4}$$

This formula was given in [15] for symmetric Z when $p = q = 0.5$.

Doukhan [19], Theorem 6 on p. 99, shows that (X_t) is β-mixing, hence strongly mixing, with geometric rate if Z has a positive Lebesgue density in some neighborhood of the expected value of Z (provided it exists) and Pham and Tran [41] proved the same statement under the condition that Z has a Lebesgue density and a finite pth moment for some $p > 0$. Hence (2.2.3), (M1) and (M2) are satisfied under these conditions; see Remarks 2.2.1 and 2.2.2. Next we verify condition (2.2.4). We observe that it trivially holds for an s-dependent sequence for any integer $s \geq 1$. Hence it is satisfied for any moving average of order s, in particular for the truncated sequence $(X_t^{(s)})$. For ease of presentation, we assume $\epsilon = 1$. Since $X_h^{(h-1)}$ and X_0 are independent we have

$$\begin{aligned} P(|X_h| > a_m \mid |X_0| > a_m) &\leq P(|X_h^{(h-1)}| > 0.5\, a_m) + P(|X_h - X_h^{(h-1)}| > 0.5\, a_m, \mid |X_0| > a_m) \\ &\leq I_1 + I_2\,. \end{aligned}$$

Recall that there exist $\varphi \in (0,1)$ such that $|\psi_i| \leq \varphi^i$ for i sufficiently large; see Brockwell and Davis [9], Chapter 3. We have for a positive constant $c > 0$, for every $k \geq 1$,

$$\begin{aligned} \sum_{h=k+1}^{r_n} I_1 &\leq r_n P\Big(\sum_{i=0}^{\infty} |\psi_i||Z_i| > 0.5\, a_m\Big) \\ &\sim c\, r_n P(|Z| > a_m) = o(1) \quad \text{as } n \to \infty. \end{aligned}$$

(Here and in what follows, c denotes any constant whose value is not of interest.) For sufficiently large k, we have in view of the uniform convergence theorem for regularly varying functions (see Bingham et al. [6], Section 1.2),

$$\begin{aligned} \sum_{h=k+1}^{r_n} I_2 &\leq c\, m_n \sum_{h=k+1}^{r_n} P\Big(\sum_{i=h+1}^{\infty} |\psi_i||Z_i| > 0.5\, a_m\Big) \\ &\leq c\, m_n \sum_{h=k+1}^{r_n} P\Big(\varphi^h \sum_{i=0}^{\infty} \varphi^i |Z_i| > 0.5\, a_m\Big) \\ &\leq c \sum_{h=k+1}^{r_n} \varphi^{\alpha h} \leq c\, \varphi^{\alpha(k+1)}/(1-\varphi^{\alpha})\,, \end{aligned}$$

and the right-hand side converges to zero as $k \to \infty$. Thus we proved that (M), (M1) and (M2) hold for ARMA processes if the noise has some Lebesgue density.

If $\mathrm{var}(X) < \infty$ relation (2.3.4) bears some similarity with the autocorrelation function of (X_t) given by $\rho(h) = \sum_{i=1}^{\infty} \psi_i \psi_{i+h} / \sum_{i=1}^{\infty} \psi_i^2$. Replacing ρ_A in (2.1.2) by ρ, one obtains the well-known spectral density of a causal ARMA process (up to a constant multiple): $f_X(\lambda) = (2\pi)^{-1} |\theta(\mathrm{e}^{-i\lambda})|^2 / |\phi(\mathrm{e}^{-i\lambda})|^2$, $\lambda \in [0,\pi]$. Such a compact formula can in general not be derived for f_A. An exception is a causal ARMA(1,1) process; see Section 2.7.2. There are various analogies between the functions ρ and ρ_A for causal invertible ARMA processes. In this case, $\psi_h \to 0$ as $h \to \infty$ at an exponential rate and therefore both $\rho(h)$ and $\rho_A(h)$ decay exponentially fast to zero as well. The latter property also makes the spectral densities f_X and f_A analytical functions bounded away from infinity. We also mention that for an MA(q) process, $\rho(h) = \rho_A(h) = 0$ for $h > q$.

2.3.4 Max-moving averages

Consider a regularly varying iid sequence (Z_t) with index $\alpha > 0$ and tail balance parameters p, q; see (2.3.1). For a real-valued sequence (ψ_j), the process

$$X_t = \bigvee_{i=0}^{\infty} \psi_i Z_{t-i}, \quad t \in \mathbb{Z}, \tag{2.3.5}$$

is a *max-moving average*. We will also assume that $|\psi_j| \le c$, $j \ge 0$, for some constant c and $\psi_0 = 1$. Obviously, if X is finite a.s., (X_t) constitutes a strictly stationary process. The random variable X does not assume the value ∞ if $\lim_{x\to\infty} P(X > x) = 0$. We have

$$P(X > x) \quad = \quad P\Big(\bigvee_{i=0}^{\infty} \psi_i Z_i > x\Big) = 1 - \lim_{n\to\infty} \prod_{i=0}^{n} P(\psi_i Z \le x)\,.$$

The product $\prod_{i=0}^{\infty} P(\psi_i Z \le x)$ converges if $\sum_{i=0}^{\infty} P(\psi_i Z > x) < \infty$. By regular variation of Z, this amounts to the condition

$$\psi_+ = \sum_{i=0}^{\infty} [p\,(\psi_i^+)^\alpha + q\,(\psi_i^-)^\alpha] < \infty\,.$$

A Taylor expansion and regular variation of Z yield

$$P(X > x) = 1 - \mathrm{e}^{\,-(1+o(1))P(|Z|>x)\psi_+} \sim P(|Z| > x)\psi_+ \to 0\,, \quad x \to \infty\,. \tag{2.3.6}$$

We also have $P(X \le -x) = O(P(|Z| > x))$. Hence X is regularly varying with index α if $0 < \psi_+ < \infty$. We always assume the latter condition.

We show that (X_t) is regularly varying. Consider the truncated max-moving average process for $s \ge 0$,

$$X_t^{(s)} = \bigvee_{i=0}^{s} \psi_i Z_{t-i}\,, \quad t \in \mathbb{Z}\,.$$

Regular variation of $(X_0^{(s)}, \ldots, X_h^{(s)})$ is a consequence of regular variation of (Z_t) and the fact that regular variation is preserved under the max-operation acting on independent components. Moreover,

$$\lim_{s\to\infty} \limsup_{n\to\infty} n\,P\Big(a_n^{-1} \bigvee_{i=s+1}^{\infty} \psi_i Z_{t-i} > x\Big) \quad = \quad c \lim_{s\to\infty} \sum_{i=s+1}^{\infty} [p\,(\psi_i^+)^\alpha + q\,(\psi_i^-)^\alpha] = 0\,.$$

Then an application of Lemma 3.6 in Jessen and Mikosch [32] shows that $(X_0, \ldots, X_h)$ is regularly varying with index α for every $h \ge 0$.

Next we determine the extremogram ρ_A corresponding to the set $A = (1, \infty)$. For $h \ge 1$, we have

$$\begin{aligned} P(X_h > x, X_0 > x) &= P\Big(\bigvee_{i=0}^{\infty} \psi_i Z_{-i} > x\,,\ \bigvee_{i=-h}^{-1} \psi_{i+h} Z_{-i} \vee \bigvee_{i=0}^{\infty} \psi_{i+h} Z_{-i} > x\Big) \\ &= P\Big(\bigvee_{i=0}^{\infty} (\psi_i Z_{-i}) \wedge (\psi_{i+h} Z_{-i}) > x\Big) + o(P(|Z| > x)) \\ &\sim P(|Z| > x) \sum_{i=0}^{\infty} \Big[p\,(\min(\psi_i^+, \psi_{i+h}^+))^\alpha + q\,(\min(\psi_i^-, \psi_{i+h}^-))^\alpha\Big]\,. \end{aligned}$$

Finally, in view of (2.3.6), $\rho_A(h)$ is given by (2.3.4), i.e., the linear process (2.3.2) and the max-moving average (2.3.5) have the same extremogram provided the coefficients (ψ_j) and the distribution of Z are the same. Hence their spectral densities f_A are the same as well.

As for ARMA processes, mixing conditions for infinite max-moving processes are not easily verified and additional conditions on the noise (Z_t) are needed. Assume that (Z_t) is iid with common Fréchet distribution $\Psi_\alpha(x) = \mathrm{e}^{-x^{-\alpha}}$, $x > 0$, for some $\alpha > 0$. Then (X_t) constitutes a stationary max-stable process. For such processes, Dombry and Eyi-Minko [18] proved rather general sufficient conditions for β-mixing, implying strong mixing. An application of their Corollary 2.2 implies that the condition $|\psi_h| \le c_0\,\mathrm{e}^{-c_1 h}$, $h \ge 1$, for suitable constants $c_1, c_2 > 0$ implies strong mixing of (X_t) with geometric rate function (ξ_h). In this situation, (M), (M1) and (M2) are satisfied.

2.4 Basic properties of the periodogram

In this section we study some basic properties of the periodogram $I_{nA}(\lambda)$ for extremal events defined in (2.2.5). Notice that

$$I_{nA}(\lambda) = \frac{1}{2}\Big[(\alpha_n(\lambda))^2 + (\beta_n(\lambda))^2\Big]$$

where $\alpha_n(\lambda)$ and $\beta_n(\lambda)$ denote the normalized and centered cosine and sine transforms of $(I_t)_{t=1,\dots,n}$:

$$\begin{aligned} \alpha_n(\lambda) &= \Big(\frac{2m_n}{n}\Big)^{1/2} \sum_{t=1}^{n} \widetilde{I}_t \cos(\lambda t)\,, \\ \beta_n(\lambda) &= \Big(\frac{2m_n}{n}\Big)^{1/2} \sum_{t=1}^{n} \widetilde{I}_t \sin(\lambda t)\,. \end{aligned}$$

Here we suppress the dependence of α_n and β_n on a_m and the set A which is bounded away from zero. For practical purposes, the periodogram will typically be evaluated at some Fourier frequencies $\lambda = 2\pi j/n$ for some integer j. If $\lambda \in (0, \pi)$ is such a *Fourier frequency,* then

$$\sum_{t=1}^{n} \mathrm{e}^{\,i\,\lambda\,t} = 0\,,$$

and therefore the I_t's in $\alpha_n(\lambda)$ and $\beta_n(\lambda)$ are automatically centered by their (in general unknown) expectations $EI_t = p_0 = P(a_m^{-1} X \in A)$.

2.4.1 The periodogram ordinates at distinct frequencies are asymptotically uncorrelated

Our first result is an analog of the fact that the sine and cosine transforms of a stationary sequence at distinct fixed or Fourier frequencies in $(0,\pi)$ are asymptotically uncorrelated.

Proposition 2.4.1. *Consider a strictly stationary $\mathbb{R}^d$-valued sequence (X_t) which is regularly varying with index $\alpha>0$ and satisfies the mixing condition* (M). *Let $A\subset\overline{\mathbb{R}}_0^d$ be bounded away from zero such that A is a continuity set with respect to μ_0 and $A\times\overline{\mathbb{R}}_0^{dh}$ and $A\times\overline{\mathbb{R}}_0^{d(h-1)}\times A$ are continuity sets with respect to the limiting measures μ_h for every $h\ge 1$; see Section 2.2.1. Also assume that $\sum_{h\ge1}\rho_A(h)<\infty$. Let λ,ω be either any two Fourier or fixed frequencies in $(0,\pi)$.*

(1) If λ,ω are distinct then the covariances of the pairs $(\alpha_n(\lambda),\beta_n(\omega))$, $(\alpha_n(\lambda),\alpha_n(\omega))$, $(\beta_n(\lambda),\beta_n(\omega))$ converge to zero as $n\to\infty$.

(2) The covariance of $(\alpha_n(\lambda),\beta_n(\lambda))$ converges to zero as $n\to\infty$.

(3) If $\lambda\in(0,\pi)$ is fixed and if (λ_n) are Fourier frequencies such that $\lambda_n\to\lambda$ then the asymptotic variances are given by

$$\begin{aligned}&\operatorname{var}(\alpha_n(\lambda_n))\sim\operatorname{var}(\alpha_n(\lambda))\sim\operatorname{var}(\beta_n(\lambda_n))\sim\operatorname{var}(\beta_n(\lambda))\\ \sim\ &\mu_0(A)\Big[1+2\sum_{h=1}^{\infty}\cos(\lambda h)\rho_A(h)\Big]=\mu_0(A)f_A(\lambda)\,.\end{aligned}$$

Remark 2.4.2. The smoothness condition on the set A ensures that the extremogram ρ_A with respect to A is well defined; see Section 2.2.1.

Remark 2.4.3. Since $E\alpha_n(\lambda)=E\beta_n(\lambda)=0$ an immediate consequence of part (3) is that

$$EI_{nA}(\lambda)=\frac{1}{2}\big[\operatorname{var}(\alpha_n(\lambda))+\operatorname{var}(\beta_n(\lambda))\big]\sim\mu_0(A)\Big[1+2\sum_{h=1}^{\infty}\cos(\lambda h)\rho_A(h)\Big]=\mu_0(A)f_A(\lambda)\,.$$

Following the lines of the proof below, one can see that the error one encounters in the above approximation is uniform for $\lambda\in[a,b]\subset(0,\pi)$. The same remark applies to the quantities $EI_{nA}(\lambda_n)$ evaluated at Fourier frequencies $\lambda_n\to\lambda\in(0,\pi)$.

Proof. We start by calculating the asymptotic covariances. Any of the covariances can be written in the form

$$\begin{aligned}J&=\frac{2m_n}{n}E\Big[\sum_{s=1}^{n}\sum_{t=1}^{n}(I_sI_t-p_0^2)f_1(\lambda s)f_2(\omega t)\Big]\\&=\frac{2m_n}{n}\Big(\sum_{1\le t=s\le n}+\sum_{1\le s\ne t\le n}\Big)(p_{|s-t|}-p_0^2)\,f_1(\lambda s)f_2(\omega t)\\&=J_1+J_2\,,\end{aligned}$$

where f_1 and f_2 are cosine or sine functions and

$$p_{|t-s|}=P(a_m^{-1}X_s\in A,a_m^{-1}X_t\in A)\quad\text{for any } s,t.$$

We estimate J_1 separately for each possible combination of sine and cosine functions f_1, f_2. We start with $f_1(x) = \cos x$ and $f_2(x) = \sin x$. Then, if λ, ω are Fourier frequencies, so are $\lambda \pm \omega$ and therefore

$$\begin{aligned} J_1 &= (p_0 - p_0^2)\, \frac{2m_n}{n} \sum_{t=1}^{n} \cos(\lambda t) \sin(\omega t) \\ &= (p_0 - p_0^2)\, \frac{m_n}{n} \sum_{t=1}^{n} \big[\sin((\lambda+\omega)t) - \sin((\omega-\lambda)t) \big] = 0\,. \end{aligned}$$

If λ, ω are fixed frequencies, we conclude from (2.7.2) that the sum on the right-hand side is bounded. Hence $J_1 = O(n^{-1})$.

For $f_1(x) = f_2(x) = \cos x$ we get

$$\begin{aligned} J_1 &= (p_0 - p_0^2) \frac{2m_n}{n} \sum_{t=1}^{n} \cos(\lambda t) \cos(\omega t) \\ &= (p_0 - p_0^2) \frac{m_n}{n} \sum_{t=1}^{n} \big[\cos((\lambda+\omega)t) + \cos((\omega-\lambda)t) \big]\,. \end{aligned}$$

If λ, ω are Fourier frequencies, so are $\lambda \pm \omega$ and then the right-hand side vanishes unless $\lambda + \omega = \pi$. However, if $\lambda + \omega = \pi$ the second sum vanishes and the first sum is bounded. Therefore $J_1 = O(n^{-1})$. If $\lambda \neq \omega$ are fixed it follows from (2.7.1) that the sum on the right-hand side is bounded and therefore $J_1 = O(n^{-1})$.

For $f_1(x) = f_2(x) = \sin x$ we have

$$\begin{aligned} J_1 &= (p_0 - p_0^2)\, \frac{2m_n}{n} \sum_{t=1}^{n} \sin(\lambda t)\, \sin(\omega t) \\ &= (p_0 - p_0^2)\, \frac{m_n}{n} \sum_{t=1}^{n} \big[\cos((\lambda-\omega)t) - \cos((\lambda+\omega)t) \big]\,. \end{aligned}$$

The same arguments as above show that $J_1 = O(n^{-1})$ both for Fourier and fixed frequencies $\lambda \neq \omega$.

Next we consider J_2. We start with $\operatorname{cov}(\alpha_n(\lambda), \beta_n(\lambda))$. If λ is a Fourier frequency, we have $\sin(\lambda n) = 0$. Hence, by (2.7.7),

$$\begin{aligned} J_2 &= \frac{2m_n}{n} \sum_{h=1}^{n-1} (p_h - p_0^2) \sum_{s=1}^{n-h} \big[\sin(\lambda s) \cos(\lambda(s+h)) + \cos(\lambda s)\, \sin(\lambda(s+h)) \big] \\ &= -\frac{2m_n}{n} \sum_{h=1}^{n-1} (p_h - p_0^2) \sin(\lambda h) \end{aligned}$$

By definition of strong mixing, $|p_h - p_0^2| \leq \xi_h$. Then, by condition (M),

$$|J_2| \;\leq\; \frac{2m_n}{n} \sum_{h=1}^{\infty} \xi_h = O(m_n/n)\,.$$

The same argument applies for a fixed frequency λ since the expressions in (2.7.7) are bounded for every n and $h < n$.

If $\lambda \neq \omega$ are fixed frequencies we conclude from (2.7.8)–(2.7.10) and condition (M) that there exist constants $c(\lambda, \omega)$ such that

$$\begin{aligned} |J_2| &= \Big|\frac{2m_n}{n}\sum_{h=1}^{n-1}(p_h - p_0^2)\sum_{s=1}^{n-h}\big(f_1(\lambda s)f_2(\omega(s+h)) + f_1(\lambda(s+h))f_2(\omega s)\big)\Big| \\ &\leq c(\lambda,\omega)\,\frac{m_n}{n}\sum_{h=1}^{n-1}|p_h - p_0^2| \leq c(\lambda,\omega)\frac{m_n}{n}\sum_{h=1}^{\infty}\xi_h = O(m_n/n)\,. \end{aligned}$$

Now we consider the case of two distinct Fourier frequencies λ, ω. We start with $f_1(x) = \cos x$ and $f_2(x) = \sin x$. If $\lambda+\omega$ and $|\lambda-\omega|$ are bounded away from zero we can use the argument for general distinct frequencies. Now assume that $\lambda+\omega \leq 0.1$ say. Since λ, ω are Fourier frequencies a glance at (2.7.8)–(2.7.10) shows that one has to find suitable bounds for

$$\frac{|\sin((n-h+1)(\lambda+\omega)/2)|}{|\sin((\lambda+\omega)/2)|} = \frac{|\sin((-h+1)(\lambda+\omega)/2)|}{|\sin((\lambda+\omega)/2)|}\,.$$

If $h(\lambda+\omega) \leq 0.1$ Taylor expansions for the nominator and the denominator show that the right-hand side is bounded by $c\,h$. If $h(\lambda+\omega) > 0.1$ bound the nominator by 1 and Taylor expand the denominator to conclude that the right-hand side is bounded by $c\,h$ for some constant $c > 0$ as well. Then, by (2.7.8), for fixed k,

$$|J_2| \leq c\Big[\frac{m_n}{n}\sum_{h=1}^{k}|p_h - p_0^2| + m_n\sum_{h=k+1}^{r_n}|p_h - p_0^2| + m_n\sum_{h=r_n+1}^{\infty}\xi_h\Big]\,.$$

The right-hand side vanishes by virtue of condition (M), first letting $n \to \infty$ and then $k \to \infty$. The case of small $|\lambda-\omega|$, $|\lambda-\omega| \leq 0.1$ say, can be treated analogously.

The remaining cases $f_1(x) = f_2(x) = \cos x$ and $f_1(x) = f_2(x) = \sin x$ can be treated in the same way by exploiting (2.7.9) and (2.7.10).

Now we turn to the asymptotic variances. We restrict ourselves to $\alpha_n(\lambda)$ for fixed $\lambda \in (0,\pi)$; the variance of $\beta_n(\lambda)$ and the case of Fourier frequencies can be treated analogously. Write

We have

$$\mathrm{var}(\alpha_n(\lambda)) = \frac{2m_n}{n}\Big[(p_0 - p_0^2)\sum_{t=1}^{n}(\cos(\lambda t))^2 + 2\sum_{h=1}^{n-1}(p_h - p_0^2)\sum_{t=1}^{n-h}\cos(\lambda t)\cos(\lambda(t+h))\Big]\,.$$

For any frequency $\lambda \in (0,\pi)$ bounded away from zero and π, the relation $n^{-1}\sum_{t=1}^{n}(\cos(\lambda t))^2 \sim 0.5$ holds. Moreover, $\cos(\lambda t)\cos(\lambda(t+h)) = 0.5[\cos(\lambda h) + \cos(\lambda(2t+h))]$. Similar calculations as above yield

$$\mathrm{var}(\alpha_n(\lambda)) \sim m_n p_0 + 2m_n\sum_{h=1}^{n-1}(p_h - p_0^2)(1-h/n)\cos(\lambda h) \sim \mu_0(A)\Big[1 + 2\sum_{h=1}^{\infty}\rho_A(h)\cos(\lambda h)\Big]\,.$$

This concludes the proof. □

2.4.2 Central limit theorem

Our next result shows that the periodogram ordinates at distinct frequencies are asymptotically independent and exponentially distributed.

Theorem 2.4.4. *Consider a strictly stationary $\mathbb{R}^d$-valued sequence (X_t) which is regularly varying with index $\alpha > 0$. Let $A \subset \overline{\mathbb{R}}_0^d$ be bounded away satisfying the smoothness conditions of Proposition 2.4.1. Assume that the conditions* (M), (M1) *and $\sum_{h\geq 1} \rho_A(h) < \infty$ hold. Consider any fixed frequencies $0 < \lambda_1 < \cdots < \lambda_N < \pi$ for some $N \geq 1$. Then the following central limit theorem holds:*

$$\mathbf{Z}_n = \big(\alpha_n(\lambda_i), \beta_n(\lambda_i)\big)_{i=1,\ldots,N} \stackrel{d}{\to} \big((\alpha(\lambda_i), \beta(\lambda_i)\big)_{i=1,\ldots,N}\,, \quad n \to \infty\,, \tag{2.4.1}$$

where the limiting vector has $N(\mathbf{0}, \Sigma_N)$ distribution with

$$\Sigma_N = \mu_0(A)\,\mathrm{diag}\big(f_A(\lambda_1), f_A(\lambda_1), \ldots, f_A(\lambda_N), f_A(\lambda_N)\big)\,.$$

The limit relation (2.4.1) *remains valid if the frequencies λ_i, $i = 1, \ldots, N$, are replaced by distinct Fourier frequencies $\omega_i(n) \to \lambda_i \in (0, \pi)$ as $n \to \infty$. The limits λ_i do not have to be distinct.*

Then the following result is immediate.

Corollary 2.4.5. *Assume the conditions of Theorem 2.4.4. Let (E_i) be a sequence of iid standard exponential random variables.*

1. *Consider any fixed frequencies $0 < \lambda_1 < \cdots < \lambda_N < \pi$ for some $N \geq 1$. Then the following relations hold:*

$$\begin{aligned}
\big(I_{nA}(\lambda_i)\big)_{i=1,\ldots,N} &\stackrel{d}{\to} \mu_0(A)\,\big(f_A(\lambda_i)E_i\big)_{i=1,\ldots,N}\,, \quad n \to \infty\,,\\
\big(\widetilde{I}_{nA}(\lambda_i)\big)_{i=1,\ldots,N} &\stackrel{d}{\to} \big(f_A(\lambda_i)E_i\big)_{i=1,\ldots,N}\,, \quad n \to \infty\,.
\end{aligned}$$

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

$$\begin{aligned}
\big(I_{nA}(\omega_i(n))\big)_{i=1,\ldots,N} &\stackrel{d}{\to} \mu_0(A)\,\big(f_A(\lambda_i)E_i\big)_{i=1,\ldots,N}\,, \quad n \to \infty\,,\\
\big(\widetilde{I}_{nA}(\omega_i(n))\big)_{i=1,\ldots,N} &\stackrel{d}{\to} \big(f_A(\lambda_i)E_i\big)_{i=1,\ldots,N}\,, \quad n \to \infty\,.
\end{aligned}$$

Proof of the Theorem 2.4.4. We will prove (2.4.1) by applying the Cramér-Wold device, i.e., we will show that for any choice of constants $\mathbf{c} \in \mathbb{R}^{2N}$,

$$\mathbf{c}'\mathbf{Z}_n \stackrel{d}{\to} N(0, \mathbf{c}'\Sigma_N\mathbf{c})\,. \tag{2.4.2}$$

The proof of the result for distinct converging Fourier frequencies is analogous and therefore omitted. We will prove (2.4.2) by applying the method of small and large blocks. The difficulty we encounter here is that, due to the presence of sine and cosine functions, we are dealing with partial sums of non-stationary sequences. For $t = 1, \dots, n$, we write

$$Y_{nt} = \Big(\frac{2m_n}{n}\Big)^{1/2} \widetilde{I}_t \sum_{j=1}^{N} [c_{2j-1}\cos(\lambda_j t) + c_{2j}\sin(\lambda_j t)]\,, \quad t = 1, \dots, n\,. \tag{2.4.3}$$

For ease of presentation, we always assume that $n/m_n = k_n$ is an integer; the general case can be treated in a similar way. Consider the large blocks

$$K_{ni} = \{(i-1)m_n + 1, \dots, im_n\}, \qquad i = 1, \dots, k_n\,,$$

the index sets $\widetilde{K}_{ni}$, which consist of all but the first r_n elements of K_{ni}, and the small blocks $J_{ni} = K_{ni} \setminus \widetilde{K}_{ni}$. In view of condition (M), $r_n/m_n \to 0$ and $m_n \to \infty$, the sets $\widetilde{K}_{ni}$ and J_{ni} are non-empty for large n. For any set $B \subset \{1, \dots, n\}$, we write $S_n(B) = \sum_{t\in B} Y_{nt}$. First we show that the joint contribution of the sums over the small blocks to $\mathbf{c}'\mathbf{Z}_n$ is asymptotically negligible.

Lemma 2.4.6. *Under the conditions of Theorem 2.4.4, the following relation holds:*

$$\mathrm{var}\Big(\sum_{i=1}^{k_n} S_n(J_{ni})\Big) \to 0, \quad n \to \infty\,. \tag{2.4.4}$$

Proof. We have

$$\begin{aligned}\mathrm{var}\Big(\sum_{i=1}^{k_n} S_n(J_{ni})\Big) &\leq \sum_{i=1}^{k_n} \mathrm{var}(S_n(J_{ni})) + 2\sum_{1\leq i_1<i_2\leq k_n} |\mathrm{cov}\big(S_n(J_{ni_1}), S_n(J_{ni_2})\big)| \\ &= P_1 + P_2.\end{aligned}$$

Due to the sum structure of Y_{nt} given in (2.4.3) each of the sums $S_n(J_{ni})$ can be written as a sum of $2N$ subsums where each of these subsums only involves either the functions $\cos(\lambda_j t)$ or $\sin(\lambda_j s)$ for some $j \leq N$. Then each of the terms $\mathrm{var}(S_n(J_{ni}))$ and $|\mathrm{cov}\big(S_n(J_{ni_1}), S_n(J_{ni_2})\big)|$ is bounded by a linear combination of the variances/covariances of such subsums. In other words, it suffices to prove (2.4.4) for $N = 1$. We give the corresponding calculations only for the functions $\cos(\lambda t)$ where λ stands for any of the frequencies λ_j. The calculations are similar to those in the proof of Proposition 2.4.1. For any $i \leq k_n$ and fixed $k \geq 1$, condition (M) ensures that there is a constant $c(k)$ such that for large n,

$$\begin{aligned}\mathrm{var}(S_n(J_{ni})) &= \frac{2m_n}{n}\Big[\sum_{s=(i-1)m_n+1}^{(i-1)m_n+r_n} \mathrm{var}(I_s)(\cos(\lambda s))^2 \\ &\quad +2\sum_{h=1}^{r_n-1}\sum_{s=(i-1)m_n+1}^{(i-1)m_n+r_n-h} \mathrm{cov}(I_s, I_{s+h})\cos(s\lambda)\cos(\lambda(s+h))\Big] \\ &\leq \frac{2m_n}{n}\Big(r_n(p_0 - p_0^2) + 2\sum_{h=1}^{r_n-1}(r_n - h)|p_h - p_0^2|\Big) \\ &\leq c\frac{r_n}{n}\Big(m_n\sum_{h=0}^{k} p_h + m_n\sum_{h=k+1}^{r_n} p_h\Big) \leq c(k)(r_n/n)\,,\end{aligned}$$

and the right-hand side does not depend on i. Consequently, $P_1 \le c(k)k_n r_n/n = c(k)r_n/m_n \to 0$ for every fixed k. Similarly, for $i_1 < i_2$,

$$\begin{aligned}
|\mathrm{cov}(S_n(J_{ni_1}), S_n(J_{n,i_2}))| &= \frac{2m_n}{n}\Big|\Big[\sum_{s=(i_1-1)m_n+1}^{(i_1-1)m_n+r_n}\ \sum_{t=(i_2-1)m_n+1}^{(i_2-1)m_n+r_n} \mathrm{cov}(I_t, I_s)\cos(\lambda s)\cos(\lambda t)\Big]\Big| \\
&\le c\frac{m_n}{n}\sum_{q=(i_2-i_1)m_n-(r_n-1)}^{(i_2-i_1)m_n+r_n-1} (r_n - |q-(i_2-i_1)m_n|)\,|p_q - p_0^2| \\
&\le c\frac{m_n r_n}{n}\sum_{q=(i_2-i_1)m_n-(r_n-1)}^{(i_2-i_1)m_n+r_n-1} \xi_q\,,
\end{aligned}$$

where (ξ_t) is the mixing rate function. Hence for large n, in view of condition (M),

$$\begin{aligned}
|P_2| &\le c\frac{m_n r_n}{n}\sum_{i_1=1}^{k_n}\sum_{i_2=i_1+1}^{k_n}\ \sum_{q=(i_2-i_1)m_n-(r_n-1)}^{(i_2-i_1)m_n+r_n-1}\xi_q \\
&\le c\frac{m_n r_n}{n}\sum_{i_1=1}^{k_n-1}\ \sum_{q=m_n+1-r_n}^{\infty}\xi_q \le c r_n \sum_{q=r_n+1}^{\infty}\xi_q = o(1)\,.
\end{aligned}$$

This proves (2.4.4). □

Relation (2.4.4) implies that $\mathbf{c}'\mathbf{Z}_n$ and $\sum_{i=1}^{k_n} S_n(\widetilde{K}_{ni})$ have the same limit distribution provided such a limit exists. Let $\widetilde{S}_n(\widetilde{K}_{ni}) \stackrel{d}{=} S_n(\widetilde{K}_{ni})$ for $i=1,\ldots,k_n$ and assume that $(\widetilde{S}_n(\widetilde{K}_{ni}))_{i=1,\ldots,k_n}$ has independent components. A telescoping sum argument yields

$$\begin{aligned}
&\Big|E\prod_{l=1}^{k_n}\mathrm{e}^{\,itS_n(\widetilde{K}_{nl})} - E\prod_{s=1}^{k_n}\mathrm{e}^{\,it\widetilde{S}_n(\widetilde{K}_{ns})}\Big| \\
&= \Big|\sum_{l=1}^{k_n} E\Big[\Big(\mathrm{e}^{\,itS_n(\widetilde{K}_{nl})} - \mathrm{e}^{\,it\widetilde{S}_n(\widetilde{K}_{nl})}\Big)\prod_{s=1}^{l-1}\mathrm{e}^{\,it\widetilde{S}_n(\widetilde{K}_{ns})}\prod_{s=l+1}^{k_n}\mathrm{e}^{\,itS_n(\widetilde{K}_{ns})}\Big]\Big| \\
&\le \sum_{l=1}^{k_n}\Big|E\Big(\prod_{s=1}^{l-1}\mathrm{e}^{\,it\widetilde{S}_n(\widetilde{K}_{ns})}\Big(\mathrm{e}^{\,itS_n(\widetilde{K}_{nl})} - \mathrm{e}^{\,it\widetilde{S}_n(\widetilde{K}_{nl})}\Big)\prod_{s=l+1}^{k_n}\mathrm{e}^{\,itS_n(\widetilde{K}_{ns})}\Big)\Big| \\
&\le 4k_n\,\xi_{r_n} \to 0.
\end{aligned}$$

In the last step we used Theorem 17.2.1 in Ibragimov and Linnik [29] and condition (M1). Hence, $\sum_{l=1}^{k_n} S_n(\widetilde{K}_{nl})$ and $\sum_{l=1}^{k_n} \widetilde{S}_n(\widetilde{K}_{nl})$ have the same limits in distribution provided these limits exist. In view of (2.4.4) and the last conclusion the central limit theorem (2.4.2) holds if and only if the same limit relation holds for $\sum_{i=1}^{k_n}\widetilde{S}_n(K_{ni})$, where $\widetilde{S}_n(K_{ni}) \stackrel{d}{=} S_n(K_{ni})$ and $(\widetilde{S}_n(K_{ni}))_{i=1,\ldots,k_n}$ has independent components. Thus we may apply a classical central limit theorem for triangular arrays of independent random variables; see e.g. Theorem 4.1 in Petrov [40].

According to this result, the central limit theorem

$$Z_n = \sum_{i=1}^{k_n} \widetilde{S}_n(K_{ni}) \xrightarrow{d} N(0, \mathbf{c}'\Sigma_N \mathbf{c}),$$

holds if and only if the following three conditions are satisfied: $EZ_n = 0$, $\mathrm{var}(Z_n) \to \mathbf{c}'\Sigma_N\mathbf{c}$ and for every $\varepsilon > 0$,

$$\sum_{i=1}^{k_n} E\Big[(S_n(K_{ni}))^2 I_{\{|S(K_{ni})|>\varepsilon\}}\Big] \to 0. \tag{2.4.5}$$

The condition $EZ_n = 0$ holds since $E\widetilde{I}_t = 0$, hence $E\widetilde{S}_n(K_{ni}) = 0$ for every i. As for (6.8) in Davis and Mikosch [15], a trivial bound of the left-hand side in (2.4.5) is given by

$$c\,\frac{m_n^3}{n} \sum_{i=1}^{k_n} P(|S_n(K_{ni})| > \varepsilon) \le c\,\frac{m_n^3}{n} \sum_{i=1}^{k_n} I_{\{c\,(m_n^3/n)^{0.5} > \varepsilon\}}\,.$$

In view of (M1), $m_n^3/n = o(1)$, and therefore the right-hand side vanishes for sufficiently large n. Therefore (2.4.5) holds.

Lemma 2.4.7. *Under the conditions of Theorem 2.4.4,*

$$\mathrm{var}(Z_n) = \sum_{i=1}^{k_n} \mathrm{var}(S_n(K_{ni})) \to \mathbf{c}'\Sigma_N\mathbf{c}\,.$$

Proof. We proceed in a similar way as for Proposition 2.4.1. It will be convenient to introduce the following notation for $\lambda \in (0, \pi)$,

$$\begin{aligned}\widetilde{\alpha}_n(\lambda) &= \Big(\frac{2m_n}{n}\Big)^{1/2} \sum_{i=1}^{k_n} \sum_{t\in K_{ni}} \cos(\lambda t)\widetilde{I}_t(i)\,,\\ \widetilde{\beta}_n(\lambda) &= \Big(\frac{2m_n}{n}\Big)^{1/2} \sum_{i=1}^{k_n} \sum_{t\in K_{ni}} \sin(\lambda t)\widetilde{I}_t(i)\,,\end{aligned}$$

where for each $i \le k_n$,

$$(I_1, \ldots, I_{m_n}) \stackrel{d}{=} (I_{(i-1)m_n+1}(i), \ldots, I_{im_n}(i))$$

the vectors on the right-hand side are mutually independent for $i \le k_n$ and the quantities $\widetilde{I}_t(i)$ are the mean corrected versions of $I_t(i)$, i.e., $\widetilde{I}_t(i) = I_t(i) - p_0$. The statement of the lemma is proved if we can show that the pairs $(\widetilde{\alpha}_n(\lambda), \widetilde{\beta}_n(\omega)$, $(\widetilde{\alpha}_n(\lambda), \widetilde{\alpha}_n(\omega)$, $(\widetilde{\beta}_n(\lambda), \widetilde{\beta}_n(\omega)$, $(\widetilde{\alpha}_n(\lambda), \widetilde{\beta}_n(\lambda))$, are asymptotically uncorrelated for $\lambda \ne \omega$ and that

$$\mathrm{var}(\widetilde{\alpha}_n(\lambda)) \sim \mathrm{var}(\widetilde{\beta}_n(\lambda)) \sim \mu_0(A)\Big[1 + 2\sum_{h=1}^{\infty} \rho_A(h)\cos(\lambda h)\Big]\,. \tag{2.4.6}$$

We check the asymptotic variance of $\widetilde{\alpha}_n(\lambda)$ and omit similar calculations for $\mathrm{var}(\widetilde{\beta}_n(\lambda))$. By independence of the sums over the blocks K_{ni} we have for fixed $k \geq 1$,

$$
\begin{aligned}
\mathrm{var}(\widetilde{\alpha}_n(\lambda)) &= 2\frac{m_n}{n}\sum_{i=1}^{k_n}\mathrm{var}\Big(\sum_{t\in K_{ni}}\cos(\lambda t)\widetilde{I}_t\Big)\\
&= 2\frac{m_n}{n}\Big[\sum_{i=1}^{k_n}\sum_{t\in K_{ni}}\mathrm{var}(I_t)(\cos(\lambda t))^2+\sum_{i=1}^{k_n}\sum_{(i-1)m_n+1\leq t\neq s\leq im_n}\mathrm{cov}(I_t,I_s)\cos(\lambda t)\cos(\lambda s)\Big]\\
&= 2\frac{m_n}{n}(p_0-p_0^2)\sum_{t=1}^{n}(\cos(\lambda t))^2\\
&\quad +2\frac{m_n}{n}\sum_{i=1}^{k_n}\sum_{h=1}^{m_n-1}\sum_{t=1}^{m_n-h}(p_h-p_0^2)\big(\cos(\lambda h)+\cos\big(\lambda h+2\lambda(t+(i-1)m_n)\big)\big)\\
&= P_1+P_{21}+P_{22}\,.
\end{aligned}
$$

Then we have by (M) and regular variation of (X_t),

$$
P_1+P_{21}\sim\mu_0(A)+2\sum_{h=1}^{m_n-1}(p_h-p_0^2)(m_n-h)\cos(\lambda h)\sim\mu_0(A)f_A(\lambda)\,.
$$

We have for fixed $k \geq 1$,

$$
2\frac{m_n}{n}\Big|\sum_{i=1}^{k_n}\sum_{h=k+1}^{m_n-1}\sum_{t=1}^{m_n-h}(p_h-p_0^2)\cos(\lambda h+2\lambda(t+(i-1)m_n))\Big| \;\leq\; c\,m_n\sum_{h=k+1}^{m_n-1}|p_h-p_0^2|,
$$

and the right-hand side is negligible in view of (M) by first letting $n\to\infty$ and then $k\to\infty$. Thus it suffices to consider only finitely many h-terms in P_{22}. In view of (2.7.1), for fixed k as $n\to\infty$,

$$
\Big|2\frac{m_n}{n}\sum_{i=1}^{k_n}\sum_{h=1}^{k}(p_h-p_0^2)\sum_{t=1}^{m_n-h}\cos(\lambda h+2\lambda(t+(i-1)m_n))\Big|\leq c\sum_{h=1}^{k}|p_h-p_0^2|=o(1)\,.
$$

This proves (2.4.6).

Next we consider the case of two different frequencies $\lambda,\omega\in(0,\pi)$ and show that the following covariances vanish as $n\to\infty$:

$$
\begin{aligned}
&\mathrm{cov}(\widetilde{\alpha}_n(\lambda),\widetilde{\alpha}_n(\omega))\\
&= \frac{2m_n}{n}\sum_{i=1}^{k_n}\mathrm{cov}\Big(\sum_{t=1}^{m_n}\widetilde{I}_t\cos(\lambda(t+(i-1)m_n)),\sum_{t=1}^{m_n}\widetilde{I}_t\cos(\omega(t+(i-1)m_n))\Big)\\
&= \frac{2m_n}{n}\sum_{t=1}^{n}(p_0-p_0^2)\cos(\lambda t)\cos(\omega t)\\
&\quad +\frac{2m_n}{n}\sum_{i=1}^{k_n}\sum_{h=1}^{m_n-1}\sum_{t=1}^{m_n-h}(p_h-p_0^2)\big[\cos(\lambda(t+(i-1)m_n+h))\cos(\omega(t+(i-1)m_n)\\
&\quad +\cos(\lambda(t+(i-1)m_n))\cos(\omega(t+(i-1)m_n+h))\big]=Q_1+Q_2\,.
\end{aligned}
$$

In view of (2.7.1) and since $\lambda \neq \omega$,

$$|Q_1| = \frac{m_n}{n}(p_0 - p_0^2)\Big|\sum_{t=1}^{n}(\cos((\lambda+\omega)t) + \cos((\lambda-\omega)t))\Big| \le c\frac{m_n}{n}(p_0 - p_0^2) = O(n^{-1})\,.$$

Similarly, multiple application of (2.7.1), first summing over t, then over l, yields

$$\begin{aligned} |Q_2| &= \frac{m_n}{n}\Big|\sum_{h=1}^{m_n}(p_h - p_0^2)\sum_{l=0}^{k_n-1}\sum_{t=1}^{m_n-h}\Big(\cos((\lambda+\omega)(t+h+lm_n)+\lambda h) \\ &\quad + \cos((\lambda-\omega)(t+h+lm_n)+\lambda h) + \cos((\lambda+\omega)(t+h+lm_n)+\omega h) \\ &\quad + \cos((\lambda-\omega)(t+h+lm_n)-\omega h)\Big)\Big| \\ &\le c_0\sum_{h=1}^{m_n}|p_h - p_0^2| \le c\frac{r_n}{m_n}(m_n p_0) + c\frac{r_n}{m_n^2}(m_n p_0)^2 + c\sum_{h=r_n+1}^{m_n}\xi_h \to 0\,, \end{aligned}$$

where $c_0 = 4\max\{1/\sin((\lambda+\omega)/2), 1/\sin(|\lambda-\omega|/2)\}+4$. Thus $\mathrm{cov}(\widetilde{\alpha}_n(\lambda), \widetilde{\alpha}_n(\omega)) = o(1)$. Using similar arguments, it also follows that the covariances of the pairs $(\widetilde{\alpha}_n(\lambda), \widetilde{\beta}_n(\omega))$, $(\widetilde{\beta}_n(\lambda), \widetilde{\beta}_n(\omega))$ and $(\widetilde{\alpha}_n(\lambda), \widetilde{\beta}_n(\lambda))$ are asymptotically negligible. This proves the lemma. □

2.5 Smoothing the periodogram

Corollary 2.4.5 is analogous to the asymptotic theory for the periodogram of a stationary sequence; see Brockwell and Davis [9], Section 10.4, where the corresponding results are proved for the periodogram ordinates of a general linear processes with iid innovations. These results are then employed for showing that smoothed versions of the periodogram are consistent estimators of the spectral density at a given frequency. Our next goal is to prove a similar result.

We start by introducing the smoothed periodogram. For a fixed frequency $\lambda \in (0,\pi)$ define

$$\lambda_0 = \min\{2\pi j/n : 2\pi j/n \ge \lambda\}\,, \quad \text{and} \quad \lambda_j = \lambda_0 + 2\pi j/n\,, \quad |j| \le s_n\,.$$

Here we suppress the dependence of λ_j on n. In what follows, we will assume that $s_n \to \infty$ and $s_n/n \to 0$ as $n \to \infty$. For a given set $A \subset \overline{\mathbb{R}}_0^d$ bounded away from zero and any non-negative weight function $w = (w_n(j))_{|j|\le s_n}$ satisfying the conditions

$$\textstyle\sum_{|j|\le s_n} w_n(j) = 1 \quad \text{and} \quad \sum_{|j|\le s_n} w_n^2(j) \to 0 \text{ as } n \to \infty\,, \tag{2.5.1}$$

we introduce the *smoothed periodogram*

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

Theorem 2.5.1. *Assume the conditions of Theorem 2.4.4,* (2.5.1) *on the weight function* w *and* (M2)*. Then for every fixed frequency* $\lambda \in (0,\pi)$*, as* $n \to \infty$*,*

$$\widetilde{f}_{nA}(\lambda) \stackrel{L^2}{\to} \mu_0(A)\, f_A(\lambda) \quad \text{and} \quad \frac{\widetilde{f}_{nA}(\lambda)}{\widehat{P}_m(A)} \stackrel{P}{\to} f_A(\lambda)\,.$$

In Figures 2.1 and 2.2 we show the extremogram, the standardized periodogram and the corresponding smoothed periodogram for some simulated and real-life data. The data underlying Figure 2.1 are simulated from an ARMA(1,1) process (X_t) with parameters $\phi = 0.8$ and $\theta = 0.1$ and iid t-distributed noise (Z_t) with 3 degrees of freedom, hence (X_t) is regularly varying with $\alpha = 3$. The top-left graph shows the sample extremogram based on a sample of size $n = 31,757$ and the threshold is chosen as the 98% empirical quantile of the data. The top-right graph visualizes the theoretical spectral density f_A for $A = (1, \infty)$ (see Appendix 2.7.2 for an expression) and the raw periodogram which exhibits rather erratic behavior. The bottom graph shows the smoothed periodogram with Daniell window $w_n(i) = 1/(2s_n + 1)$, $|i| \le s_n = 50$. We also show the curves $f_A(\lambda)(1 \pm 1.96/\sqrt{2s_n + 1})$, which constitute a confidence band based on the following heuristic argument. In the proof of Theorem 2.5.1, we show that $\mathrm{var}(\widetilde{f}_{nA}(\lambda)) \sim \sum_{|j| \le s_n} w_n^2(j)\mu_0^2(A) f_{nA}^2(\lambda)$ for every $\lambda \in (0, \pi)$. Furthermore, we know that $\widehat{P}_m(A) \stackrel{P}{\to} \mu_0(A)$. Based on these calculations, we take $\sum_{|j| \le s_n} w_n^2(j) f_{nA}^2(\lambda)$ as a surrogate quantity for the unknown variance of $\widetilde{f}_n(\lambda)/\widehat{P}_m(A)$.

The data underlying Figure 2.2 are 5-min returns for the stock price of Bank of America (BAC) with the sample size $n = 31,757$, and a_m is chosen as the 98% empirical quantile of the data. We provide the same type of analysis as in Figure 2.1 for these data. The largest peak in the periodogram at the frequency 0.29 corresponds to an extremal cycle length of 6 hours, this is roughly the length of a trading day. We also show 95% pointwise confidence bands for the smoothed periodogram. They are not asymptotic since we do not have a central limit theorem for the smoothed periodogram yet. They are constructed from the distribution of the corresponding smoothed periodograms based on 99 random permutations of the data. If the data were iid, any permutation would not change the dependence structure of the data and one would expect that the estimated spectral density stays inside the band, but this is obviously not the case, indicating that the data exhibit some significant extremal dependence.

Proof. We mentioned in Remark 2.4.3 that

$$EI_{nA}(\lambda) \to \mu_0(A) f_A(\lambda) \quad \text{as } n \to \infty \text{ uniformly on sets } [a, b] \subset (0, \pi). \tag{2.5.2}$$

Therefore, since $\max_{|j| \le s_n} |\lambda_j - \lambda| \to 0$ and f_A is continuous, we have

$$E\widetilde{f}_{nA}(\lambda) = \sum_{|j| \le s_n} w_n(j)\, EI_{nA}(\lambda_j) \to \mu_0(A)\, f_A(\lambda)\,, \quad n \to \infty\,.$$

The statement of the theorem then follows if we can show that $\mathrm{var}(\widetilde{f}_n(\lambda)) \to 0$. We observe that

$$\mathrm{var}\big(\widetilde{f}_{nA}(\lambda)\big) = \sum_{|j| \le s_n} w_n^2(j)\, c_{jj} + \sum_{-s_n \le j_1 \ne j_2 \le s_n} w_n(j_1)\, w_n(j_2)\, c_{j_1 j_2}\,.$$

In view of condition (2.5.1) it suffices to show that $c_{j_1 j_2} = \mathrm{cov}(I_{nA}(\lambda_{j_1}), I_{nA}(\lambda_{j_2})) \to 0$ and

$$c_{jj} = \mathrm{var}(I_{nA}(\lambda_j)) \to (\mu_0(A)\, f_A(\lambda))^2 \quad \text{uniformly for } j, j_1, j_2 \in [-s_n, s_n],\ j_1 \ne j_2. \tag{2.5.3}$$

We will only show (2.5.3); the proof of $c_{j_1, j_2} \to 0$ for $j_1 \ne j_2$ is similar and therefore omitted. Since (2.5.2) holds we have to show that

$$E(I_{nA}^2(\lambda_j)) \to 2\, (\mu_0(A)\, f_A(\lambda))^2\,. \tag{2.5.4}$$

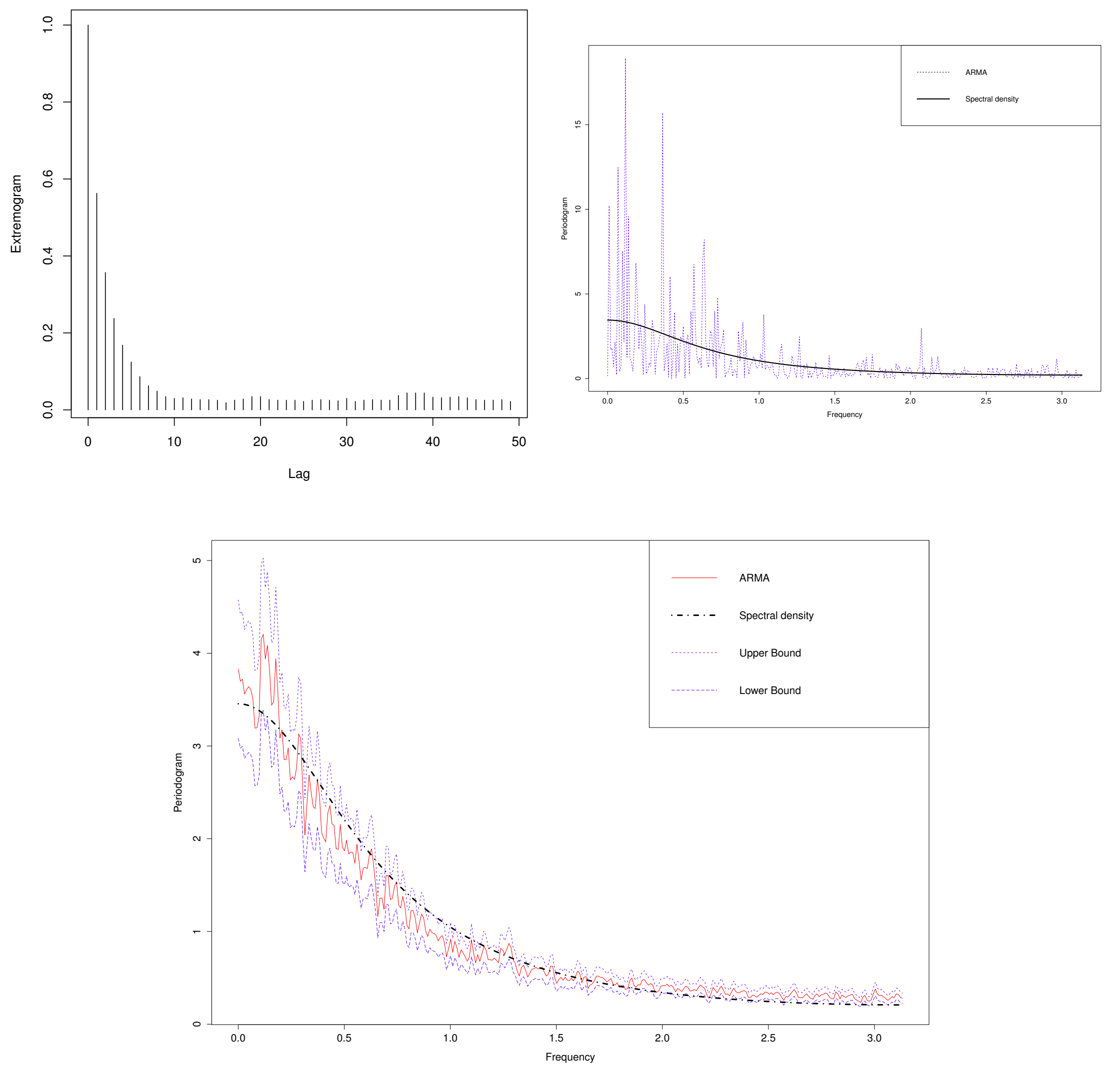

Figure 2.1: Top-Left: The sample extremogram of an ARMA(1,1) process with parameters $\phi = 0.8$, $\theta = 0.1$ and iid t-distributed noise with 3 degrees of freedom. We choose $A = (1, \infty)$. Top-Right: The corresponding raw periodogram and the theoretical spectral density f_A (solid line). Bottom: The smoothed periodogram with Daniell window, $s_n = 50$.

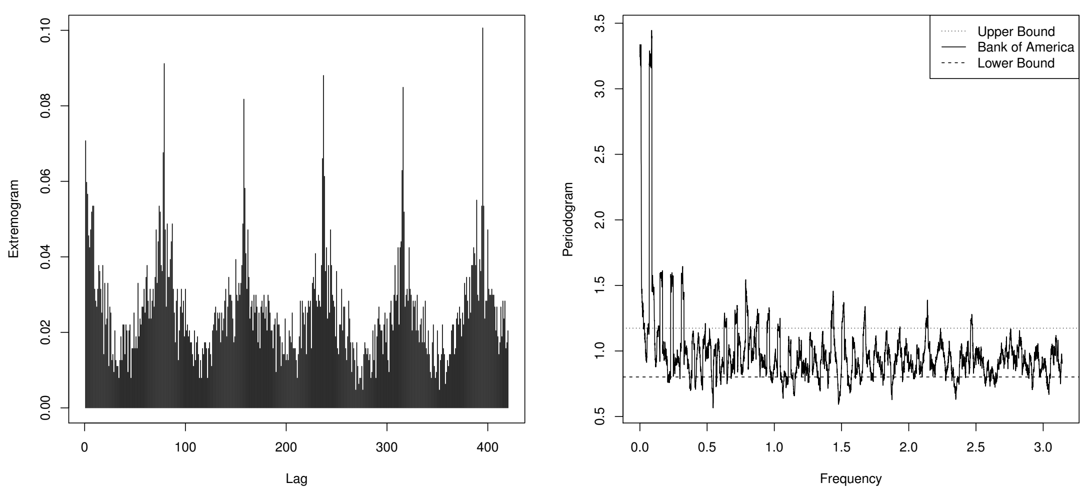

Figure 2.2: Left: The sample extremogram of 5-min returns of BAC stock price for $A = (1, \infty)$. Right: The smoothed periodogram with Daniell window, $s_n = 50$. The confidence bands are constructed from the smoothed periodograms of 99 permutations of the data.

Recall $\widehat{f}_{nA}(\lambda)$ from (2.2.7) and define

$$\widehat{g}_{nA}(\lambda) = 2 \sum_{h=r_n+1}^{n-1} \cos(\lambda h) \widetilde{\gamma}_n(h) \,.$$

We will study the decomposition

$$E(I_{nA}^2(\lambda_j)) = E\widehat{f}_{nA}^2(\lambda_j) + 2\, E\big(\widehat{f}_{nA}(\lambda_j)\widehat{g}_{nA}(\lambda_j)\big) + E\widehat{g}_{nA}^2(\lambda_j).$$

Following the lines of the proof of Theorem 5.1 in [15], we conclude that

$$E\widehat{f}_{nA}^2(\lambda_j) \to (\mu_0(A)\, f_A(\lambda))^2 \,, \tag{2.5.5}$$

uniformly for the considered frequencies λ_j. Then (2.5.4) is proved if we can show that

$$\begin{aligned} E(\widehat{f}_{nA}(\lambda_j)\widehat{g}_{nA}(\lambda_j)) &\to 0 \,, & (2.5.6)\\ E\widehat{g}_{nA}^2(\lambda_j) &\to (\mu_0(A)\, f_A(\lambda))^2 \,. & (2.5.7) \end{aligned}$$

Throughout we will use the notation, for $h_1, h_2, h_3 \geq 0$,

$$\begin{aligned} p_{h_1 h_2 h_3} &= P(X_0 > a_m, X_{h_1} > a_m, X_{h_1+h_2} > a_m, X_{h_1+h_2+h_3} > a_m) \,, \\ p_{h_1 h_2} &= p_{h_1 h_2 0} \,, \quad p_{h_1} = p_{h_1 0} \,, \end{aligned}$$

and we observe that

$$\begin{aligned} p_h &= (p_h - p_0^2) + p_0^2\,, & (2.5.8)\\ p_{h_1h_2} &= (p_{h_1h_2} - p_{h_1}p_0) + p_{h_1}p_0 = p_{h_1h_2} - p_0p_{h_2} + p_0p_{h_2} & (2.5.9)\\ &= (p_{h_1h_2} - p_0p_{h_2}) + p_0(p_{h_2} - p_0^2) + p_0^3\,, \\ p_{h_1h_2h_3} &= (p_{h_1h_2h_3} - p_0p_{h_2h_3}) + p_0p_{h_2h_3} & (2.5.10)\\ &= (p_{h_1h_2h_3} - p_0p_{h_2h_3}) + p_0(p_{h_2h_3} - p_0p_{h_3}) + p_0^2p_{h_3}\,. \end{aligned}$$

Proof of (2.5.6)

We have

$$\begin{aligned} E(\widehat{f}_{nA}(\lambda_j)\widehat{g}_{nA}(\lambda_j)) &= E\Big[2\widetilde{\gamma}_n(0)\widehat{g}_{nA}(\lambda_j) + 4\,\widehat{g}_{nA}(\lambda_j)\sum_{h=1}^{r_n}\cos(\lambda_j h)\widetilde{\gamma}_n(h))\Big]\\ &= J_1 + J_2\,, \end{aligned}$$

where

$$\begin{aligned} J_1 &= 4\frac{m_n^2}{n^2}\sum_{t_1=1}^{n}\sum_{h=r_n+1}^{n-1}\sum_{t_2=1}^{n-h} E[I_{t_1}I_{t_2}I_{t_2+h}]\cos(\lambda_j h)\,,\\ J_2 &= 8\frac{m_n^2}{n^2}\sum_{t_1=1}^{n-1}\sum_{h_1=1}^{r_n}\sum_{h_2=r_n+1}^{n-1}\sum_{t_2=1}^{n-h_2} E[I_{t_1}I_{t_1+h_1}I_{t_2}I_{t_2+h_2}]\cos(\lambda_j h_1)\cos(\lambda_j h_2)\,. \end{aligned}$$

Proof that J_1 is negligible.

We observe, that depending on the values h, t_1, t_2, $E[I_{t_1}I_{t_2}I_{t_2+h}]$ may simplify: if $t_1 = t_2$ or $t_1 = t_2 + h$, $E[I_{t_1}I_{t_2}I_{t_2+h}] = p_h$; if $t_1 < t_2$, $E[I_{t_1}I_{t_2}I_{t_2+h}] = p_{t_2-t_1,h}$; if $t_2 < t_1 < t_2 + h$, $E[I_{t_1}I_{t_2}I_{t_2+h}] = p_{t_1-t_2,h-t_1+t_2}$; if $t_1 > t_2 + h$, $E[I_{t_1}I_{t_2}I_{t_2+h}] = p_{h,t_1-h-t_2}$. If we take into account these different cases, we obtain

$$\begin{aligned} J_1 &= 4\frac{m_n^2}{n^2}\sum_{h=r_n+1}^{n-1}(n-h)(2p_h)\cos(\lambda_j h) + 4\frac{m_n^2}{n^2}\sum_{h_2=r_n+1}^{n-2}\sum_{h_1=1}^{n-h_2-1}(n-h_1-h_2)p_{h_1h_2}\cos(\lambda_j h_2)\\ &\quad +4\frac{m_n^2}{n^2}\sum_{h_2=r_n+1}^{n-1}\sum_{h_1=1}^{h_2-1}(n-h_2)p_{h_1,h_2-h_1}\cos(\lambda_j h_2)\\ &\quad +4\frac{m_n^2}{n^2}\sum_{h_2=r_n+1}^{n-2}\sum_{h_1=1}^{n-h_2-1}(n-h_1-h_2)p_{h_2h_1}\cos(\lambda_j h_2) = \sum_{i=1}^{4} J_{1i}\,. \end{aligned}$$

Applying (2.5.8), the mixing condition (M2) and Lemma 2.7.1 imply that

$$J_{11} \le c\,m_n\sum_{h=r_n+1}^{\infty}\xi_h + c\,\frac{(m_np_0)^2}{n(\sin(\lambda_j/2))^2} = o(1)\,.$$

As regards J_{12}, apply (2.5.9) and split the h_1-index set into $h_1 \le r_n$ and $h_1 > r_n$. Then (M2) and Lemma 2.7.1 imply that

$$\begin{aligned}
|J_{12}| &\le c\, m_n^2 \sum_{h_2=r_n+1}^{n-1} \xi_{h_2} \\
&\quad + c\Big|\frac{m_n^2}{n^2} \sum_{h_2=r_n+1}^{n-2} \Big(\sum_{h_1=1}^{\min(r_n,n-h_2-1)} + \sum_{h_1=r_n+1}^{n-h_2-1} \Big)(n-h_1-h_2)(p_{h_1} \pm p_0^2)p_0 \cos(\lambda_j h_2)\Big| \\
&\le o(1) + c\frac{r_n}{n}(m_n p_0)^2 (\sin(\lambda_j/2))^{-2} + c\,(m_n p_0) m_n \sum_{h_1=r_n+1}^{n-1} \xi_{h_1} + c\,\frac{(m_n p_0)^3}{m_n} = o(1)\,.
\end{aligned}$$

Now consider J_{13}. Abusing notation, we will write h_2 instead of $h_2 - h_1$. Introduce the index sets

$$\begin{aligned}
K_1 &= \{(h_1,h_2) : 1 \le h_i \le r_n\,, i = 1,2\}\,, \\
K_2 &= \{(h_1,h_2) : 1 \le h_1 \le r_n\,, r_n < h_2 < n - h_1\}\,, \\
K_3 &= \{(h_1,h_2) : r_n < h_1 \le n-1\,, 1 \le h_2 \le \min(r_n, n-h_1-1)\}\,, \\
K_4 &= \{(h_1,h_2) : r_n < h_1 \le n-1\,, r_n < h_2 < n - h_1\}\,.
\end{aligned}$$

Now introduce the mixing coefficients ξ_h and use Lemma 2.7.1:

$$\begin{aligned}
|J_{13}| &\le c\,\frac{m_n^2}{n^2}\Big| \sum_{h_1=1}^{n-2} \sum_{h_2=\max(1,r_n+1-h_1)}^{n-h_1-1} (n-h_1-h_2) p_{h_1 h_2} \cos(\lambda_j(h_1+h_2))\Big| \\
&\le c\,\frac{m_n^2}{n^2} \sum_{i=1}^{4} \Big| \sum_{K_i} (n-h_1-h_2) p_{h_1 h_2} \cos(\lambda_j(h_1+h_2))\Big| \\
&\le c\,\frac{m_n r_n^2}{n}(m_n p_0) + c\,\Big[\frac{m_n r_n}{n} m_n \sum_{h_2=r_n+1}^{n-1} \xi_{h_2} + \frac{r_n}{n}(m_n p_0)^2 (\sin(\lambda_j/2))^{-2}\Big] \\
&\quad + c\Big[\frac{m_n r_n}{n} m_n \sum_{h_1=r_n+1}^{n-1} \xi_{h_1} + \frac{r_n}{n}(m_n p_0)^2 (\sin(\lambda_j/2))^{-2}\Big] \\
&\quad + c\,\Big[m_n^2 \sum_{h_1=r_n+1}^{n-1} \xi_{h_1} + (m_n p_0) m_n \sum_{h_2=r_n+1}^{n-1} \xi_{h_2} + \frac{1}{m_n}(m_n p_0)^3 (\sin(\lambda_j/2))^{-2}\Big]\,.
\end{aligned}$$

The right-hand side vanishes as $n \to \infty$ by virtue of (M2). The same idea of proof applies to the relation $J_{14} = o(1)$. Thus we showed that $J_1 = o(1)$.

Proof that J_2 is negligible.

We split the summation over disjoint index sets, depending on the ordering of $\{t_1, t_1+h_1, t_2, t_2+h_2\}$: $t_1 = t_2$, $t_1 + h_1 = t_2 + h_2$, $t_1 + h_1 = t_2$, $t_1 = t_2 + h_2$, $t_1 < t_2 < t_1 + h_1 < t_2 + h_2$,

$t_2 < t_1 < t_1 + h_1 < t_2 + h_2$, $t_2 < t_1 < t_2 + h_2 < t_1 + h_1$, $t_1 + h_1 < t_2$ and $t_2 + h_2 < t_1$. Consider the index sets (we recycle the notation h_1, h_2 here)

$$\begin{aligned}
L_1 &= \{(h_1,h_2) : 1 \le h_1 \le r_n\,, r_n < h_2 < n\}\,,\\
L_2 &= \{(h_1,h_2) : 1 \le h_1 \le r_n\,, r_n < h_2 < n - h_1\}\,,\\
L_3 &= \{(h_1,h_2,h_3) : 2 \le h_1 \le r_n\,, r_n < h_2 < n - h_1 - 1\,, 1 \le h_3 < h_1\}\,,\\
L_4 &= \{(h_1,h_2,h_3) : 1 \le h_1 \le r_n\,, r_n < h_2 < n, 1 \le h_3 < h_2\}\,,\\
L_5 &= \{(h_1,h_2,h_3) : 1 \le h_1 \le r_n\,, r_n < h_2 < n-1, h_2 - h_1 < h_3 \le \min(n, h_2 + h_1 - 1)\}\,,\\
L_6 &= \{(h_1,h_2,h_3) : 1 \le h_1 \le r_n\,, r_n < h_2 < n - h_1 - 1, 1 \le h_3 < n - h_1 - h_2\}\,.
\end{aligned}$$

We write for short $f_{h_1h_2} = \cos(\lambda_{j_1} h_1)\cos(\lambda_{j_1} h_2)$. Then

$$\begin{aligned}
J_2 &= 8\frac{m_n^2}{n^2}\Big[\sum_{L_1}(n-h_2)\big(p_{h_1,h_2-h_1} + p_{h_2-h_1,h_1}\big)\, f_{h_1h_2}\\
&\quad + \sum_{L_2}(n-h_1-h_2)\big(p_{h_2h_1} + p_{h_1h_2}\big)\, f_{h_1h_2} + \sum_{L_3}(n-h_2-h_3)p_{h_3,h_1-h_3,h_2-h_1+h_3}\, f_{h_1h_2}\\
&\quad + \sum_{L_4}(n-h_2)p_{h_3,h_1,h_2-h_1}\, f_{h_1h_2} + \sum_{L_5}(n-h_1-h_3)p_{h_3,h_2-h_3,h_1-h_2+h_3}\, f_{h_1h_2}\\
&\quad + \sum_{L_6}(n-h_1-h_2-h_3)\big(p_{h_1h_3h_2} + p_{h_2h_3h_1}\big)\, f_{h_1h_2}\Big] = \sum_{i=1}^{6} J_{2i}
\end{aligned}$$

The terms J_{2i}, $i = 1, 2$, involve probabilities of the form p_{kl}. These terms can be treated in the same way as J_1 and shown to be negligible. We omit details.

The remaining J_{2i}'s contain probabilities of the form p_{kls}. We illustrate how one can deal with these pieces. We start with

$$\begin{aligned}
|J_{23}| &= 8\Big|\frac{m_n^2}{n^2}\sum_{h_1=1}^{r_n-1}\sum_{h_3=1}^{r_n-h_1}\Big(\sum_{h_2=r_n+1-h_3}^{r_n} + \sum_{h_2=r_n+1}^{n-h_1-h_3-1}\Big)(n-h_1-h_2-h_3)p_{h_1h_3h_2} f_{h_1+h_3,h_2+h_3}\Big|\\
&\le c\,\frac{m_n r_n^3}{n}(m_n p_0) + \Big[c\frac{m_n^2}{n}\sum_{h_1=1}^{r_n-1}\sum_{h_3=1}^{r_n-h_1}\sum_{h_2=r_n+1}^{n-h_1-h_3-1}|p_{h_1h_3h_2} - p_{h_1h_3}p_0|\\
&\qquad + c\,\frac{m_n^2}{n^2}p_0\sum_{h_1=1}^{r_n-1}\sum_{h_3=1}^{r_n-h_1} p_{h_1h_3}\cos(\lambda_{j_1}(h_1+h_3))\times\\
&\qquad \times\sum_{h_2=r_n+1}^{n-h_1-h_3-1}(n-h_1-h_2-h_3)\cos(\lambda_{j_1}(h_1+h_2))\Big]\\
&\le c\,\frac{m_n r_n^3}{n} + c\,\frac{m_n^2 r_n^2}{n}\sum_{h_2=r_n+1}^{\infty}\xi_{h_2} + c\,\frac{r_n^2}{n}(m_n p_0)^2\,(\sin(\lambda_j/2))^{-2}\,.
\end{aligned}$$

In the last step we used Lemma 2.7.1. The right-hand side in the latter relation converges to zero in view of the assumptions on r_n, m_n and (M2). The remaining expressions J_{2i} which contain probabilities p_{kls} over index sets such that $k,l > r_n, s \le r_n$ or $k > r_n, l, s \le r_n$ can be shown to be negligible by using similar arguments. We omit details. Those sums which contain probabilities p_{kls} over index sets such that $k,l,s > r_n$ are most difficult to deal with. The corresponding bounds follow from the next lemma.

Lemma 2.5.2. *Let $\lambda, \omega \in [a,b]$, $0 < a < b < \pi$, possibly depending on n, and x_1, x_2 be real numbers. Assume that*

$$m_n^2 n \sum_{h=r_n+1}^{n} \xi_h \to 0\,, \quad n \to \infty\,, \tag{2.5.11}$$

where (ξ_t) is the mixing rate function. Then

$$Q_0 = \frac{m_n^2}{n^2} \sum_{h_1,h_2,h_3>r_n} (n-h_1-h_2-h_3)_+ p_{h_1h_2h_3} \cos(\lambda h_1 + x_1)\cos(\omega h_3 + x_2) \to 0\,, \tag{2.5.12}$$

$$\frac{m_n^2}{n^2} \sum_{h_1,h_2,h_3>r_n} (n-h_1-h_2-h_3)_+ p_{h_1h_2h_3} \sin(\lambda h_1 + x_1)\sin(\lambda h_3 + x_2) \to 0\,. \tag{2.5.13}$$

Proof. Recall (2.5.10). Write $g_{h_1h_3} = \cos(\lambda h_1 + x_1)\cos(\omega h_3 + x_2)$. Then we have

$$\begin{aligned}
|Q_0| \le\ & \frac{m_n^2}{n^2} \sum_{h_1,h_2,h_3>r_n} (n-h_1-h_2-h_3)_+ |p_{h_1h_2h_3} - p_{h_1}p_{h_3}| \\
& + \Big|\frac{m_n^2}{n^2} \sum_{h_3=r_n+1}^{n-2r_n-3} \sum_{h_1=r_n+1}^{n-h_3-r_n-2} \sum_{h_2=r_n+1}^{n-h_1-h_3-1} (n-h_1-h_2-h_3)(p_{h_1}-p_0^2)(p_{h_3}-p_0^2) g_{h_1h_3}\Big| \\
& + \Big|\frac{m_n^2}{n^2} \sum_{h_2=r_n+1}^{n-2r_n-3} \sum_{h_3=r_n+1}^{n-h_2-r_n-2} \sum_{h_1=r_n+1}^{n-h_2-h_3-1} (n-h_1-h_2-h_3) p_0^2 (p_{h_3}-p_0^2) g_{h_1h_3}\Big| \\
& + \Big|\frac{m_n^2}{n^2} \sum_{h_2=r_n+1}^{n-2r_n-3} \sum_{h_1=r_n+1}^{n-h_2-r_n-2} \sum_{h_3=r_n+1}^{n-h_1-h_2-1} (n-h_1-h_2-h_3) p_0^2 (p_{h_1}-p_0^2) g_{h_1h_3}\Big| \\
& + \Big|\frac{m_n^2}{n^2} \sum_{h_2=r_n+1}^{n-2r_n-3} \sum_{h_1=r_n+1}^{n-h_2-r_n-2} \sum_{h_3=r_n+1}^{n-h_1-h_2-1} (n-h_1-h_2-h_3) p_0^4 g_{h_1h_3}\Big| = \sum_{i=1}^{5} Q_i\,.
\end{aligned}$$

By virtue of (2.5.11), Q_1 is negligible. Similarly, $Q_2 \le m_n^2 (\sum_{h=r_n+1}^n \xi_h)^2 \to 0$. As to Q_3, Lemma 2.7.1 and mixing imply that

$$\begin{aligned}
Q_3 \le\ & c\,\frac{(m_n p_0)^2}{n^2} \sum_{h_2=r_n+1}^{n-2r_n-3} \sum_{h_3=r_n+1}^{n-h_2-r_n-2} (n-h_2-h_3)|p_{h_3}-p_0^2|\,(\sin(\lambda/2))^{-2} \le cn \sum_{h_3=r_n+1}^{n} \xi_{h_3} \\
\to\ & 0\,.
\end{aligned}$$

A similar bound applies to Q_4. A double application of Lemma 2.7.1 yields

$$Q_5 \le c\,\frac{(m_n p_0)^4}{m_n^2}\,(\sin(\omega/2)\sin(\lambda/2))^{-2} \to 0\,.$$

Collecting these bounds, we proved (2.5.12). Similar arguments apply to (2.5.13). □

Thus we showed that J_1 and J_2 are negligible as $n \to \infty$. Hence (2.5.6) holds.

Proof of (2.5.7)

Following the steps for showing that J_2 is negligible, we decompose $E\widehat{g}_{nA}^2(\lambda_j)$ into sums over disjoint index sets depending on the ordering of $\{t_1, t_1+h_1, t_2, t_2+h_2\}$: $t_1 = t_2$ and $h_1 = h_2$; $t_1 = t_2$ and $h_1 > h_2$; $t_1 = t_2$ and $h_1 < h_2$; $t_1+h_1 = t_2+h_2$ and $t_1 > t_2$; $t_1+h_1 = t_2+h_2$ and $t_1 < t_2$; $t_1 = t_2+h_2$; $t_2 = t_1+h_1$; $t_1 < t_2 < t_1+h_1 < t_2+h_2$; $t_2 < t_1 < t_2+h_2 < t_1+h_1$; $t_1 < t_2 < t_2+h_2 < t_1+h_1$; $t_2 < t_1 < t_1+h_1 < t_2+h_2$; $t_1 > t_2+h_2$; $t_2 > t_1+h_1$. Consider the index sets (we recycle the notation h_1, h_2 here)

$$
\begin{aligned}
B_1 &= \{h : r_n < h < n\}\,,\\
B_2 &= \{(h_1,h_2) : r_n < h_1 < n-r_n, 1 \le h_2 < n-h_1\}\,,\\
B_3 &= \{(h_1,h_2) : r_n < h_1 < n-r_n-1, r_n < h_2 < n-h_1\}\,,\\
B_4 &= \{(h_1,h_2,h_3) : 1 \le h_1 < n-r_n-2, r_n < h_2 < n-h_1-1, 1 \le h_3 < n-h_1-h_2\}\,,\\
B_5 &= \{(h_1,h_2,h_3) : 1 \le h_1 < n-r_n-1, \max(1, r_n+1-h_1) \le h_2 < n-h_1-1,\\
&\qquad \max(1, r_n+1-h_2) \le h_3 < n-h_1-h_2\}\,,\\
B_6 &= \{(h_1,h_2,h_3) : r_n < h_1 < n-r_n-2, r_n < h_3 < n-1-h_1, 1 \le h_2 < n-h_1-h_3\}\,.
\end{aligned}
$$

Then we have

$$
\begin{aligned}
E\widehat{g}_{nA}^2(\lambda_j) &= 4\frac{m_n^2}{n^2}\sum_{B_1}(n-h)p_h f_{hh} + 4\frac{m_n^2}{n^2}\sum_{B_2}(n-h_1-h_2)(p_{h_1h_2}+p_{h_2h_1})f_{h_1+h_2,h_1}\\
&\quad +4\frac{m_n^2}{n^2}\sum_{B_3}(n-h_1-h_2)(p_{h_1h_2}+p_{h_2h_1})f_{h_1h_2}\\
&\quad +8\frac{m_n^2}{n^2}(n-h_1-h_2-h_3)\sum_{B_4}p_{h_1h_2h_3}f_{h_1+h_2+h_3,h_2}\\
&\quad +8\frac{m_n^2}{n^2}\sum_{B_5}(n-h_1-h_2-h_3)p_{h_1h_2h_3}f_{h_1+h_2,h_2+h_3}\\
&\quad +8\frac{m_n^2}{n^2}\sum_{B_6}(n-h_1-h_2-h_3)p_{h_1h_2h_3}f_{h_1h_3} = \sum_{i=1}^{6} G_i\,.
\end{aligned}
$$

Proof that G_3 and G_6 are negligible.

Using mixing and Lemma 2.7.1, we have as $n \to \infty$,

$$
\begin{aligned}
|G_3| &= 8\frac{m_n^2}{n^2}\Big|\sum_{B_3}(n-h_1-h_2)\big((p_{h_1h_2}-p_0p_{h_2}) + p_0(p_{h_2}-p_0^2) + p_0^3\big)\, f_{h_1h_2}\Big|\\
&\le cm_n^2\sum_{h_1=r_n+1}^{n}\xi_{h_1} + c\frac{m_n}{n}\sum_{h_2=r_n+1}^{n}\xi_{h_2} + c\frac{(m_np_0)^3}{m_n(\sin(\lambda_j/2))^2} = G_3' \to 0\,.
\end{aligned}
$$

We also have

$$\begin{aligned} |G_6| &\leq c\,\frac{m_n^2}{n^2}\Big|\sum_{h_1=r_n+1}^{n-r_n-3}\sum_{h_3=r_n+1}^{n-h_1-2}\Big(\sum_{h_2=1}^{r_n}+\sum_{h_2=r_n+1}^{n-h_1-h_3-1}\Big)(n-h_1-h_2-h_3)\,p_{h_1h_2h_3}f_{h_1h_2}\Big| \\ &= G_{61}+G_{62}\,. \end{aligned}$$

By (2.5.12), G_{62} is negligible and the same arguments as for G_3 show that $G_{61}\leq r_nG_3'\to 0$. Thus G_6 is negligible as $n\to\infty$.

The non-negligible contributions of G_1, G_2, G_4, G_5.

First, observe that

$$\begin{aligned} E\widehat{f}^2_{nA}(\lambda_{j_1}) &= (m_np_0)^2+4m_n^2p_0\sum_{h=1}^{r_n}\frac{n-h}{n}p_h\cos(\lambda_jh) \\ &\quad +4\frac{m_n^2}{n^2}\sum_{h_1=1}^{r_n}\sum_{h_2=1}^{r_n}(n-h_1)(n-h_2)p_{h_1}p_{h_2}f_{h_1h_2}=P_1+P_2+P_3\,, \end{aligned}$$

and we also know that (2.5.5) holds. Thus, (2.5.7) is proved if we can show that G_1-P_1, G_2-P_2 and $G_4+G_5-P_3$ are negligible. Observe that $\cos^2\lambda=0.5(1+\cos(2\lambda))$. Then by mixing and Lemma 2.7.1,

$$\begin{aligned} |G_1-P_1| &= 4\frac{m_n^2}{n^2}\Big|\sum_{h=r_n+1}^{n-1}(n-h)\big((p_h-p_0^2)+p_0^2\big)\,0.5\,(1+\cos(2\lambda_jh))-(m_np_0)^2\Big| \\ &\leq c\frac{m_n}{n}\,m_n\sum_{h=r_n+1}^{n-1}\xi_h+c(m_np_0)^2\Big|\frac{1}{2n^2}\sum_{h=r_n+1}^{n-1}(n-h)-1\Big|+c\,\frac{1}{n}(m_np_0)^2\to 0\,. \end{aligned}$$

As to G_2, we split the index set B_2 into the disjoint parts for $h_2\leq r_n$ and $h_2>r_n$. The sum over B_2 restricted to $h_2>r_n$ can be shown to be bounded by cG_3'. Recall that $2f_{h_1+h_2,h_1}=\cos(\lambda_jh_2)+\cos(\lambda_j(2h_1+h_2))$. Then

$$\begin{aligned} &|G_2-P_2|\leq c\,G_3' \\ &\quad+\Big|2\frac{m_n^2}{n^2}\sum_{h_2=1}^{r_n}\sum_{h_1=r_n+1}^{n-h_2-1}(n-h_1-h_2)(p_{h_1h_2}+p_{h_2h_1})\big(\cos(\lambda_jh_1)+\cos(\lambda_j(2h_2+h_1))\big) \\ &\quad-4m_n^2p_0\sum_{h=1}^{r_n}\frac{n-h}{n}p_h\cos(\lambda_jh)\Big| \\ &\leq c\,G_3'+c\frac{r_n^2}{n}(m_np_0)^2+c\frac{m_nr_n}{n}m_n\sum_{h_2=r_n+1}^{n-1}\xi_{h_2}+c(m_np_0)^2\frac{r_n}{n(\sin(\lambda_j))^2}\to 0\,. \end{aligned}$$

Here we used (2.5.9) to rewrite $p_{h_1h_2}$ such that the mixing condition and Lemma 2.7.1 can be applied.

Finally we turn to G_4 and G_5. By virtue of (2.5.12) and (2.5.13), we can neglect those parts of G_4+G_5 which contain (h_1,h_2,h_3)-indices with $h_1,h_2,h_3>r_n$. Those parts of G_4+G_5 for which two indices out of (h_1,h_2,h_3) exceed r_n we can deal with like J_{23}, and a similar argument applies when either $h_1>r_n$ or $h_3>r_n$. Thus we need to study those summands in G_4+G_5 indexed on $\{1\le h_1,h_3\le r_n, r_n<h_2<n-h_1-h_3\}$. We write G_{4+5} for the remaining sum. Recall that

$$f_{h_1+h_2+h_3,h_2}+f_{h_1+h_2,h_2+h_3} \quad = \quad f_{h_1h_3}+\cos(\lambda_j(h_1+2h_2+h_3))\,.$$

Then we have

$$\begin{aligned}|G_{4+5}-P_3| \quad &= \quad \Big|4\frac{m_n^2}{n^2}\sum_{h_1=1}^{r_n}\sum_{h_3=1}^{r_n}\sum_{h_2=r_n+1}^{n-h_1-h_3-1}(n-h_1-h_2-h_3)\big((p_{h_1h_2h_3}-p_{h_1}p_{h_3})+p_{h_1}p_{h_3}\big)\times\\ &\quad\times\big(2f_{h_1h_3}+2\cos(\lambda_{j_1}(h_1+2h_2+h_3))\big)\\ &\quad-4\frac{m_n^2}{n^2}\sum_{h_1=r_n+1}^{n-1}\sum_{h_2=r_n+1}^{n-1}(n-h_1)(n-h_2)p_{h_1}p_{h_2}f_{h_1h_2}\Big|\\ &\le \quad c\frac{r_n^3}{n}(m_np_0)^2+c\frac{m_nr_n^2}{n}m_n\sum_{h_3=r_n+1}^{n-1}\xi_{h_3}+c\frac{r_n^2}{n(\sin(\lambda_{j_1}))^2}(m_np_0)^2\,.\end{aligned}$$

Thus we also proved that $G_4+G_5-P_3$ is negligible.

Collecting all the arguments above, we finally proved the theorem. □

2.6 A discussion of related results and possible extensions

Extremogram-type quantities for time series have been introduced by various authors. Ledford and Tawn [35] discussed $\rho_{(1,\infty)}$ as a possible measure of extremal dependence for univariate stationary processes with unit Fréchet marginals under the regular variation condition $P(X_0>x,X_t>x)=L_t(x)x^{-1/\eta_t}$, for slowly varying L_t and $\eta_t\in(0,1]$. They were particularly interested in the case of asymptotic independence when $\rho_{(1,\infty)}(t)=0$ and $P(X_0>x,X_t>x)/[P(X>x)]^2\to 1$ as $x\to\infty$ and also suggested diagnostic conditions in this situation. Hill [28] proposed the quantities $\lim_{x\to\infty}[P(X_0>x,X_t>x)/[P(X>x)]^2-1]$ as alternative measure of serial extremal dependence in the case when the extremogram vanishes. Fasen et al. [21] considered lag-dependent tail dependence coefficients under regular variation conditions on the process (X_t). These coefficients can be interpreted as special extremograms. Hill [27] showed a pre-asymptotic functional central limit theorem for the sample extremogram of univariate time series over classes of upper quadrants. His mixing and domain of maximum domain of attraction are not directly comparable with strong mixing and regular variation od stationary sequences but the results are similar in spirit to Theorem 3.2 in Davis and Mikosch [15], where multivariate time series can be treated but uniform convergence over certain classes of sets was not considered.

Recently, various articles on the spectral analysis of indicator functions and their covariances based on a strictly stationary time series were written; see e.g. Dette et al. [17] and the references therein, Hagemann [25], Lee and Subba Rao [36]. The results are similar to those of classical time series analysis. The aforementioned papers do not deal with the spectral analysis of serial extremal dependence. In particular, they do not involve sequences of indicator functions of the form $(I_{\{a_m^{-1}X_t\in A\}})$ for sets A bounded away from zero. Therefore

these papers do not need additional conditions such as regular variation of (X_t) which are typical for extreme value theory and they do not require to consider the normalization m/n of the periodogram but use the classical $1/n$ constants.

The present paper focuses on the basic properties of the extremal periodogram. These properties parallel the results of classical time series analysis, but the proofs are different because of the triangular array nature of the stochastic processes $(I_{\{a_m^{-1}X_t\in A\}})$. In particular, the calculation of sufficiently high moments necessary to prove central limit theorems becomes rather technical. The central limit theorem for the smoothed periodogram is still an open question.

The (smoothed) periodogram as such contains information about the length of random cycles between extremal events $\{a_m^{-1}X_t \in A\}$. But it also opens the door to the methods and procedures of classical time series analysis, including the rich theory related to the integrated periodogram with applications to parameter estimation (e.g. Whittle estimation), good-of-fit tests, change point analysis, detection of long-range dependence effects and other problems. The solution to these problems is again rather technical and will be treated in future work.

2.7 Appendix

2.7.1 Some trigonometric sum formulas

Equations (2.7.1) and (2.7.2) are given in Gradshteyn and Ryzhik [24], 1.341 on p. 29; (2.7.3) and (2.7.4) are 1.352 on p. 31; and (2.7.5) and (2.7.6) are listed as 1.353 on p. 31. For any λ, x and $n \geq 1$, the following identities hold

$$\sum_{k=0}^{n-1}\cos(x+k\lambda) = \frac{\cos(x+(n-1)\lambda/2)\sin(n\lambda/2)}{\sin(\lambda/2)}, \tag{2.7.1}$$

$$\sum_{k=0}^{n-1}\sin(x+k\lambda) = \frac{\sin(x+(n-1)\lambda/2)\sin(n\lambda/2)}{\sin(\lambda/2)}, \tag{2.7.2}$$

$$\sum_{k=1}^{n-1}k\cos(k\lambda) = \frac{n\sin((2n-1)\lambda/2)}{2\sin(\lambda/2)} - \frac{1-\cos n\lambda}{4(\sin(\lambda/2))^2}, \tag{2.7.3}$$

$$\sum_{k=1}^{n-1}k\ \sin(k\lambda) = \frac{\sin(n\lambda)}{4(\sin(\lambda/2))^2} - \frac{n\cos((2n-1)\lambda/2)}{2\sin(\lambda/2)}, \tag{2.7.4}$$

$$\sum_{k=1}^{n-1}p^k\sin(k\lambda) = \frac{p\sin(\lambda)-p^n\sin(n\lambda)+p^{n+1}\sin\big((n-1)\lambda\big)}{1-2p\cos(\lambda)+p^2}, \tag{2.7.5}$$

$$\sum_{k=0}^{n-1}p^k\cos(k\lambda) = \frac{1-p\cos(\lambda)-p^n\cos(n\lambda)+p^{n+1}\cos((n-1)\lambda)}{1-2p\cos(\lambda)+p^2}. \tag{2.7.6}$$

Using these formulas, direct calculation yields for any frequency λ,

$$\sum_{s=1}^{n-h}\Big[\cos(\lambda s)\sin(\lambda(s+h))+\cos(\lambda(s+h))\sin(\lambda s)\Big] = \sum_{s=1}^{n-h}\sin(2\lambda s+\lambda h) = \frac{\sin(\lambda n)\sin(\lambda(n-h+1))}{\sin\lambda}-\sin(\lambda h)\,. \tag{2.7.7}$$

For distinct frequencies λ,ω we then obtain

$$\begin{aligned}
&\sum_{s=1}^{n-h}\Big[\cos(\lambda s)\sin(\omega(s+h))+\cos(\lambda(s+h))\sin(\omega s)\Big]\\
&= 0.5\sum_{s=1}^{n-h}[\sin((\lambda+\omega)s+\omega h)-\sin((\lambda-\omega)s-\omega h)]\\
&\quad +0.5\sum_{s=1}^{n-h}[\sin((\lambda+\omega)s+\lambda h)-\sin((\lambda-\omega)s+\lambda h)]\\
&= -\sin(\omega h)\\
&\quad +0.5\frac{\sin((n-h+1)(\lambda+\omega)/2)}{\sin((\lambda+\omega)/2)}\Big[\sin(\omega h+(n-h)(\lambda+\omega)/2)+\sin(\lambda h+(n-h)(\lambda+\omega)/2)\Big]\\
&\quad -0.5\frac{\sin((n-h+1)(\lambda-\omega)/2)}{\sin((\lambda-\omega)/2)}\Big[\sin(-\omega h+(n-h)(\lambda-\omega)/2)+\sin(\lambda h+(n-h)(\lambda-\omega)/2)\Big]\,,
\end{aligned} \tag{2.7.8}$$

$$\begin{aligned}
&\sum_{s=1}^{n-h}\Big[\cos(\lambda s)\cos(\omega(s+h))+\cos(\lambda(s+h))\cos(\omega s)\Big]\\
&= 0.5\sum_{s=1}^{n-h}\Big[\cos((\lambda+\omega)s+\omega h)+\cos((\lambda-\omega)s-\omega h)+\cos((\lambda+\omega)s+\lambda h)+\cos((\lambda-\omega)s+\lambda h)\Big]\\
&= -\cos(\omega h)-\cos(\lambda h)\\
&\quad +0.5\frac{\sin((n-h+1)(\lambda+\omega)/2)}{\sin((\lambda+\omega)/2)}\Big[\cos(\omega h+(n-h)(\lambda+\omega)/2)+\cos(\lambda h+(n-h)(\lambda+\omega)/2)\Big]\\
&\quad +0.5\frac{\sin((n-h+1)(\lambda-\omega)/2)}{\sin((\lambda-\omega)/2)}\Big[\cos(-\omega h+(n-h)(\lambda-\omega)/2)+\cos(\lambda h+(n-h)(\lambda-\omega)/2)\Big]\,,
\end{aligned} \tag{2.7.9}$$

$$
\begin{aligned}
&\sum_{s=1}^{n-h}\Big[\sin(\lambda s)\sin(\omega(s+h))+\sin(\lambda(s+h))\sin(\omega s)\Big]\\
=\;& 0.5\sum_{s=1}^{n-h}\Big[\cos((\lambda+\omega)s+\omega h)-\cos((\lambda-\omega)s-\omega h)+\cos((\lambda+\omega)s+\lambda h)-\cos((\lambda-\omega)s+\lambda h)\Big]\\
=\;& 0.5\frac{\sin((n-h+1)(\lambda-\omega)/2)}{\sin((\lambda-\omega)/2)}\Big[\cos(-\omega h+(n-h)(\lambda-\omega)/2)+\cos(\lambda h+(n-h)(\lambda-\omega)/2)\Big]\\
& -0.5\frac{\sin((n-h+1)(\lambda+\omega)/2)}{\sin((\lambda+\omega)/2)}\Big[\cos(\omega h+(n-h)(\lambda+\omega)/2)+\cos(\lambda h+(n-h)(\lambda+\omega)/2)\Big]\,.
\end{aligned}
\tag{2.7.10}
$$

Next assume the conditions of Theorem 2.5.1. Then a direct application of (2.7.1)–(2.7.4) yields for $\lambda \in (0,\pi)$ the following relations:

$$
\begin{aligned}
&\sum_{s=r_n+1}^{n}(n-s)\sin(\lambda s+x)\\
=\;& n\Big(\frac{\sin(x+(n-1))\lambda/2\sin(n\lambda/2)}{\sin(\lambda/2)}-\frac{\sin(x+r_n\lambda/2)\sin((r_n+1)\lambda/2)}{\sin(\lambda/2)}\Big)\\
&+\sin(x)\Big(\frac{n\sin((2n-1)\lambda/2)}{2\sin(\lambda/2)}-\frac{(r_n+1)\sin((2r_n-1)\lambda/2)}{2\sin(\lambda/2)}-\frac{\cos((r_n+1)\lambda)-\cos(n\lambda)}{4(\sin(\lambda/2))^2}\Big)\\
&+\cos(x)\Big(\frac{n\cos((2n-1)\lambda/2)}{2\sin(\lambda/2)}-\frac{(r_n+1)\cos((2r_n-1)\lambda/2)}{2\sin(\lambda/2)}-\frac{\sin(n\lambda)-\sin((r_n+1)\lambda)}{4(\sin(\lambda/2))^2}\Big)\\
&\sum_{s=r_n+1}^{n}(n-s)\cos(\lambda s+x)\\
=\;& n\Big(\frac{\cos(x+(n-1)\lambda/2)\sin(n\lambda/2)}{\sin(\lambda/2)}-\frac{\cos(x+r_n\lambda/2)\sin((r_n+1)\lambda/2)}{\sin(\lambda/2)}\Big)\\
&-\cos(x)\Big(\frac{n\sin((2n-1)\lambda/2)}{2\sin(\lambda/2)}-\frac{(r_n+1)\sin((2r_n-1)\lambda/2)}{2\sin(\lambda/2)}-\frac{\cos((r_n+1)\lambda)-\cos(n\lambda)}{4(\sin(\lambda/2))^2}\Big)\\
&+\sin(x)\Big(\frac{n\cos((2n-1)\lambda/2)}{2\sin(\lambda/2)}-\frac{(r_n+1)\cos((2r_n-1)\lambda/2)}{2\sin(\lambda/2)}-\frac{\sin(n\lambda)-\sin((r_n+1)\lambda)}{4(\sin(\lambda/2))^2}\Big)
\end{aligned}
$$

Lemma 2.7.1. *Under the assumptions of Theorem 2.5.1 the following relations hold uniformly* *for* $\lambda \in (0,2\pi)$, *as* $n \to \infty$,

$$\begin{aligned}
&\sum_{h=r_n+1}^{n-1} (n-h)\cos(\lambda h + x) \\
&= \frac{n\cos(x+(n-1)\lambda/2)\sin(n\lambda/2)}{\sin(\lambda/2)} - n - \frac{n\sin((2n-1)\lambda/2)}{2\sin(\lambda/2)} + \frac{1-\cos(n\lambda)}{4(\sin(\lambda/2))^2} \\
&\quad -\frac{n\cos(x+r_n\lambda/2)\sin((r_n+1)\lambda/2)}{\sin(\lambda/2)} - n + \frac{(r_n+1)\sin((2r_n+1)\lambda/2)}{2\sin(\lambda/2)} - \frac{1-\cos(r_n+1)\lambda}{4(\sin(\lambda/2))^2} \\
&= O(n/(\sin(\lambda/2))^2)\,, \\
&\sum_{h=r_n+1}^{n-1} (n-h)\sin(\lambda h + x) \\
&= \frac{n\sin(x+(n-1)\lambda/2)\sin(n\lambda/2)}{\sin(\lambda/2)} - \frac{\sin(n\lambda)}{4(\sin(\lambda/2))^2} + \frac{n\cos((2n-1)\lambda/2)}{2\sin(\lambda/2)} \\
&\quad -\frac{n\sin(x+r_n\lambda/2)\sin((r_n+1)\lambda/2)}{\sin(\lambda/2)} + \frac{\sin(r_n\lambda)}{4(\sin(\lambda/2))^2} - \frac{n\cos((2r_n+1)\lambda/2)}{2\sin(\lambda/2)} \\
&= O(n/(\sin(\lambda/2))^2)\,.
\end{aligned}$$

2.7.2 The spectral density f_A of an ARMA(1,1) process

In this section we calculate the spectral density f_A for an ARMA(1,1) process and the set $A=(1,\infty)$. The process (X_t) is given as the stationary causal solution to the difference equation

$$X_t = \phi X_{t-1} + Z_t + \theta Z_{t-1}\,, \quad t \in \mathbb{Z}\,,$$

where $0<|\phi|<1$ and $\theta \in \mathbb{R}$. From Brockwell and Davis [10], (2.3.3), we obtain the coefficients (ψ_j) of the linear process representation of (X_t) (cf. (2.3.2)):

$$\psi_0 = 1\,, \quad \psi_j = \phi^{j-1}(\phi+\theta)\,, \quad j \ge 1\,.$$

We assume that (Z_t) is an iid regularly varying sequence with index $\alpha > 0$.

The case $\phi \in (0,1)$, $\theta+\phi>0$, $p>0$. A direct application of (2.3.4) yields that

$$\begin{aligned}
\rho_A(h) &= \frac{\min(1,\psi_h^\alpha) + \sum_{i=h+1}^\infty \psi_i^\alpha}{\sum_{i=0}^\infty \psi_i^\alpha} \\
&= \frac{\min(1,\phi^{\alpha(h-1)}(\theta+\phi)^\alpha) + \phi^{\alpha h}(\theta+\phi)^\alpha(1-\phi^\alpha)^{-1}}{1+(\theta+\phi)^\alpha(1-\phi^\alpha)^{-1}}\,, \quad h \ge 1\,.
\end{aligned}$$

Define $h_0 = \min\{h \geq 0 : \phi^{\alpha h}(\theta+\phi)^\alpha < 1\}$ and write (see Appendix 2.7.1)

$$L^{(1)}(n,x,\lambda) = \sum_{h=1}^{n} \cos(x+h\lambda) = \begin{cases} \dfrac{\cos(x+n\lambda)\sin((n+1)\lambda/2)}{\sin(\lambda/2)} - 1\,, & n \geq 1\,, \\ 0\,, & n = 0\,; \end{cases}$$

$$L^{(2)}(n,x,\alpha,\lambda) = \sum_{h=1}^{n} |\phi|^{\alpha h}\cos(x+h\lambda)$$

$$= \begin{cases} \dfrac{|\phi|^\alpha\cos(x+\lambda) - |\phi|^{2\alpha}\cos(x) - |\phi|^{\alpha(n+1)}\cos(x+(n+1)\lambda) + |\phi|^{\alpha(n+2)}\cos(x+n\lambda)}{|1-|\phi|^\alpha \mathrm{e}^{-i\lambda}|^2}\,, & n \geq 1\,, \\ 0\,, & n = 0\,, \\ \dfrac{|\phi|^\alpha\cos(x+\lambda) - |\phi|^{2\alpha}\cos(x)}{|1-|\phi|^\alpha \mathrm{e}^{-i\lambda}|^2}\,, & n = \infty\,. \end{cases}$$

Then

$$\rho_A(h) \quad = \quad \begin{cases} c_\alpha^{(1)}(\phi,\theta) + \phi^{\alpha\, h}\, c_\alpha^{(2)}(\phi,\theta)\,, & h \leq h_0\,, \\ \phi^{\alpha\,(h-1)} c_\alpha^{(2)}(\phi,\theta)\,, & h > h_0\,, \end{cases}$$

where

$$c_\alpha^{(1)}(\phi,\theta) = \frac{1-\phi^\alpha}{1-\phi^\alpha+(\phi+\theta)^\alpha} \quad \text{and} \quad c_\alpha^{(2)}(\phi,\theta) \quad = \quad \frac{(\phi+\theta)^\alpha}{1-\phi^\alpha+(\phi+\theta)^\alpha}\,.$$

The corresponding spectral density is given by

$$\begin{aligned} f_A(\lambda) \quad &= \quad 1 + 2c_\alpha^{(1)}(\phi,\theta)\sum_{h=1}^{h_0}\cos(h\lambda) + 2(1-\phi^{-\alpha})c_\alpha^{(2)}(\phi,\theta)\sum_{h=1}^{h_0}\phi^{\alpha h}\cos(h\lambda) \\ &\quad\; +2\phi^{-\alpha}c_\alpha^{(2)}(\phi,\theta)\sum_{h=1}^{\infty}\phi^{\alpha h}\cos(h\lambda) \\ &= \quad 1 + 2c_\alpha^{(1)}(\phi,\theta)L^{(1)}(h_0,0,\lambda) + 2(1-\phi^{-\alpha})c_\alpha^{(2)}(\phi,\theta)L^{(2)}(h_0,0,\alpha,\lambda) \\ &\quad\; +2\phi^{-\alpha}c_\alpha^{(2)}(\phi,\theta)L^{(2)}(\infty,0,\alpha,\lambda)\,. \end{aligned}$$

The case $\phi \in (0,1)$, $\theta+\phi < 0$, $q > 0$. In view of (2.3.4) we have

$$\begin{aligned} \rho_A(h) \quad &= \quad \frac{q\sum_{i=0}^{\infty}\phi^{\alpha h+\alpha i}|\phi+\theta|^\alpha}{p+q\sum_{i=0}^{\infty}\phi^{\alpha i}|\phi+\theta|^\alpha} = \frac{q\phi^{\alpha h}|\phi+\theta|^\alpha}{p(1-\phi^\alpha)+q|\phi+\theta|^\alpha} = \phi^{\alpha h}c_\alpha^{(3)}(\phi,\theta), \quad h \geq 1\,, \\ f_A(\lambda) \quad &= \quad 1 + 2\,c_\alpha^{(3)}(\phi,\theta)L^{(2)}(\infty,0,\alpha,\lambda)\,. \end{aligned}$$

The case $\phi \in (-1,0)$, $\theta+\phi > 0$, $p > 0$. If $h = 2k+1$ for integer $k \geq 0$ the summand $p\,(\min(\psi_i^+,\psi_{i+h}^+))^\alpha + q\,(\min(\psi_i^-,\psi_{i+h}^-))^\alpha$ in (2.3.4) vanishes for $i \geq 1$. Thus

$$\rho_A(h) \quad = \quad \frac{p\min(1,|\psi_h|^\alpha)}{p+\sum_{i=1}^{\infty}\Big[p|\psi_{2i-1}|^\alpha + q|\psi_{2i}|^\alpha\Big]}\,.$$

For $h = 2k > 0$,

$$\rho_A(h) \quad = \quad \frac{\sum_{i=1}^{\infty}\Big[p|\psi_{2i+h-1}|^{\alpha} + q|\psi_{2i+h}|^{\alpha}\Big]}{p + \sum_{i=1}^{\infty}\Big[p|\psi_{2i-1}|^{\alpha} + q|\psi_{2i}|^{\alpha}\Big]}\,.$$

Define $k_1 = \min\{k \geq 0 : |\phi|^{2k}(\theta+\phi) < 1\}$. Then,

$$\rho_A(h) \quad = \quad \begin{cases} c_{\alpha}^{(4)}(\phi,\theta)\,, & h = 2k-1\,, 1 \leq k \leq k_1\,, \\ \phi^{\alpha\,(h-1)} c_{\alpha}^{(5)}(\phi,\theta)\,, & h = 2k-1\,, k > k_1\,, \\ \phi^{\alpha\,h} c_{\alpha}^{(6)}(\phi,\theta)\,, & h = 2k\,, k \geq 1\,. \end{cases}$$

where

$$\begin{aligned} c_{\alpha}^{(4)} &= \frac{p(1-|\phi|^{2\alpha})}{p(1-|\phi|^{2\alpha}+(\phi+\theta)^{\alpha}) + q|\phi|^{\alpha}\,(\phi+\theta)^{\alpha}}\,, \\ c_{\alpha}^{(5)} &= \frac{p(\phi+\theta)^{\alpha}(1-|\phi|^{2\alpha})}{p(1-|\phi|^{2\alpha}+(\phi+\theta)^{\alpha}) + q|\phi|^{\alpha}\,(\phi+\theta)^{\alpha}}\,, \\ c_{\alpha}^{(6)} &= \frac{p(\phi+\theta)^{\alpha} + q|\phi|^{\alpha}(\phi+\theta)^{\alpha}}{p(1-|\phi|^{2\alpha}+(\phi+\theta)^{\alpha}) + q|\phi|^{\alpha}\,(\phi+\theta)^{\alpha}}\,. \end{aligned}$$

The corresponding spectral density is

$$\begin{aligned} f_A(\lambda) \quad &= \quad 1 + 2c_{\alpha}^{(4)}(\phi,\theta)\sum_{k=1}^{k_1}\cos((2k-1)\lambda) + 2|\phi|^{-2\alpha}c_{\alpha}^{(5)}(\phi,\theta)\sum_{k=k_1+1}^{\infty}|\phi|^{\alpha(2k)}\cos((2k-1)\lambda) \\ &\qquad + 2c_{\alpha}^{(6)}(\phi,\theta)\sum_{k=1}^{\infty}|\phi|^{2k\alpha}\cos(2k\lambda) \\ &= \quad 1 + 2c_{\alpha}^{(4)}(\phi,\theta)L^{(1)}(k_1,-\lambda,2\lambda) + 2|\phi|^{-2\alpha}c_{\alpha}^{(5)}\big[L^{(2)}(\infty,-\lambda,2\alpha,2\lambda) - L^{(2)}(k_1,-\lambda,2\alpha,2\lambda)\big] \\ &\qquad + 2c_{\alpha}^{(6)}(\phi,\theta)L^{(2)}(\infty,0,\alpha,2\lambda)\,. \end{aligned}$$

The case $\phi \in (-1,0)$, $\theta+\phi < 0$, $p > 0$. If $h = 2k+1$ for integer $k \geq 0$ the summand $p\,(\min(\psi_i^+,\psi_{i+h}^+))^{\alpha} + q\,(\min(\psi_i^-,\psi_{i+h}^-))^{\alpha}$ in (2.3.4) vanishes for $i \geq 0$. Thus

$$\rho_A(h) = 0\,.$$

For $h = 2k > 0$,

$$\rho_A(h) \quad = \quad \frac{p\min(1,|\psi_h|^{\alpha}) + \sum_{i=0}^{\infty}\Big[p|\psi_{2i+h+2}|^{\alpha} + q|\psi_{2i+h+1}|^{\alpha}\Big]}{\sum_{i=0}^{\infty}\Big[p|\psi_{2i}|^{\alpha} + q|\psi_{2i+1}|^{\alpha}\Big]}\,.$$

Define $k_2 = \min\{k \geq 0 : |\phi|^{2k+1}|\theta+\phi| < 1\}$. Then

$$\rho_A(2k) = \begin{cases} c_{\alpha}^{(7)} + |\phi|^{2\alpha k}\,c_{\alpha}^{(8)}\,, & k \leq k_2\,, \\ |\phi|^{2\alpha k}\,c_{\alpha}^{(9)}\,, & k > k_2\,. \end{cases}$$

where

$$c_\alpha^{(7)} = \frac{p(1-|\phi|^{2\alpha})}{p(1-|\phi|^{2\alpha}) + p|\phi|^\alpha|\phi+\theta|^\alpha + q|\phi+\theta|^\alpha}\,,$$
$$c_\alpha^{(8)} = \frac{p|\phi|^\alpha|\phi+\theta|^\alpha + q|\phi+\theta|^\alpha}{p(1-|\phi|^{2\alpha}) + p|\phi|^\alpha|\phi+\theta|^\alpha + q|\phi+\theta|^\alpha}\,,$$
$$c_\alpha^{(9)} = \frac{p|\phi|^{-\alpha}|\phi+\theta|^\alpha + q|\phi+\theta|^\alpha}{p(1-|\phi|^{2\alpha}) + p|\phi|^\alpha|\phi+\theta|^\alpha + q|\phi+\theta|^\alpha}\,.$$

The corresponding spectral density is

$$\begin{aligned} f_A(\lambda) &= 1 + 2c_\alpha^{(7)}\sum_{k=1}^{k_2}\cos(2k\lambda) + 2(c_\alpha^{(8)} - c_\alpha^{(9)})\sum_{k=1}^{k_2}|\phi|^{2k\alpha}\cos(2k\lambda) + 2c_\alpha^{(9)}\sum_{k=1}^{\infty}|\phi|^{2k\alpha}\cos(2k\lambda) \\ &= 1 + 2c_\alpha^{(7)}L^{(1)}(k_2,0,2\lambda) + 2(c_\alpha^{(8)} - c_\alpha^{(9)})L^{(2)}(k_2,0,2\alpha,2\lambda) + 2c_\alpha^{(9)}L^{(2)}(\infty,0,2\alpha,2\lambda)\,. \end{aligned}$$

Acknowledgment. We would like to thank the reviewers of our paper for careful reading and comments, in particular for pointing out several useful references.

Bibliography

[1] Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.) (2009) *The Handbook of Financial Time Series.* Springer, Heidelberg.

[2] Bartkiewicz, K., Jakubowski, A., Mikosch, T. and Wintenberger, O. (2011) Stable limits for sums of dependent infinite variance random variables. *Probab. Th. Rel. Fields* **150**, 337–372.

[3] Basrak, B., Davis, R.A. and Mikosch. T. (2002) Regular variation of GARCH processes. *Stoch. Proc. Appl.* **99**, 95–116.

[4] Basrak, B., Krizmanić, D. and Segers, J. (2012) A functional limit theorem for dependent sequences with infinite variance stable limits. *Ann. Probab.* **40**, 2008–2033.

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

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

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

[8] Brillinger, D.R. (1981) *Time Series. Data Analysis and Theory.* 2nd Edition. Holden-Day, Inc., Oakland (California).

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

[10] Brockwell. P. and Davis., R.A. (2002) *An Introduction to Time Series and Forecasting.* 2nd Edition. Springer, New York.

[11] 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.

[12] Davis, R.A. and Mikosch, T. (2001) Point process convergence of stochastic volatility processes with application to sample autocorrelation. *J. Appl. Probab* **38A**, 93–104.

[13] Davis, R.A. and Mikosch, T. (2009) Extreme value theory for GARCH processes. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.) *Handbook of Financial Time Series.* Springer, pp. 187–200.

[14] Davis, R.A. and Mikosch, T. (2009) Extremes of stochastic volatility models. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.) *Handbook of Financial Time Series.* Springer (2009), pp. 355–364.

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

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

[17] Dette, H., Hallin, M., Kleya, T. and Volgusheva, S. (2011) On copulas, quantiles, ranks and spectra. An L_1-approach to spectral analysis. Working paper.

[18] Dombry, C. and Eyi-Minko, F. (2012) Strong mixing properties of max-infinitely divisible random fields.

[19] Doukhan, P. (1994) *Mixing. Properties and Examples. Lecture Notes in Statistics* **85**. Springer, New York.

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

[21] Fasen, V., Klppelberg, C. and Schlather, M. (2010) High-level dependence in time series models. *Extremes* **13**, 1–33. High-level dependence in time series models. *Extremes* ,.

[22] Feller, W. (1971) *An Introduction to Probability Theory II.* Wiley, New York.

[23] Grenander, U. and Rosenblatt, M. (1984) *Statistical Analysis of Stationary Time Series.* 2nd Edition. Chelsea Publishing Co., New York,

[24] Gradshteyn, I.S. and Ryzhik, I.M. (1980) *Table of Integrals, Series, and Products.* Academic Press, New York.

[25] Hagemann, A. (2011) Roubust spectral analysis. Working paper UICU.

[26] Hannan, E.J. (1960) *Time Series Analysis.* Wiley, New York.

[27] Hill, J.B. (2009) On functional central limit theorems for dependent, heterogeneous arrays with applications to tail index and tail dependence conditions. *J. Statist. Plan. Inf.* **139**, 2091–2110.

[28] Hill, J.B. (2011) Extremal memory of stochastic volatility with an application to tail shape inference. *J. Statist. Plan. Inf.* **141**, 663–676.

[29] Ibragimov. I.A. and and Linnik, Y.V. (1971) *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff, Groningen.

[30] Jakubowski, A. (1993) Minimal conditions in p-stable limit theorems. *Stoch. Proc. Appl.* **44**, 291–327.

[31] Jakubowski, A. (1997) Minimal conditions in p-stable limit theorems - II. *Stoch. Proc. Appl.* **68**, 1–20.

[32] Jessen, A.H. and Mikosch, T. (2006) Regularly varying functions. *Publ. Inst. Math.* Nouvelle Série **80(94)**, 171–192.

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

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

[35] Ledford, A.W. and Tawn, J.A. (2003) Diagnostics for dependence within time series extremes. *J. Royal Statist. Soc.* **65**, 521–543.

[36] Lee, J. and Subba Rao, S. (2012) The quantile spectral density and comparison based tests for nonlinear time series. Working paper.

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

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

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

[40] Petrov, V.V. (1995) *Limit Theorems of Probability Theory.* Oxford University Press, Oxford (UK).

[41] Pham, T. and Tran, L.T. (1985) Some mixing properties of time series models. *Stoch. Proc. Applic.* **19**, 297–303.

[42] Priestley, M.B. (1981) *Spectral Analysis and Time Series.* Academic Press, London, New York.

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

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

[45] Resnick, S.I. (2006) *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling.* Springer, New York.

Chapter 3

Measures of serial extremal dependence and their estimation

Abstract

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

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

3.1 Introduction

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

3.1.1 The extremal index as reciprocal of the expected extremal cluster size

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

The intuition about extremal clusters in a time series can be made precise using point process theory. In the classical monograph by Leadbetter, Lindgren and Rootzén [40] the *point process of exceedances* of u was used to describe clusters of extremes as an asymptotic phenomenon when the threshold u_n converges to the right endpoint of the distribution F of X. (Here and in what follows, Y denotes a generic element of any strictly stationary sequence (Y_t).) To be more precise, (u_n) has to satisfy the condition $n\,\overline{F}(u_n) = n\,(1 - F(u_n)) \to \tau$ for some $\tau \in (0, \infty)$. Under this condition and mixing assumptions, the point processes of exceedances converge weakly to a compound Poisson process (see Hsing et al. [31]):

$$N_n = \sum_{i=1}^{n} \varepsilon_{i/n}\, I_{\{X_i > u_n\}} \stackrel{d}{\to} N = \sum_{i=1}^{\infty} \xi_i\, \varepsilon_{\Gamma_i}\,, \tag{3.1.1}$$

where the state space of the point processes is $(0, 1]$, the points $0 < \Gamma_1 < \Gamma_2 < \cdots$ constitute a homogeneous Poisson process with intensity $\theta\tau$ on $(0, 1]$ which is independent of an iid positive integer-valued sequence (ξ_i). Here $\theta \in [0, 1]$ is the *extremal index* of the sequence (X_t). Thus, in an asymptotic way, a cluster of extremes is located at the Poisson points Γ_i with corresponding size ξ_i. The cluster size distribution $P(\xi = k)$, $k \geq 1$, contains plenty of information about the distribution of the extremal clusters. However, most attention has been given to determine the expected cluster size $E\xi$ which can be interpreted as reciprocal of θ as the following heuristic argument illustrates. Applying (3.1.1) on the set $(0, 1]$ and taking expectations on both sides of the limit relation, we observe that

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

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

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

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

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

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

If F belongs to the *maximum domain of attraction of an extreme value distribution* H ($F \in \mathrm{MDA}(H)$) there exist constants $c_n > 0, d_n \in \mathbb{R}$, $n \geq 1$, such that $P(c_n^{-1}(\widetilde{M}_n - d_n) \leq x) \to H(x)$

for every $x \in \operatorname{supp}(H)$ (the support of H); cf. Embrechts et al. [20], Chapter 3. Thus, writing $u_n(x) = c_n\, x + d_n$ and $\tau = \tau(x) = -\log H(x)$ for $x \in \operatorname{supp}(H)$, the existence of an extremal value index θ of the sequence (X_t) implies that

$$P(c_n^{-1}(M_n - d_n) \le x) \to H^\theta(x)\,, \quad x \in \mathbb{R}\,. \tag{3.1.2}$$

The concrete form of the extremal index is known for various standard time series models, including linear processes with iid subexponential noise (cf. [20], Section 5.5), Markov processes (see Leadbetter and Rootzén [41], Perfekt [53]) and financial time series models such as GARCH (generalized autoregressive conditionally heteroscedastic) and SV (stochastic volatility) models; cf. [11, 13, 14]. Expressions of the extremal index for regularly varying sequences (X_t) (see Section 3.1.2 for a definition) in terms of the points of the limiting point process were given in Davis and Hsing [8] and in terms of the limiting tail process in Basrak and Segers [2]; see (3.1.10) below. However, for most models these concrete expressions of θ are too complex to be useful in practice.

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

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

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

[1]This remark also indicates that long range dependence for extremes should not be defined via the covariance function γ_X. As explained above (see (3.1.2)) the existence of a positive extremal index θ ensures that the type of the limiting extreme value distribution H remains the same as in the iid case. This is easily checked since the only possible non-degenerate limit distributions H are the types of the *Fréchet distribution* $\Phi_\alpha(x) = \mathrm{e}^{-x^{-\alpha}}$, $x, \alpha > 0$, the *Weibull distribution* $\Psi_\alpha(x) = \mathrm{e}^{-(-x)^\alpha}$, $x < 0, \alpha > 0$, and the *Gumbel distribution* $\Lambda(x) = \mathrm{e}^{-\mathrm{e}^{-x}}$, $x \in \mathbb{R}$. This is a consequence of the Fisher-Tippett theorem; cf. Embrechts et al. [20], Theorem 3.2.3. The notion of long range dependence in an extreme value sense would be reasonable if in (3.1.2) a limit distribution occurred which does not belong to the type of any of the three mentioned standard extreme value distributions. This, however, can only be expected if a given stationary sequence (X_t) with $F \in \mathrm{MDA}(H)$ does not have an extremal index or if $\theta = 0$. Examples of sequences with zero extremal index are given in Leadbetter et al. [40] and Leadbetter [39], but such examples are often considered pathological; see also the discussion in Samorodnitsky [63] who studied infinite variance stable stationary sequences with zero extremal index and the boundary between short and long range extremal dependence for these sequence.

length depending on the local extremal behavior. These estimators often exhibit a rather large uncertainty.

In Figures 3.1.1 and 3.1.2 we illustrate the estimation of θ for real and simulated data. We choose the simple *blocks estimator* $\widehat{\theta} = K_n/N_n$ of the extremal index θ, where N_n is the number of exceedances of the threshold $u = u_n$ in the sample $X_1, \ldots, X_n$ and K_n is the number of blocks of size $s = s_n$, $X_{(i-1)s+1}, \ldots, X_{is}$, $i = 1, \ldots, [n/s]$, with at least one exceedance of u.

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

3.1.2 The extremogram: an asymptotic correlogram for extreme events

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

$$\gamma_{AB}(h) = \lim_{n\to\infty} n \operatorname{cov}(I_{\{a_n^{-1}X_0 \in A\}}, I_{\{a_n^{-1}X_h \in B\}}), \quad h \geq 0. \tag{3.1.3}$$

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

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

$$\rho(h) = \lim_{x\to\infty} P(X_h > x \mid X_0 > x). \tag{3.1.4}$$

(Here we assume that X has infinite right endpoint.) These pairwise tail dependence coefficients have attracted some attention in the literature on quantitative risk management; see for example McNeil et al. [42]. Notice that $\rho(h)$ coincides with $\gamma_{AA}(h)$ if we choose (a_n) such that $nP(X_0 > a_n) \sim 1$ as $n \to \infty$. Indeed,

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

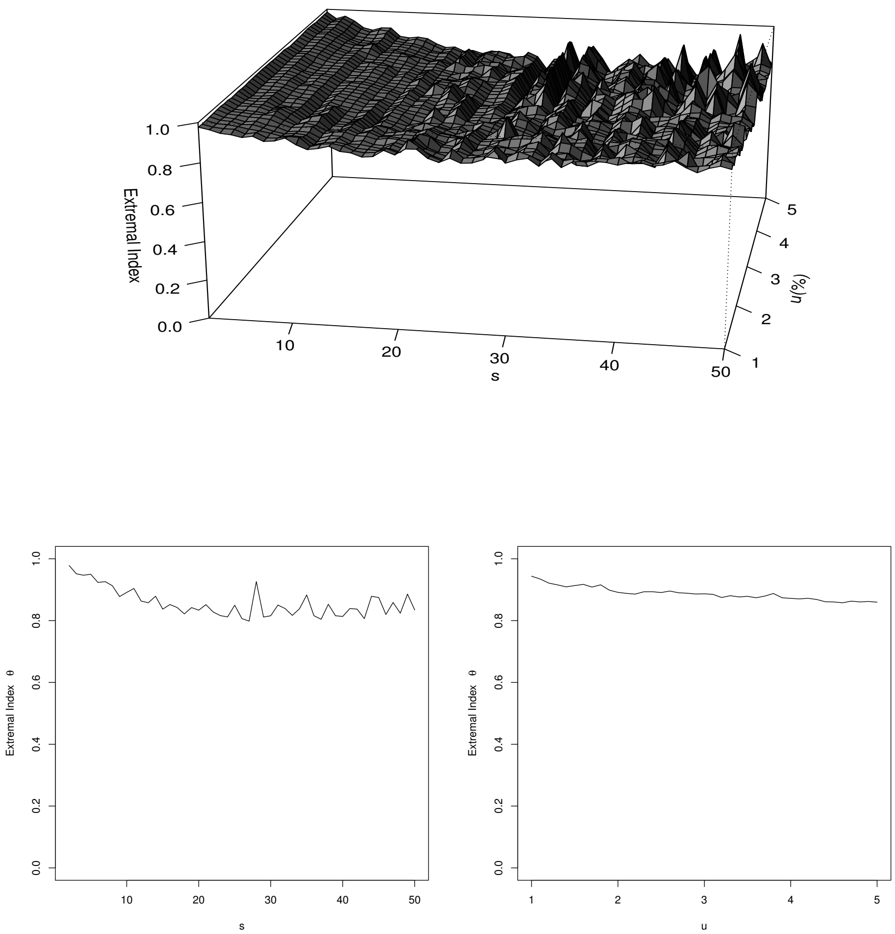

Figure 3.1.1. Blocks estimator $\widehat{\theta}$ of the extremal index θ for 31,757 5-minute log-returns of Bank of America stock prices. The blocks estimator as a function of the block size s and the $u\%$ upper order statistics (top), for fixed $u = 1.9\%$ and running s (bottom left) and for fixed $s = 10$ and running u (bottom right).

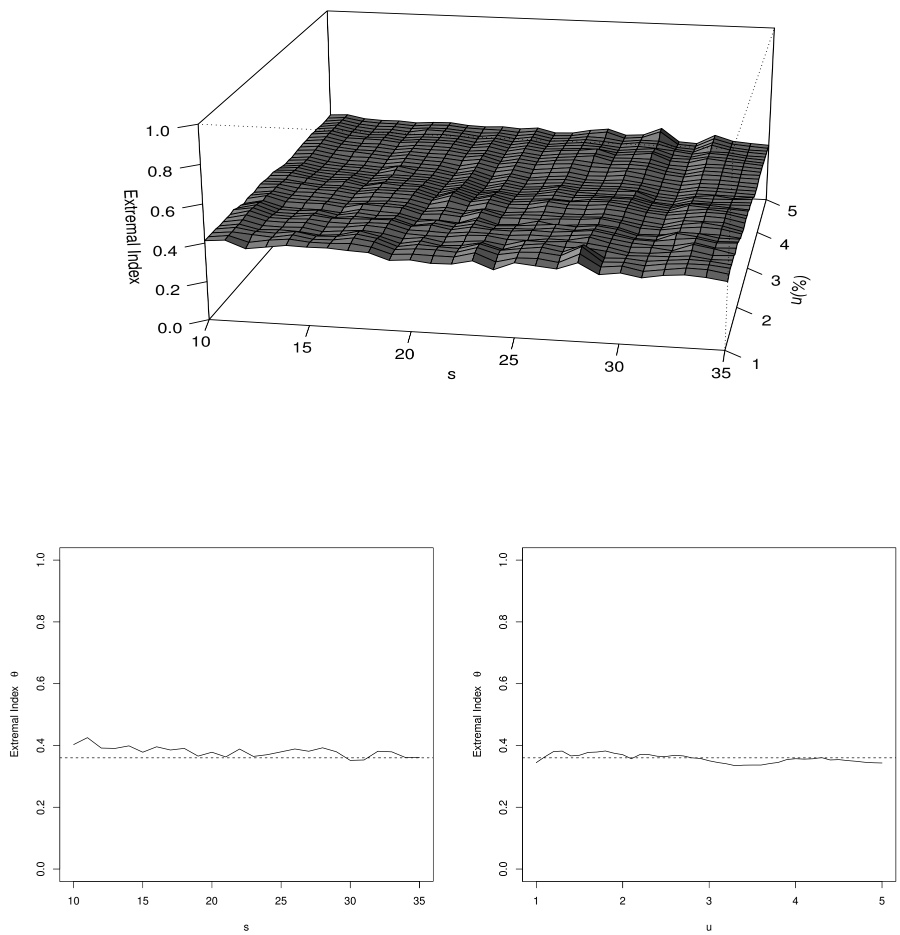

Figure 3.1.2. Blocks estimator $\widehat{\theta}$ of the extremal index θ for a sample of size 20 000 from the AR(1) process $X_t = 0.8\, X_{t-1} + Z_t$. The iid noise (Z_t) has a common student distribution with $\alpha = 2$ degrees of freedom. The extremal index $\theta = 0.37$ is known (indicated by dashed line); see [20], Section 8.1. The blocks estimator as a function of the block size s and $u\%$ of the upper order statistics (top), for fixed $u = 2\%$ and running s (bottom left) and for fixed $s = 24$ and running u (bottom right).

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

$$\begin{pmatrix} \gamma_{AA}(h) & \gamma_{AB}(h) \\ \gamma_{BA}(h) & \gamma_{BB}(h) \end{pmatrix}, \quad h \geq 0\,, \tag{3.1.5}$$

inherits the properties of a matrix covariance function. Notice that the entries of these matrices cannot be negative. The interpretation of (3.1.5) as covariance function allows one to use the classical notions of time series analysis *in an asymptotic sense.* For example, notions such as long or short range dependence of extremal events can be made precise by specifying the rate of decay of (3.1.5) as $h \to \infty$. Davis and Mikosch [16], Mikosch and Zhao [51] introduced an analog of the spectral density as a Fourier transform of the sequence (3.1.5). They showed that the periodogram of the sequence of indicators $I_{\{a_n^{-1} X_t \in A\}}$ of the extremal events $\{a_n^{-1} X_t \in A\}$, $t \in \mathbb{Z}$, has properties similar to the classical periodogram of a stationary sequence. In particular, weighted averages of the periodogram are consistent estimators of the spectral density.

In the literature, the pairwise tail dependence coefficients (3.1.4) are mostly considered for concrete examples of distributions, such as elliptical ones, including the multivariate t- and Gaussian distributions; see e.g. McNeil et al. [42]. In these cases, one can verify that the limits in (3.1.4) exist. For a Gaussian stationary sequence, $\rho(h) = 0$, $h \geq 1$, unless $X_t = X$ a.s. for all $t \in \mathbb{Z}$. The case $\rho(h) = 0$ for some $h \geq 1$ is (again) referred to as *asymptotic extremal dependence* in the vector (X_0, X_h) although no extremal index is in view.

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

Regularly varying random vectors

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

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

$$\frac{P(|X| > tx\,, X/|X| \in \cdot)}{P(|X| > x)} \stackrel{w}{\to} t^{-\alpha}\, P(\Theta \in \cdot)\,, \quad x \to \infty\,. \tag{3.1.6}$$

Relation (3.1.6) can be written in an equivalent form as a pair of conditions:

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

$$P(X/|X| \in \cdot \mid |X| > x) \stackrel{w}{\to} P(\Theta \in \cdot)\,, \quad x \to \infty\,. \tag{3.1.7}$$

In any of these limit relations, it is possible to replace the converging parameter x by a sequence (a_n) such that $P(|X| > a_n) \sim n^{-1}$. Then (3.1.6) and (3.1.7), respectively, read as

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

The convergence relation (3.1.6) can be understood as convergence on the particular Borel sets $\{\mathbf{x} \in \mathbb{R}^d : |\mathbf{x}| > t, \mathbf{x}/|\mathbf{x}| \in S\}$ for Borel sets $S \subset \mathbb{S}^{d-1}$ with a smooth boundary. This convergence can be extended to the Borel σ-field on $\overline{\mathbb{R}}_0^d = \overline{\mathbb{R}}^d \setminus \{\mathbf{0}\}$, $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$:

$$\mu_x(\cdot) = \frac{P(x^{-1}X \in \cdot)}{P(|X| > x)} \stackrel{v}{\to} \mu(\cdot)\,, \quad x \to \infty\,. \tag{3.1.8}$$

Here $\stackrel{v}{\to}$ refers to vague convergence of measures on the Borel σ-field on $\overline{\mathbb{R}}_0^d$, i.e. $\int_{\overline{\mathbb{R}}_0^d} f\, d\mu_x \to \int_{\overline{\mathbb{R}}_0^d} f\, d\mu$ as $x \to \infty$ for any continuous and compactly supported f on $\overline{\mathbb{R}}_0^d$; see Kallenberg [36], Resnick [56]. This means in particular, that the support of f is bounded away from zero. In view of (3.1.6), $\mu(\{\mathbf{x} \in \overline{\mathbb{R}}_0^d : |\mathbf{x}| > t, \mathbf{x}/|\mathbf{x}| \in S\}) = t^{-\alpha}P(\Theta \in S)$, and therefore μ is a Radon measure (i.e. finite on sets bounded away from zero) satisfying $\mu(tA) = t^{-\alpha}\mu(A)$, $t > 0$. In particular, μ does not charge points containing infinite components. Again, the parameter x in (3.1.8) can be replaced by a sequence (a_n) satisfying $P(|X| > a_n) \sim n^{-1}$ and then we get

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

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

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

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

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

Regularly varying stationary sequences

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

$$n\,P(a_n^{-1}(X_1,\ldots,X_h) \in \cdot) \stackrel{v}{\to} \mu_h(\cdot)\,, \quad n \to \infty\,. \tag{3.1.9}$$

Here and in what follows, we will choose the normalizing sequence (a_n) such that $P(|X| > a_n) \sim n^{-1}$, where we use the notation (a_n) in a way different from Section 3.1.2. Indeed, in the latter section we defined the normalization $(a_n^{(h)})$ such that $P(|(X_1,\ldots,X_h)| > a_n^{(h)}) \sim n^{-1}$, but then the normalization would depend on the dimension h. This is not desirable. However, notice that

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

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

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

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

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

For example, transforming the marginals of a Gaussian stationary sequence to unit Fréchet, the resulting sequence is regularly varying with index $\alpha = 1$. We mentioned before that the tail

[3]The notion of regular variation is essentially dimensionless; see for example relation (3.1.6) which immediately extends to normed spaces and, more generally, to metric spaces. An account of the corresponding theory can be found in Hult and Lindskog [32]. Applications of regular variation in function spaces to extreme value theory can be found in de Haan and Tao [26], Davis and Mikosch [12], Meinguet and Segers [43], to large deviations in Hult et al. [33], Mikosch and Wintenberger [49, 50], and to random sets in Mikosch et al. [44, 45].

dependence coefficient $\rho(h) = 0$, $h \geq 1$, for any non-trivial Gaussian stationary sequence. The quantities $\rho(h)$ remain invariant under monotone increasing transformations of the marginals. Hence, the transformed Gaussian distribution with unit Fréchet marginals exhibits asymptotic independence in the sense that the limit measures μ_h are concentrated on the axes.

The tail process

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

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

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

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

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

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

$$\theta = E\big[\sup_{t\geq 0} |\Theta_t|^\alpha - \sup_{t\geq 1} |\Theta_t|^\alpha\big]\,. \qquad (3.1.10)$$

The extremogram revisited

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

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

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

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

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

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

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

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

Examples of regularly varying sequences and their extremograms

In this section, we will introduce some important classes of real-valued strictly stationary regularly varying stationary sequences with index $\alpha > 0$. We will also give the values of the extremogram $\rho(h)$, $h \geq 1$, in (3.1.4). For the calculation of ρ in these examples, we refer to [16, 17, 51].

IID sequence

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

Linear process

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

$$X_t = \sum_{t=0}^{\infty} \psi_j Z_{t-j}\,, \quad t \in \mathbb{Z}\,, \tag{3.1.11}$$

inherits regular variation under conditions on the deterministic sequence (ψ_i) which are close to those dictated by the 3-series theorem, ensuring the a.s. convergence of the series in (3.1.11). This fact was proved in Mikosch and Samorodnitsky [47] for the distribution of X. The regular variation of the finite-dimensional distributions of (X_t) follows since regular variation is preserved under affine transformations of regularly varying vectors. The class (3.1.11) includes causal ARMA processes which are relevant for applications. We refer to Chapter 7 of Embrechts et al. [20] for various applications of regularly varying linear processes.

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

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

Stochastic recurrence equations

Next to linear processes, solutions to the stochastic recurrence equation

$$X_t = A_t\, X_{t-1} + B_t\,, \quad t \in \mathbb{Z}\,, \tag{3.1.12}$$

have attracted some attention. Here (A_t, B_t), $t \in \mathbb{Z}$, is an iid $\mathbb{R}^2$-valued sequence. An a.s unique causal solution to (3.1.12) exists under the moment conditions $E \log A^+ < 0$ and $E \log |B| < \infty$. It follows from work by Kesten [37] and Goldie [37] that X is regularly varying in the precise sense that

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

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

$$E|A|^\kappa = 1 \tag{3.1.13}$$

has a positive solution α (which is unique due to convexity), $EB^\alpha < \infty$ and further regularity conditions on the distribution of A are satisfied. Iteration of (3.1.12) shows that the finite-dimensional distributions of (X_t) are regularly varying with index α. This fact is rather surprising since the distributions of A and B do not need to be heavy-tailed, in contrast to linear processes , where the noise (Z_t) itself has to be heavy-tailed to ensure regular variation of (X_t). We mention that the case of multivariate B and matrix-valued A has also been studied, starting with Kesten [37]; see the recent paper Buraczewski et al. [7].

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

$$\rho(h) = E[\min(1, A_1 \cdots A_h)^\alpha]\,, \quad h \ge 1\,. \tag{3.1.14}$$

Models for returns

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

Stochastic volatility models

The most simple approach is to assume that (σ_t) and (Z_t) be independent. The resulting time series model is frequently referred to as *stochastic volatility model*. Its probabilistic properties are rather simple; see Davis and Mikosch [15]. In particular, regular variation of (X_t) results if $E\sigma^{\alpha+\delta} < \infty$ for some $\delta > 0$ and (Z_t) is iid and regularly varying with index α. The corresponding limit measures μ_h in (3.1.9) are then concentrated on the axes; see Davis and Mikosch [11, 14],[4] and then also $\rho(h) = 0$, $h \ge 1$, as in the iid case. The situation changes if $E|Z|^{\alpha+\delta} < \infty$ for some $\delta > 0$ and (σ_t) is regularly varying with index α. Then (X_t) is regularly varying with index α and extremal clustering for this sequence is possible; see Mikosch and Rezapur [46].

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

GARCH model

Among the models for returns $X_t = \sigma_t Z_t$, $t \in \mathbb{Z}$, the GARCH family gained most popularity. The simplest model of its kind (ARCH) was introduced by Engle [21] and the more sophisticated GARCH model by Bollerslev [4]. For simplicity, we consider the GARCH$(1,1)$ case given by $\sigma_t^2 = \alpha_0 + \sigma_{t-1}^2(\alpha_1 Z_{t-1}^2 + \beta_1)$, $t \in \mathbb{Z}$, where $\alpha_0, \alpha_1, \beta_1$ are positive constants with certain restrictions on the values of $\alpha_1 + \beta_1, \beta_1 < 1$, to ensure strict stationarity. Typical choices are standard normal or unit variance t-distributed Z. Notice that we can write $\sigma_t^2 = A_t \sigma_{t-1}^2 + B_t$, where $B_t = \alpha_0$ and $A_t = \alpha_1 Z_{t-1}^2 + \beta_1$, $t \in \mathbb{Z}$. Therefore regular variation of (σ_t) follows from the corresponding theory for stochastic recurrence equations, and (X_t) inherits the same property; see Davis and Mikosch [10] for the ARCH(1) case, Mikosch and Stărică [48] for the GARCH$(1,1)$, Basrak et al. [1] for GARCH(p,q) and the review paper Davis and Mikosch [13]. Real-life log-returns are typically heavy-tailed. The GARCH model captures this property and this was one of the reasons that it became a benchmark model in financial time series analysis from which numerous other models where derived. An expression of ρ for (X_t^2) is given by (3.1.14) with $A_t = \alpha_1 Z_t^2 + \beta_1$.

Infinite variance stable sequence

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

Max-stable processes

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

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

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

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

$$\Phi_\alpha(x) = \mathrm{e}^{-x^{-\alpha}}\,, \quad x > 0\,. \tag{3.1.15}$$

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

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

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

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

$$P\big(\int_E^\vee f\,dM_\nu^\alpha\le x\big)=\exp\Big\{-x^{-\alpha}\int_E f^\alpha(y)\,\nu(dy)\Big\}\,,\quad x>0\,. \tag{3.1.16}$$

Moreover, for any non-negative $f_i\in L^\alpha(E,\mathcal{E},\nu)$, $i=1,\ldots,d$, by (3.1.16),

$$\begin{aligned}&P\Big(\int_E^\vee f_i\,dM_\nu^\alpha\le x_i\,,i=1,\ldots,d\Big)\\&=P\Big(\int_E\max_{i=1,\ldots,d}\frac{f_i}{x_i}\,dM_\nu^\alpha\le 1\Big)=\exp\Big\{-\int_E\max_{i=1,\ldots,d}\big(\frac{f_i(y)}{x_i}\big)^\alpha\nu(dy)\Big\}\,,\ x_i>0\,.\end{aligned}\tag{3.1.17}$$

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

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

$$X_t=\int_E^\vee f_t\,dM_\nu^\alpha\,,\quad t\in\mathbb{R}\,,\quad f_t\ge 0\,,\quad f_t\in L^\alpha(E,\mathcal{E},\nu)\,. \tag{3.1.18}$$

As in the case of α-stable stationary ergodic processes (Rosiński [61]), the choice of (f_t) is rather sophisticated; see Stoev [66], Kabluchko [34] for details. De Haan [23] showed that any max-stable process with countable index set $T\subset\mathbb{R}$ and stochastically continuous sample paths has representation (3.1.18) and Kabluchko [34] proved this fact for any max-stable process on T for sufficiently rich measure spaces $(E,\mathcal{E},\nu)$. In what follows, we will always assume that the considered max-stable processes have representation (3.1.18).

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

Proposition 3.1.3. *The following statements hold for the skeleton process $(X_t)_{t\in\mathbb{Z}}$ of the process (3.1.18).*

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

$$\mu_h\big((\mathbf{0},\mathbf{x}]^c\big)=\frac{\int_E\max_{i=1,\ldots,h}\big(\dfrac{f_i(y)}{x_i}\big)^\alpha\,\nu(dy)}{\int_E f_0^\alpha(y)\,\nu(dy)}\,. \tag{3.1.19}$$

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

$$\theta = \lim_{n\to\infty} \frac{1}{n} \frac{\int_E \max_{t=1,\ldots,n} f_t^\alpha(y)\,\nu(dy)}{\int_E f_0^\alpha(y)\,\nu(dy)} \tag{3.1.20}$$

exists.

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

$$\gamma_{AB}(h) = \frac{\int_E f_0^\alpha(y) \wedge \big(\frac{a}{b} f_h(y)\big)^\alpha\,\nu(dy)}{a^\alpha \int_E f_0^\alpha(y)\nu(dy)}\,, \quad h\ge 0\,, \tag{3.1.21}$$

and, for $a=b=1$*,*

$$\rho(h) = \frac{\int_E f_0^\alpha(y) \wedge f_h^\alpha(y)\,\nu(dy)}{\int_E f_0^\alpha(y)\,\nu(dy)}\,, \quad h\ge 1\,. \tag{3.1.22}$$

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

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

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

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

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

$$\alpha_h \le c \sum_{s_1=-\infty}^{0} \sum_{s_2=0}^{\infty} \int_E f_0^\alpha(y) \wedge f_{h+s_2}^\alpha(y)\,\nu(dy)\,, \quad h\ge 1\,. \tag{3.1.23}$$

Remark 3.1.4. If the limit in (3.1.20) exists it belongs to the interval $[0,1]$. Indeed, by stationarity of (X_t),

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

Remark 3.1.5. Part (4) is a consequence of Corollary 2.2 in Dombry and Eyi-Minko [19] proved for β-mixing. If $\int_E f_0^\alpha(y)\wedge f_h^\alpha(y)\,\nu(dy) \le c_0\,\mathrm{e}^{-c_1 h}$, $h\ge 1$, for some constants $c_0,c_1>0$, then we conclude that $\alpha_h \le C\,\mathrm{e}^{-c_1 h}$, for some $C>0$.

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

$$n\,P\big(a_n^{-1}(X_1,\ldots,X_h)) \in [\mathbf{0},\mathbf{x}]^c\big) \to \mu_h\big([\mathbf{0},\mathbf{x}]^c\big)\,, \tag{3.1.24}$$

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

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

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

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

An application of (3.1.17) and a Taylor expansion yield (3.1.24) with limit as specified in (3.1.19).
Part (2) Applying (3.1.17) for $x_i = x > 0$, we obtain

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

By definition of the extremal index, the right-hand side must converge to $\Phi_\alpha^\theta(x)$ for some $\theta \in [0,1]$. Equivalently, the limit θ in (3.1.20) exists.
Part (3) As regards the extremogram for sets $A = (a,\infty)$, $B = (b,\infty)$, $a, b > 0$, we have the relation

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

In view of stationarity, $\int_E f_h^\alpha(y)\,\nu(dy) = \int_E f_0^\alpha(y)\,\nu(dy)$. Using a Taylor expansion as $x \to \infty$, we obtain the desired formulas (3.1.21) and (3.1.22)
Part (4) We obtain from Corollary 2.2 in Dombry and Eyi-Minko [19] for any disjoint closed countable subsets S_1, S_2 of $\mathbb{R}$

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

Then (3.1.23) is immediate. □

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

Example 3.1.6. The *Brown-Resnick process* (see [6]) has representation

$$X_t = \sup_{i\ge 1} \Gamma_i^{-1/\alpha}\,\mathrm{e}^{\,W_i(t) - 0.5\,\sigma^2(t)}\,, \quad t \in \mathbb{R}\,, \tag{3.1.25}$$

where (Γ_i) is an enumeration of the points of a unit rate homogeneous Poisson process on $(0,\infty)$ independent of the iid sequence (W_i) of sample continuous mean zero Gaussian processes on $\mathbb{R}$ with stationary increments and variance function σ^2. The max-stable process (3.1.25) is stationary (Theorem 2 in Kabluchko et al. [35]) and its distribution only depends on the variogram $V(h) = \mathrm{var}(W(t+h) - W(t))$, $t \in \mathbb{R}, h \geq 0$. It follows from Example 2.1 in Dombry and Eyi-Minko [19] that the functions (f_t) in representation (3.1.18) satisfy the condition

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

where Φ is the standard normal distribution. For example, if W is standard Brownian motion, $V(h) = h$, $\overline{\Phi}(0.5\sqrt{h}) \sim c\,\mathrm{e}^{-h/8}h^{-0.5}$, as $h \to \infty$. An application of Remark 3.1.5 shows that (α_h) decays at an exponential rate.

Recently, the Brown-Resnick process has attracted some attention for modeling spatio-temporal extremes; see [34, 35, 66, 52]. The processes (3.1.25) can be extended to random fields on $\mathbb{R}^d$. These fields found various applications for modeling spatio-temporal extremal effects; see Kabluchko et al. [35]. For further spatio-temporal applications of max-stable random fields, see also Davis et al. [9].

Example 3.1.7. We consider de Haan and Pereira's [25] *max-moving process*

$$X_t = \sup_{i\geq 1} \Gamma_i^{-1/\alpha} f(t - U_i)\,, \quad t \in \mathbb{R}\,, \tag{3.1.26}$$

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

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

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

For example, if f is the standard normal density, this implies that (α_h) decays to zero faster than exponentially, i.e. the memory in this sequence is very short. In Figure 3.1.2 a simulation of the corresponding process (3.1.26) for $\alpha = 5$ is shown.

3.2 Estimation of the extremogram

3.2.1 Asymptotic theory

Natural estimators of the extremogram are obtained by replacing the probabilities in the limit relations (3.1.3) and (3.1.4) with their empirical counterparts. In this context, one works with quantities which are derived from the *tail empirical process*; see the monographs de Haan and Ferreira [24], Resnick [57] for the underlying theory. For the introduction of the sample extremogram, consider an $\mathbb{R}^d$-valued strictly stationary regularly varying process (X_t) and a Borel set $C \subset \overline{\mathbb{R}}_0^d$ bounded away from zero. Then, for any sequence $m = m_n \to \infty$ with $m_n/n \to 0$ as $n \to \infty$, we define the following estimator of $P_m(C) = m\,P(a_m^{-1}X \in C)$:

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

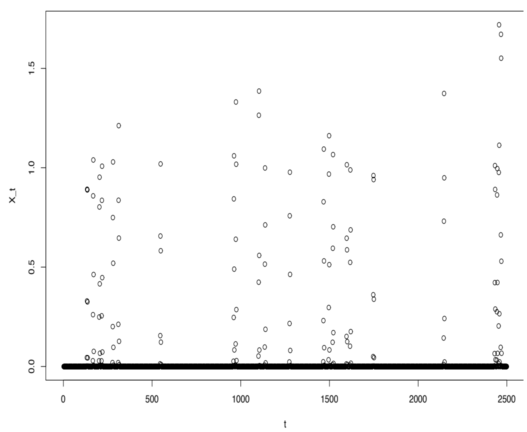

Figure 3.1.8. Max-stable process (3.1.26) where f is the standard normal density and $\alpha = 5$. Extremal clusters are clearly visible.

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

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

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

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

$$\lim_{n\to\infty} m_n \sum_{h=r_n}^{\infty} \alpha_h = 0 \tag{3.2.1}$$

and for every $\epsilon > 0$,

$$\lim_{k\to\infty} \limsup_{n\to\infty} m_n \sum_{h=k}^{r_n} P(|X_h| > \epsilon\, a_m\,, |X_0| > \epsilon\, a_m) = 0\,. \tag{3.2.2}$$

This condition is technical: (3.2.1) imposes some rate on the mixing function (α_h) and (3.2.2) avoids "extremal long range dependence"; (3.2.2) is an asymptotic independence condition in the spirit of the classical condition D'; see Leadbetter et al. [40], Embrechts et al. [20]. The quantities m_n and r_n have some straightforward interpretation as size in large-small block scheme: the sample $X_1, \ldots, X_n$ consists of roughly $[n/m_n]$ large disjoint blocks of size m_n. After chopping off the first r_n elements in each large block one aims at ensuring the asymptotic independence of the resulting large blocks.

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

$$\text{var}[\widehat{P}_m(C)] \sim \frac{m}{n} V(C)\,, \tag{3.2.3}$$

where

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

and we also assume that the infinite series is finite. The asymptotic relation (3.2.3) indicates that the condition $m_n/n \to 0$ is needed to ensure the consistency of the estimator $\widehat{P}_m(C)$. Under additional conditions, $(\widehat{P}_m(C))$ is asymptotically normal and this property also holds jointly for finitely many sets $C_1, \ldots, C_h$. The complicated form of the asymptotic variance in (3.2.3) suggests that it is difficult to apply this central limit theorem for constructing asymptotic confidence bands.

The motivating examples of extremograms are limits of conditional probabilities

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

for sets A, B bounded away from zero. In Section 3.1.2 we mentioned the close relationship of ρ_{AB} with a cross-correlation function. Replacing the probabilities in these conditional probabilities by estimators of the type $\widehat{P}_m(C)$ and applying the corresponding central limit theory from [16], Section 3, and the continuous mapping theorem, one obtains an asymptotic theory for the *ratio estimators*

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

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

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

Theorem 3.2.1. *Let (X_t) be an $\mathbb{R}^d$-valued strictly stationary regularly varying sequence with index $\alpha > 0$. Assume that the following conditions are satisfied.*

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

$$\frac{m_n^4}{n} \sum_{j=r_n}^{m_n} \alpha_j \to 0 \quad \text{and} \quad \frac{m_n r_n^3}{n} \to 0 \quad \text{as } n \to \infty. \tag{3.2.4}$$

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

$$\sqrt{\frac{n}{m_n}}\left[\widehat{\rho}_{AB}(h) - \rho_{AB,m}(h)\right]_{h=0,\ldots,m} \xrightarrow{d} N(\mathbf{0}, (\mu_1(A))^{-4}\mathbf{\Sigma})\,. \tag{3.2.5}$$

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

Some comments are here in place.

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

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

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

- The convergence rate $\sqrt{n/m}$ in the central limit theorem (3.2.5) is due to the triangular array structure; it can be significantly slower than standard $\sqrt{n}$-rates.

- We call $\rho_{AB,m}$ in (3.2.5) the *pre-asymptotic extremogram* since, in general, one cannot replace the centering constants $\rho_{AB,m}(h)$ by their limits $\rho_{AB}(h)$; see Example 3.2.2 below. Moreover, it is in general very difficult to show that

$$\sqrt{\frac{n}{m_n}}|\rho_{AB,m}(h) - \rho_{AB}(h)| \to 0\,, \quad \text{as } n \to \infty, \tag{3.2.6}$$

  even for "nice" models such as the GARCH$(1,1)$. For this well studied model one lacks precise information about the tail behavior. The central limit theorem (3.2.5) (and related bootstrap procedures) are then used to approximate the conditional probabilities $\rho_{AB,m}(h)$. These have a very concrete interpretation in contrast to their less intuitive limits $\rho_{AB}(h)$.

- If (3.2.6) holds with a rather slow convergence rate one faces a bias problem. This problem can be observed e.g. for a simulated stochastic volatility process (X_t) and $A = B = (1, \infty)$. Then $\rho(h) = \rho_{AA}(h) = 0$ for $h \geq 1$. If (X_t) is α-mixing with geometric rate it can be verified that (3.2.6) and (3.2.4) hold for $m_n = n^\gamma$, $\gamma \in (1/3, 1)$, and then (3.2.5) applies with $\rho_{AA,m}(h)$ replaced by $\rho(h) = 0$; see [16], Section 4.2. Also notice that $\widehat{\rho}_{AA}(h)$ is of the order $1/m$ in the iid case and hence greater than zero.

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

- The formulation of the results in [16, 17, 18] related to Theorem 3.2.1 involve various other continuity conditions on sets. These conditions can be avoided as a close inspection of the proofs shows: these conditions follow from continuity of A and B with respect to μ_1. Indeed, one needs that sets of the form $\bigotimes_{i=1}^k C_i$ are μ_k-continuity sets, where $C_i \in \{A, B, \overline{\mathbb{R}}_0^d, A \cap B\}$, $i = 1, \dots, k$, and at least one of the sets C_i does not coincide with $\overline{\mathbb{R}}_0^d$. Let S be the set of indices i such that C_i does not coincide with $\overline{\mathbb{R}}_0^d$. Then

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

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

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

- In applications one needs to choose the threshold a_m in some reasonable way. The choice of a threshold is inherent to extreme value statistics to which no easy solution exists. In [16] we advocated to choose a_m as a fixed high/low empirical quantile of the absolute values of the data and to experiment with several quantile values. If the plot of the sample extremogram is robust for a range of such quantiles one can choose a quantile from that region. The theory in [16, 17, 51] is based on *deterministic* values a_m. The heuristic method described above advocates the choice of a *data dependent* threshold. Recent work by Kulik and Soulier [38] yields a theory for a modified sample extremogram with data dependent threshold in the case of short and long memory stochastic volatility processes.

Example 3.2.2. In [16] we did not provide a concrete example of a sequence (X_t) for which (3.2.6) does not hold. Such counterexamples can easily be constructed from max-stable strictly stationary processes with Fréchet marginals; see Section 3.1.2. We assume the conditions of Proposition 3.1.3. For simplicity choose $\alpha = 1$, $A = B = (1, \infty)$ and $a_m = m \int_E f_0(x)\,\nu(dx)$. The function ρ is given in (3.1.22). By Taylor expansion,

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

for some constant $c_h \neq 0$. Hence

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

and the right-hand side converges to zero if and only if $n^{1/3} = o(m)$ and the rate of convergence to zero can be arbitrarily small. The latter condition is, of course, in contradiction with $m = o(n^{1/3})$ which is one possible sufficient condition for (3.2.5). Fortunately, the other sufficient condition (3.2.4) can still be satisfied if $m = o(n^{1/3})$ does not hold. For example, if α_h decays to zero at a geometric rate and one chooses $r_n = [C \log n]$ for some $C > 0$ and $m_n = n^\gamma$ for some $\gamma \in (1/3, 1)$. Particular cases with geometric decay of (α_h) were mentioned in Examples 3.1.6 and 3.1.7.

3.2.2 Cross-extremogram for bivariate time series

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

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

The asymptotic theory of Section 3.2.1 is applicable to the sets $\widetilde{A} = A \times \mathbb{R}$ and $\widetilde{B} = \mathbb{R} \times B$. In this case, no extremal dependence between the two series would be measured. To avoid these rather uninteresting cases and obtain a more meaningful measure of extremal dependence, we transform the two series so that they have the same marginals. In extreme value theory, the transformation to the unit Fréchet distribution is standard. For the sake of argument, assume that both X_t and Y_t are positive so that the focus of attention will be on extremal dependence in the upper tails. The case of extremal dependence in the lower tails or upper and lower tails is similar. Under the positivity constraint, if F_1 and F_2 denote the distribution functions of X_t and Y_t, respectively, and are continuous, then the two transformed series, $\widetilde{X}_t = G_1(X_t)$ and $\widetilde{Y}_t = G_2(Y_t)$ with $G_i(z) = -1/\log(F_i(z))$, $i = 1, 2$, have unit Fréchet marginals Φ_1; see (3.1.15). Now assuming that the bivariate time series $((\widetilde{X}_t, \widetilde{Y}_t))_{t\in\mathbb{Z}}$ is regularly varying, we define the *cross-extremogram* by

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

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

$$\rho_{\widetilde{A}\widetilde{B}}(h) \quad = \quad \lim_{n\to\infty} P(Y_h \in a_{Y,n}B \mid X_0 \in a_{X,n}\,A)\,. \tag{3.2.7}$$

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

3.3 An example: Equity indices

We consider daily log-returns equity indices of four countries: S&P 500 for US, FTSE 100 for UK, DAX for Germany, Nikkei 225 for Japan. Figure 3.3.1 shows the sample extremogram $\widehat{\rho} = \widehat{\rho}_{AA}$ for the negative tails ($A = B = (-\infty, -1)$ with a_m estimated as the 96% empirical quantiles of the absolute values of the negative data) applied to 6,443 log-returns of the FTSE and S&P (April 4, 1984 to October 2, 2009), to 4,848 log-returns of the DAX (November 13, 1990 to October 2, 2009) and to 6,333 log-returns of the Nikkei (August 23, 1984 to October 2, 2009).[7] The solid horizontal lines in the plots represent 98% confidence bands. They correspond to the maximum and minimum of the sample extremogram at lag 1 based on 99 random permutations of the data. If the data were independent random permutations would not change the dependence structure: values $\widehat{\rho}_{AB}(h)$ which are outside the confidence bands indicate that there is significant extremal dependence at lag h. The sample extremograms for all four indices decay rather slowly to zero, with S&P the slowest. Among the four indices, the Nikkei displays the least amount of extremal dependence as measured by the extremogram. The top graphs in Figure 3.3.1 indicate extremal dependence in the lower tail over a period of 40 days.

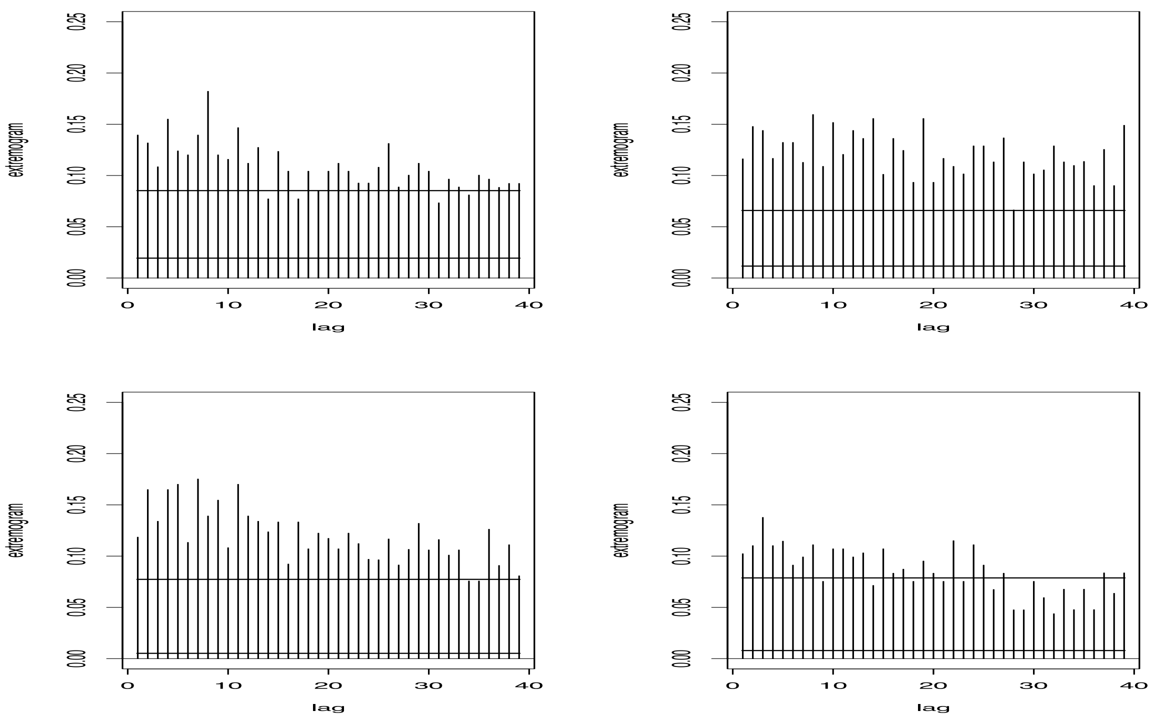

Figure 3.3.1. *The sample extremogram for the lower tails of the FTSE* (top left)*, S&P* (top right)*, DAX* (bottom left) *and Nikkei. The solid lines represent 98% confidence bands based on* 99 *random permutations of the data.*

We assume that the log-return series are modeled by a GARCH$(1,1)$ process (see Section 3.1.2 for its definition). Then we can estimate its parameters, calculate the volatility sequence $(\widehat{\sigma}_t)$ and

[7]As noted in the literature, the lower tails of returns tend to be heavier than the upper tails. Similar plots (not shown here) of the sample extremogram for the upper tails also reveal extremal dependence, but to a lesser extent than seen in the lower tails.

the *filtered* sequence $\widehat{Z}_t = X_t/\widehat{\sigma}_t$, $t = 1, \ldots, n$. Figure 3.3.2 shows the sample extremograms $\widehat{\rho}$ for the filtered FTSE and S&P sequences. These plots confirm that much of the extremal dependence (as measured by the extremogram) has been removed. Hence the extremal dependence in the log-returns is due to the volatility sequences (σ_t), as suggested by the GARCH model.

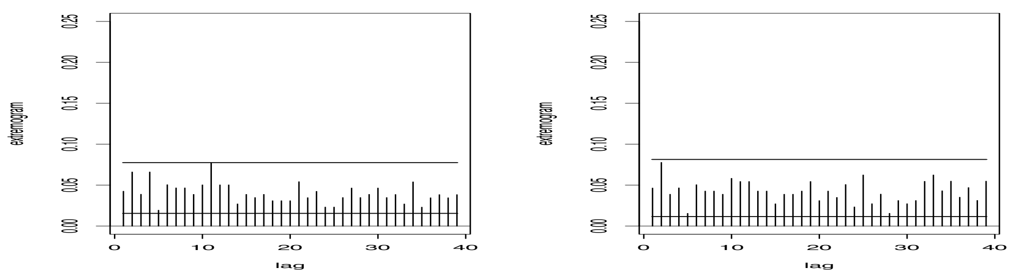

Figure 3.3.2. *The sample extremograms for the filtered FTSE* (left) *and filtered S&P* (right) *series. The bold lines represent 98% confidence bounds based on* 99 *random permutations of the series.*

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

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

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

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

The sample cross-extremogram of any pair of series exhibits a similar pattern of slow decay as seen in the univariate sample extremograms (we do not include these figures). Figure 3.3.3 shows the sample cross-extremograms for the filtered series. For example, in the first row of graphs, (X_t) is the filtered FTSE and (Y_t) are the filtered S&P, DAX and Nikkei, respectively. There are signs of various types of cross-extremal dependence in the filtered series. The spikes at lag zero (except between the Nikkei and S&P) indicate the strong extremal dependence of the multiplicative shocks, In the second row, there is evidence of significant extremal dependence at lag one for each sample cross-extremogram: given the S&P has an extreme left tail event in a shock at time t there will be a corresponding large left tail shock in the FTSE, the DAX and the Nikkei at time $t = 1$. Given the dominance of the US stock market, one might expect a carry-over effect of the shocks on the other exchanges on the next day. Since only marginal GARCH models were fitted to the data, it may not seem all that surprising that the filtered series exhibit serial dependence. We should note, however, that the dependence in the shocks does not appear to last beyond one time lag.

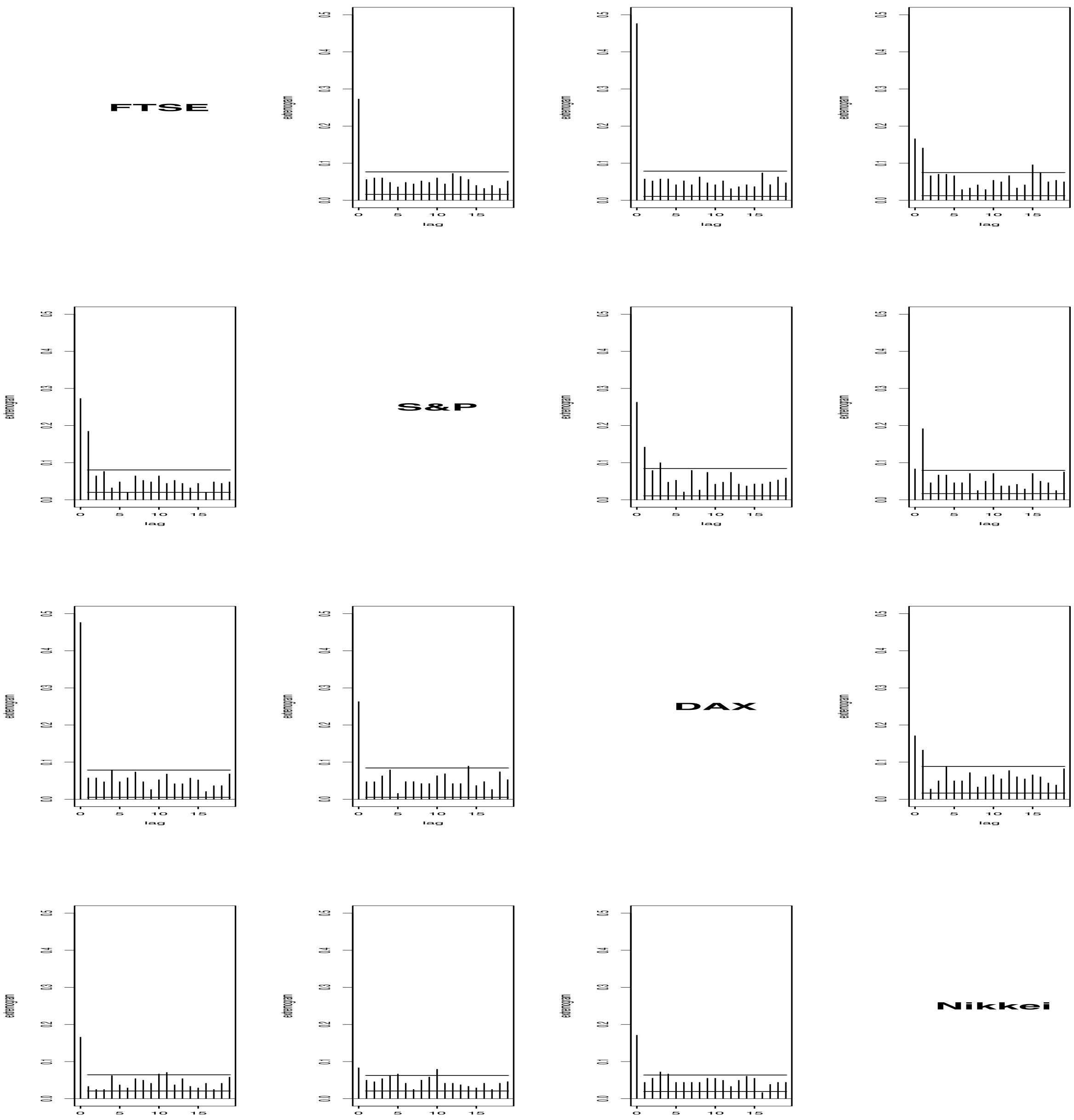

Figure 3.3.3. *The sample cross-extremograms for the filtered FTSE, S&P, DAX and Nikkei series. For the first row, (X_t) is the filtered FTSE and (Y_t) are the filtered S&P, DAX and Nikkei (from left to right). For the second, third and fourth rows, the X_t's are the filtered S&P, DAX and Nikkei series, respectively.*

3.4 A Fourier analysis of extreme events

Classical time series analysis studies the second order properties of stationary processes in the time and frequency domains. The latter approach refers to spectral (or Fourier) analysis of the time series. We mentioned in Section 3.1.2 that the extremograms $\gamma_A = \gamma_{AA}$ and $\rho_A = \rho_{AA}$ for a set A bounded away from zero are covariance and correlations functions of some stationary sequence. Then it is possible to study the corresponding spectral properties of the extremogram. Research in this direction was started in [16] and continued in [51]. We recall some of the results. Throughout we assume that (X_t) is a $\mathbb{R}^d$-valued strictly stationary regularly varying sequence with index $\alpha > 0$ and A is a μ_1-continuity set. In the comments following Theorem 3.2.1 we mentioned the latter property also implies that the sets $A \times (\overline{\mathbb{R}}_0^d)^h \times A$ are μ_h-continuity sets for $h \geq 0$, so the limits $\rho_A(h)$ exist.

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

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

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

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

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

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

We will refer to I_{nA} and its standardized version $\widetilde{I}_{nA}$ as *periodogram of the extreme event* $a_m\,A$. Indeed, if we replaced the normalization m/n by $1/n$, I_{nA} is the periodogram of the sequence of centered indicators $(I_{\{a_m^{-1}X_t\in A\}} - p_0)$. These sequences constitute a triangular array of row-wise strictly stationary sequences for which standard asymptotic theory for the periodogram does not apply; for an asymptotic theory of the periodogram of a strictly stationary linear process, see Brockwell and Davis [5], Chapter 10. However, the periodogram of extreme events shares some of the basic properties of the periodogram, as the following results from Mikosch and Zhao [51] show.

Theorem 3.4.1. *Let (X_t) be an $\mathbb{R}^d$-valued strictly stationary regularly varying sequence with index $\alpha > 0$ satisfying condition* (M)*, $A \subset \overline{\mathbb{R}}_0^d$ be a μ_1-continuity set and $\sum_{h\geq 1}\rho_A(h) < \infty$.*

[8] The centering of the indicator functions with their expectation p_0 is crucial for deriving asymptotic theory. In applications, $I_{nA}(\lambda)$ is typically evaluated at the Fourier frequencies $\omega_j(n) = 2\pi j/n \in (0,\pi)$ and since $\sum_{h=1}^{n}\mathrm{e}^{-i\,h\,\omega_j(n)} = 0$, centering in $I_{nA}(\omega_j(n))$ is not needed.

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

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

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

$$\begin{aligned}\big(I_{nA}(\lambda_i)\big)_{i=1,\ldots,N} &\stackrel{d}{\to} \mu_1(A)\,\big(f_A(\lambda_i)E_i\big)_{i=1,\ldots,N}\,, \quad n\to\infty\,,\\ \big(\widetilde{I}_{nA}(\lambda_i)\big)_{i=1,\ldots,N} &\stackrel{d}{\to} \big(f_A(\lambda_i)E_i\big)_{i=1,\ldots,N}\,, \quad n\to\infty\,.\end{aligned}$$

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

$$\begin{aligned}\big(I_{nA}(\omega_i(n))\big)_{i=1,\ldots,N} &\stackrel{d}{\to} \mu_1(A)\,\big(f_A(\lambda_i)E_i\big)_{i=1,\ldots,N}\,, \quad n\to\infty\,,\\ \big(\widetilde{I}_{nA}(\omega_i(n))\big)_{i=1,\ldots,N} &\stackrel{d}{\to} \big(f_A(\lambda_i)E_i\big)_{i=1,\ldots,N}\,, \quad n\to\infty\,.\end{aligned}$$

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

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

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

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

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

Introduce the corresponding *smoothed periodogram*

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

Under the conditions of the second item in Theorem 3.4.1 and some further restrictions on the growth of (m_n) and (α_h) the following limit relations hold for a fixed frequency $\lambda \in (0,\pi)$,

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

In Figure 3.4 we illustrate how the smoothed periodogram $\widehat{f}_{n,(-\infty,-1)}$ works for 5-minute log-returns of Bank of America stock prices. We choose a simple Daniell window with $w_n(j) = 1/(2s_n+1)$ and $s_n = 52$.

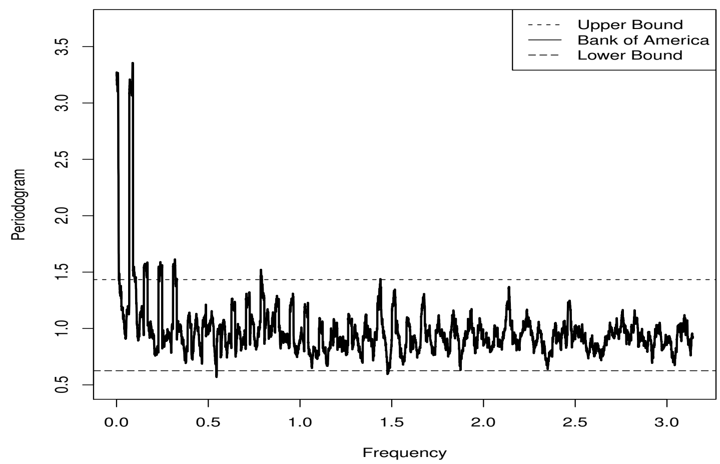

Figure 3.4.2. The smoothed periodogram (corresponding to the losses) for 31, 757 5-minute log-returns of Bank of America stock prices with Daniell window, $s_n = 52$. The simultaneous confidence bands are constructed by taking the 97.5% quantile of the maxima and the 2.5% quantile of the minima over the Fourier frequencies calculated from the smoothed periodograms of 10 000 random permutations of the data. If the data were iid, permutations would not change the dependence structure. The fact that the periodogram is outside the confidence bands at various frequencies indicates that there is significant extremal dependence in the data. The peaks at various frequencies show that there are cycles of extremal behavior in the data. These cycles cannot be detected by autocorrelation plots of the data, their absolute values or squares.

Bibliography

[1] Basrak, B., Davis, R.A. and Mikosch. T. (2002) Regular variation of GARCH processes. *Stoch. Proc. Appl.* **99**, 95–116.

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

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

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

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

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

[7] Buraczewski, D., Damek, E., Guivarc'h, Y., Hulanicki, A. and Urban, R. (2009) Tail-homogeneity of stationary measures for some multidimensional stochastic recursions. *Probab. Th. Rel. Fields* **145**, 385–420.

[8] Davis, R.A. and Hsing, T. (1995) Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Prob.* **23**, 879–917.

[9] Davis, R.A., Klüppelberg, C. and Steinkohl, C. (2012) Statistical inference for max-stable processes in space and time. *J. Royal Statist. Soc., Series B*, to appear.

[10] Davis, R.A. and Mikosch, T. (1998) Limit theory for the sample ACF of stationary process with heavy tails with applications to ARCH. *Ann. Statist.* **26**, 2049–2080.

[11] Davis, R.A. and Mikosch, T. (2001) Point process convergence of stochastic volatility processes with application to sample autocorrelation. *J. Appl. Probab* **38A**, 93–104.

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

[13] Davis, R.A. and Mikosch, T. (2009) Extreme value theory for GARCH processes. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.) *Handbook of Financial Time Series.* Springer, pp. 187–200.

[14] Davis, R.A. and Mikosch, T. (2009) Extremes of stochastic volatility models. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.) *Handbook of Financial Time Series.* Springer (2009), pp. 355–364.

[15] Davis, R.A. and Mikosch, T. (2009) Probabilistic properties of stochastic volatility models. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.) *Handbook of Financial Time Series.* Springer (2009), pp. 255–268.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[31] HSING, T., HÜSLER, J. AND LEADBETTER, M.R. (1988) On the exceedance point process for a stationary process. *Probab. Th. Rel. Fields* **78**, 97–112.

[32] HULT, H. AND LINDSKOG, F. (2006) Regular variation for measures on metric spaces. *Publ. Inst. Math. (Beograd) (N.S.)* **80(94)**, 121–140.

[33] HULT, H., LINDSKOG, F., MIKOSCH, T. AND SAMORODNITSKY, G. (2005) Functional large deviations for multivariate regularly varying random walks. *Ann. Appl. Probab.* **15**, 2651–2680..

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

[35] KABLUCHKO, Z., SCHLATHER, M. AND HAAN, L. DE (2009) Stationary max-stable fields associated to negative definite functions. *Ann. Probab.* **37**, 2042–2065.

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

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

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

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

[40] LEADBETTER, M.R., LINDGREN, G. AND ROOTZÉN, H. (1983) *Extremes and Related Properties of Random Sequences and Processes.* Springer, Berlin.

[41] LEADBETTER, M.R. AND ROOTZÉN, H. (1988) Extremal theory for stochastic processes. *Ann. Probab.* **16**, 431–478.

[42] MCNEIL, A., FREY, R. AND EMBRECHTS, P. (2005) *Quantitative Risk Management: Concepts, Techniques, and Tools.* Princeton University Press, Princeton.

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

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

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

[46] MIKOSCH, T. AND REZAPUR, M. (2012) Stochastic volatility models with possible extremal clustering. *Bernouli*, to appear.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[64] SAMORODNITSKY, G. AND TAQQU, M.S. (1994) *Stable Non-Gaussian Random Processes.* Chapman & Hall, New York.

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

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

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

Chapter 4

The integrated periodogram of a dependent extremal event sequence

Abstract

We investigate the asymptotic properties of the integrated periodogram calculated from a sequence of indicator functions of dependent extremal events. An event in Euclidean space is extreme if it occurs far away from the origin. We use a regular variation condition on the underlying stationary sequence to make these notions precise. Our main result is a functional central limit theorem for the integrated periodogram of the indicator functions of dependent extremal events. The limiting process is a continuous Gaussian process whose covariance structure is in general unfamiliar, but in the iid case a Brownian bridge appears. We indicate how the developed theory can be used to detect periodic cycles of extremes in a stationary sequence.

4.1 Introduction

4.1.1 Regularly varying sequences

We consider a strictly stationary $\mathbb{R}^d$-valued sequence (X_t) for some $d \geq 1$ with a generic element X and assume that its finite-dimensional distributions are regularly varying. This means that for every $h \geq 1$, there exists a non-null Radon measure μ_h on the Borel σ-field $\overline{\mathcal{B}}_0^{dh}$ of $\overline{\mathbb{R}}_0^{dh} = \overline{\mathbb{R}}^{dh} \backslash \{0\}$, $\overline{\mathbb{R}} = \{-\infty, \infty\}$, such that

$$\frac{P(x^{-1}(X_1, \ldots, X_h) \in \cdot)}{P(|X| > x)} \stackrel{v}{\rightarrow} \mu_h(\cdot)\,, \tag{4.1.1}$$

where $\stackrel{v}{\rightarrow}$ denotes vague convergence in $\overline{\mathcal{B}}_0^{dh}$; cf. Resnick [23, 24], Kallenberg [18]. The limiting measure μ_h necessarily has the property $\mu_h(t\cdot) = t^{-\alpha}\mu_h(\cdot)$, $t > 0$, for some $\alpha \geq 0$, the index of regular variation. In what follows, we assume that $\alpha > 0$. Relation (4.1.1) is equivalent to the sequential definition

$$n\,P(a_n^{-1}(X_1,\ldots,X_h)\in\cdot)\stackrel{v}{\to}\mu_h(\cdot)\,,\quad n\to\infty\,,\qquad(4.1.2)$$

where (a_n) is chosen such that $P(|X|>a_n)\sim n^{-1}$ as $n\to\infty$. We will say that the sequence (X_t) and any of the vectors $(X_1,\ldots,X_h)$, $h\ge 1$, are *regularly varying with index* α.

Examples of regularly varying strictly stationary sequences are linear and stochastic volatility processes with iid regularly varying noise, GARCH processes, infinite variance stable processes and max-stable processes with Fréchet marginals. These examples are discussed in Davis et al. [7, 9, 10], Mikosch and Zhao [19].

4.1.2 The extremogram

Consider a μ_1-continuity Borel set $D_0 = A\subset\overline{\mathbb{R}}_0^d$ bounded away from zero and such that $\mu_1(A)>0$. Then the sets $D_h = A\times\overline{\mathbb{R}}^{d(h-1)}\times A$ are bounded away from zero as well and are continuity sets with respect to the corresponding limiting measures μ_{h+1}, $h\ge 1$. We conclude from (4.1.2) that the limits

$$\gamma_A(h)=\lim_{n\to\infty}n\,P(a_n^{-1}X_0\in A\,,a_n^{-1}X_h\in A)\to\mu_{h+1}(D_h)\,,\quad h\ge 0\,,\qquad(4.1.3)$$

exist. For $t\in\mathbb{Z}$, it is not difficult to see that

$$\begin{aligned} n\,\mathrm{cov}(I_{\{a_n^{-1}X_t\in A\}},I_{\{a_n^{-1}X_{t+h}\in A\}}) &\sim nEI_{\{a_n^{-1}X_t\in A\,,a_n^{-1}X_{t+h}\in A\}}\\ &= n\,P(a_n^{-1}X_0\in A\,,a_n^{-1}X_h\in A)\\ &\to \gamma_A(h)\,,\quad n\to\infty\,. \end{aligned}$$

Hence γ_A constitutes the covariance function of a stationary process. We refer to γ_A as the *extremogram relative to the set A*. We will also consider the *standardized extremogram* given as the limiting sequence

$$\rho_A(h)=\lim_{n\to\infty}P(a_n^{-1}X_h\in A\mid a_n^{-1}X_0\in A)=\frac{\mu_{h+1}(D_h)}{\mu_1(D_0)}\,,\quad h\ge 0\,.$$

The quantities $\rho_A(h)$ have an intuitive interpretation as limiting conditional probabilities. Moreover, ρ_A is the autocorrelation function of a stationary process. The quantities $\rho_A(h)$ are generalizations of the upper tail dependence coefficient of a two-dimensional vector (Y_1,Y_2) with identical marginals given as the limit $\lim_{x\to\infty}P(Y_2>x\mid Y_1>x)$.

The extremogram was introduced in Davis and Mikosch [7] as a measure of serial extremal dependence in a strictly stationary sequence. There and in Davis et al. [9, 10] various aspects of the estimation of the extremogram were discussed, including asymptotic theory and the use of the stationary bootstrap for the construction of confidence bands.

4.1.3 The sample extremogram

Natural estimators of the extremograms γ_A and ρ_A are given by their respective sample analogs

$$\widetilde{\gamma}_A(h)\;=\;\frac{m}{n}\sum_{t=1}^{n-h}\widetilde{I}_t\widetilde{I}_{t+h}\quad\text{and}\quad\widetilde{\rho}_A(h)=\frac{\widetilde{\gamma}_A(h)}{\widetilde{\gamma}_A(0)}\,,\quad h\ge 0\,.$$

Here $m = m_n$ is any integer sequence satisfying the conditions $m_n \to \infty$ and $m_n/n = o(1)$ and

$$I_t = I_{\{a_m^{-1} X_t \in A\}}\,, \quad \widetilde{I}_t = I_t - p_0\,, \quad \text{and} \quad p_0 = EI_t = P(a_m^{-1} X \in A)\,, \quad t \in \mathbb{Z}\,.$$

It is shown in Davis and Mikosch [7] that the conditions $m_n \to \infty$ and $m_n/n = o(1)$ are needed for the validity of the asymptotic properties $E\widetilde{\gamma}_A(h) \to \gamma_A(h)$ and $\mathrm{var}(\widetilde{\gamma}_A(h)) \to 0$ as $n \to \infty$. Moreover, under a mixing condition, the finite-dimensional distributions of $\widetilde{\gamma}_A$ and $\widetilde{\rho}_A$ satisfy a central limit theorem with normalization $(n/m)^{1/2}$; cf. Lemma 4.2.4 below.

4.1.4 Spectral density and periodogram

Since γ_A and ρ_A are the autocovariance and autocorrelation functions of a stationary process, respectively, it is possible to enter the corresponding frequency domain. If γ_A is square summable one can define the spectral densities

$$h_A(\lambda) = \sum_{h \in \mathbb{Z}} \gamma_A(h)\, \mathrm{e}^{-ih\lambda} \quad \text{and} \quad f_A(\lambda) = \sum_{h \in \mathbb{Z}} \rho_A(h)\, \mathrm{e}^{-ih\lambda}\,, \quad \lambda \in [0,\pi] = \Pi\,.$$

A natural estimator of the spectral density is the periodogram. Since the sample autocovariances $\widetilde{\gamma}_A(h)$ are derived from the triangular array of the stationary sequences $(\widetilde{I}_t)$, an analog of the classical periodogram for h_A is given by

$$I_{nA}(\lambda) \quad = \quad \frac{m}{n}\Big|\sum_{t=1}^{n} \widetilde{I}_t\, \mathrm{e}^{-i\,t\,\lambda}\Big|^2 = \widetilde{\gamma}_A(0) + 2\sum_{h=1}^{n-1} \widetilde{\gamma}_A(h)\cos(h\lambda)\,, \quad \lambda \in \Pi\,,$$

and the periodogram for the standardized spectral density f_A is obtained as the scaled periodogram $I_{nA}/\widetilde{\gamma}_A(0)$. Mikosch and Zhao [19] showed under mixing conditions that the *extremal periodogram ordinates* $I_{nA}(\lambda)$ share various of the classical properties of the periodogram ordinates for a stationary sequence (cf. Brockwell and Davis [3]): consistency in the mean, convergence in distribution to independent exponential random variables with expectation $h_A(\lambda_j)$ at distinct fixed frequencies $\lambda_j \in (0,\pi)$ and at distinct Fourier frequencies $\omega_n(j) = 2\pi j/n \in (0,\pi)$ provided these frequencies converge to a limit $\lambda_j \in (0,\pi)$ as $n \to \infty$. The latter property ensures that weighted versions of the periodogram I_{nA} at fixed frequencies $\lambda \in (0,\pi)$ converge in mean square to $h_A(\lambda)$.

For practical purposes, one will mostly work with the periodogram at the Fourier frequencies $\omega_n(j) \in (0,\pi)$. Then

$$I_{nA}(\omega_n(j)) = \frac{m}{n}\Big|\sum_{t=1}^{n} I_t\, \mathrm{e}^{-i\,t\,\omega_n(j)}\Big|^2\,,$$

i.e., centering of the indicator functions I_t is not needed. However, for proving asymptotic theory it will be convenient to work with the extremal periodogram I_{nA} based on the centered quantities $\widetilde{I}_t$, $t = 1,\dots,n$.

4.1.5 The integrated periodogram

The integrated periodogram of a stationary sequence has a long history in time series analysis, starting with classical work of Grenander and Rosenblatt [13], and was extensively used in the monographs Hannan [14], Priestley [22], Brockwell and Davis [3], to name a few references. Dahlhaus [4] discovered a close relationship of the integrated periodogram, considered as a process indexed by functions, and empirical process theory. Under entropy conditions, he proved uniform convergence results over suitable classes of index functions; see also the survey paper Dahlhaus and Polonik [5]. These papers gave some general theoretical background for various periodogram based techniques such as Whittle estimation of the parameters of a FARIMA process and goodness of fit tests for linear processes as mentioned in Grenander and Rosenblatt [13] and Priestley [22].

In this paper, we will consider the integrated periodogram

$$J_{nA}(g) = \int_\Pi I_{nA}(\lambda)\, g(\lambda)\, d\lambda = c_0(g)\, \widetilde{\gamma}_A(0) + 2\sum_{h=1}^{n-1} c_h(g)\, \widetilde{\gamma}_A(h)\,, \tag{4.1.4}$$

and its standardized version

$$J^\circ_{nA}(g) = \frac{1}{\widetilde{\gamma}_A(0)} \int_\Pi I_{nA}(\lambda)\, g(\lambda)\, d\lambda = c_0(g) + 2\sum_{h=1}^{n-1} c_h(g)\, \widetilde{\rho}_A(h)\,,$$

where g is non-negative and square integrable with respect to Lebesgue measure on Π (we write $g \in L^2_+(\Pi)$) with corresponding Fourier coefficients

$$c_h(g) = \int_\Pi \cos(h\lambda)\, g(\lambda)\, d\lambda\,, \quad h \in \mathbb{Z}\,.$$

We will understand $J_{nA}(g)$ and $J^\circ_{nA}(g)$ as natural estimators of

$$J_A(g) = \int_\Pi h_A(\lambda)\, g(\lambda)\, d\lambda = c_0(g)\, \gamma_A(0) + 2\sum_{h=1}^{\infty} c_h(g)\, \gamma_A(h)\,, \tag{4.1.5}$$

$$J^\circ_A(g) = \int_\Pi f_A(\lambda)\, g(\lambda)\, d\lambda = c_0(g) + 2\sum_{h=1}^{\infty} c_h(g)\, \rho_A(h)\,,$$

respectively. The latter identities holds if $\sum_{h=0}^\infty \gamma_A(h) < \infty$, a condition we assume throughout this paper.

The main results of this paper (see Section 4.3) are functional central limit theorems for the integrated periodogram J_{nA} with $g = hI_{[0,\cdot]}$ for a sufficiently smooth function h on Π. The limit processes are Gaussian whose covariance structure strongly depends on the limit measures (μ_h). The rate of convergence in these results is typically slower than $\sqrt{n}$. However, in the case of an iid sequence, the limiting process is a Brownian bridge and the convergence rates are much faster than in the case of a dependent sequence. These results differ from classical theory for the periodogram of a stationary sequence (X_t) (see e.g. Dahlhaus [4], Klüppelberg and Mikosch [17]), where the limiting process is completely determined by the covariance structure of (X_t).

The methods of proof combine classical techniques of weak convergence and strong mixing (e.g. Billingsley [1]) with extreme value theory for dependent sequences (e.g. Davis and Mikosch [7]). The proofs are rather technical.

The paper is organized as follows. We start in Section 4.2 with some moment calculations and we also introduce the relevant mixing conditions and central limit theory for the sample extremogram. In Section 4.2.4 we provide a result about the mean square consistency of the integrated periodogram; the proof is given in Section 4.5. The main results (Theorems 4.3.1 and 4.3.3) are functional central limit theorems for the integrated periodogram. They are given in Section 4.3; the corresponding proofs are provided in Sections 4.6 and 4.7. The covariance structure of the limiting Gaussian processes in Theorem 4.3.1 is rather complicated. Therefore in Section 4.4 we supplement the asymptotic theory by consistency results for the stationary bootstrap applied to the integrated periodogram of extremal events in a strictly stationary sequence. The corresponding proofs are given in Section 4.8.

4.2 Preliminaries

4.2.1 Some moment calculations

Recall the notation and conditions of Section 4.1. We write

$$p_0 = P(a_m^{-1} X_0 \in A) \quad \text{and} \quad p_h = P(a_m^{-1} X_0 \in A, a_m^{-1} X_h \in A)\,, \quad h \geq 1\,,$$

where as above, $m_n \to \infty$ and $m_n/n = o(1)$ as $n \to \infty$. For integers $s, t, u, v \geq 0$, we set

$$\begin{aligned} \Gamma(s,t,u,v) &= E\widetilde{I}_s\widetilde{I}_t\widetilde{I}_u\widetilde{I}_v\,, \\ \Gamma(s,t,u) &= E\widetilde{I}_s\widetilde{I}_t\widetilde{I}_u\,, \\ \Gamma(s,t) &= E\widetilde{I}_s\widetilde{I}_t = p_{|s-t|} - p_0^2\,. \end{aligned}$$

We will often have to calculate variances and covariances of the sample extremogram $\widetilde{\gamma}_A$. We provide some of these formulas for further use.

Lemma 4.2.1. *Let (X_t) be a strictly stationary sequence. Then, for $1 \leq h \leq n-1$,*

$$(n/m)^2 E\widetilde{\gamma}_A^2(h) = (n-h)E(\widetilde{I}_0\widetilde{I}_h)^2 + 2\sum_{t=1}^{n-h-1}(n-h-t)\Gamma(0,h,t,t+h)$$

and for $1 \leq h < h+u \leq n-1$,

$$\begin{aligned} (n/m)^2\, E\widetilde{\gamma}_A(h)\widetilde{\gamma}_A(h+u) &= (n-h-u)\Gamma(0,h,0,h+u) \\ &\quad + \sum_{t=1}^{n-h-u-1}(n-h-u-t)\Gamma(0,h,t,t+h+u) \\ &\quad + \sum_{t=1}^{n-h-1}\min(n-h-u,n-h-t)\Gamma(0,h+u,t,t+h)\,. \end{aligned}$$

4.2.2 Mixing conditions

The following two mixing conditions were introduced in Davis and Mikosch [7] for a strongly mixing $\mathbb{R}^d$-valued sequence (X_t) with rate function (ξ_h).

Condition (M)

There exist integer sequences $m = m_n \to \infty$ and $r_n \to \infty$ such that $m_n/n \to 0$, $r_n/m_n \to 0$ and

$$\lim_{n\to\infty} m_n \sum_{h=r_n}^{\infty} \xi_h \quad = \quad 0\,, \tag{4.2.1}$$

Moreover, an anti-clusterng condition holds:

$$\lim_{k\to\infty} \limsup_{n\to\infty} \sum_{h=k}^{r_n} P(|X_h| > \epsilon\, a_m \mid |X_0| > \epsilon\, a_m) \quad = \quad 0\,, \quad \epsilon > 0\,. \tag{4.2.2}$$

Condition (M1)

Assume (M) and that the sequences (m_n), (r_n), $k_n = [n/m_n]$ from (M) also satisfy the growth conditions $k_n \xi_{r_n} \to 0$, and $m_n = o(n^{1/3})$.

Remark 4.2.2. The condition $m_n = o(n^{1/3})$ in (M1) can be replaced by

$$\frac{m_n^4}{n} \sum_{j=r_n}^{m_n} \xi_j \to 0 \quad \text{and} \quad \frac{m_n r_n^3}{n} \to 0\,,$$

which is often much weaker.

Condition (4.2.1) is easily satisfied if the mixing rate (ξ_h) is geometric, i.e. exponentially decaying to zero. Under mild conditions, the popular classes of ARMA, max-stable, GARCH and stochastic volatility processes are strongly mixing with geometric rate; cf. Davis et al. [7, 9, 10, 19] for discussions of these examples. Condition (4.2.2) is similar to (2.8) in Davis and Hsing [6]. It serves the purpose of establishing the convergence of a sequence of point processes to a limiting cluster point process. This condition is much weaker than the anti-clustering condition $D'(\epsilon a_n)$ of Leadbetter; cf. Section 5.3.2 in Embrechts et al. [12].

The mixing rate (ξ_h) in conditions (M) and (M1) is useful for finding bounds on the moments $\Gamma(s,t,u,v)$ introduced above. *In what follows, c will denote any (possibly different) constants whose value is not of interest.*

Lemma 4.2.3. *Let (X_t) be a strongly mixing sequence with mixing rate (ξ_h). Then for integers $h, l, u \ge 1$ and for some constants $c > 0$,*

$$|\Gamma(0,h,h+l,h+l+u)| \le c\, \min(\xi_h, \xi_u)\,, \tag{4.2.3}$$

$$|\Gamma(0,h,h+l,h+l+u) - (p_h - p_0^2)(p_u - p_0^2)| \le c\, \xi_l\,, \tag{4.2.4}$$

$$|\Gamma(0,h,h+l)| \le c\, \min(\xi_h, \xi_l)\,, \tag{4.2.5}$$

$$|\Gamma(0,h)| \le \xi_h\,. \tag{4.2.6}$$

The proof of Lemma 4.2.3 follows by a direct application of Theorem 17.2.1 in Ibragimov and Linnik [16]. Relation (4.2.3) combined with (4.2.1) will ensure that sums of $\Gamma(0,h,h+l,h+l+u)$ are asymptotically negligible if h or u exceed r_n.

4.2.3 Central limit theory for the sample extremogram

In this section we recall a central limit theorem for the extremogram from Davis and Mikosch [7], Section 3.

Lemma 4.2.4. *Assume that (X_t) is an $\mathbb{R}^d$-valued strictly stationary regularly varying sequence with index $\alpha > 0$ and that the Borel set A satisfies the conditions of Section 4.1.2. If the mixing conditions* (M), (M1) *hold and $\sum_{l=1}^{\infty} \gamma_A(l) < \infty$ then for $h \geq 0$,*

$$\widetilde{\gamma}_A(h) \stackrel{P}{\rightarrow} \gamma_A(h)\,, \tag{4.2.7}$$

$$(n/m)^{1/2}\big(\widetilde{\gamma}_A(i) - E\widetilde{\gamma}_A(i)\big)_{i=0,\dots,h} \stackrel{d}{\rightarrow} (Z_i)_{i=0,\dots,h}\,, \tag{4.2.8}$$

where $(Z_i)_{i=0,\dots,h}$ is mean zero Gaussian with covariance matrix $\Sigma_h = (\sigma_{ij})_{i,j=0,\dots,h}$ given by

$$\sigma_{ij} = \gamma_A(i,j) + \sum_{l=1}^{\infty} \big[\gamma_A(i,l,l+j) + \gamma_A(j,l,l+i)\big]\,, \quad i,j = 0,\dots,h\,,$$

and for $u,s,t \geq 0$,

$$\gamma_A(u,s,t) = \lim_{n\to\infty} n\,P(a_n^{-1}X_0 \in A\,, a_n^{-1}X_u \in A, a_n^{-1}X_s \in A, a_n^{-1}X_t \in A)\,,$$

with the convention that $\gamma_A(u,t) = \gamma_A(u,u,t)$. Moreover, we have for $h \geq 1$

$$\widetilde{\rho}_A(h) \stackrel{P}{\rightarrow} \rho_A(h)\,, \tag{4.2.9}$$

$$(n/m)^{1/2}\Big(\widetilde{\rho}_A(i) - \frac{p_i}{p_0}\Big)_{i=1,\dots,h} \stackrel{d}{\rightarrow} \frac{1}{\gamma_A(0)}\Big(Z_i - \rho_A(i)Z_0\Big)_{i=1,\dots,h}\,. \tag{4.2.10}$$

Proof. The proof of (4.2.7) was given in Section 3 of Davis and Mikosch [7]. There we can also find the proof of (4.2.8) in a more general context. Here we will calculate the covariance matrix Σ_h explicitly. The expressions for σ_{ii}, $i \geq 0$, were derived in Davis and Mikosch [7] for $i = 0$ and $i \geq 1$ in Theorem 3.1 and Lemma 5.2, respectively. We notice that $\gamma_A(i,l,l+i) \leq \gamma_A(l)$ and therefore the infinite series in σ_{ij} are finite.

For $i \neq j$, similar calculations as for Lemma 4.2.1 yield for $k \geq 1$ and $r_n/m_n \to 0$,

$$\begin{aligned}
&\frac{m}{n}\operatorname{cov}\Big(\sum_{t=1}^{n} \widetilde{I}_t\widetilde{I}_{t+i}, \sum_{s=1}^{n} \widetilde{I}_s\widetilde{I}_{s+j}\Big)\\
&= m\,\Gamma(0,0,i,j) + m\sum_{l=1}^{n}\Big[(1-l/n)\big[\Gamma(0,i,l,l+j) + \Gamma(0,j,l,l+i)\big] - (p_i - p_0^2)(p_j - p_0^2)\Big]\\
&= m\,\Gamma(0,0,i,j)\\
&\quad + m\Big(\sum_{l=1}^{k} + \sum_{l=k+1}^{r_n} + \sum_{l=r_n+1}^{n}\Big)\Big[(1-l/n)\big[\Gamma(0,i,l,l+j) + \Gamma(0,j,l,l+i)\big] - (p_i - p_0^2)(p_j - p_0^2)\Big]\\
&= Q_1 + Q_2 + Q_3 + Q_4\,.
\end{aligned}$$

By regular variation, for fixed $k \geq 1$ as $n \to \infty$,

$$Q_1 + Q_2 \to \gamma_A(i,j) + \sum_{l=1}^{k} \big[\gamma_A(i,l,l+j) + \gamma_A(j,l,l+i)\big]\,,$$

and the right-hand side converges to σ_{ij} as $k \to \infty$. By (4.2.2) we have

$$\lim_{k\to\infty} \limsup_{n\to\infty} |Q_3| = 0\,.$$

Using (4.2.4) and (4.2.1), we also have

$$|Q_4| \le c m_n \sum_{l=r_n+1}^{\infty} \xi_l \to 0\,, \quad n \to \infty\,.$$

This proves (4.2.7) and (4.2.8). Relations (4.2.9) and (4.2.10) follow by a continuous mapping argument, observing that for $1 \le i \le h$,

$$\begin{aligned} \Big(\frac{n}{m}\Big)^{1/2}\Big(\widetilde{\rho}_A(i) - p_i/p_0\Big) &= \Big(\frac{n}{m}\Big)^{1/2}\frac{\widetilde{\gamma}_A(i) - E\widetilde{\gamma}_A(i)}{\widetilde{\gamma}_A(0)} - E\widetilde{\gamma}_A(i)\frac{(n/m)^{1/2}\big(\widetilde{\gamma}_A(0) - E\widetilde{\gamma}_A(0)\big)}{\widetilde{\gamma}_A(0)E\widetilde{\gamma}_A(0)} + o_P(1) \\ &\stackrel{d}{\to} \frac{1}{\gamma_A(0)}\Big(Z_i - \rho_A(i)Z_0\Big)\,. \end{aligned}$$

□

Recall that a strictly stationary process (X_t) is η-dependent for some integer $\eta \ge 0$ if $(X_t)_{t\le 0}$ and $(X_t)_{t>\eta}$ are independent. For such a process we observe that $\sigma_{hh} = 0$ for $h > \eta$ and hence (4.2.8) collapses into $(n/m)^{0.5}\widetilde{\gamma}_A(h) \stackrel{P}{\to} 0$ for $h > \eta$. In particular, for an iid sequence (X_t), $Z_h = 0$ a.s. for $h \ge 1$, while $(n/m)^{0.5}\widetilde{\gamma}_A(0) \stackrel{d}{\to} Z_0$ and Z_0 is $N(0, \gamma_A(0))$ distributed.

In these cases, the rate of convergence in (4.2.8) can be improved.

Lemma 4.2.5. *Assume that (X_t) is an $\mathbb{R}^d$-valued η-dependent regularly varying strictly stationary sequence with index $\alpha > 0$ for some $\eta \ge 0$, and the Borel set A satisfies the conditions of Section 4.1.2. Additionally, assume that for $j \ge i > \eta$ and $1 \le t \le \eta - (j - i)$, the following limits exist:*

$$\overline{\gamma}_A(t, i, t+j) = \lim_{n\to\infty} m^2 P(a_m^{-1}X_0 \in A, a_m^{-1}X_t \in A, a_m^{-1}X_i \in A, a_m^{-1}X_{t+j} \in A)\,, \quad (4.2.11)$$

Then for $h \ge 1$,

$$n^{0.5}\big(\widetilde{\gamma}_A(\eta + i)\big)_{i=1,\dots,h} \stackrel{d}{\to} (Z_i)_{i=1,\dots,h}\,,$$

where $(Z_i)_{i=1,\dots,h}$ is Gaussian $N(0, \overline{\Sigma}_h)$ whose covariance matrix $\overline{\Sigma}_h = (\sigma_{ij})_{i,j=1,\dots,h}$ is given by

$$\sigma_{ij} = \gamma_A(0)\gamma_A(j - i) + 2\sum_{t=1}^{\eta-(j-i)} \overline{\gamma}(t, i, t+j)\,, \quad 1 \le i \le j\,. \quad (4.2.12)$$

Remark 4.2.6. Condition (4.2.11) is an additional asymptotic independence condition. Indeed, regular variation of (X_t) only implies that the limits

$$\lim_{n\to\infty} mP(a_m^{-1}X_0 \in A, a_m^{-1}X_t \in A, a_m^{-1}X_i \in A, a_m^{-1}X_{t+j} \in A)$$

exist and are finite. Then (4.2.11) implies that the latter limits must be zero. In Example 4.2.8 we consider some simple cases when (4.2.11) is satisfied.

Remark 4.2.7. Assume $j-i>\eta$. Then, by η-dependence, $\gamma_A(j-i)=0$ and the index set in (4.2.12) is empty. Hence $\sigma_{ij}=0$ for $j-i>\eta$. In particular, if (X_t) is iid, $\sigma_{ij}=0$ for $i\neq j$ and $\sigma_{ii}=\gamma_A^2(0)$.

Proof. We start by calculating the asymptotic covariances. Assume $j\geq i>\eta$. Then, using the independence of I_0 and $(I_jI_i, I_iI_tI_{t+j})$ for $t>\eta$ and of I_{t+j} and $I_0I_tI_i$ for $t\leq\eta$ and $t\geq \eta-(j-i)$, we obtain

$$\begin{aligned}\operatorname{cov}\big(n^{0.5}\widetilde{\gamma}_A(i), n^{0.5}\widetilde{\gamma}_A(j)\big) &= m^2E\widetilde{I}_0^2E\widetilde{I}_i\widetilde{I}_j+m^2\sum_{t=1}^{\eta}E\widetilde{I}_0\widetilde{I}_i\widetilde{I}_t\widetilde{I}_{t+j}+o(1)\\ &= \gamma_A(0)\gamma_A(j-i)+m^2\sum_{t=1}^{\eta-(j-i)}E\widetilde{I}_0\widetilde{I}_i\widetilde{I}_t\widetilde{I}_{t+j}+o(1)\\ &= \gamma_A(0)\gamma_A(j-i)+2\sum_{t=1}^{\eta-(j-i)}\overline{\gamma}_A(t,i,t+j)\,.\end{aligned}$$

This completes the calculation of $\overline{\Sigma}_h$. Furthermore, we observe that for $h\geq 1$,

$$n^{0.5}\big(\widetilde{\gamma}_A(i)\big)_{i=\eta+1,\ldots,\eta+h} = (m/n^{0.5})\sum_{t=1}^{n}\big(\widetilde{I}_t\widetilde{I}_{t+i}\big)_{i=\eta+1,\ldots,\eta+h}+o_P(1)\,. \qquad (4.2.13)$$

The vector sequence $(\widetilde{I}_t\widetilde{I}_{t+i})_{i=\eta+1,\ldots,\eta+h}$, $t=1,2,\ldots$, is strictly stationary and $(h+\eta)$-dependent. Now an application of the central limit theorem for strongly mixing triangular arrays in Rio [25] and the Cramér-Wold device to (4.2.13) conclude the proof. □

The following examples fulfill the conditions of Lemma 4.2.5.

Example 4.2.8. An iid regularly varying sequence (X_t) is 0-dependent, and thus (4.2.11) holds. Its limiting covariance matrix $\overline{\Sigma}_h$ is a diagonal matrix with entries $\gamma_A^2(0)=(\mu_1(A))^2$ on the main diagonal.

We consider the stochastic volatility model $X_t=\sigma_tZ_t$ where (σ_t) is independent of (Z_t), (σ_t) is a positive η-dependent strictly stationary sequence and (Z_t) is a regularly varying iid sequence with index $\alpha>0$; see Davis and Mikosch [8]. Assume that $E\sigma^{\alpha+\varepsilon}<\infty$ for some $\varepsilon>0$. In this case, (X_t) is η-dependent, strictly stationary and regularly varying with index α. We will show that (4.2.11) holds with $\overline{\gamma}_A(u,s,t)=0$ for $0<u<s<t$. Since A is bounded away from zero, there exists a $\delta>0$ such that

$$\begin{aligned}\overline{\gamma}_A(u,s,t) &\leq \limsup_{n\to\infty} m^2P(a_m^{-1}\min(|X_0|,|X_u|,|X_s|,|X_t|)>\delta)\\ &\leq \limsup_{n\to\infty} m^2P(a_m^{-1}\max(\sigma_0,\sigma_u,\sigma_s,\sigma_t)\min(|Z_0|,|Z_u|,|Z_s|,|Z_t|)>\delta)\\ &\leq \limsup_{n\to\infty} 4m^2P(a_m^{-1}\sigma_0\min(|Z_0|,|Z_u|,|Z_s|,|Z_t|)>\delta)\\ &\leq \limsup_{n\to\infty} cm^2(E\sigma^{\alpha})^4(P(|Z_0|>a_m\delta))^4=0\,,\end{aligned}$$

where we used that $P(\sigma_0|Z_0|>a_m)\sim E\sigma^{\alpha}P(|Z_0|>a_m\delta)$ by virtue of Breiman's lemma; see [2].

4.2.4 Mean square consistency of the integrated periodogram

Recall the definitions of $J_{nA}(g)$ and $J_A(g)$ for $g \in L_+^2(\Pi)$ from (4.1.4) and (4.1.5), respectively. The following elementary result deals with the convergence of the first and second moments of $J_{nA}(g)$ for a given function g.

Lemma 4.2.9. *Consider an $\mathbb{R}^d$-valued strictly stationary regularly varying sequence (X_t) with index $\alpha > 0$. Assume that the Borel set $A \subset \overline{\mathbb{R}}_0^d$ satisfies the conditions of Section 4.1.2, $\sum_{l=1}^{\infty} \gamma_A(l) < \infty$ and (M) holds. Then the following asymptotic relations hold for $g \in L_+^2(\Pi)$.*

1. *$EJ_{nA}(g) \to J_A(g)$ as $n \to \infty$.*
2. *If in addition, $m \log^2 n/n = O(1)$ as $n \to \infty$, and there exists a constant $c > 0$ such that*

$$|c_h(g)| \le c/h\,, \quad h \in \mathbb{Z}\,, \tag{4.2.14}$$

then $E(J_{nA}(g) - J_A(g))^2 \to 0$ and $J_{nA}^{\circ}(g) \xrightarrow{P} J_A^{\circ}(g)$ as $n \to \infty$.

The proof of the lemma is given in Section 4.5.

Remark 4.2.10. Condition (4.2.14) holds under mild smoothness conditions on g, e.g. if g is Lipschitz or has bounded variation on Π; see Theorem 4.7 on p. 46 and Theorem 4.12 on p. 47 in Zygmund [27].

4.3 Functional central limit theorem for the integrated periodogram

Recall the definition of the spectral density h_A from Section 4.1.4. In this section, we assume that the weight function g is a non-negative continuous function. Abusing notation, we define the empirical spectral distribution function with weight function g by

$$J_{nA}(x) = J_{nA}(gI_{[0,x]}) = \int_0^x I_{nA}(\lambda)\, g(\lambda)\, d\lambda\,, \quad x \in \Pi\,. \tag{4.3.1}$$

Under the conditions of Lemma 4.2.9, again abusing notation, we have

$$J_{nA}(x) \xrightarrow{P} J_A(x) = J_A(gI_{[0,x]}) = \int_0^x h_A(\lambda)\, g(\lambda)\, d\lambda\,, \quad x \in \Pi\,.$$

In view of the monotonicity and continuity of the functions J_{nA} and J_A we also have

$$\sup_{x \in \Pi} |J_{nA}(x) - J_A(x)| \xrightarrow{P} 0\,. \tag{4.3.2}$$

Our next goal is to complement this consistency result by a functional central limit theorem of the type

$$(n/m)^{0.5}(J_{nA} - J_A) \xrightarrow{d} G\,,$$

in $\mathbb{C}(\Pi)$, the space of continuous functions on Π equipped with the uniform topology, for a suitable Gaussian limit process G.

However, this result is unlikely to hold in general, due to asymptotic bias problems. It is mentioned in Davis and Mikosch [7] in relation with the central limit theorem for the sample extremogram (see Lemma 4.2.4 above) that the pre-asymptotic centerings $E\widetilde{\gamma}_A(i) = ((n-i)/n)m(p_i - p_0^2)$ can in general not be replaced by their limits $\gamma_A(i)$ due to the failure of the relation $(n/m)^{0.5}|m(p_i - p_0^2) - \gamma_A(i)| \to 0$ as $n \to \infty$. Therefore we will equip the empirical spectral distribution function J_{nA} with the pre-asymptotic centering EJ_{nA}. It follows from Lemma 4.2.9 that under (M), $EJ_{nA}(x) \to J_A(x)$ for every $x \in \Pi$, and again using monotonicity of EJ_{nA} and J_A, we have $\sup_{x\in\Pi} |EJ_{nA}(x) - J_A(x)| \to 0$.

We observe that

$$\begin{aligned} J_{nA}(x) &= \psi_0(x)\,\widetilde{\gamma}_A(0) + 2\sum_{h=1}^{n-1} \psi_h(x)\,\widetilde{\gamma}_A(h)\,, \\ J_{nA}^{\circ}(x) &= \psi_0(x) + 2\sum_{h=1}^{n-1} \psi_h(x)\,\widetilde{\rho}_A(h)\,, \end{aligned}$$

where

$$\psi_h(x) = \int_0^x \cos(h\lambda)\,g(\lambda)\,d\lambda\,, \quad x \in \Pi\,.$$

We also consider a Riemann sum approximation of the coefficients $\psi_h(x)$ at the Fourier frequencies $\omega_n(i) = 2i\pi/n \in \Pi$ given by

$$\widehat{\psi}_h(x) = \frac{2\pi}{n}\sum_{i=1}^{x_n} g(\omega_n(i))\cos(h\omega_n(i))\,, \quad x \in \Pi\,,$$

where $x_n = [nx/2\pi]$. The corresponding analogs of J_{nA} and J_{nA}° are then given by

$$\begin{aligned} \widehat{J}_{nA}(x) &= \widehat{\psi}_0(x)\widetilde{\gamma}_A(0) + 2\sum_{h=1}^{n-1} \widehat{\psi}_h(x)\,\widetilde{\gamma}_A(h)\,, \\ \widehat{J}_{nA}^{\circ}(x) &= \widehat{\psi}_0(x) + 2\sum_{h=1}^{n-1} \widehat{\psi}_h(x)\,\widetilde{\rho}_A(h)\,, \end{aligned}$$

Now we are ready to formulate the main result of this paper.

Theorem 4.3.1. *Assume that (X_t) is an $\mathbb{R}^d$-valued strictly stationary regularly varying sequence with index $\alpha > 0$ and the Borel set $A \subset \overline{\mathbb{R}}_0^d$ is bounded away from zero, $\mu_1(\partial A) = 0$ and $\mu_1(A) > 0$. Let g be a non-negative β-Hölder continuous function with $\beta \in (3/4, 1]$. If the conditions* (M), (M1) *and $\sum_{l=1}^{\infty} \gamma_A(l) < \infty$ hold then in $\mathbb{C}(\Pi)$,*

$$(n/m)^{0.5}(J_{nA} - EJ_{nA}) \xrightarrow{d} G\,, \quad n \to \infty\,, \tag{4.3.3}$$

$$(n/m)^{0.5}(\widehat{J}_{nA} - E\widehat{J}_{nA}) \xrightarrow{d} G\,, \quad n \to \infty\,, \tag{4.3.4}$$

where the limit process is given by the infinite series

$$G = \psi_0 Z_0 + 2\sum_{h=1}^{\infty} \psi_h \, Z_h \,, \tag{4.3.5}$$

which converges in distribution in $\mathbb{C}(\Pi)$, (Z_h) *is a mean zero Gaussian sequence such that* $(Z_0,\ldots,Z_h)$ *has the covariance matrix* (Σ_h), $h \geq 0$, *given in Lemma 4.2.4. Moreover, the following limit relations hold*

$$(n/m)^{0.5}\big(J^\circ_{nA} - EJ_{nA}/(mp_0)\big) \quad \stackrel{d}{\to} \quad G^\circ \,, \quad n\to\infty \,, \tag{4.3.6}$$

$$(n/m)^{0.5}\big(\widehat{J}^\circ_{nA} - E\widehat{J}_{nA}/(mp_0)\big) \quad \stackrel{d}{\to} \quad G^\circ \,, \quad n\to\infty \,, \tag{4.3.7}$$

where the limit process is given by the infinite series

$$G^\circ = \frac{2}{\gamma_A(0)} \sum_{h=1}^{\infty} \psi_h (Z_h - \rho_A(h) Z_0) \,.$$

The proof of this result is given in Section 4.6.

Remark 4.3.2. For practical purposes, the discretized version $\widehat{J}_{nA}$ will be preferred to J_{nA} since it does not involve the calculation of integrals. Moreover, since $\sum_{t=1}^n \mathrm{e}^{\,i\omega_n(j)t} = 0$ for $\omega_n(j) \in (0,\pi)$, centering of the indicators I_t with the unknown parameter p_0 in the periodogram ordinates $I_{nA}(\omega_n(j)) = (m/n)|\sum_{t=1}^n I_t \mathrm{e}^{\,i\omega_n(j)t}|^2$ is not needed.

For an η-dependent sequence (X_t), we know that $Z_h = 0$ a.s. for $h > \eta$. Then we conclude from Theorem 4.3.1 and Lemma 4.2.4 that the limit process G collapses into

$$G = \psi_0 Z_0 + 2\sum_{h=1}^{\eta} \psi_h Z_h \,.$$

However, taking into account Lemma 4.2.5, a more sophisticated result with a different convergence rate can be derived. The corresponding result for J°_{nA} is similar nd therefore omitted.

Theorem 4.3.3. *Assume that* (X_t) *is an* $\mathbb{R}^d$*-valued* η*-dependent regularly varying sequence with index* $\alpha > 0$ *for some* $\eta \geq 0$ *and the Borel set* $A \subset \overline{\mathbb{R}}_0^d$ *is bounded away from zero,* $\mu_1(\partial A) = 0$ *and* $\mu_1(A) > 0$. *Also assume that the limits in* (4.2.11) *exist. Let* g *be a non-negative* β*-Hölder continuous function with* $\beta \in (3/4, 1]$. *Then the relations*

$$\sqrt{n}\big(J_{nA} - \psi_0 \widetilde{\gamma}_A(0) - 2\sum_{h=1}^{\eta} \psi_h \widetilde{\gamma}_A(h)\big) \quad \stackrel{d}{\to} \quad \overline{G} \,,$$

$$\sqrt{n}\big(\widehat{J}_{nA} - \widehat{\psi}_0 \widetilde{\gamma}_A(0) - 2\sum_{h=1}^{\eta} \widehat{\psi}_h \widetilde{\gamma}_A(h)\big) \quad \stackrel{d}{\to} \quad \overline{G} \,,$$

hold in $\mathbb{C}(\Pi)$, *where the limit process is given by the a.s. converging infinite series*

$$\overline{G} = 2\sum_{h=1}^{\infty} \psi_{\eta+h} \, Z_h \,,$$

and (Z_h) *is a mean zero Gaussian sequence such that* $(Z_1, \ldots, Z_h)$ *has covariance matrix* $\overline{\Sigma}_h$*,* $h \geq 1$*, given in Lemma 4.2.5.*

The proof is given in Section 4.7.

Example 4.3.4. Assume that (X_t) is an iid regularly varying sequence with index $\alpha > 0$. Then (Z_h) is an iid mean zero Gaussian sequence with $\mathrm{var}(Z) = \gamma_A^2(0) = (\mu_1(A))^2$. If we choose the function $g \equiv 1$ we obtain

$$\psi_h(x) = \int_0^x \cos(h\lambda) d\lambda = \frac{\sin(hx)}{h}\,, \quad h \geq 0\,, \quad x \in \Pi\,,$$

and

$$\overline{G}(x) = 2 \sum_{h=1}^{\infty} \frac{\sin(hx)}{h} Z_h\,, \quad x \in \Pi\,.$$

We notice that $\overline{G}$ is a series representation of a Brownian bridge; see Hida [15].

As in classical limit theory for the empirical spectral distribution (see Grenander and Rosenblatt [13], Dahlhaus [4]), an application of the continuous mapping theorem to Theorems 4.3.1 and 4.3.3 yields limit theory for functionals of the integrated periodogram. These functionals can be used for testing the goodness of fit of the spectral density of the time series model underlying the data, under the null hypothesis that the model is correct. From Theorem 4.3.1 we get the following limit results for the corresponding test statistics.

- *Grenander-Rosenblatt test:*

$$(n/m)^{0.5} \sup_{x \in \Pi} \Big| J_{nA}(x) - EJ_{nA}(x) \Big| \quad \stackrel{d}{\to} \quad \sup_{x \in \Pi} |G(x)|\,.$$

- ω^2*- or Cramér-von Mises test:*

$$(n/m) \int_{x \in \Pi} \Big(J_{nA}(x) - EJ_{nA}(x) \Big)^2 dx \stackrel{d}{\to} \int_{x \in \Pi} G^2(x)\, dx\,.$$

If (X_t) is an η-dependent sequence satisfying the conditions of Theorem 4.3.3, the corresponding limit results read as follows:

- *Grenander-Rosenblatt test:*

$$\sqrt{n} \sup_{x \in \Pi} \Big| J_{nA}(x) - \psi_0(x)\widetilde{\gamma}_A(0) - 2 \sum_{h=1}^{\eta} \psi_h(x) \widetilde{\gamma}_A(h) \Big| \quad \stackrel{d}{\to} \quad \sup_{x \in \Pi} |\overline{G}(x)|\,. \qquad (4.3.8)$$

- ω^2-statistic or Cramér-von Mises test:

$$\sqrt{n} \int_{x \in \Pi} \big(J_{nA}(x) - \psi_0(x)\widetilde{\gamma}_A(0) - 2 \sum_{h=1}^{\eta} \psi_h(x) \widetilde{\gamma}_A(h) \big)^2 dx \stackrel{d}{\to} \int_{x \in \Pi} \overline{G}^2(x)\, dx\,. \qquad (4.3.9)$$

In Figures 4.1 we show the estimated densities in (4.3.8) and (4.3.9) when $\eta = 0$ and compare them with the corresponding theoretical densities of the limits.

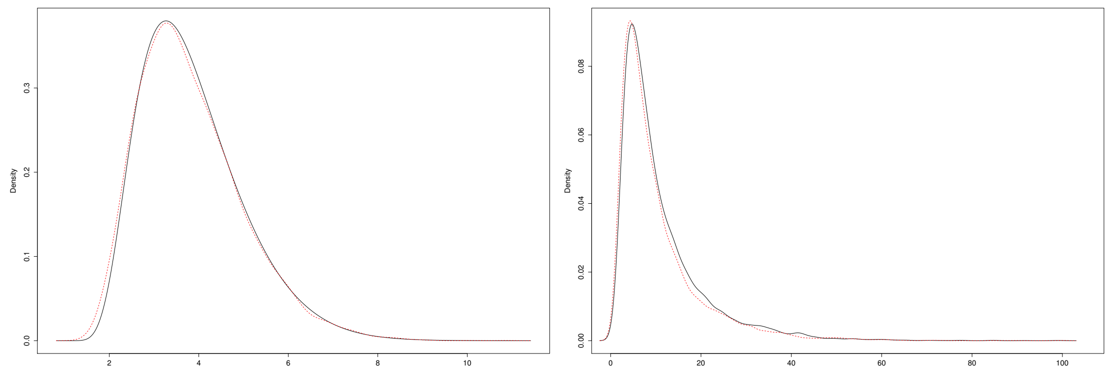

Figure 4.1: We choose the set $A = (1, \infty)$, the threshold a_m such that $p_0 = P(X > a_m) = 0.05$, the sample size $n = 10,000$ and $g \equiv 1$. The underlying sequence (X_t) is iid t-distributed with $\alpha = 3$ degrees of freedom. **Left**: Densities of the left-hand side in (4.3.8) with $\eta = 0$ (dotted line) and its limit $\sup_{x\in\Pi} |\overline{G}(x)|$ (solid line). The density of $\sup_{x\in\Pi} |\overline{G}(x)|$ is given by $4\pi^{-2} \sum_{j=1}^{\infty} (-1)^{j+1} y \exp\big(-j^2y^2/(\pi^2)\big)$, $y > 0$; see Shorack and Wellner [26]. **Right**: The density of the left-hand side in (4.3.9) with $\eta = 0$ (dotted line) and its limit $\int_{x\in\Pi} \overline{G}^2(x)\,dx$ (solid line). We use the identity in law $\int_{x\in\Pi} \overline{G}^2(x)\,dx \stackrel{d}{=} 2\pi \sum_{j=1}^{\infty} (2/j^2) N_j^2$ for an iid standard normal sequence (N_j) (see [26]) for the simulation of the limiting random variable.

4.4 The bootstrapped integrated periodogram

With a few exceptions, the limit processes G and $\overline{G}$ in Theorem 4.3.1 and 4.3.3 have an unfamiliar dependence structure and then it is impossible to give confidence bands for the test statistics mentioned in the previous section. One faces a similar problem when dealing with the sample extremograms whose asymptotic covariance matrix is a complicated function of the measures μ_h in (4.1.2). Davis et al. [9] proposed to apply the stationary bootstrap for constructing confidence bands for the sample extremogram. The stationary bootstrap can also be used for the integrated periodogram, as we will show below.

4.4.1 Stationary bootstrap

The stationary bootstrap was introduced by Politis and Romano [21] as an alternative block bootstrap method. First, we describe this procedure for a strictly stationary sequence (Y_t). Given a sample $Y_1, \ldots, Y_n$, consider the bootstrapped sequence

$$Y_{K_1}, \ldots, Y_{K_1+L_1-1}, \ldots, Y_{K_N}, \ldots, Y_{K_N+L_N-1}, \ldots, \tag{4.4.10}$$

where (Y_i), (K_i), (L_i) are independent sequences, (K_i) is an iid sequence of random variables uniformly distributed on $\{1, \ldots, n\}$, (L_i) is an iid sequence of geometrically distributed random variables with distribution $P(L_1 = i) = \theta(1-\theta)^{i-1}$, $i = 1, 2, \ldots$, for some $\theta = \theta_n \in (0, 1)$ such that $\theta_n \to 0$ as $n \to \infty$ and $N = N_n = \inf\{i \geq 1 : \sum_{j=1}^{i} L_j \geq n\}$. If any element Y_t in (4.4.10) has an index $t > n$, we replace it by $Y_{t \bmod n}$. As a matter of fact, $(Y_t)_{t\geq 1}$ constitutes a strictly stationary sequence. The stationary bootstrap sample is now chosen as the block of the first n elements in (4.4.10). In what follows, we write $(Y_{t^*})_{t\geq 1}$ for the bootstrap sequence (4.4.10), indicating that this sequence is nothing but the original Y-sequence sampled at the random indices $(K_1,\ldots,K_1+L_1-1,K_2,\ldots,K_2+L_2-1,\ldots)$ with the convention that indices larger than n are taken modulo n.

In what follows, the probability measure generated by the bootstrap procedure is denoted by P^*, i.e. $P^*(\cdot)=P(\cdot\mid (X_t))$. The corresponding expected value, variance and covariance are denoted by E^*, var^* and cov^*.

4.4.2 The bootstrapped sample extremogram

Davis et al. [9] applied the stationary bootstrap to the sequence of lagged vectors

$$I_t(h)=(I_t^2, I_tI_{t+1},\ldots,I_tI_{t+h})\,,\quad t=1,2,\ldots\,,$$

for fixed $h\geq 0$ and showed consistency of the bootstrapped sample extremogram. In particular, they showed the following result which we cite for further reference. A close inspection of the proof in [9] shows that the results remain true if in $I_t(h)$ we replace the quantities I_s by $\widetilde{I}_s$, $s=t,\ldots,t+h$. We denote the corresponding vector by $\widetilde{I}_t(h)$. Consider the stationary bootstrap sequence $(\widetilde{I}_{t^*}(h))$ and write

$$\widetilde{\gamma}_A^*(i)=\frac{m}{n}\sum_{t=1}^{n-i}\widetilde{I}_{t^*}\widetilde{I}_{t^*+i},\quad i=0,\ldots,h\,.$$

Theorem 4.4.1. *Consider an $\mathbb{R}^d$-valued strictly stationary regularly varying sequence (X_t) with index $\alpha>0$ and assume the following conditions:*

1. *The mixing conditions* (M), (M1) *and in addition $\sum_{h=1}^\infty h\xi_h<\infty$.*
2. *The growth conditions $\theta=\theta_n\to 0$ and $n\theta^2/m\to\infty$.*
3. *The set A is bounded away from zero, $\mu_1(\partial A)=0$ and $\mu_1(A)>0$.*

Then the following bootstrap consistency results hold for $h\geq 0$:

$$\begin{aligned} E^*\big(\widetilde{\gamma}_A^*(h)\big) &\stackrel{P}{\to} \gamma_A(h)\,,\\ \mathrm{var}^*\big((n/m)^{0.5}\widetilde{\gamma}_A^*(h)\big) &\stackrel{P}{\to} \sigma_{hh}\,,\end{aligned}$$

where the covariance matrix $\Sigma_h=(\sigma_{ij})$ is given in Lemma 4.2.4. Moreover, writing d_{P^*} for any metric describing weak convergence in Euclidean space relative to the probability measure P^* and $(Z_i)_{i=0,\ldots,h}$ for an $N(0,\Sigma_h)$ Gaussian vector, we also have

$$d_{P^*}\Big((n/m)^{1/2}\big(\widetilde{\gamma}_A^*(i)-\widetilde{\gamma}_A(i)\big)_{i=0,\ldots,h},(Z_i)_{i=0,\ldots,h}\Big)\stackrel{P}{\to}0\,,\quad n\to\infty\,.$$

In what follows, we will abuse the notation d_{P^*} for a metric describing weak convergence relative to P^* in any space of interest.

4.4.3 The bootstrapped integrated periodogram

Bootstrapping the sequence $(I_t(h))$ has the advantage that we preserve the neighbors I_{t^*+i} of I_{t^*} from the original sequence (I_s). However, this method depends on the lag h and creates problems if the number of lags increases with the sample size n. In what follows, we will apply the stationary bootstrap directly to (I_t). Then we have to re-define the bootstrap sample extremogram at any lag $h < n$. Write

$$\overline{I}_n = n^{-1} \sum_{t=1}^{n} I_t \quad \text{and} \quad \widehat{I}_t = I_t - \overline{I}_n \,, \quad t \in \mathbb{Z}\,,$$

and define the corresponding bootstrap sample extremogram

$$\widehat{\gamma}_A^*(h) = \frac{m}{n} \sum_{t=1}^{n-h} \widehat{I}_{t^*} \widehat{I}_{(t+h)^*}\,, \quad h = 0, \ldots, n-1\,,$$

and the bootstrap periodogram

$$I_{nA}^*(\lambda) = \frac{m}{n} \Big| \sum_{t=1}^{n} \widehat{I}_{t^*} \mathrm{e}^{-it\lambda} \Big|^2\,, \quad \lambda \in \Pi\,.$$

Note the crucial difference: in general, $I_{t^*} I_{(t+h)^*} \neq I_{t^*} I_{t^*+h}$, but, as we will see in Lemma 4.8.1, the quantities $\widetilde{\gamma}_A^*(h)$ and $\widehat{\gamma}_A^*(h)$ are asymptotically close for fixed $h \geq 0$.

In what follows, we focus on the bootstrap for the continuous version J_{nA} of the integrated periodogram for a given smooth weight function g; bootstrap consistency can also be shown for the discretized version $\widehat{J}_{nA}$; we omit further details. In the definition of J_{nA} in (4.3.1), we simply replace (I_t) by $(\widehat{I}_{t^*})$, resulting in its bootstrap version

$$J_{nA}^*(\lambda) \quad = \quad \int_0^\lambda I_{nA}^*(x)\, g(x)\, dx = \psi_0\, \widehat{\gamma}_A^*(0) + 2 \sum_{h=1}^{n-1} \psi_h\, \widehat{\gamma}_A^*(h)\,, \quad \lambda \in \Pi\,.$$

Now we can formulate a bootstrap analog of Theorem 4.3.1 which shows the consistency of the stationary bootstrap procedure.

Theorem 4.4.2. *Assume the conditions of Theorem 4.3.1 and 4.4.1. Then*

$$d_{P^*}\Big((n/m)^{1/2} \big(J_{nA}^* - E^* J_{nA}^* \big), G \Big) \stackrel{P}{\to} 0\,, \quad n \to \infty\,,$$

where the Gaussian process G is defined in Theorem 4.3.1 and d_{P^*} is any metric which describes weak convergence in $\mathbb{C}(\Pi)$ relative to the probability measure P^*.

Remark 4.4.3. Recall that, in general, it is not possible to replace the centering EJ_{nA} of J_{nA} in the functional central limit theorem of Theorem 4.3.1 by its limit $\int_0^{\cdot} h_A(\lambda)\, g(\lambda)\, d\lambda$. A similar remark applies to Theorem 4.4.2. Although $\sup_{\lambda \in \Pi} |E^* J_{nA}^*(\lambda) - J_{nA}(\lambda)| \stackrel{P}{\to} 0$, under the conditions of Theorem 4.4.2, it is in general not possible to replace the centering $E^* J_{nA}^*$ by J_{nA}; see Lemma 4.8.4.

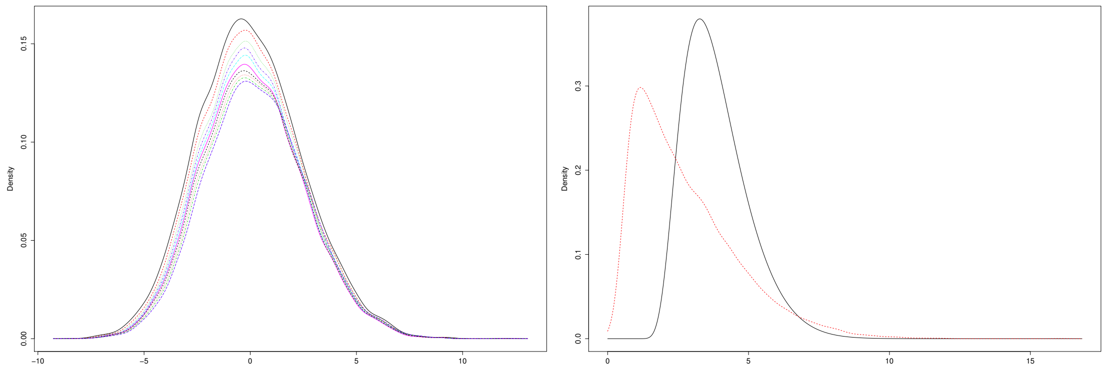

Figure 4.2: We choose the set $A = (1, \infty)$, the threshold a_m such that $p_0 = P(X > a_m) = 0.05$ and the sample size $n = 5,000$ and $g \equiv 1$. We apply the stationary bootstrap $10,000$ times to an ARMA$(1,1)$ process $X_t - 0.8X_{t-1} = N_t + 0.1N_{t-1}$ where the iid t-distributed noise (N_t) has 3 degrees of freedom. **Left**: Densities of $\sqrt{n/m}(J^*_{nA} - E^* J^*_{nA})$ at the fixed frequencies $\{6\pi/20, 7\pi/20, \ldots, 15\pi/20\}$. **Right**: A comparison of the densities of $\sup_{x\in\Pi} \sqrt{n/m}|J^*_{nA}(x) - E^* J^*_{nA}(x)|$ (dotted line) and the right-hand side in (4.3.8) (solid line).

In Figure 4.2 we present the results of applying the stationary bootstrap to an ARMA$(1,1)$ process: the density of $\sqrt{n/m}(J^*_{nA}(x) - E^* J^*_{nA}(x))$ at fixed frequencies have shapes similar to the normal distribution. Moreover, we illustrate that the Grenander-Rosenblatt statistic of this process has a distribution which significantly differs from the distribution of the Grenander-Rosenblatt statistic of an iid sequence defined in (4.3.8).

4.5 Proof of Lemma 4.2.9

Part 1. Recall the series representations of $J_{nA}(g)$ and $J_A(g)$ from (4.1.4) and (4.1.5), respectively. Then for every fixed $k \geq 1$, large n,

$$\begin{aligned} J_{nA}(g) - J_A(g) &= \Big(c_0(g)[\widetilde{\gamma}_A(0) - \gamma_A(0)] + 2\sum_{h=1}^{k} c_h(g)\,[\widetilde{\gamma}_A(h) - \gamma_A(h)]\Big) \\ &\quad + 2\sum_{h=k+1}^{n-1} c_h(g)\,[\widetilde{\gamma}_A(h) - \gamma_A(h)] - 2\sum_{h=n}^{\infty} c_h(g)\,\gamma_A(h) \\ &= I_1(k) + I_2(k) - I_3\,. \end{aligned}$$

Then $I_3 \to 0$ as $n \to \infty$ since $(\gamma_A(h))$ is summable and $EI_1(k)$ converges to zero as $n \to \infty$ due to regular variation, for every k. In view of (4.2.1) in (M),

$$\begin{aligned}\Big|E\sum_{h=r_n+1}^{n-1}\widetilde{\gamma}_A(h)c_h(g)\Big| &= \Big|\frac{m}{n}\sum_{h=r_n+1}^{n-1}(n-h)\,c_h(g)\,(p_h-p_0^2)\Big| \\ &\leq c\,m\sum_{h=r_n+1}^{\infty}\xi_h\to 0\,,\quad n\to\infty\,,\end{aligned}$$

and (4.2.2) in (M) implies

$$\lim_{k\to\infty}\limsup_{n\to\infty}\Big|E\sum_{h=k+1}^{r_n}\widetilde{\gamma}_A(h)c_h(g)\Big|\leq c\lim_{k\to\infty}\limsup_{n\to\infty}m\sum_{h=k+1}^{r_n}p_h=0\,.$$

Since $\lim_{k\to\infty}\sum_{h=k+1}^{\infty}\gamma_A(h)=0$, we have $\lim_{k\to\infty}\limsup_{n\to\infty}|EI_2(k)|=0$. This proves Part 1.

Part 2. It follows from Theorem 3.1 in Davis and Mikosch [7] that $\widetilde{\gamma}_A(h)\stackrel{L^2}{\to}\gamma_A(h)$, $h\geq 1$. Hence $I_1(k)\stackrel{L^2}{\to}0$ as $n\to\infty$ for fixed $k\geq 1$. It remains to show that

$$\lim_{k\to\infty}\limsup_{n\to\infty}\mathrm{var}(I_2(k))=0\,.$$

We have

$$I_2(k)=2\Big(\sum_{h=k+1}^{r_n}+\sum_{h=r_n+1}^{n-1}\Big)c_h(g)\big[\widetilde{\gamma}_A(h)-\gamma_A(h)\big]=2I_{21}(k)+2I_{22}\,.$$

In view of Lemma 4.2.1 we get the bound

$$\begin{aligned}\mathrm{var}(I_{21}(k)) &\leq \frac{m^2}{n}\sum_{h=k+1}^{r_n}\sum_{l=0}^{r_n-h}|c_h(g)c_{h+l}(g)|\times \\ &\quad\Big(|\Gamma(0,h,0,h+l)|+\sum_{t=1}^{n-h-l}|\Gamma(0,h,t,t+h+l)|+\sum_{t=1}^{n-h}|\Gamma(0,h+l,t,t+h)|\Big) \\ &= Q_1+Q_2+Q_3\,.\end{aligned}$$

Since $|c_h(g)|\leq c/h$ (see (4.2.14)),

$$\begin{aligned}|Q_1| &\leq c\,\frac{m^2}{n}\sum_{h=k+1}^{r_n}|c_h(g)|\sum_{s=h}^{r_n}|c_s(g)|p_s \\ &= c\frac{m^2}{n}\sum_{s=k+1}^{r_n}|c_s(g)|p_s\sum_{h=k+1}^{s}|c_h(g)| \\ &\leq c\frac{m^2}{n}\sum_{s=k+1}^{r_n}p_s s^{-1}\log s\,,\end{aligned}$$

and the right-hand side converges to 0 by first letting $n \to \infty$ and then $k \to \infty$, using (4.2.2), Since the structures of Q_2 and Q_3 are similar we restrict ourselves to showing $Q_2 \to 0$ as $n \to \infty, k \to \infty$. We observe that

$$\begin{aligned} |Q_2| &\leq c\frac{m^2}{n} \sum_{h=k+1}^{r_n} \sum_{s=h}^{r_n} \frac{1}{hs} \Big(\sum_{t=1}^{2r_n} + \sum_{t=2r_n+1}^{n} \Big) |\Gamma(0,h,t,t+s)| \\ &\leq c\frac{m \log^2 r_n}{n} m \sum_{h=k+1}^{3r_n} p_h + c\,\frac{m \log^2 r_n}{n} m \sum_{h=r_n+1}^{n} \xi_h + cn^{-1}\Big(m \sum_{h=k+1}^{r_n} p_h/h\Big)^2 . \end{aligned}$$

In the last step, we used (4.2.4). The right-hand side vanishes as $n \to \infty$ and $k \to \infty$. Finally, we conclude that $\lim_{k\to\infty} \limsup_{n\to\infty} \operatorname{var}(I_{21}(k)) = 0$.

Now we turn to bounding $\operatorname{var}(I_{22})$. In view of Lemma 4.2.1 we have

$$\begin{aligned} \operatorname{var}(I_{22}) &\leq \frac{m^2}{n} \sum_{h=r_n+1}^{n-1} \sum_{s=h}^{n-1} |c_h(g)c_s(g)| \Big(|\Gamma(0,h,0,s)| + \sum_{t=1}^{n-s} |\Gamma(0,h,t,t+s)| + \sum_{t=1}^{n-h} |\Gamma(0,s,t,t+h)| \Big) \\ &= Q_4 + Q_5 + Q_6 \,. \end{aligned}$$

We have by (4.2.14),

$$\begin{aligned} Q_4 &\leq c\frac{m^2}{n} \sum_{h=r_n+1}^{n-1} \sum_{s=h}^{n-1} |c_h(g)c_s(g)|\, |E\widetilde{I}_0 \widetilde{I}_s| \\ &\leq c\frac{m^2}{n} \sum_{h=r_n+1}^{n-1} h^{-2} \sum_{s=h}^{n-1} [(p_s - p_0^2) + p_0^2] \\ &\leq c\Big[\frac{m}{nr_n} m \sum_{h=r_n+1}^{\infty} \xi_h + \frac{(p_0 m)^2}{r_n}\Big] = o(1)\,, \quad n \to \infty\,. \end{aligned}$$

The terms Q_5 and Q_6 can be treated in a similar way; we focus on Q_5. By (4.2.14),

$$\begin{aligned} Q_5 &\leq \frac{cm^2}{n} \sum_{h=r_n+1}^{n-1} \sum_{s=h}^{h+r_n} (hs)^{-1} \sum_{t=1}^{r_n} |\Gamma(0,h,t,t+s)| \\ &\quad + \frac{cm^2}{n} \sum_{h=r_n+1}^{n-1} \sum_{s=h+1}^{n-1} \sum_{t=r_n+1}^{n-s} (hs)^{-1} |\Gamma(0,h,t,t+s)| \\ &\quad + \frac{cm^2}{n} \sum_{h=r_n+1}^{n-1} \sum_{s=h+r_n+1}^{n-1} \sum_{t=1}^{r_n} (hs)^{-1} |\Gamma(0,h,t,t+s)| \\ &= Q_{51} + Q_{52} + Q_{53}\,, \end{aligned}$$

and

$$\begin{aligned} Q_{51} &\leq c\frac{m^2}{n} \sum_{h=r_n+1}^{n-1} \sum_{s=h}^{h+r_n} (hs)^{-1} \sum_{t=1}^{r_n} \big[(p_h - p_0^2) + p_0^2\big] \\ &\leq c\Big(\frac{m}{n} m \sum_{h=r_n+1}^{\infty} \xi_h + (mp_0)^2 \frac{r_n}{n}\Big) \to 0\,, \quad n \to \infty\,. \end{aligned}$$

Next we consider Q_{52} and Q_{53}. By (4.2.3), we have

$$Q_{52} \le c\frac{2m^2}{n}\sum_{h=r_n+1}^{n-1}\sum_{s=h}^{n-1}(hs)^{-1}\sum_{t=r_n+1}^{n}\xi_t \le c\frac{m\log^2 n}{n}m\sum_{t=r_n+1}^{\infty}\xi_t\,.$$

The right-hand side converges to zero by using the assumption $m\log^2 n/n = O(1)$ and the condition (4.2.1). Similarly, using (4.2.3), we obtain

$$Q_{53} \quad\le\quad c\frac{m}{n}m\sum_{h=r_n+1}^{\infty}\xi_h\,.$$

We conclude that $\mathrm{var}(I_{22}) \to 0$ as $n \to \infty$.

We proved above that $E(J_{nA} - J_A(g))^2 \to 0$, hence $J_{nA}(g) \stackrel{P}{\to} J_A(g)$, combined with (4.2.7), yields $J^\circ_{nA}(g) \stackrel{P}{\to} J^\circ_A(g)$.

4.6 Proof of Theorem 4.3.1

We start by proving (4.3.3). An application of the continuous mapping theorem in $\mathbb{C}(\Pi)$ and Lemma 4.2.4 yield in $\mathbb{C}(\Pi)$ for every $k \ge 1$,

$$(m/n)^{0.5}\Big(\psi_0\left(\widetilde{\gamma}_A(0) - E\widetilde{\gamma}_A(0)\right) + 2\sum_{h=1}^{k}\psi_h\left(\widetilde{\gamma}_A(h) - E\widetilde{\gamma}_A(h)\right)\Big) \stackrel{d}{\to} \psi_0 Z_0 + 2\sum_{h=1}^{k}\psi_h Z_h\,.$$

Here (Z_h) is mean zero Gaussian process with covariance structure specified in Lemma 4.2.4. In view of Theorem 2 in Dehling et al. [11] relation (4.3.3) will follow if we can prove the following result.

Lemma 4.6.1. *Assume that the conditions of Theorem 4.3.1 hold. Then for any $\varepsilon > 0$,*

$$\lim_{k\to\infty}\limsup_{n\to\infty} P\Big((n/m)^{0.5}\sup_{\lambda\in\Pi}\Big|\sum_{h=k+1}^{n-1}\psi_h(\lambda)\left(\widetilde{\gamma}_A(h) - E\widetilde{\gamma}_A(h)\right)\Big| > \varepsilon\Big) = 0\,.$$

Proof of Lemma 4.6.1. We borrow the techniques of the proof of Theorem 3.2 in Klüppelberg and Mikosch [17]. Without loss of generality we assume that $k = 2^a - 1$ and $n = 2^{b+1}$ where $a < b$ are integers; if k or n do not have this representation we have to modify the proof slightly but we omit details. For integer $q > 0$ and some constant $\kappa > 0$ to be chosen later, let $\varepsilon_q = 2^{-2q/\kappa}$.

We have for $\varepsilon > 0$,

$$
\begin{aligned}
Q = P\Big((n/m)^{0.5} \sup_{\lambda\in\Pi}\Big|\sum_{h=k+1}^{n-1}(\widetilde{\gamma}_A(h)-E\widetilde{\gamma}_A(h))\,\psi_h(\lambda)\Big| > \varepsilon\Big) \\
&\le P\Big((n/m)^{0.5}\sum_{q=a}^{b}\sup_{\lambda\in\Pi}\Big|\sum_{h=2^q}^{2^{q+1}-1}(\widetilde{\gamma}_A(h)-E\widetilde{\gamma}_A(h))\,\psi_h(\lambda)\Big| > \varepsilon) \\
&\le P\Big(\sum_{q=a}^{b}\varepsilon_q > \varepsilon\Big) + P\Big(\bigcup_{q=a}^{b}\Big\{(n/m)^{0.5}\sup_{\lambda\in\Pi}\Big|\sum_{h=2^q}^{2^{q+1}-1}(\widetilde{\gamma}_A(h)-E\widetilde{\gamma}_A(h))\psi_h(\lambda)\Big| > \varepsilon_q\Big\}\Big) \\
&\le \sum_{q=a}^{b}P\Big((n/m)^{0.5}\sup_{\lambda\in\Pi}\Big|\sum_{h=2^q}^{2^{q+1}-1}(\widetilde{\gamma}_A(h)-E\widetilde{\gamma}_A(h))\,\psi_h(\lambda)\Big| > \varepsilon_q\Big) \\
&= \sum_{q=a}^{b}Q_q\,.
\end{aligned}
$$

In the last steps we used that $P(\sum_{q=a}^{b}\varepsilon_q > \varepsilon)$ vanishes for fixed ε and sufficiently large a. Next we will bound the expressions Q_q. Write $J_{qv} = \{(v-1)2^q+1,\ldots,v2^q\}$ and

$$
Y_{qj}(\lambda) = (n/m)^{0.5}\sum_{h=2^q}^{2^{q+1}-1}(\widetilde{\gamma}_A(h)-E\widetilde{\gamma}_{nA}(h))\psi_h(\lambda+(j-1)\pi 2^{-2q})\,,\quad j\in J_{qv}\,,\ \lambda\in[0,2^{-2q}\pi]\,.
$$

Then

$$
\begin{aligned}
Q_q &= P\Big((n/m)^{0.5}\max_{v=1,\ldots,2^q}\max_{j\in J_{qv}}\sup_{\lambda\in[(j-1)\pi 2^{-2q+1},j\pi 2^{-2q+1}]}\Big|\sum_{h=2^q}^{2^{q+1}-1}(\widetilde{\gamma}_A(h)-E\widetilde{\gamma}_A(h))\,\psi_h(\lambda)\Big| > \varepsilon_q\Big) \\
&\le \sum_{v=1}^{2^q}P\Big(((n/m)^{0.5}\max_{j\in J_{qv}}\sup_{\lambda\in[0,2^{-2q+1}\pi]}|Y_{qj}(\lambda)| > \varepsilon_q\Big) = \sum_{v=1}^{2^q}Q_{qv}\,.
\end{aligned}
$$

We will bound each of the terms Q_{qv} by twice applying the maximal inequality of Theorem 10.2 in Billingsley [1]. For this reason we have to control the variance of the increments of the process Y_{qj} both as a function of λ and j. In particular, we will derive the following bound

$$
\frac{n}{m}\,E\Big(\sum_{h=2^q}^{2^{q+1}-1}(\widetilde{\gamma}_A(h)-E\widetilde{\gamma}_A(h))\,d_h(\omega,\lambda,j,j')\Big)^2 \le c\,|j-j'|^2|\lambda-\omega|^{2\beta}\,K_{k,n}\,, \tag{4.6.1}
$$

where β is the Hölder coefficient of the function g,

$$
K_{k,n} \le c\Big[m\sum_{h=r_n+1}^{\infty}\xi_h + m\sum_{h=k+1}^{r_n}p_h + r_n/m\Big]
$$

and for $j < j'$ in J_{qv}, $h \in \{2^q, \dots, 2^{q+1}-1\}$ and $\omega < \lambda$ in $[0, 2^{-2q+1}\pi]$,

$$
\begin{aligned}
& d_h(\omega,\lambda,j,j') \\
= & \ \big(\psi_h(\lambda+(j'-1)\pi 2^{-2q+1}) - \psi_h(\lambda+(j-1)\pi 2^{-2q+1})\big) \\
& -\big(\psi_h(\omega+(j'-1)\pi 2^{-2q+1}) - \psi_h(\omega+(j-1)\pi 2^{-2q+1})\big) \\
= & \int_{\lambda+(j-1)\pi 2^{-2q+1}}^{\lambda+(j'-1)\pi 2^{-2q+1}} g(x)\cos(hx)dx - \int_{\omega+(j-1)\pi 2^{-2q+1}}^{\omega+(j'-1)\pi 2^{-2q+1}} g(x)\cos(hx)dx \\
= & \int_{(j-1)\pi 2^{-2q+1}}^{(j'-1)\pi 2^{-2q+1}} \Big(g(x+\lambda)[\cos(h(x+\lambda)) - \cos(h(x+\omega))] \\
& \qquad -[g(x+\lambda)-g(x+\omega)]\cos(h(x+\omega))\Big)\,dx\,.
\end{aligned}
\tag{4.6.2}
$$

Since g is β-Hölder continuous we have

$$
\Big|\int_{(j-1)\pi 2^{-2q+1}}^{(j'-1)\pi 2^{-2q+1}} [g(x+\lambda)-g(x+\omega)]\cos(h(\omega+x))dx\Big| \le c(\lambda-\omega)^\beta (j'-j)2^{-2q}\,.
$$

Similarly,

$$
\begin{aligned}
& \Big|\int_{(j-1)\pi 2^{-2q}}^{(j'-1)\pi 2^{-2q}} g(x+\lambda)[\cos(h(\lambda+x)) - \cos(h(\omega+x))]dx\Big| \\
= & \ \Big|\int_{(j-1)\pi 2^{-2q}}^{(j'-1)\pi 2^{-2q}} g(x+\lambda)(2\sin(h(\lambda-\omega)/2)\sin(h(\lambda+\omega+2x)/2))dx\Big| \\
\le & \ ch(\lambda-\omega)(j'-j)2^{-2q} \le c(\lambda-\omega)(j'-j)2^{-q}\,.
\end{aligned}
$$

The last two inequalities yield for a constant c only depending on g,

$$
|d_h(\omega,\lambda,j,j')| \le c|\lambda-\omega|^\beta\,|j'-j|\,2^{-q}\,. \tag{4.6.3}
$$

Using this bound, we have

$$
\begin{aligned}
& \frac{n}{m}\,E\Big(\sum_{h=2^q}^{2^{q+1}-1} (\widetilde{\gamma}_A(h) - E\widetilde{\gamma}_A(h))\,d_h(\omega,\lambda,j,j')\Big)^2 \\
\le & \ \ c\,|j-j'|^2|\lambda-\omega|^{2\beta}\,2^{-2q}\,\frac{n}{m}\sum_{h=2^q}^{2^{q+1}-1}\sum_{s=h}^{2^{q+1}-1} \big|\mathrm{cov}\big(\widetilde{\gamma}_A(h),\widetilde{\gamma}_A(s)\big)\big|\,.
\end{aligned}
\tag{4.6.4}
$$

In what follows, it will be convenient to write $\sum_{h,l}^{(q)} = \sum_{h=2^q}^{2^{q+1}-1}\sum_{l=0}^{2^{q+1}-h-1}$. In view of Lemma 4.2.1

we can bound the last term in (4.6.4) as follows:

$$
\begin{aligned}
&\frac{n}{m}\sum_{h,l}^{(q)}\big|\mathrm{cov}(\widetilde{\gamma}_A(h),\widetilde{\gamma}_A(h+l))\big| \\
= \quad & \frac{m}{n}\sum_{h,l}^{(q)}\Big|(n-h-l)\Gamma(0,h,0,h+l)+\sum_{t=1}^{n-h-l-1}(n-h-l-t)\Gamma(0,h,t,t+h+l) \\
& +\sum_{t=1}^{n-h-1}\min(n-h-l,n-h-t)\Gamma(0,h+l,t,t+h)-(n-h)(n-h-l)(p_h-p_0^2)(p_{h+l}-p_0^2)\Big| \\
\leq \quad & m\sum_{h,l}^{(q)}\Big[|\Gamma(0,h,0,h+l)|+\sum_{t=1}^{h+r_n}|\Gamma(0,t,h,t+h+l)|+\sum_{t=1}^{h+l+r_n}|\Gamma(0,h+l,t,t+h)| \\
& +\frac{1}{n}\Big|\sum_{t=h+r_n+1}^{n-h-l-1}(n-t-h-l)\Gamma(0,h,t,t+h+l)+\sum_{t=h+l+r_n+1}^{n-h-1}(n-t-h)\Gamma(0,h+l,t,t+h) \\
& -(n-h)(n-h-l)(p_h-p_0^2)(p_{h+l}-p_0^2)\Big|\Big] \\
= \quad & W_1+W_2+W_3+W_4\,.
\end{aligned}
$$

We will treat two cases of interest for the sums $\sum_{h,l}^{(q)}$: when $2^{q+1}-1\leq r_n$ and $2^q > r_n$. If $2^q \leq r_n < 2^{q+1}-1$ the sums $\sum_{h,l}^{(q)}$ can be split into two sums corresponding to $h\leq r_n$ and $h>r_n$ and these can be treated in a similar fashion.

We start by studying the case $2^{q+1}-1\leq r_n$. Then $r_n \geq 2^{q+1}-1\geq h\geq 2^q > k$ and consequently $2^{q+1}-h-1\leq 2^q$. Thus,

$$
W_1 \quad \leq \quad c2^q m\sum_{h=k+1}^{r_n} p_h\,.
$$

The terms W_2, W_3 have a similar structure and can be treated in the same way; we focus on W_2. Then we get the following bound from Lemma 4.2.1

$$
W_2 \leq c\,2^{2q}\Big[m\sum_{h=k+1}^{r_n}p_h+m\sum_{h=r_n+1}^{2r_n}\xi_h+(r_n/m)\Big]\,.
$$

In view of (4.2.4), we also have

$$
\begin{aligned}
W_4 \quad \leq \quad & \frac{m}{n}\sum_{h,l}^{(q)}\Bigg[\sum_{t=h+r_n+1}^{n-h-l}(n-t-h-l)\big|\Gamma(0,h,t,t+h+l)-(p_h-p_0^2)(p_{h+l}-p_0^2)\big| \\
& +\sum_{t=h+r_n+l+1}^{n-h}(n-t-h)\big|\Gamma(0,h+l,t,t+h)-(p_h-p_0^2)(p_{h+l}-p_0^2)\big| \\
& +cnr_n\big|(p_h-p_0^2)(p_{h+l}-p_0^2)\big|\Bigg] \\
\leq \quad & c2^{2q}m\sum_{h=r_n+1}^{n}\xi_h+c\,\frac{r_n}{m}\Big(m\sum_{h=k+1}^{r_n}p_h\Big)^2\,.
\end{aligned}
$$

Next we assume that $2^q > r_n$. By (4.2.3) and (4.2.4),

$$
\begin{aligned}
W_1 &\le m \sum_{h=2^q}^{2^{q+1}-1} \sum_{l=0}^{r_n} |\Gamma(0,0,h,h+l) - (p_0 - p_0^2)(p_l - p_0^2)| + 2^q m \sum_{l=0}^{r_n} (p_0 - p_0^2)(p_l - p_0^2) \\
&\quad + m \sum_{h=2^q}^{2^{q+1}-1} \sum_{l=r_n+1}^{2^{q+1}-h-1} |\Gamma(0,0,h,h+l)| \\
&\le c\, 2^q m \sum_{h=r_n+1}^{\infty} \xi_h + \frac{2^q r_n}{m} (m p_0)^2 .
\end{aligned}
$$

We again focus on W_2; W_3 can be treated in a similar way.

$$
\begin{aligned}
W_2 &\le m \sum_{h=2^q}^{2^{q+1}-1} \Big(\sum_{l=1}^{r_n} \Big(\sum_{t=1}^{r_n} + \sum_{t=r_n+1}^{h} + \sum_{t=h+1}^{h+r_n} \Big) + \sum_{l=r_n+1}^{2^{q+1}-1} \sum_{t=1}^{h+r_n} \Big) \Big) |\Gamma(0,t,h,t+h+l)| \\
&\le c\, 2^{2q} m \sum_{h=r_n+1}^{\infty} \xi_h + c\, 2^{2q} \frac{r_n}{m} (m p_0)^2
\end{aligned}
$$

To obtain the bounds for W_4 we use (4.2.4):

$$
\begin{aligned}
W_4 &\le \frac{m}{n} \sum_{h=2^q}^{2^{q+1}-1} \sum_{l=0}^{2^{q+1}-h-1} \Bigg[\sum_{t=h+r_n+1}^{n-h-l} (n-t-h-l) \big|\Gamma(0,h,t,t+h+l) - (p_h - p_0^2)(p_{h+l} - p_0^2)\big| \\
&\quad + \sum_{t=h+r_n+l+1}^{n-h} (n-t-h) \big|\Gamma(0,h+l,t,t+h) - (p_h - p_0^2)(p_{h+l} - p_0^2)\big| \\
&\quad + cn 2^q \big|(p_h - p_0^2)(p_{h+l} - p_0^2)\big| \Bigg] \\
&\le c 2^{2q} m \sum_{t=r_n+1}^{\infty} \xi_t + c(2^q/m) \Bigg(m \sum_{h=r_n+1}^{\infty} \xi_h \Bigg)^2 .
\end{aligned}
$$

Collecting the bounds for W_i, $i \le 4$, and using (4.6.4), we finally proved (4.6.1).

Using this bound, we can apply the maximal inequality of Theorem 10.2 in Billingsley [1] with respect to the variable $\lambda \le 2^{-2q}\pi$ and for fixed j, j':

$$
\begin{aligned}
P(\max_{0 \le \lambda \le 2^{-2q}\pi} |Y_j(\lambda) - Y_{j'}(\lambda)| > \varepsilon_q) &\le c\varepsilon_q^{-2} (2^{-2q}\pi)^{2\beta} \,(j-j')^2\, K_{k,n} \\
&\le c\, 2^{4q(1-\beta+\kappa^{-1})} \,((j-j')2^{-2q})^2\, K_{k,n} .
\end{aligned}
$$

Another application of this maximal inequality to $\max_{0 \le \lambda \le 2^{-2q}\pi} |Y_j(\lambda)|$ with respect to the variable $j \in J_{qv}$ yields

$$
Q_{qv} = P\Big(\max_{j \in \{(v-1)2^q+1, \ldots, v2^q\}} \max_{0 \le \lambda \le 2^{-2q}\pi} |Y_j(\lambda)| > \varepsilon_q \Big) \le c 2^{4q(2^{-1}-\beta+\kappa^{-1})} K_{k,n} .
$$

Then we also have

$$Q_q \le \sum_{v=1}^{2^q} Q_{qv} \le c\, 2^{4q(3/4-\beta+\kappa^{-1})} K_{k,n}\,.$$

The right-hand side converges to zero as $q \to \infty$ provided $\beta \in (3/4, 1]$ and κ is chosen sufficiently large. Therefore we conclude for every $\varepsilon > 0$,

$$Q \le \sum_{q=a}^{b} Q_q \le c K_{k,n} \sum_{q=a}^{\infty} 2^{4q(3/4-\beta+\kappa^{-1})}\,. \tag{4.6.5}$$

The right-hand side converges to zero by first letting $n \to \infty$ and then $k \to \infty$. This concludes the proof of (4.3.3).

Next we turn to the proof of (4.3.4). It will follow from (4.3.3) once we prove the following lemma.

Lemma 4.6.2. *Assume that the conditions of Theorem 4.3.1 hold. Then for any $\varepsilon > 0$,*

$$\limsup_{n\to\infty} P\Big((n/m)^{0.5} \sup_{\lambda\in\Pi} \big|\big(\widehat{J}_{nA}(\lambda) - E\widehat{J}_{nA}(\lambda)\big) - \big(J_{nA}(\lambda) - EJ_{nA}(\lambda)\big)\big| > \varepsilon\Big) = 0\,.$$

Proof of Lemma 4.6.2: For any fixed $k \ge 1$ we have

$$\begin{aligned}
& P\Big((n/m)^{0.5} \sup_{\lambda\in\Pi} \big|\big(\widehat{J}_{nA}(\lambda) - E\widehat{J}_{nA}(\lambda)\big) - \big(J_{nA}(\lambda) - EJ_{nA}(\lambda)\big)\big| > \varepsilon\Big) \\
\le\;& P\Big((n/m)^{0.5} \sup_{\lambda\in\Pi} \Big|\sum_{h=0}^{k} \big(\widetilde{\gamma}_A(h) - E\widetilde{\gamma}_A(h)\big)\big(\psi_h(\lambda) - \widehat{\psi}_h(\lambda)\big)\Big| > \varepsilon/3\Big) \\
& + P\Big((n/m)^{0.5} \sup_{\lambda\in\Pi} \Big|\sum_{h=k+1}^{n} \big(\widetilde{\gamma}_A(h) - E\widetilde{\gamma}_A(h)\big)\psi_h(\lambda)\Big| > \varepsilon/3\Big) \\
& + P\Big((n/m)^{0.5} \sup_{\lambda\in\Pi} \Big|\sum_{h=k+1}^{n} \big(\widetilde{\gamma}_A(h) - E\widetilde{\gamma}_A(h)\big)\widehat{\psi}_h(\lambda)\Big| > \varepsilon/3) \\
=\;& V_1 + V_2 + V_3\,.
\end{aligned}$$

An application of Chebyshev's and Hölder's inequalities yields,

$$\begin{aligned}
V_1 \;\le\;& 9\varepsilon^{-2}\frac{n}{m} E \sup_{\lambda\in\Pi} \Big|\sum_{h=0}^{k} \big(\widetilde{\gamma}_A(h) - E\widetilde{\gamma}_A(h)\big)\big(\psi_h(\lambda) - \widehat{\psi}_h(\lambda)\big)\Big|^2 \\
\le\;& c\,\frac{n}{m}\, E \sup_{\lambda\in\Pi} \sum_{h=0}^{k} \big(\widetilde{\gamma}_A(h) - E\widetilde{\gamma}_A(h)\big)^2 \big|\psi_h(\lambda) - \widehat{\psi}_h(\lambda)\big| \sum_{s=0}^{k} \big|\psi_s(\lambda) - \widehat{\psi}_s(\lambda)\big| \\
\le\;& c\, k\frac{n}{m} \sum_{h=0}^{k} \operatorname{var}\big(\widetilde{\gamma}_A(h)\big) \sup_{x\in\Pi} \big|\psi_h(x) - \widehat{\psi}_h(x)\big|\,.
\end{aligned}$$

Next we will study $\sup_{\lambda\in\Pi}|\psi_h(\lambda)-\widehat{\psi}_h(\lambda)|$. Trivially, for $x\in\Pi$, $|\int_{\omega_n(x_n)}^x \cos(h\lambda)\,g(\lambda)\,d\lambda|\le c/n$, where the constant c only depends on g. We also have for the frequencies $x\in\Pi$,

$$\begin{aligned}
&|\psi_h(\omega_n(x_n))-\widehat{\psi}_h(\omega_n(x_n))|\\
=\;& \Big|\sum_{i=1}^{x_n}\Big(\int_{\omega_n(i-1)}^{\omega_n(i)}\cos(h\lambda)\,g(\lambda)\,d\lambda-\omega_n(1)\cos(h\omega_n(i))\,g(\omega_n(i))\Big)\Big|\\
\le\;& \sum_{i=1}^{x_n}\Big|\int_{\omega_n(i-1)}^{\omega_n(i)}\cos(h\lambda)\,(g(\lambda)-g(\omega_n(i)))\,d\lambda\Big| \qquad (4.6.6)\\
&+\Big|\sum_{i=1}^{x_n} g(\omega_n(i))\Big(\frac{\sin(h\omega_n(i))-\sin(h\omega_n(i-1))}{h}-\omega_n(1)\cos(h\omega_n(i))\Big)\Big|\,. \qquad (4.6.7)
\end{aligned}$$

Since g is β-Hölder continuous there exists a constant $c>0$ such that

$$|g(\lambda)-g(\omega_n(i))|\le cn^{-\beta}\,,\quad \lambda\in[\omega_n(i-1),\omega_n(i)]\,.$$

Therefore the term in (4.6.6) is bounded by $c\,n^{-\beta}$. A Taylor expansion as $z\to 0$ yields $\sin(z)=z-z^3/3!+o(z^3)$. Then we have for $h\le n$,

$$\begin{aligned}
&\Big|\frac{\sin(h\omega_n(i))-\sin(h\omega_n(i-1))}{h}-\omega_n(1)\cos(h\omega_n(i))\Big|\\
=\;&\Big|2h^{-1}\sin(h\omega_n(0.5))\cos(h\theta(i+0.5))-\omega_n(1)\cos(h\omega_n(i))\Big|\\
=\;&\Big|2h^{-1}(\sin(h\omega_n(0.5))-h\omega_n(0.5))\cos(h\theta(i+0.5))+\omega_n(1)(\cos(h\omega_n(i+0.5))-\cos(h\omega_n(i)))\Big|\\
\le\;& c(h\omega_n(1))^3+\omega_n(1)\Big|2\sin(h\omega_n(0.25)\sin(h\omega_n(i+0.25))\Big|\le c\,(h^3n^{-3}+hn^{-2})\,.
\end{aligned}$$

Consequently, we have the bound $c(k/n)(1+k^2/n)$ for (4.6.7) uniformly for $x\in\Pi$ and $h\le k$, Thus, uniformly for $h\le k$,

$$\sup_{x\in\Pi}|\widetilde{\psi}_h(x)-\widehat{\psi}_h(x)|\le c\big[n^{-\beta}+(k/n)(1+k^2/n)]\,.$$

As we have shown in Lemma 4.2.4, $(n/m)\sum_{h=0}^k \operatorname{var}\big(\widetilde{\gamma}_A(h)\big)\le c\,k$; see also Davis and Mikosch [7], Lemma 5.2. Thus, as $n\to\infty$,

$$V_1\le c\big[k^2n^{-\beta}+(k^3/n)(1+k^2/n)]\to 0\,.$$

It follows from Lemma 4.6.1 that $\lim_{k\to\infty}\limsup_{n\to\infty}V_2=0$. We adapt the proof of Lemma 4.6.1 for the case V_3. Abusing notation, consider

$$\begin{aligned}
d_h(\omega,\lambda,j,j') \;=\;& \big(\widehat{\psi}_h(\lambda+(j'-1)\pi 2^{-2q+1})-\widehat{\psi}_h(\lambda+(j-1)\pi 2^{-2q+1})\big)\\
&-\big(\widehat{\psi}_h(\omega+(j'-1)\pi 2^{-2q+1})-\widehat{\psi}_h(\omega+(j-1)\pi 2^{-2q+1})\big)\,.
\end{aligned}$$

Recall that we assume $n=2^b$ for some integer b and $x_n=[nx/(2\pi)]$. Therefore for $\lambda\in\Pi$ and integer j,

$$\begin{aligned}
(\lambda+(j-1)\pi 2^{-2q+1})_n \;=\;& [n\lambda/(2\pi)+(j-1)2^{-2q+b}]\\
=\;& [n\lambda/(2\pi)]+(j-1)2^{-2q+b}\\
=\;& \lambda_n+(j-1)2^{-2q+b}\,.
\end{aligned}$$

Thus we can write

$$\begin{aligned}
d_h(\omega,\lambda,j,j') &= \frac{2\pi}{n}\sum_{i=\lambda_n+(j-1)2^{b-2q}}^{\lambda_n+(j'-1)2^{b-2q}} g(\omega_n(i))\cos(h\omega_n(i)) - \frac{2\pi}{n}\sum_{i=\omega_n+(j-1)2^{b-2q}}^{\omega_n+(j'-1)2^{b-2q}} g(\omega_n(i))\cos(h\omega_n(i)) \\
&= \frac{2\pi}{n}\sum_{i=(j-1)2^{b-2q}}^{(j'-1)2^{b-2q}} \Big(g(\omega_n(\lambda_n+i))[\cos(h\omega_n(\lambda_n+i)) - \cos(h\omega_n(\omega_n+i))] \\
&\qquad - [g(\omega_n(\lambda_n+i)) - g(\omega_n(\omega_n+i))]\cos(h\omega_n(\omega_n+i))\Big) \\
&= T_1 + T_2\,.
\end{aligned}$$

Calculation yields

$$\begin{aligned}
|T_1| &\le c|\omega_n(\lambda_n)-\omega_n(\omega_n)|\big|(j'-j)2^{-2q}\big|2^q \le c\big|(\lambda_n-\omega_n)/n\big|\big|(j'-j)2^{-2q}\big|2^q\,, \\
|T_2| &\le c|\omega_n(\lambda_n)-\omega_n(\omega_n)|^\beta\big|(j'-j)2^{-2q}\big| \le c\big|(\lambda_n-\omega_n)/n\big|^\beta\big|(j'-j)2^{-2q}\big|\,.
\end{aligned}$$

Combining these bounds, we have,

$$|d_h(\omega,\lambda,j,j')| \le c|(\lambda_n-\omega_n)/n|^\beta\big|(j'-j)2^{-2q}\big|2^q\,.$$

In the remaining argument we can follow the proof of Lemma 4.6.1; the only difference is that we have to replace the supremum over $\lambda,\omega \in [0, j2^{-2q+1}]$ by the corresponding quantities $\lambda_n/n, \omega_n/n \in [0, j2^{-2q+1}]$. This proves $\lim_{k\to\infty}\limsup_{n\to\infty} V_3 = 0$ and concludes the proof of the lemma.

The proofs of (4.3.6) and (4.3.7) are completely analogous. Instead of the relations (4.2.8) one has to use (4.2.10).

4.7 Proof of Theorem 4.3.3

We adapt the proof of Theorem 4.3.1. We need to prove that

$$n\sum_{h,l}^{(q)} |\mathrm{cov}(\widetilde{\gamma}_A(h),\widetilde{\gamma}_A(h+l))| \le c2^q\,,$$

where $\sum_{h,l}^{(q)}$ is defined in the proof of Lemma 4.6.1. Here $h > \eta$.

$$\begin{aligned}
& n\sum_{h,l}^{(q)} \big|\mathrm{cov}(\widetilde{\gamma}_A(h), \widetilde{\gamma}_A(h+l))\big| \\
= & \frac{m^2}{n}\sum_{h,l}^{(q)} \Big|(n-h-l)\Gamma(0,h,0,h+l) + \sum_{t=1}^{n-h-l-1}(n-h-l-t)\Gamma(0,h,t,t+h+l) \\
& + \sum_{t=1}^{n-h-1}\min(n-h-l, n-h-t)\Gamma(0,h+l,t,t+h)\Big| \\
\le & m^2\sum_{h=2^q}^{2^{q+1}-1}|\Gamma(0,0,h,h)| + m^2\sum_{h=2^q}^{2^{q+1}-1}\sum_{l=1}^{2^{q+1}-h-1}|\Gamma(0,0,h,h+l)| \\
& +m^2\sum_{h,l}^{(q)}\sum_{t=1}^{n-h-l-1}|\Gamma(0,h,t,t+h+l)| + m^2\sum_{h,l}^{(q)}\sum_{t=1}^{n-h-l-1}|\Gamma(0,h+l,t,t+h)| \\
= & m^2\sum_{h=2^q}^{2^{q+1}-1}|\Gamma(0,0,h,h)| + m^2\sum_{h=2^q}^{2^{q+1}-1}\sum_{l=1}^{\eta}|\Gamma(0,0,h,h+l)| \\
& +m^2\sum_{h=2^q}^{2^{q+1}-1}\sum_{l=1}^{\eta}\sum_{t=1}^{\eta}|\Gamma(0,h,t,t+h+l)| + m^2\sum_{h=2^q}^{2^{q+1}-1}\sum_{l=1}^{\eta}\sum_{t=1}^{\eta}|\Gamma(0,h+l,t,t+h)| \\
\le & c2^q
\end{aligned}$$

In the above calculation, we use the facts that for $s \le t \le u \le v$, $\Gamma(s,t,u,v) = 0$ where $t-s > \eta$ or $v-u > \eta$.

In the remaining argument we can follow the proof of Theorem 4.3.1; instead of Lemma 4.2.4 we use the central limit theory of Lemma 4.2.5. □

4.8 Proof of Theorem 4.4.2

We will mimic the proof of Theorem 4.3.1. We start by proving a result for the bootstrapped sample extremogram $\widehat{\gamma}_A^*$ analogous to Theorem 4.4.1.

Lemma 4.8.1. *Under the conditions and with the notation of Theorem 4.4.1, for $h \ge 0$,*

$$d_{P^*}\Big((n/m)^{0.5}\big(\widehat{\gamma}_A^*(i) - E^*\widehat{\gamma}_A^*(i)\big)_{i=0,\dots,h}, (Z_i)_{i=0,\dots,h}\Big) \xrightarrow{P} 0\,, \quad n\to\infty\,.$$

Proof. We start by observing (see Lemma 4.8.3) that for $h \geq 0$

$$\begin{aligned} E^*\widetilde{\gamma}_A^*(h) &= \frac{m}{n}\,(n-h)E^*\widetilde{I}_{1^*}\widetilde{I}_{1^*+h} \\ &= (1-h/n)\Big[\widetilde{\gamma}_A(h)+\frac{m}{n}\sum_{t=n-h+1}^{n}\widetilde{I}_t\widetilde{I}_{t+h}\Big]\,, \\ E^*\widehat{\gamma}_A^*(h) &= \frac{m}{n}\,(n-h)E^*\widehat{I}_{1^*}\widehat{I}_{(1+h)^*} \\ &= (1-h/n)\,(1-\theta)^h\Big[\widehat{\gamma}_A(h)+\frac{m}{n}\sum_{t=n-h+1}^{n}\widehat{I}_t\widehat{I}_{t+h}\Big]\,, \end{aligned}$$

where we interpret indices larger than n modulo n, and therefore

$$(n/m)^{0.5}\big[(1-\theta)^h E^*\widetilde{\gamma}_A^*(h)-E^*\widehat{\gamma}_A^*(h)\big]=O_P(m^{-1})\stackrel{P}{\to}0\,, \tag{4.8.1}$$

where we used that $\overline{I}_n^2-p_0^2=O_P(1/\sqrt{mn})$. By virtue of Theorem 4.4.1 it suffices to show that for any $\varepsilon>0$ and $h\geq 0$,

$$P^*\Big((n/m)^{0.5}\Big|(1-\theta)^h\big(\widetilde{\gamma}_A^*(h)-E^*\widetilde{\gamma}_A^*(h)\big)-\big(\widehat{\gamma}_A^*(h)-E^*\widehat{\gamma}_A^*(h)\big)\Big|>\varepsilon\Big)\stackrel{P}{\to}0\,,\quad n\to\infty\,.$$

Markov's inequality ensures that it suffices to prove that

$$\frac{n}{m}\mathrm{var}^*\big((1-\theta)^h\widetilde{\gamma}_A^*(h)-\widehat{\gamma}_A^*(h)\big)\ \stackrel{P}{\to}0\,,\quad n\to\infty\,.$$

We observe that

$$\begin{aligned} &\frac{n}{m}\mathrm{var}^*\big((1-\theta)^h\widetilde{\gamma}_A^*(h)-\widehat{\gamma}_A^*(h)\big) \\ &= m\big(1-\frac{h}{n}\big)\mathrm{var}^*\Big(\widehat{I}_{1^*}\widehat{I}_{(1+h)^*}-(1-\theta)^h\widetilde{I}_{1^*}\widetilde{I}_{1^*+h}\Big) \\ &\quad +2m\sum_{s=1}^{n-h-1}\big(1-\frac{h+s}{n}\big)\times \\ &\quad \mathrm{cov}^*\Big(\widehat{I}_{1^*}\widehat{I}_{(1+h)^*}-(1-\theta)^h\widetilde{I}_{1^*}\widetilde{I}_{1^*+h},\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*}-(1-\theta)^h\widetilde{I}_{(1+s)^*}\widetilde{I}_{(1+s)^*+h}\Big) \\ &= m\big(1-\frac{h}{n}\big)\mathrm{var}^*\Big(\widehat{I}_{1^*}\widehat{I}_{(1+h)^*}-(1-\theta)^h\widetilde{I}_{1^*}\widetilde{I}_{1^*+h}\Big) \\ &\quad +2m\sum_{s=1}^{n-h-1}\big(1-\frac{h+s}{n}\big)\Big[\mathrm{cov}^*(\widehat{I}_{1^*}\widehat{I}_{(1+h)^*},\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*}) \\ &\quad -(1-\theta)^h\mathrm{cov}^*(\widehat{I}_{1^*}\widehat{I}_{(1+h)^*},\widetilde{I}_{(1+s)^*}\widetilde{I}_{(1+s)^*+h})-(1-\theta)^h\mathrm{cov}^*(\widetilde{I}_{1^*}\widetilde{I}_{1^*+h},\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*}) \\ &\quad +(1-\theta)^{2h}\mathrm{cov}^*(\widetilde{I}_{1^*}\widetilde{I}_{1^*+h},\widetilde{I}_{(1+s)^*}\widetilde{I}_{(1+s)^*+h})\Big] \\ &= Q_1+Q_2\,. \end{aligned}$$

We will show that the right-hand side converges to zero in P-probability, where we focus on Q_2 and omit the details for Q_1. We start by looking at the summands in Q_2 for fixed $s\leq h$, using the structure of the covariances in Lemma 4.8.3. The expressions for the covariances in Lemma 4.8.3 contain terms with normalization n^{-2}. For example, by (4.8.7) a corresponding term in Q_2 is of the order

$$m\Big(n^{-1}\sum_{i=1}^n \widetilde{I}_i\widetilde{I}_{i+h}\Big)^2 = m^{-1}\Big(\frac{m}{n}\sum_{i=1}^n \widetilde{I}_i\widetilde{I}_{i+h}\Big)^2 = O_P(m^{-1})\,,$$

since $\frac{m}{n}\sum_{i=1}^n \widetilde{I}_i\widetilde{I}_{i+h} \stackrel{P}{\to} \gamma_A(h)$; see Lemma 4.2.4. In the latter sums, the $\widetilde{I}_i$'s can be exchanged by the I_i's or the $\widehat{I}_i$'s. Therefore all other terms in Q_2 with normalization mn^{-2} converge to zero in P-probability. Another appeal to Lemma 4.8.3 shows that it remains to consider those expressions in Q_2 that are normalized by mn^{-1} again for fixed $s \le h$. From (4.8.9) and (4.8.10) we see that, on one hand, we have to deal with the differences

$$(1-\theta)^{s+h}\ \frac{m}{n}\sum_{i=1}^n \widehat{I}_i\widehat{I}_{i+s}\widehat{I}_{i+h}\widehat{I}_{i+s+h} - (1-\theta)^{s+2h}\frac{m}{n}\sum_{i=1}^n \widetilde{I}_i\widetilde{I}_{i+h}\widehat{I}_{i+s}\widehat{I}_{i+s+h}\,, \tag{4.8.2}$$

but both sums are consistent estimators of $\lim_{n\to\infty} mP(a_m^{-1}X_0 \in A, a_m^{-1}X_s \in A, a_m^{-1}X_h \in A, a_m^{-1}X_{s+h} \in A)$ (see [7], Theorem 3.1). Therefore (4.8.2) converges to zero in P-probability. On the other hand, in view of (4.8.7) and (4.8.8) we have to deal with the differences, for $s \le h$,

$$(1-\theta)^{s+2h}\frac{m}{n}\sum_{i=1}^n \widetilde{I}_i\widetilde{I}_{i+s}\widetilde{I}_{i+h}\widetilde{I}_{i+s+h} - (1-\theta)^{2h}\ \frac{m}{n}\sum_{i=1}^n \widehat{I}_i\widehat{I}_{i+h}\widetilde{I}_{i+s}\widetilde{I}_{i+s+h}\,,$$

which again converge to zero in P-probability. These arguments finish the proof for $s \le h$.

An inspection of the covariances in Lemma 4.8.3 shows that for $s > h$ all expressions with normalization n^{-2} do not depend s. The corresponding aggregated terms in Q_2 are then given by

$$\begin{aligned}
&2m\sum_{s=h+1}^{n-h-1}\big(1-\frac{h+s}{n}\big)\Big[-(1-\theta)^{s+h}\Big(n^{-1}\sum_{i=1}^n \widehat{I}_i\widehat{I}_{i+h}\Big)^2\\
&+(1-\theta)^{s+h}\Big(n^{-1}\sum_{i=1}^n \widehat{I}_i\widehat{I}_{i+h}\Big)\Big(n^{-1}\sum_{i=1}^n \widetilde{I}_i\widetilde{I}_{i+h}\Big)\Big) + (1-\theta)^{s+2h}\ \Big(n^{-1}\sum_{i=1}^n \widehat{I}_i\widehat{I}_{i+h}\Big)\Big(n^{-1}\sum_{i=1}^n \widetilde{I}_i\widetilde{I}_{i+h}\Big)\\
&-(1-\theta)^{s+2h}\ \Big(n^{-1}\sum_{i=1}^n \widetilde{I}_i\widetilde{I}_{i+h}\Big)^2\Big]\\
= \quad & -2m^{-1}\Big(\frac{m}{n}\sum_{i=1}^n \widehat{I}_i\widehat{I}_{i+h} - \frac{m}{n}\sum_{i=1}^n \widetilde{I}_i\widetilde{I}_{i+h}\Big)\Big(\frac{m}{n}\sum_{i=1}^n \widehat{I}_i\widehat{I}_{i+h}\Big)\sum_{s=h+1}^{n-h-1}\big(1-\frac{h+s}{n}\big)\ (1-\theta)^{s+h}\\
&-2m^{-1}\Big(\frac{m}{n}\sum_{i=1}^n \widetilde{I}_i\widetilde{I}_{i+h} - \frac{m}{n}\sum_{i=1}^n \widehat{I}_i\widehat{I}_{i+h}\Big)\Big(\frac{m}{n}\sum_{i=1}^n \widetilde{I}_i\widetilde{I}_{i+h}\Big)\sum_{s=h+1}^{n-h-1}\big(1-\frac{h+s}{n}\big)\ (1-\theta)^{s+2h}\\
= \quad & O_P(1/(\theta\sqrt{mn})) = o_P(1)\,.
\end{aligned}$$

In the last step we used (4.8.1) and the assumption $n\theta^2/m \to \infty$. Finally, we deal with the remaining terms in Q_2. In view of Lemma 4.8.3 they are given by

$$
\begin{aligned}
& 2m \sum_{s=h+1}^{n-h-1} \big(1-\frac{h+s}{n}\big)\Big[(1-\theta)^{s+h}\, n^{-1}\sum_{i=1}^{n} \widehat{I}_i \widehat{I}_{i+s}\widehat{I}_{i+h}\widehat{I}_{i+s+h} \\
& -(1-\theta)^{s+h}\, n^{-1}\sum_{i=1}^{n} \widehat{I}_i \widehat{I}_{i+h}\widetilde{I}_{i+s}\widetilde{I}_{i+s+h} \\
& -(1-\theta)^{s+2h}\, n^{-1}\sum_{i=1}^{n} \widetilde{I}_i \widetilde{I}_{i+h}\widehat{I}_{i+s}\widehat{I}_{i+s+h} + (1-\theta)^{s+2h}\, n^{-1}\sum_{i=1}^{n} \widetilde{I}_i \widetilde{I}_{i+s}\widetilde{I}_{i+h}\widetilde{I}_{i+s+h}\Big] \\
= \;& 2m \sum_{s=h+1}^{n-h-1} \big(1-\frac{h+s}{n}\big)(1-\theta)^{s+h} n^{-1}\sum_{i=1}^{n} \widehat{I}_i \widehat{I}_{i+h}\big(\widehat{I}_{i+s}\widehat{I}_{i+s+h} - \widetilde{I}_{i+s}\widetilde{I}_{i+s+h}\big) \\
+ \;& 2m \sum_{s=h+1}^{n-h-1} \big(1-\frac{h+s}{n}\big)(1-\theta)^{s+2h}\, n^{-1}\sum_{i=1}^{n} \widetilde{I}_i \widetilde{I}_{i+h}\big(\widetilde{I}_{i+s}\widetilde{I}_{i+s+h} - \widehat{I}_{i+s}\widehat{I}_{i+s+h}\big) \\
= \;& J_0\,.
\end{aligned}
$$

Using the assumption $n\theta^2/m \to \infty$, we have

$$
\begin{aligned}
E|J_0| &\le c\,m \sum_{s=h+1}^{n-h-1} (1-\theta)^{s+h} E|\widehat{I}_0\widehat{I}_h - \widetilde{I}_0\widetilde{I}_h| \\
&\le c\,mE|p_0 - \overline{I}_n| \sum_{s=h+1}^{n-h-1} (1-\theta)^{s+h} \\
&\le c(m/n)^{0.5}\theta^{-1} = o(1)\,.
\end{aligned}
$$

This finishes the proof of the lemma. □

We conclude from Lemma 4.8.1 that for any $k \ge 1$, as $n \to \infty$,

$$
d_{P^*}\Big((n/m)^{0.5}\Big(\psi_0\left(\widehat{\gamma}^*_A(0) - E^*\widehat{\gamma}^*_A(0)\right) + 2\sum_{h=1}^{k}\psi_h\left(\widehat{\gamma}^*_A(h) - E^*\widehat{\gamma}^*_A(h)\right)\Big), \psi_0\, Z_0 + 2\sum_{h=1}^{k}\psi_h\, Z_h\Big) \stackrel{P}{\to} 0\,,
$$

where the dependence structure of (Z_h) is defined in Lemma 4.2.4.

The proof of the theorem is finished by the following result which parallels Lemma 4.6.1.

Lemma 4.8.2. *Assume the conditions of Theorem 4.4.2. Then the following relation holds for* $\delta > 0$,

$$
\lim_{k\to\infty}\limsup_{n\to\infty} P\Big(d_{P^*}\Big((n/m)^{0.5}\sup_{\lambda\in\Pi}\Big|\sum_{h=k+1}^{n-1}\psi_h(\lambda)\left(\widehat{\gamma}^*_A(h) - E^*\widehat{\gamma}^*_A(h)\right)\Big|, 0\Big) > \delta\Big) = 0\,. \quad (4.8.3)
$$

Proof. We follow the lines of the proof of Lemma 4.6.1 and use the same notation. We again assume without loss of generality that $k = 2^a - 1$ and $n = 2^{b+1}$ for integers $a < b$, a chosen sufficiently large, and we write $\varepsilon_q = 2^{-2q/\kappa}$ for $\kappa > 0$ to be chosen later. Then, for large a depending on $\varepsilon > 0$, the steps of the proof lead to the inequality (cf. (4.6.5))

$$\begin{aligned} Q^* &= P^*\Big((n/m)\sup_{\lambda\in\Pi}\Big|\sum_{h=k+1}^{n-1}\big(\widetilde{\gamma}_A^*(h)-E^*\widetilde{\gamma}_A^*(h)\big)\,\psi_h(\lambda)\Big|>\varepsilon\Big) \\ &\le c\sum_{q=a}^{b}2^{4q(0.25-\beta+\kappa^{-1})}K_q\,, \end{aligned}$$

where $\beta \in (3/4, 1]$ is the Hölder coefficient of the function g, the number $\kappa > 0$ can be chosen arbitrarily large and

$$K_q = \frac{n}{m}\sum_{h=2^q}^{2^{q+1}-1}\sum_{s=h}^{2^{q+1}-1}|\mathrm{cov}^*(\widehat{\gamma}_A^*(h),\widehat{\gamma}_A^*(s)|\,.$$

By the Cauchy-Schwarz inequality, for $s, h \in [2^q, 2^{q+1})$ and $h \le s$,

$$(n/m)^2|\mathrm{cov}^*(\widehat{\gamma}_A^*(h),\widehat{\gamma}_A^*(s)|^2 \le (n/m)\mathrm{var}^*(\widehat{\gamma}_A^*(h))\;(n/m)\mathrm{var}^*(\widehat{\gamma}_A^*(s))\,.$$

We will show that

$$(n/m)E\mathrm{var}^*(\widehat{\gamma}_A^*(h)) \le c \tag{4.8.4}$$

for some constant c, uniformly for $k \le h \le n$ and n. Then

$$EQ^* \le c\sum_{q=a}^{b}2^{4q(3/4-\beta+\kappa^{-1})} \le c\sum_{q=a}^{\infty}2^{4q(3/4-\beta+\kappa^{-1})}\,.$$

The right-hand side converges since $\beta \in (3/4, 1]$ and κ can be chosen arbitrarily large. Moreover, the right-hand side converges to zero as $k \to \infty$.

Thus it remains to show (4.8.4). In view of Lemma 4.8.3 we have

$$
\begin{aligned}
&(n/m)E\mathrm{var}^*(\widehat{\gamma}_A^*(h))\\
&= (m/n)\Big[(n-h)E\mathrm{var}^*(\widehat{I}_{1^*}\widehat{I}_{(1+h)^*}) + 2\sum_{t=1}^{n-h-1}(n-h-t)E\mathrm{cov}^*(\widehat{I}_{1^*}\widehat{I}_{(1+h)^*},\widehat{I}_{(1+t)^*}\widehat{I}_{(1+t+h)^*})\Big]\\
&= \Big[m(1-h/n)(1-\theta)^{2h}\Big[E(\widehat{I}_1\widehat{I}_{1+h})^2 - E\Big(n^{-1}\sum_{i=1}^{n}\widehat{I}_i\widehat{I}_{i+h}\Big)^2\Big]\Big]\\
&\quad +2m\sum_{t=1}^{n-h-1}(1-(h+t)/n)\Big[n^{-1}\sum_{i=1}^{n}E\widehat{I}_i\widehat{I}_{i+h}\widehat{I}_{i+t}\widehat{I}_{i+t+h}\Big](1-\theta)^{t+h}\\
&\quad +2m\sum_{t=1}^{\min(h-1,n-h-1)}(1-(h+t)/n)E\Big(n^{-1}\sum_{i=1}^{n}\widehat{I}_i\widehat{I}_{i+t}\Big)^2((1-\theta)^{2t}-(1-\theta)^{t+h})\\
&\quad -2m\sum_{t=1}^{\min(h-1,n-h-1)}(1-(h+t)/n)E\Big(n^{-1}\sum_{i=1}^{n}\widehat{I}_i\widehat{I}_{i+h}\Big)^2(1-\theta)^{2h}\\
&\quad -2m\sum_{t=h}^{n-h-1}(1-(h+t)/n)E\Big(n^{-1}\sum_{i=1}^{n}\widehat{I}_i\widehat{I}_{i+h}\Big)^2(1-\theta)^{t+h}\\
&\le m\,E(\widehat{I}_1\widehat{I}_{1+h})^2\\
&\quad +2m\sum_{t=1}^{n-h-1}(1-(h+t)/n)\Big(n^{-1}\sum_{i=1}^{n}E\widehat{I}_i\widehat{I}_{i+h}\widehat{I}_{i+t}\widehat{I}_{i+t+h}\Big)(1-\theta)^{t+h}\\
&\quad +2m\sum_{t=1}^{\min(h-1,n-h-1)}(1-(h+t)/n)E\Big(n^{-1}\sum_{i=1}^{n}\widehat{I}_i\widehat{I}_{i+t}\Big)^2(1-\theta)^{2t}\\
&= V_1+V_2+V_3\,.
\end{aligned}
$$

We observe that, for some constant $c_0>0$,

$$V_1 \le m\,E(\widehat{I}_1\widehat{I}_{1+h})^2 \le c\,m\,\Big[EI_1I_{1+h}+(E\overline{I}_n)^2\Big] \le cm\,p_0 \le c_0\,.$$

For V_2, we observe that for $i\le n$,

$$m\theta^{-1}\big|E\big[\widehat{I}_i\widehat{I}_{i+h}\widehat{I}_{i+t}\widehat{I}_{i+t+h}-\widetilde{I}_i\widetilde{I}_{i+h}\widetilde{I}_{i+t}\widetilde{I}_{i+t+h}\big]\big| \le c\,m\theta^{-1}E\big|\overline{I}_n-p_0\big| = O(\sqrt{m/n}\theta^{-1}) = o(1)\,,$$

by virtue of the condition $n\theta^2/m\to\infty$. Therefore, for showing that $|V_2|\le c$ uniformly for h,n, it suffices to show that $|\widetilde{V}_2|\le c$, where $\widetilde{V}_2$ is obtained from V_2 by replacing the $\widehat{I}_t$'s by the corresponding $\widetilde{I}_t$'s. Taking into account $E\widetilde{I}_1\widetilde{I}_{1+t}=p_t-p_0^2$ and the Cauchy-Schwarz inequality, we have for a fixed integer $M>0$,

$$
\begin{aligned}
|\widetilde{V}_2| &\le c\,m\sum_{t=1}^{n-h-1}\Big|n^{-1}\sum_{i=1}^{n}E\widetilde{I}_i\widetilde{I}_{i+h}\widetilde{I}_{i+t}\widetilde{I}_{i+t+h}\Big|\\
&= c\,m\sum_{t=1}^{n-h-1}|E\widetilde{I}_1\widetilde{I}_{1+h}\widetilde{I}_{1+t}\widetilde{I}_{1+t+h}|\\
&\le (mp_0)M + c\,m\sum_{t=M+1}^{r_n}(p_t+p_0^2) + c\,m\sum_{t=r_n+1}^{\infty}\xi_t \le c\,,
\end{aligned}
$$

in view of condition (M) and regular variation. A similar argument as for V_2 shows that one may replace the $\widehat{I}_t$'s in V_3 by the corresponding $\widetilde{I}_t$'s. We denote the resulting quantity by $\widetilde{V}_3$. Then we have

$$\begin{aligned}
\widetilde{V}_3 &\leq m\sum_{t=1}^{n}(1-\theta)^t E\Big(n^{-1}\sum_{i=1}^{n-t}\widetilde{I}_i\widetilde{I}_{i+t}+n^{-1}\sum_{i=n-t+1}^{n}\widetilde{I}_i\widetilde{I}_{i+t-n}\Big)^2\\
&\leq c\,m\sum_{t=1}^{n}(1-\theta)^t E\Big(n^{-1}\sum_{i=1}^{n-t}\widetilde{I}_i\widetilde{I}_{i+t}\Big)^2+c\,m\sum_{t=1}^{n}(1-\theta)^t E\Big(n^{-1}\sum_{i=n-t+1}^{n}\widetilde{I}_i\widetilde{I}_{i+t-n}\Big)^2\\
&= \widetilde{V}_{31}+\widetilde{V}_{32}\,.
\end{aligned}$$

We will only deal with $\widetilde{V}_{31}$, the other term can be bounded in a similar way. We observe that for fixed $M>1$, using condition (M),

$$\begin{aligned}
\widetilde{V}_{31} &\leq c\,\frac{m}{n}\sum_{t=1}^{n}(1-\theta)^t\Big(E(\widetilde{I}_1\widetilde{I}_{1+t})^2+2\sum_{s=1}^{n-t-1}|E\widetilde{I}_1\widetilde{I}_{1+t}\widetilde{I}_{1+s}\widetilde{I}_{1+s+t}|\Big)\\
&\leq o(1)+c\frac{m}{n}\sum_{t=1}^{n}(1-\theta)^t\sum_{s=1}^{n-t-1}|E\widetilde{I}_1\widetilde{I}_{1+t}\widetilde{I}_{1+s}\widetilde{I}_{1+s+t}|\\
&\leq o(1)+c\frac{m}{n}\sum_{t=1}^{n}(1-\theta)^t\sum_{s=M+1}^{r_n}(p_s+p_0^2)+c\frac{m}{n}\sum_{t=1}^{n}(1-\theta)^t\sum_{r_n+1\leq s\leq n-t-1,s\leq t}\xi_s\\
&\quad +c\frac{m}{n}\sum_{t=1}^{n}(1-\theta)^t\sum_{r_n+1\leq s\leq n-t-1,s>t}\big(|E\widetilde{I}_1\widetilde{I}_{1+t}\widetilde{I}_{1+s}\widetilde{I}_{1+s+t}-(p_t-p_0^2)^2|+(p_t-p_0^2)^2\big)\,.
\end{aligned}$$

In view of condition (M), the first two terms on the right-hand side are negligible as $n\to\infty$. The third term is bounded by

$$c\frac{m}{n}\sum_{t=1}^{n}(1-\theta)^t\sum_{r_n+1\leq s\leq n-t-1,s>t}\xi_{s-t}+cm\sum_{t=1}^{n}(1-\theta)^t(p_t-p_0^2)^2\,.$$

Multiple use of (M) again shows that the right-hand side is negligible. This proves (4.8.4). $\square$

Lemma 4.8.3. *Under the conditions of Theorem 4.4.2 the following relations hold for* $s,h\geq 0$*:*[1]

$$E^*\widehat{I}_{1^*}=0\,, \tag{4.8.5}$$

$$E^*\widehat{I}_{1^*}\widehat{I}_{(1+h)^*}=(1-\theta)^h\,n^{-1}\sum_{i=1}^n\widehat{I}_i\widehat{I}_{i+h}\,,\quad E^*\widetilde{I}_{1^*}\widetilde{I}_{1^*+h}=n^{-1}\sum_{i=1}^n\widetilde{I}_i\widetilde{I}_{i+h}\,, \tag{4.8.6}$$

$$\begin{aligned}&\operatorname{cov}^*(\widetilde{I}_{1^*}\widetilde{I}_{1^*+h},\widetilde{I}_{(1+s)^*}\widetilde{I}_{(1+s)^*+h}) \\ =\;&(1-\theta)^s\Big(n^{-1}\sum_{i=1}^n\widetilde{I}_i\widetilde{I}_{i+s}\widetilde{I}_{i+h}\widetilde{I}_{i+s+h}-\Big(n^{-1}\sum_{i=1}^n\widetilde{I}_i\widetilde{I}_{i+h}\Big)^2\Big)\,,\end{aligned} \tag{4.8.7}$$

$$\begin{aligned}&\operatorname{cov}^*\big(\widehat{I}_{1^*}\widehat{I}_{(1+h)^*},\widetilde{I}_{(1+s)^*}\widetilde{I}_{(1+s)^*+h}\big) \\ =\;&(1-\theta)^{\max(s,h)}\,\Big(n^{-1}\sum_{i=1}^n\widehat{I}_i\widehat{I}_{i+h}\widetilde{I}_{i+s}\widetilde{I}_{i+s+h}-\Big(n^{-1}\sum_{i=1}^n\widehat{I}_i\widehat{I}_{i+h}\Big)\Big(n^{-1}\sum_{i=1}^n\widetilde{I}_i\widetilde{I}_{i+h}\Big)\Big)\,,\end{aligned} \tag{4.8.8}$$

$$\begin{aligned}&\operatorname{cov}^*\big(\widetilde{I}_{1^*}\widetilde{I}_{1^*+h},\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*}\big) \\ =\;&(1-\theta)^{s+h}\,\Big(n^{-1}\sum_{i=1}^n\widetilde{I}_i\widetilde{I}_{i+h}\widehat{I}_{i+s}\widehat{I}_{i+s+h}-\Big(n^{-1}\sum_{i=1}^n\widehat{I}_i\widehat{I}_{i+h}\Big)\Big(n^{-1}\sum_{i=1}^n\widetilde{I}_i\widetilde{I}_{i+h}\Big)\Big)\,,\end{aligned} \tag{4.8.9}$$

$$\begin{aligned}&\operatorname{cov}^*\big(\widehat{I}_{1^*}\widehat{I}_{(1+h)^*},\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*}\big) \\ =&\begin{cases}(1-\theta)^{s+h}\,\Big[n^{-1}\sum_{i=1}^n\widehat{I}_i\widehat{I}_{i+s}\widehat{I}_{i+h}\widehat{I}_{i+s+h}-\Big(n^{-1}\sum_{i=1}^n\widehat{I}_i\widehat{I}_{i+s}\Big)^2\Big]+\\ \Big(n^{-1}\sum_{i=1}^n\widehat{I}_i\widehat{I}_{i+s}(1-\theta)^s\Big)^2-\Big(n^{-1}\sum_{i=1}^n\widehat{I}_i\widehat{I}_{i+h}(1-\theta)^h\Big)^2, & s<h\,,\\ (1-\theta)^{s+h}\,\Big(n^{-1}\sum_{i=1}^n\widehat{I}_i\widehat{I}_{i+s}\widehat{I}_{i+h}\widehat{I}_{i+s+h}-\Big(n^{-1}\sum_{i=1}^n\widehat{I}_i\widehat{I}_{i+h}\Big)^2\Big), & s\geq h\,.\end{cases}\end{aligned} \tag{4.8.10}$$

Proof. Relations (4.8.5) and (4.8.6) follow from the defining properties of the stationary bootstrap; see Politis and Romano [21].

We will only show that (4.8.10) holds; (4.8.7)–(4.8.9) can be proved in a similar (and even simpler) way but we omit further details. First assume $s<h$. Recall L_1 from the construction of the stationary bootstrap scheme. Consider the following decomposition

$$\begin{aligned}&E^*[\widehat{I}_{1^*}\widehat{I}_{(1+h)^*}\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*}]\\ &\quad=\;E^*[\widehat{I}_{1^*}\widehat{I}_{(1+h)^*}\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*}\mid L_1\leq s]\,P(L_1\leq s)\\ &\qquad+E^*[\widehat{I}_{1^*}\widehat{I}_{(1+h)^*}\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*}\mid s<L_1\leq h]\,P(s<L_1\leq h)\\ &\qquad+E^*[\widehat{I}_{1^*}\widehat{I}_{(1+h)^*}\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*}\mid h<L_1\leq s+h]\,P(h<L_1\leq s+h)\\ &\qquad+E^*[\widehat{I}_{1^*}\widehat{I}_{(1+h)^*}\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*}\mid L_1>s+h]\,P(L_1>s+h)\\ &\quad=\;Q_1+Q_2+Q_3+Q_4\,.\end{aligned}$$

We start with Q_1. For $L_1\leq s<h$, $\widehat{I}_{1^*}$ is independent of $(\widehat{I}_{(1+h)^*},\widehat{I}_{(1+s)^*},\widehat{I}_{(1+s+h)^*})$, given (X_t), but $E^*\widehat{I}_{1^*}=0$ by (4.8.5) and therefore $Q_1=0$. Similarly, for $h<L_1\leq s+h$, $\widehat{I}_{(1+s+h)^*}$ is independent of $(\widehat{I}_{1^*},\widehat{I}_{(1+h)^*},\widehat{I}_{(1+s)^*})$, given (X_t), and since $E^*\widehat{I}_{(1+s+h)^*}=0$, $Q_3=0$. Each of the values $i=1,\ldots,n$ has the same chance to be chosen by the bootstrap, i.e. $P^*(\widehat{I}_1^*=$

[1] If indices in the sums below exceed the value n they are interpreted in the circular sense, i.e. mod n.

$\widehat{I}_i) = n^{-1}$ for $i = 1,\dots,n$. Thus, for $L_1 > s+h$ and the chosen i, the natural ordering $(1^*,(1+h)^*,(1+s)^*,(1+s+h)^*) = (i,i+h,i+s,i+s+h)$ is preserved and therefore

$$\begin{aligned} Q_4 &= n^{-1}\sum_{i=1}^{n} \widehat{I}_i\widehat{I}_{i+s}\widehat{I}_{i+h}\widehat{I}_{i+s+h}\; P(L_1 > s+h) \\ &= n^{-1}\sum_{i=1}^{n} \widehat{I}_i\widehat{I}_{i+s}\widehat{I}_{i+h}\widehat{I}_{i+s+h}\; (1-\theta)^{s+h}\,. \end{aligned}$$

By a similar argument, (4.8.6) and using stationarity, we have

$$\begin{aligned} Q_2 &= E^*[\widehat{I}_{1^*}\widehat{I}_{(1+s)^*} \mid s < L_1 \le h]\, E^*[\widehat{I}_{(1+h)^*}\widehat{I}_{(1+h+s)^*}]\; P(s < L_1 \le h) \\ &= n^{-2}\Big(\sum_{i=1}^{n} \widehat{I}_i\widehat{I}_{i+s}\Big)^2 (1-\theta)^s\, \Big((1-\theta)^s - (1-\theta)^h\Big)\,. \end{aligned}$$

Combining the above expressions and taking into account (4.8.6), we arrive at (4.8.10) for $s < h$.

We proceed with the case $s > h$. Then we have the corresponding decomposition

$$\begin{aligned} &E^*[\widehat{I}_{1^*}\widehat{I}_{(1+h)^*}\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*}] \\ &\quad= E^*[\widehat{I}_{1^*}\widehat{I}_{(1+h)^*}\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*} \mid L_1 \le h]\; P(L_1 \le h) \\ &\qquad + E^*[\widehat{I}_{1^*}\widehat{I}_{(1+h)^*}\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*} \mid h < L_1 \le s]\; P(h < L_1 \le s) \\ &\qquad + E^*[\widehat{I}_{1^*}\widehat{I}_{(1+h)^*}\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*} \mid s < L_1 \le s+h]\; P(s < L_1 \le s+h) \\ &\qquad + E^*[\widehat{I}_{1^*}\widehat{I}_{(1+h)^*}\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*} \mid L_1 > s+h]\; P(L_1 > s+h) \\ &\quad= Q_1' + Q_2' + Q_3' + Q_4'\,. \end{aligned}$$

We observe that the left-hand side is symmetric in h, s and therefore the same arguments as above show that $Q_1' = Q_3' = 0$, $Q_4 = Q_4'$ and

$$\begin{aligned} Q_2' &= E^*[\widehat{I}_{1^*}\widehat{I}_{(1+h)^*} \mid h < L_1 \le s]\, E^*[\widehat{I}_{(1+s)^*}\widehat{I}_{(1+s+h)^*}]\; P(h < L_1 \le s) \\ &= n^{-2}\Big(\sum_{i=1}^{n} \widehat{I}_i\widehat{I}_{i+h}\Big)^2 (1-\theta)^h\, \Big((1-\theta)^h - (1-\theta)^s\Big) \end{aligned}$$

The case $h = s$ can be considered as a degenerate case, where $Q_2' = 0$. This completes the proof of (4.8.10). $\square$

We conclude with a short discussion of the bias problem of the bootstrapped integrated periodogram mentioned in Remark 4.4.3.

Lemma 4.8.4. *Assume the conditions of Theorem 4.4.2 and the additional condition* $\sup_{x\in\Pi}|\psi_h(x)| \le c/h$ *for* $h \ge 1$ *and a constant* c. *Then the following relation holds as* $n \to \infty$,

$$(n/m)^{0.5} \sup_{\lambda\in\Pi}\Big|\psi_0(\lambda)\big(E^*\widehat{\gamma}_A^*(0) - \widetilde{\gamma}_A(0)\big) + 2\sum_{h=1}^{n-1}\psi_h(\lambda)\,\big(E^*\widehat{\gamma}_A^*(h) - (1-\theta)^h\widetilde{\gamma}_A(h)\big)\Big| \stackrel{P}{\to} 0\,. \tag{4.8.11}$$

Proof. We observe that for $h \ge 0$,

$$\begin{aligned} E^*\widehat{\gamma}^*_A(h) - (1-\theta)^h \widetilde{\gamma}_A(h) &= (1-\theta)^h \big[(\widehat{\gamma}_A(h) + \widehat{\gamma}_A(n-h)) - \widetilde{\gamma}_A(h)\big] \\ &= (1-\theta)^h \big[\widetilde{\gamma}_A(n-h) - m(p_0 - \overline{I}_n)^2\big] . \end{aligned} \quad (4.8.12)$$

For fixed h we have $(n/m)^{0.5} m(p_0 - \overline{I}_n)^2 \stackrel{P}{\to} 0$ as $n \to \infty$ and

$$(n/m)^{0.5} E|\widetilde{\gamma}_A(n-h)| \le c\,(m/n)^{0.5} h p_0 \to 0\,, \quad n \to \infty\,.$$

Therefore it suffices to show that

$$\lim_{k\to\infty} \limsup_{n\to\infty} P\Big(\sup_{\lambda\in\Pi}\Big|\sum_{h=k+1}^{n-1} \psi_h(\lambda)\,\big(E^*\widehat{\gamma}^*_A(h) - (1-\theta)^h\widetilde{\gamma}_A(h)\big)\Big| > \delta\Big)\,, \quad \delta > 0\,.$$

Keeping in mind (4.8.12), we have

$$(n/m)^{0.5} m(p_0 - \overline{I}_n)^2 \sup_{\lambda\in\Pi}\Big|\sum_{h=k+1}^{n-1} \psi_h(\lambda)(1-\theta)^h\Big| = O_P(1/(\theta\sqrt{mn})) = o_P(1)\,,$$

where we used $\theta^2 n/m \to \infty$, and

$$\begin{aligned} &(n/m)^{0.5} \sup_{\lambda\in\Pi}\Big|\sum_{h=k+1}^{n-1} \psi_h(\lambda)\,(1-\theta)^h \widetilde{\gamma}_A(n-h)\Big| \\ \le\;& (n/m)^{0.5} \sup_{\lambda\in\Pi}\Big|\sum_{h=1}^{n-k-1} \psi_{n-h}(\lambda)\,(1-\theta)^{n-h}[\widetilde{\gamma}_A(h) - m(1-h/n)(p_h - p_0^2)]\Big| \\ &+ (n/m)^{0.5} \sup_{\lambda\in\Pi}\Big|\sum_{h=1}^{n-k-1} \psi_{n-h}(\lambda)\,(1-\theta)^{n-h} m(1-h/n)(p_h - p_0^2)\Big| \\ =\;& I_1 + I_2\,. \end{aligned}$$

Under the assumption $\sup_{x\in\Pi}|\psi_h(x)| \le c/h$ uniformly for $h \ge 1$, we have for small $\varepsilon > 0$,

$$I_2 \le (m/n)^{0.5} c \sum_{h=1}^{\infty} \xi_h \to 0\,, \quad n \to \infty\,.$$

Now we can adapt the proof of Lemma 4.6.1 to prove that

$$\lim_{k\to\infty} \limsup_{n\to\infty} P(I_1 > \delta) = 0\,, \quad \delta > 0\,.$$

This proves (4.8.11). □

However, under the assumptions of Theorem 4.3.1, it is in general not possible to replace the quantities $(1-\theta)^h \widetilde{\gamma}_A(h)$ in (4.8.11) by $\widetilde{\gamma}_A(h)$, i.e. in general we do not have the relation

$(n/m)^{0.5}(E^* J^*_{nA} - J_{nA}) \stackrel{P}{\to} 0$. Indeed, taking into account (4.8.11) and assuming η-dependence for (X_t), we have $E\widetilde{\gamma}_A(h) = 0$ for $h > \eta$ and

$$\begin{aligned}
(n/m)^{0.5}(E^* J^*_{nA} - J_{nA}) &= 2(n/m)^{0.5}\sum_{h=1}^{n-1}\psi_h(\lambda)\,[(1-\theta)^h - 1]\widetilde{\gamma}_A(h) + o_P(1) \\
&= 2(n/m)^{0.5}\sum_{h=1}^{n-1}\psi_h(\lambda)\,[(1-\theta)^h - 1]\big(\widetilde{\gamma}_A(h) - E\widetilde{\gamma}_A(h)\big) \\
&\quad + 2(n/m)^{0.5}\sum_{h=1}^{\eta}\psi_h(\lambda)\,[(1-\theta)^h - 1](1-h/n)m(p_h - p_0^2) + o_P(1)\,.
\end{aligned}$$

An argument similar to the proof of Theorem 4.3.1 shows that the first term on the right-hand side is stochastically bounded, while the second term may diverge (for example, if $\gamma_A(\eta) > 0$ and $\psi_\eta \neq 0$) since it is of the order $\theta(n/m)^{0.5}$ which converges to infinity in view of the assumption $\theta^2 n/m \to \infty$ which is vital for the proof of the consistency of the stationary bootstrap.

Bibliography

[1] Billingsley, P. (1999) *Convergence of Probability Measures.* 2nd Edition. Wiley, New York.

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

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

[4] Dahlhaus, R. (1988) Empirical spectral processes and their applications to time series analysis. *Stoch. Proc. Appl.* **30**, 69–83.

[5] Dahlhaus, R. and Polonik, W. (2002) Empirical spectral processes and nonparametric maximum likelihood estimation for time series. In: *Dehling, H.G. Mikosch, T. and Sørensen, M.* (Eds.) (2002) *Empirical Process Techniques for Dependent Data.* Birkhäuser, Boston, pp. 275–298,

[6] 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.

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

[8] Davis, R.A. and Mikosch, T. (2009) Extremes of stochastic volatility models. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.) *Handbook of Financial Time Series.* Springer (2009), pp. 355–364.

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

[10] Davis, R.A., Mikosch, T. and Zhao, Y. (2013) Measures of serial extremal dependence and their estimation. *Stoch. Proc. Appl.* **123**, 2575–2602.

[11] Dehling, H., Durieu, O. and Volny, D. (2009) New techniques for empirical processes of dependent data. *Stoch. Proc. Appl.* **119**, 3699–3718.

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

[13] Grenander, U. and Rosenblatt, M. (1984) *Statistical Analysis of Stationary Time Series.* 2nd Edition. Chelsea Publishing Co., New York,

[14] Hannan, E.J. (1960) *Time Series Analysis.* Wiley, New York.

[15] Hida, T. (1980) *Brownian Motion.* Springer, New York.

[16] Ibragimov. I.A. and and Linnik, Yu.V. (1971) *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff, Groningen.

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

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

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

[20] Petrov, V.V. (1995) *Limit Theorems of Probability Theory.* Oxford University Press, Oxford (UK).

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

[22] Priestley, M.B. (1981) *Spectral Analysis and Time Series.* Academic Press, London, New York.

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

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

[25] Rio, E. (1995) About the Lindeberg method for strongly mixing sequences. *ESAIM: Probab. and Statist.* **1**, 35–61.

[26] Shorack, G.R. and Wellner, J.A. (1986) *Empirical Processes with Applications to Statistics.* Wiley, New York.

[27] Zygmund, A. (2002) *Trigonometric Series. Vol. I, II.* 3rd Edition Cambridge University Press, Cambridge (UK).