

# A Fourier analysis of extreme events

THOMAS MIKOSCH\* and YUWEI ZHAO\*\*

University of Copenhagen, Department of Mathematics, Universitetsparken 5, DK-2100 Copenhagen, Denmark. E-mail: \*mikosch@math.ku.dk; \*\*zhao@math.ku.dk

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.

*Keywords:* ARMA; asymptotic theory; extremogram; GARCH; multivariate regular variation; periodogram; spectral density; stationary sequence; stochastic volatility process; strong mixing

## 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 [24], 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

$$\rho_A(h) = \lim_{x \rightarrow \infty} P(x^{-1}X_h \in A | x^{-1}X_0 \in A). \quad (1.1)$$

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 2.1. Since  $A$  is assumed to be bounded away from zero, the probabilities  $P(x^{-1}X \in A)$  converge to zero as  $x \rightarrow \infty$ . Then the following calculation is straightforward

for  $A$ :

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

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  $(\text{corr}(I_{\{x^{-1}X_0 \in A\}}, I_{\{x^{-1}X_h \in A\}}))$ . 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  $x$ , that is, 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 3. They studied the basic asymptotic properties of the extremogram (consistency, asymptotic normality) and also introduced a frequency domain analog of the correlation function  $\rho_A$  given as the Fourier series

$$f_A(\lambda) = \sum_{h \in \mathbb{Z}} \rho_A(h) e^{-ih\lambda}, \quad \lambda \in [0, \pi]. \tag{1.2}$$

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  $\lambda$  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, for example, 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, we introduce basic notions and conditions used throughout this paper. In Section 2.1, we define regular variation of a strictly stationary

sequence. In Section 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 2.3. In Section 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 4, we give the main results of this paper. We start in Section 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 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 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 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. Preliminaries

### 2.1. Regular variation

It was mentioned in Section 1 that one needs conditions to ensure that the limits  $\rho_A(h)$  in (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 [4,5], 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 [12–14]; cf. Andersen *et al.* [1] and the references therein. See also the examples in Section 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  $\mu$  on the Borel  $\sigma$ -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

$$\frac{P(x^{-1}X \in \cdot)}{P(|X| > x)} \xrightarrow{v} \mu(\cdot), \quad x \rightarrow \infty. \tag{2.1}$$

Here  $\xrightarrow{v}$  denotes *vague convergence* on the Borel  $\sigma$ -field of  $\overline{\mathbb{R}}_0^d$ ; for definitions see Kallenberg [33], Resnick [43,44]. In this context, bounded sets are those which are bounded away from zero and the Radon measure  $\mu$  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  $\sigma$ -field of  $\overline{\mathbb{R}}_0^d$ . We refer to *regular variation of  $X$  with limiting measure  $\mu$  and index  $\alpha$* . A multivariate  $t$ -distributed random vector is regularly varying and the index  $\alpha$  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 (2.1): there exists  $(a_n)$  such that  $a_n \rightarrow \infty$  as  $n \rightarrow \infty$  and

$$nP(a_n^{-1}X \in \cdot) \xrightarrow{v} \mu(\cdot), \quad n \rightarrow \infty. \tag{2.2}$$

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  $Y_h = \text{vec}(X_0, \dots, X_h)$ ,  $h \geq 0$ , are regularly varying with index  $\alpha$ . Of course, the limiting non-null Radon measures  $\mu_h$  in (2.1) now depend on the lag  $h$  and the normalization in (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  $h \geq 0$ ,

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

where  $(a_n)$  satisfies  $nP(|X_0| > a_n) \rightarrow 1$ , as  $n \rightarrow \infty$ . These relations are equivalent to the definitions (2.1) and (2.2) of regular variation of  $Y_h$ .

Now we are in the position to verify that the limits  $\rho_A(h)$  in (1.1) exist for any Borel set  $A \subset \mathbb{R}_0^d$  bounded away from zero. Write  $\tilde{A} = A \times \mathbb{R}_0^{dh}$  and  $\tilde{B} = A \times \mathbb{R}_0^{d(h-1)} \times A$ . These sets are bounded away from zero in  $\mathbb{R}_0^{d(h+1)}$ . If these sets are continuity sets with respect to  $\mu_h$  we obtain from the sequential definition of regular variation of  $Y_h$  for  $h \geq 0$ ,

$$\begin{aligned} \rho_A(h) &= \lim_{n \rightarrow \infty} P(a_n^{-1}X_h \in A | a_n^{-1}X_0 \in A) \\ &= \lim_{n \rightarrow \infty} \frac{nP(a_n^{-1}Y_h \in \tilde{B})}{nP(a_n^{-1}Y_h \in \tilde{A})} = \frac{\mu_h(\tilde{B})}{\mu_h(\tilde{A})}. \end{aligned}$$

### 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  $(\xi_t)$ . There exist  $m = m_n \rightarrow \infty$  and  $r_n \rightarrow \infty$  such that  $m_n/n \rightarrow 0$  and  $r_n/m_n \rightarrow 0$  and

$$\lim_{n \rightarrow \infty} m_n \sum_{h=r_n}^{\infty} \xi_h = 0, \tag{2.3}$$

and for all  $\epsilon > 0$ ,

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} m_n \sum_{h=k}^{r_n} P(|X_h| > \epsilon a_m, |X_0| > \epsilon a_m) = 0. \tag{2.4}$$

Condition (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  $nP(|X| > a_n) \rightarrow 1$  as  $n \rightarrow \infty$ , (2.4) is equivalent to

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

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  $k_n \xi_{r_n} \rightarrow 0$ , and  $m_n = o(n^{1/3})$ .

**Remark 2.1.** Some of the examples in Section 3 are strongly mixing with geometric rate, that is, there exists  $a \in (0, 1)$  such that  $\xi_h \leq a^h$  for sufficiently large  $h$ . Then (2.3) is satisfied if  $m_n a^{r_n} = o(1)$ . If  $m_n = n^\gamma$  for some  $\gamma \in (0, 1)$  then (2.3) is satisfied for  $r_n = c \log n$  if  $c$  is chosen sufficiently large and (M1) trivially holds as well. If  $\xi_h \leq h^{-s}$  for some  $s > 1$  and sufficiently large  $h$  then (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 (2.3) holds. Condition (M1) is satisfied if  $(1+s)^{-1} < \gamma < 1/3$  and  $\delta \in ((1-\gamma)/s, \gamma)$ . Thus (2.3) and (M1) are always satisfied if  $s$  can be chosen arbitrarily large.

For our main result on the smoothed periodogram (see Theorem 5.1), we finally need the condition:

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

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

**Remark 2.2.** Condition (M2) is stronger than (2.3). If  $(X_t)$  is strongly mixing with geometric or polynomial rate, a similar argument as in Remark 2.1 shows that (M2) holds for suitable choices of  $(r_n)$  and  $(m_n)$ .

### 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 (1.2). Write

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

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

$$I_{nA}(\lambda) = \frac{m_n}{n} \left| \sum_{t=1}^n \tilde{I}_t e^{-it\lambda} \right|^2, \quad \lambda \in [0, \pi] \quad \text{and} \quad \hat{P}_m(A) = \frac{m_n}{n} \sum_{t=1}^n I_t. \tag{2.5}$$

It follows from Theorem 3.1 in [15] that

$$\hat{P}_m(A) = \frac{m_n}{n} \sum_{t=1}^n I_t \xrightarrow{L^2} \mu_0(A) = \lim_{n \rightarrow \infty} m_n P(a_m^{-1} X \in A), \tag{2.6}$$

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

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

$$\hat{f}_{nA}(\lambda) = \tilde{\gamma}_n(0) + 2 \sum_{h=1}^{r_n} \cos(\lambda h) \tilde{\gamma}_n(h) \tag{2.7}$$

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

$$E \hat{f}_{nA}(\lambda) \rightarrow \mu_0(A) f_A(\lambda) \quad \text{and} \quad E(\hat{f}_{nA}(\lambda) - \mu_0(A) f_A(\lambda))^2 \rightarrow 0 \tag{2.8}$$

under condition (M), if  $A$  is a  $\mu_0$ -continuity set and the sets  $A \times \overline{\mathbb{R}}_0^{k-1} \times A$  are continuity sets with respect to  $\mu_k$ ,  $k \geq 1$ , and  $m_n r_n^2 = O(n)$ . If we combine (2.6) and (2.8) we have for fixed  $\lambda \in (0, \pi)$ ,

$$\frac{\hat{f}_{nA}(\lambda)}{\hat{P}_m(A)} \xrightarrow{P} f_A(\lambda). \tag{2.9}$$

A natural self-normalized estimator of the spectral density  $f_A(\lambda)$  in (1.2) is the following analog of the periodogram

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

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

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

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 i.i.d. and has finite variance; see, for example, 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 4.

### 3. Examples

In this section, we collect some examples of regularly varying stationary time series models, give their extremograms (1.1) and, if possible, give an explicit expression of the corresponding spectral density (1.2). However, in general, the extremogram is too complicated and one cannot calculate the Fourier series (1.2). Some of the examples below are taken from Davis and Mikosch [15].

#### 3.1. IID sequence

Consider an i.i.d. real-valued sequence  $(Z_t)$  such that

$$P(Z > x) \sim px^{-\alpha}L(x) \quad \text{and} \quad P(Z \leq -x) \sim qx^{-\alpha}L(x), \quad x \rightarrow \infty, \quad (3.1)$$

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

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

where  $\varepsilon_y$  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,

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

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

#### 3.2. Stochastic volatility model

Let  $(\sigma_t)$  be a strictly stationary sequence of non-negative random variables with  $E\sigma^{\alpha+\delta} < \infty$  for some  $\delta > 0$ , independent of the i.i.d. regularly varying sequence  $(Z_t)$  with index  $\alpha > 0$ , satisfying the tail balance condition (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  $\alpha$  and limiting measures concentrated on the axes. The extremogram and the spectral density coincide with these quantities in the i.i.d. 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  $(\sigma_t)$ . In particular, if  $(\sigma_t)$  is strongly mixing with *geometric rate*,  $(X_t)$  is also strongly mixing with geometric rate, and then the conditions (2.3), (M1) and (M2) are satisfied; see Remarks 2.1 and 2.2. Condition (2.4) also holds if  $E\sigma^{4\alpha} < \infty$ ; see Davis and Mikosch [15].

The situation of a vanishing  $\rho_A$  is rather incomplete information about tail dependence. Hill [28] proposed to use an alternative lag-wise dependence measure of the form  $\lim_{x \rightarrow \infty} P(X_h > x, X_0 > x) / [P(X_0 > x)]^2 - 1$  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 i.i.d. regularly varying noise  $(Z_t)$  with index  $\alpha$  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  $\alpha$  when the sequence  $(\sigma_t)$  is regularly varying with index  $\alpha$ ,  $E|Z|^{\alpha+\delta} < \infty$  for some  $\delta > 0$  and they give examples with  $\rho_A \neq 0$  and  $f_A \neq 1$  for  $A$  bounded away from zero. The aforementioned comments about mixing also apply in this setting.

### 3.3. ARMA process

Consider the linear process

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \quad t \in \mathbb{Z}, \tag{3.2}$$

where  $(Z_t)$  is an i.i.d. one-dimensional regularly varying sequence with index  $\alpha > 0$  and tail balance condition (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

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

for integers  $r, s \geq 0$ , and the coefficients  $\theta_i, \phi_i$  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  $\alpha$ ; see, for example, 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 i.i.d. regularly varying random variables are regularly varying; see Feller [22], page 278; cf. Lemma 1.3.1 in [20]. Moreover,

$$\lim_{s \rightarrow \infty} \limsup_{n \rightarrow \infty} n P(a_n^{-1} |X_t - X_t^{(s)}| > \varepsilon) = 0, \quad \varepsilon > 0. \tag{3.3}$$

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

The vector  $(X_0^{(s)}, \dots, X_h^{(s)})$  is also regularly varying with index  $\alpha$ . 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 (3.3) holds a



straightforward multivariate extension of Lemma 3.6 in [32] yields that  $(X_0, \dots, X_h)$  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  $A = (1, \infty)$ ,

$$\rho_A(h) = \frac{\sum_{i=0}^{\infty} [p(\min(\psi_i^+, \psi_{i+h}^+))^\alpha + q(\min(\psi_i^-, \psi_{i+h}^-))^\alpha]}{\sum_{i=0}^{\infty} [p(\psi_i^+)^\alpha + q(\psi_i^-)^\alpha]}, \quad h \geq 1. \quad (3.4)$$

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

Doukhan [19], Theorem 6 on page 99, shows that  $(X_t)$  is  $\beta$ -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  $p$ th moment for some  $p > 0$ . Hence (2.3), (M1) and (M2) are satisfied under these conditions; see Remarks 2.1 and 2.2. Next, we verify condition (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

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

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  $k \geq 1$ ,

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

(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),

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

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

If  $\text{var}(X) < \infty$  relation (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  $\rho_A$  in (1.2) by  $\rho$ , one obtains the well-known spectral density of a causal ARMA process (up to a constant multiple):  $f_X(\lambda) = (2\pi)^{-1} |\theta(e^{-i\lambda})|^2 / |\phi(e^{-i\lambda})|^2$ ,  $\lambda \in [0, \pi]$ . Such a compact formula can in general not be derived for  $f_A$ . An exception is a causal ARMA(1, 1) process; see Section B. There are various analogies between the functions  $\rho$  and  $\rho_A$  for causal invertible ARMA processes. In this case,  $\psi_h \rightarrow 0$  as  $h \rightarrow \infty$  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  $h > q$ .

### 3.4. Max-moving averages

Consider a regularly varying i.i.d. sequence  $(Z_t)$  with index  $\alpha > 0$  and tail balance parameters  $p, q$ ; see (3.1). For a real-valued sequence  $(\psi_j)$ , the process

$$X_t = \bigvee_{i=0}^{\infty} \psi_i Z_{t-i}, \quad t \in \mathbb{Z}, \tag{3.5}$$

is a *max-moving average*. We will also assume that  $|\psi_j| \leq c$ ,  $j \geq 0$ , 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  $\infty$  if  $\lim_{x \rightarrow \infty} P(X > x) = 0$ . We have

$$P(X > x) = P\left(\bigvee_{i=0}^{\infty} \psi_i Z_i > x\right) = 1 - \lim_{n \rightarrow \infty} \prod_{i=0}^n P(\psi_i Z \leq x).$$

The product  $\prod_{i=0}^{\infty} P(\psi_i Z \leq x)$  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

$$P(X > x) = 1 - e^{-(1+o(1))P(|Z|>x)\psi_+} \sim P(|Z| > x)\psi_+ \rightarrow 0, \quad x \rightarrow \infty. \tag{3.6}$$

We also have  $P(X \leq -x) = O(P(|Z| > x))$ . Hence,  $X$  is regularly varying with index  $\alpha$  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  $s \geq 0$ ,

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

Regular variation of  $(X_0^{(s)}, \dots, X_h^{(s)})$  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,

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

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

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

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

Finally, in view of (3.6),  $\rho_A(h)$  is given by (3.4), that is, the linear process (3.2) and the max-moving average (3.5) have the same extremogram provided the coefficients  $(\psi_j)$  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 i.i.d. with common Fréchet distribution  $\Psi_{\alpha}(x) = e^{-x^{-\alpha}}$ ,  $x > 0$ , 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  $\beta$ -mixing, implying strong mixing. An application of their Corollary 2.2 implies that the condition  $|\psi_h| \leq c_0 e^{-c_1 h}$ ,  $h \geq 1$ , for suitable constants  $c_1, c_2 > 0$  implies strong mixing of  $(X_t)$  with geometric rate function  $(\xi_h)$ . In this situation, (M), (M1) and (M2) are satisfied.

### 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 (2.5). Notice that

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

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

$$\alpha_n(\lambda) = \left(\frac{2m_n}{n}\right)^{1/2} \sum_{t=1}^n \tilde{I}_t \cos(\lambda t),$$

$$\beta_n(\lambda) = \left(\frac{2m_n}{n}\right)^{1/2} \sum_{t=1}^n \tilde{I}_t \sin(\lambda t).$$

Here we suppress the dependence of  $\alpha_n$  and  $\beta_n$  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

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

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

### 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 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  $\mu_0$  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  $\mu_h$  for every  $h \geq 1$ ; see Section 2.1. Also assume that  $\sum_{h \geq 1} \rho_A(h) < \infty$ . Let  $\lambda, \omega$  be either any two Fourier or fixed frequencies in  $(0, \pi)$ .*

(1) *If  $\lambda, \omega$  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  $n \rightarrow \infty$ .*

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

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

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

**Remark 4.2.** The smoothness condition on the set  $A$  ensures that the extremogram  $\rho_A$  with respect to  $A$  is well defined; see Section 2.1.

**Remark 4.3.** Since  $E\alpha_n(\lambda) = E\beta_n(\lambda) = 0$  an immediate consequence of part (3) is that

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

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  $\lambda_n \rightarrow \lambda \in (0, \pi)$ .

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

$$\begin{aligned} J &= \frac{2m_n}{n} E \left[ \sum_{s=1}^n \sum_{t=1}^n (I_s I_t - p_0^2) f_1(\lambda s) f_2(\omega t) \right] \\ &= \frac{2m_n}{n} \left( \sum_{1 \leq t=s \leq n} + \sum_{1 \leq s \neq t \leq n} \right) (p_{|s-t|} - p_0^2) f_1(\lambda s) f_2(\omega t) \\ &= J_1 + J_2, \end{aligned}$$

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  $\lambda, \omega$  are Fourier frequencies, so are  $\lambda \pm \omega$  and therefore

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

If  $\lambda, \omega$  are fixed frequencies, we conclude from (A.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

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

If  $\lambda, \omega$  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 (A.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

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

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  $\text{cov}(\alpha_n(\lambda), \beta_n(\lambda))$ . If  $\lambda$  is a Fourier frequency, we have  $\sin(\lambda n) = 0$ . Hence, by (A.7),

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

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  $\lambda$  since the expressions in (A.7) are bounded for every  $n$  and  $h < n$ .

If  $\lambda \neq \omega$  are fixed frequencies, we conclude from (A.8)–(A.10) and condition (M) that there exist constants  $c(\lambda, \omega)$  such that

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

Now we consider the case of two distinct Fourier frequencies  $\lambda, \omega$ . 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  $\lambda, \omega$  are Fourier frequencies a glance at (A.8)–(A.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  $ch$ . 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  $ch$  for some constant  $c > 0$  as well. Then, by (A.8), for fixed  $k$ ,

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

The right-hand side vanishes by virtue of condition (M), first letting  $n \rightarrow \infty$  and then  $k \rightarrow \infty$ . 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 (A.9) and (A.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

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

For any frequency  $\lambda \in (0, \pi)$  bounded away from zero and  $\pi$ , 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

$$\begin{aligned} \text{var}(\alpha_n(\lambda)) &\sim m_n p_0 + 2m_n \sum_{h=1}^{n-1} (p_h - p_0^2) (1 - h/n) \cos(\lambda h) \\ &\sim \mu_0(A) \left[ 1 + 2 \sum_{h=1}^{\infty} \rho_A(h) \cos(\lambda h) \right]. \end{aligned}$$

This concludes the proof. □

## 4.2. Central limit theorem

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

**Theorem 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 4.1. Assume that the conditions (M), (M1) and  $\sum_{h \geq 1} \rho_A(h) < \infty$  hold. Consider any fixed frequencies  $0 < \lambda_1 < \dots < \lambda_N < \pi$  for some  $N \geq 1$ . Then the following central limit theorem holds:

$$\mathbf{Z}_n = (\alpha_n(\lambda_i), \beta_n(\lambda_i))_{i=1, \dots, N} \xrightarrow{d} (\alpha(\lambda_i), \beta(\lambda_i))_{i=1, \dots, N}, \quad n \rightarrow \infty, \quad (4.1)$$

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

$$\Sigma_N = \mu_0(A) \text{diag}(f_A(\lambda_1), f_A(\lambda_1), \dots, f_A(\lambda_N), f_A(\lambda_N)).$$

The limit relation (4.1) remains valid if the frequencies  $\lambda_i, i = 1, \dots, N$ , are replaced by distinct Fourier frequencies  $\omega_i(n) \rightarrow \lambda_i \in (0, \pi)$  as  $n \rightarrow \infty$ . The limits  $\lambda_i$  do not have to be distinct.

Then the following result is immediate.

**Corollary 4.5.** Assume the conditions of Theorem 4.4. Let  $(E_i)$  be a sequence of i.i.d. standard exponential random variables.

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

$$(I_{nA}(\lambda_i))_{i=1, \dots, N} \xrightarrow{d} \mu_0(A)(f_A(\lambda_i)E_i)_{i=1, \dots, N}, \quad n \rightarrow \infty,$$

$$(\tilde{I}_{nA}(\lambda_i))_{i=1, \dots, N} \xrightarrow{d} (f_A(\lambda_i)E_i)_{i=1, \dots, N}, \quad n \rightarrow \infty.$$

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

$$(I_{nA}(\omega_i(n)))_{i=1, \dots, N} \xrightarrow{d} \mu_0(A)(f_A(\lambda_i)E_i)_{i=1, \dots, N}, \quad n \rightarrow \infty,$$

$$(\tilde{I}_{nA}(\omega_i(n)))_{i=1, \dots, N} \xrightarrow{d} (f_A(\lambda_i)E_i)_{i=1, \dots, N}, \quad n \rightarrow \infty.$$

**Proof of the Theorem 4.4.** We will prove (4.1) by applying the Cramér–Wold device, that is, we will show that for any choice of constants  $\mathbf{c} \in \mathbb{R}^{2N}$ ,

$$\mathbf{c}'\mathbf{Z}_n \xrightarrow{d} N(0, \mathbf{c}'\Sigma_N\mathbf{c}). \quad (4.2)$$

The proof of the result for distinct converging Fourier frequencies is analogous and therefore omitted. We will prove (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

$$Y_{nt} = \left(\frac{2m_n}{n}\right)^{1/2} \tilde{I}_t \sum_{j=1}^N [c_{2j-1} \cos(\lambda_j t) + c_{2j} \sin(\lambda_j t)], \quad t = 1, \dots, n. \tag{4.3}$$

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

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

the index sets  $\tilde{K}_{ni}$ , which consist of all but the first  $r_n$  elements of  $K_{ni}$ , and the small blocks  $J_{ni} = K_{ni} \setminus \tilde{K}_{ni}$ . In view of condition (M),  $r_n/m_n \rightarrow 0$  and  $m_n \rightarrow \infty$ , the sets  $\tilde{K}_{ni}$  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.  $\square$

**Lemma 4.6.** *Under the conditions of Theorem 4.4, the following relation holds:*

$$\text{var}\left(\sum_{i=1}^{k_n} S_n(J_{ni})\right) \rightarrow 0, \quad n \rightarrow \infty. \tag{4.4}$$

**Proof.** We have

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

Due to the sum structure of  $Y_{nt}$  given in (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  $\text{var}(S_n(J_{ni}))$  and  $|\text{cov}(S_n(J_{ni_1}), S_n(J_{ni_2}))|$  is bounded by a linear combination of the variances/covariances of such subsums. In other words, it suffices to prove (4.4) for  $N = 1$ . We give the corresponding calculations only for the functions  $\cos(\lambda t)$  where  $\lambda$  stands for any of the frequencies  $\lambda_j$ . The calculations are similar to those in the proof of Proposition 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  $n$ ,

$$\begin{aligned} \text{var}(S_n(J_{ni})) &= \frac{2m_n}{n} \left[ \sum_{s=(i-1)m_n+1}^{(i-1)m_n+r_n} \text{var}(I_s)(\cos(\lambda s))^2 \right. \\ &\quad \left. + 2 \sum_{h=1}^{r_n-1} \sum_{s=(i-1)m_n+1}^{(i-1)m_n+r_n-h} \text{cov}(I_s, I_{s+h}) \cos(\lambda s) \cos(\lambda(s+h)) \right] \end{aligned}$$

$$\begin{aligned} &\leq \frac{2m_n}{n} \left( r_n (p_0 - p_0^2) + 2 \sum_{h=1}^{r_n-1} (r_n - h) |p_h - p_0^2| \right) \\ &\leq c \frac{r_n}{n} \left( m_n \sum_{h=0}^k p_h + m_n \sum_{h=k+1}^{r_n} p_h \right) \leq c(k)(r_n/n), \end{aligned}$$

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

$$\begin{aligned} &|\text{cov}(S_n(J_{ni_1}), S_n(J_{ni_2}))| \\ &= \frac{2m_n}{n} \left| \left[ \sum_{s=(i_1-1)m_n+1}^{(i_1-1)m_n+r_n} \sum_{t=(i_2-1)m_n+1}^{(i_2-1)m_n+r_n} \text{cov}(I_t, I_s) \cos(\lambda s) \cos(\lambda t) \right] \right| \\ &\leq c \frac{m_n}{n} \sum_{q=(i_2-i_1)m_n-(r_n-1)}^{(i_2-i_1)m_n+r_n-1} (r_n - |q - (i_2 - i_1)m_n|) |p_q - p_0^2| \\ &\leq c \frac{m_n r_n}{n} \sum_{q=(i_2-i_1)m_n-(r_n-1)}^{(i_2-i_1)m_n+r_n-1} \xi_q, \end{aligned}$$

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

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

This proves (4.4). □

Relation (4.4) implies that  $\mathbf{c}'\mathbf{Z}_n$  and  $\sum_{i=1}^{k_n} S_n(\tilde{K}_{ni})$  have the same limit distribution provided such a limit exists. Let  $\tilde{S}_n(\tilde{K}_{ni}) \stackrel{d}{=} S_n(\tilde{K}_{ni})$  for  $i = 1, \dots, k_n$  and assume that  $(\tilde{S}_n(\tilde{K}_{ni}))_{i=1, \dots, k_n}$  has independent components. A telescoping sum argument yields

$$\begin{aligned} &\left| E \prod_{l=1}^{k_n} e^{it S_n(\tilde{K}_{nl})} - E \prod_{s=1}^{k_n} e^{it \tilde{S}_n(\tilde{K}_{ns})} \right| \\ &= \left| \sum_{l=1}^{k_n} E \left[ (e^{it S_n(\tilde{K}_{nl})} - e^{it \tilde{S}_n(\tilde{K}_{nl})}) \prod_{s=1}^{l-1} e^{it \tilde{S}_n(\tilde{K}_{ns})} \prod_{s=l+1}^{k_n} e^{it S_n(\tilde{K}_{ns})} \right] \right| \end{aligned}$$

$$\begin{aligned} &\leq \sum_{l=1}^{k_n} \left| E \left( \prod_{s=1}^{l-1} e^{it\tilde{S}_n(\tilde{K}_{ns})} (e^{itS_n(\tilde{K}_{nl})} - e^{it\tilde{S}_n(\tilde{K}_{nl})}) \prod_{s=l+1}^{k_n} e^{itS_n(\tilde{K}_{ns})} \right) \right| \\ &\leq 4k_n \xi_{r_n} \rightarrow 0. \end{aligned}$$

In the last step, we used Theorem 17.2.1 in Ibragimov and Linnik [29] and condition (M1). Hence,  $\sum_{l=1}^{k_n} S_n(\tilde{K}_{nl})$  and  $\sum_{l=1}^{k_n} \tilde{S}_n(\tilde{K}_{nl})$  have the same limits in distribution provided these limits exist. In view of (4.4) and the last conclusion the central limit theorem (4.2) holds if and only if the same limit relation holds for  $\sum_{i=1}^{k_n} \tilde{S}_n(K_{ni})$ , where  $\tilde{S}_n(K_{ni}) \stackrel{d}{=} S_n(K_{ni})$  and  $(\tilde{S}_n(K_{ni}))_{i=1, \dots, k_n}$  has independent components. Thus, we may apply a classical central limit theorem for triangular arrays of independent random variables; see, for example, Theorem 4.1 in Petrov [40].

According to this result, the central limit theorem

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

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

$$\sum_{i=1}^{k_n} E[(S_n(K_{ni}))^2 I_{\{|S_n(K_{ni})|>\varepsilon\}}] \rightarrow 0. \tag{4.5}$$

The condition  $E Z_n = 0$  holds since  $E \tilde{I}_t = 0$ , hence  $E \tilde{S}_n(K_{ni}) = 0$  for every  $i$ . As for (6.8) in Davis and Mikosch [15], a trivial bound of the left-hand side in (4.5) is given by

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

In view of (M1),  $m_n^3/n = o(1)$ , and therefore the right-hand side vanishes for sufficiently large  $n$ . Therefore, (4.5) holds.

**Lemma 4.7.** *Under the conditions of Theorem 4.4,*

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

**Proof.** We proceed in a similar way as for Proposition 4.1. It will be convenient to introduce the following notation for  $\lambda \in (0, \pi)$ ,

$$\tilde{\alpha}_n(\lambda) = \left(\frac{2m_n}{n}\right)^{1/2} \sum_{i=1}^{k_n} \sum_{t \in K_{ni}} \cos(\lambda t) \tilde{I}_t(i),$$

$$\tilde{\beta}_n(\lambda) = \left(\frac{2m_n}{n}\right)^{1/2} \sum_{i=1}^{k_n} \sum_{t \in K_{ni}} \sin(\lambda t) \tilde{I}_t(i),$$

where for each  $i \leq k_n$ ,

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

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

$$\text{var}(\tilde{\alpha}_n(\lambda)) \sim \text{var}(\tilde{\beta}_n(\lambda)) \sim \mu_0(A) \left[ 1 + 2 \sum_{h=1}^{\infty} \rho_A(h) \cos(\lambda h) \right]. \tag{4.6}$$

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

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

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  $k \geq 1$ ,

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

and the right-hand side is negligible in view of (M) by first letting  $n \rightarrow \infty$  and then  $k \rightarrow \infty$ . Thus, it suffices to consider only finitely many  $h$ -terms in  $P_{22}$ . In view of (A.1), for fixed  $k$  as  $n \rightarrow \infty$ ,

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

This proves (4.6).

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

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

In view of (A.1) and since  $\lambda \neq \omega$ ,

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

Similarly, multiple application of (A.1), first summing over  $t$ , then over  $l$ , yields

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

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

### 5. Smoothing the periodogram

Corollary 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 i.i.d. 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

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

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

$$\sum_{|j| \leq s_n} w_n(j) = 1 \quad \text{and} \quad \sum_{|j| \leq s_n} w_n^2(j) \rightarrow 0 \quad \text{as } n \rightarrow \infty, \tag{5.1}$$

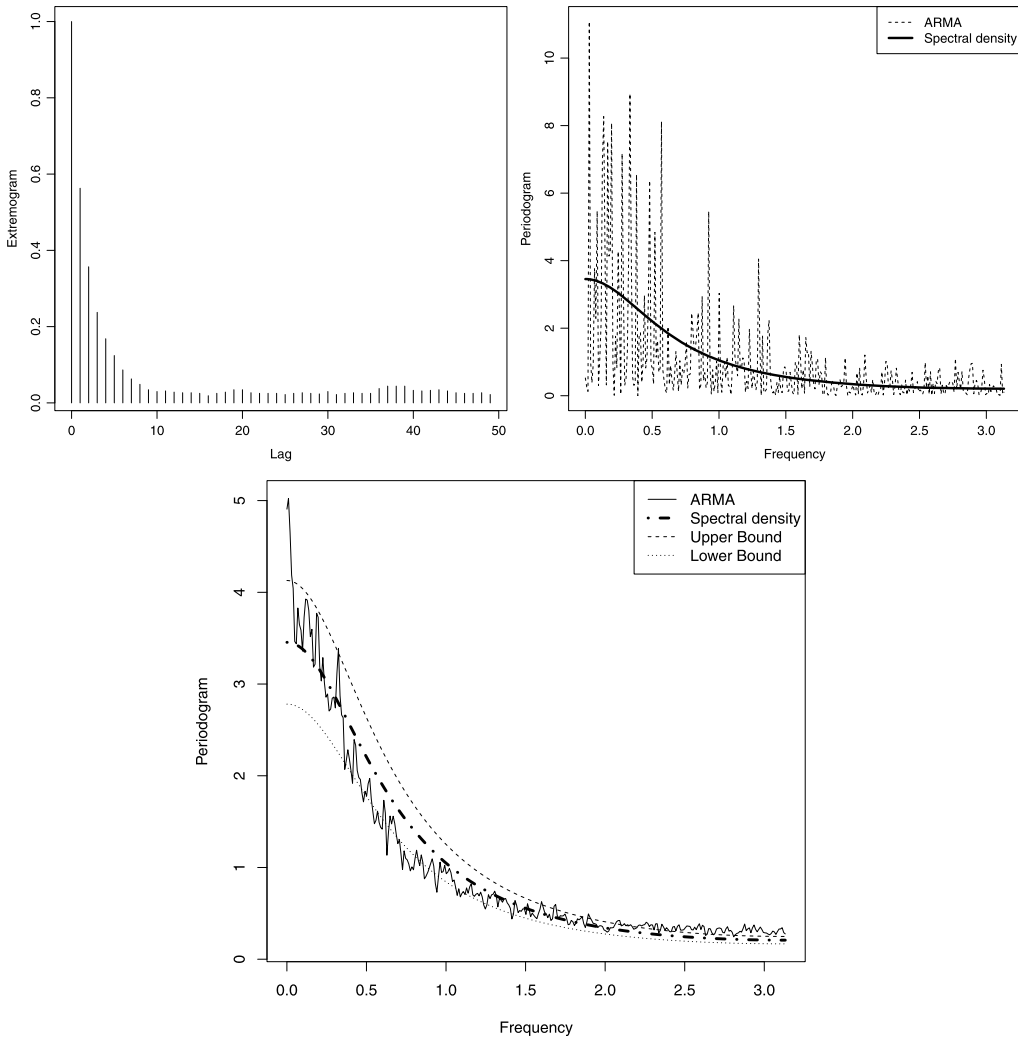
we introduce the *smoothed periodogram*

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

**Theorem 5.1.** *Assume the conditions of Theorem 4.4, (5.1) on the weight function  $w$  and (M2). Then for every fixed frequency  $\lambda \in (0, \pi)$ , as  $n \rightarrow \infty$ ,*

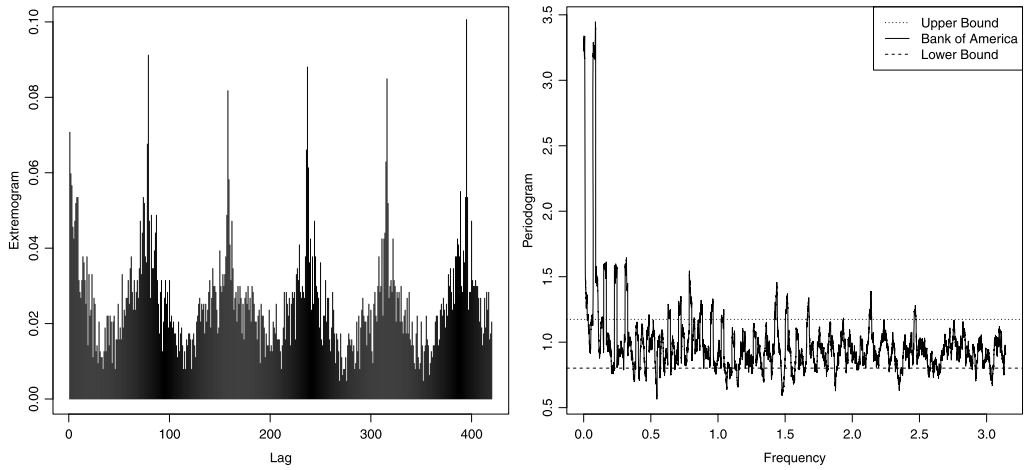
$$\tilde{f}_{nA}(\lambda) \xrightarrow{L^2} \mu_0(A) f_A(\lambda) \quad \text{and} \quad \frac{\tilde{f}_{nA}(\lambda)}{\tilde{P}_m(A)} \xrightarrow{P} f_A(\lambda).$$

In Figures 1 and 2 we show the extremogram, the standardized periodogram and the corresponding smoothed periodogram for some simulated and real-life data. The data underlying Figure 1 are simulated from an ARMA(1, 1) process  $(X_t)$  with parameters  $\phi = 0.8$  and  $\theta = 0.1$  and i.i.d.  $t$ -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 B for an expression) and the raw periodogram which exhibits rather erratic behavior. The bottom graph shows the smoothed periodogram with Daniell window  $w_n(i) = 1/(2s_n + 1)$ ,  $|i| \leq s_n = 50$ . 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 5.1, we show



**Figure 1.** Top-Left: The sample extremogram of an ARMA(1, 1) process with parameters  $\phi = 0.8, \theta = 0.1$  and i.i.d.  $t$ -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,  $s_n = 50$ .

that  $\text{var}(\tilde{f}_{nA}(\lambda)) \sim \sum_{|j| \leq s_n} w_n^2(j) \mu_0^2(A) f_{nA}^2(\lambda)$  for every  $\lambda \in (0, \pi)$ . Furthermore, we know that  $\hat{P}_m(A) \xrightarrow{P} \mu_0(A)$ . Based on these calculations, we take  $\sum_{|j| \leq s_n} w_n^2(j) f_{nA}^2(\lambda)$  as a surrogate quantity for the unknown variance of  $\tilde{f}_n(\lambda) / \hat{P}_m(A)$ .



**Figure 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.

The data underlying Figure 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 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 i.i.d., 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 of Theorem 5.1.** We mentioned in Remark 4.3 that

$$EI_{nA}(\lambda) \rightarrow \mu_0(A) f_A(\lambda) \quad \text{as } n \rightarrow \infty \text{ uniformly on sets } [a, b] \subset (0, \pi). \quad (5.2)$$

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

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

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

$$\text{var}(\tilde{f}_{nA}(\lambda)) = \sum_{|j| \leq s_n} w_n^2(j) c_{jj} + \sum_{-s_n \leq j_1 \neq j_2 \leq s_n} w_n(j_1) w_n(j_2) c_{j_1 j_2}.$$



In view of condition (5.1) it suffices to show that  $c_{j_1 j_2} = \text{cov}(I_{nA}(\lambda_{j_1}), I_{nA}(\lambda_{j_2})) \rightarrow 0$  and

$$c_{jj} = \text{var}(I_{nA}(\lambda_j)) \rightarrow (\mu_0(A) f_A(\lambda))^2 \quad \text{uniformly for } j, j_1, j_2 \in [-s_n, s_n], j_1 \neq j_2. \quad (5.3)$$

We will only show (5.3); the proof of  $c_{j_1, j_2} \rightarrow 0$  for  $j_1 \neq j_2$  is similar and therefore omitted. Since (5.2) holds, we have to show that

$$E(I_{nA}^2(\lambda_j)) \rightarrow 2(\mu_0(A) f_A(\lambda))^2. \quad (5.4)$$

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

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

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

$$E\widehat{f}_{nA}^2(\lambda_j) \rightarrow (\mu_0(A) f_A(\lambda))^2, \quad (5.5)$$

uniformly for the considered frequencies  $\lambda_j$ . Then (5.4) is proved if we can show that

$$E(\widehat{f}_{nA}(\lambda_j)\widehat{g}_{nA}(\lambda_j)) \rightarrow 0, \quad (5.6)$$

$$E\widehat{g}_{nA}^2(\lambda_j) \rightarrow (\mu_0(A) f_A(\lambda))^2. \quad (5.7)$$

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

$$p_{h_1 h_2 h_3} = P(X_0 > a_m, X_{h_1} > a_m, X_{h_1+h_2} > a_m, X_{h_1+h_2+h_3} > a_m),$$

$$p_{h_1 h_2} = p_{h_1 h_2 0}, \quad p_{h_1} = p_{h_1 0},$$

and we observe that

$$p_h = (p_h - p_0^2) + p_0^2, \quad (5.8)$$

$$\begin{aligned} p_{h_1 h_2} &= (p_{h_1 h_2} - p_{h_1} p_0) + p_{h_1} p_0 = p_{h_1 h_2} - p_0 p_{h_2} + p_0 p_{h_2} \\ &= (p_{h_1 h_2} - p_0 p_{h_2}) + p_0 (p_{h_2} - p_0^2) + p_0^3, \end{aligned} \quad (5.9)$$

$$\begin{aligned} p_{h_1 h_2 h_3} &= (p_{h_1 h_2 h_3} - p_0 p_{h_2 h_3}) + p_0 p_{h_2 h_3} \\ &= (p_{h_1 h_2 h_3} - p_0 p_{h_2 h_3}) + p_0 (p_{h_2 h_3} - p_0 p_{h_3}) + p_0^2 p_{h_3}. \end{aligned} \quad (5.10)$$

**Proof of (5.6)**

We have

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

where

$$J_1 = 4\frac{m_n^2}{n^2}\sum_{t_1=1}^n\sum_{h=r_n+1}^{n-1}\sum_{t_2=1}^{n-h}E[I_{t_1}I_{t_2}I_{t_2+h}]\cos(\lambda_j h),$$

$$J_2 = 8\frac{m_n^2}{n^2}\sum_{t_1=1}^{n-1}\sum_{h_1=1}^{r_n}\sum_{h_2=r_n+1}^{n-1}\sum_{t_2=1}^{n-h_2}E[I_{t_1}I_{t_1+h_1}I_{t_2}I_{t_2+h_2}]\cos(\lambda_j h_1)\cos(\lambda_j h_2).$$

*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

$$J_1 = 4\frac{m_n^2}{n^2}\sum_{h=r_n+1}^{n-1}(n-h)(2p_h)\cos(\lambda_j h) + 4\frac{m_n^2}{n^2}\sum_{h_2=r_n+1}^{n-2}\sum_{h_1=1}^{n-h_2-1}(n-h_1-h_2)p_{h_1, h_2}\cos(\lambda_j h_2)$$

$$+ 4\frac{m_n^2}{n^2}\sum_{h_2=r_n+1}^{n-1}\sum_{h_1=1}^{h_2-1}(n-h_2)p_{h_1, h_2-h_1}\cos(\lambda_j h_2)$$

$$+ 4\frac{m_n^2}{n^2}\sum_{h_2=r_n+1}^{n-2}\sum_{h_1=1}^{n-h_2-1}(n-h_1-h_2)p_{h_2, h_1}\cos(\lambda_j h_2)$$

$$= \sum_{i=1}^4 J_{1i}.$$

Applying (5.8), the mixing condition (M2) and Lemma A.1 imply that

$$J_{11} \leq cm_n \sum_{h=r_n+1}^{\infty} \xi_h + c \frac{(m_n p_0)^2}{n(\sin(\lambda_j/2))^2} = o(1).$$

As regards  $J_{12}$ , apply (5.9) and split the  $h_1$ -index set into  $h_1 \leq r_n$  and  $h_1 > r_n$ . Then (M2) and Lemma A.1 imply that

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

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

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

Now introduce the mixing coefficients  $\xi_h$  and use Lemma A.1:

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

The right-hand side vanishes as  $n \rightarrow \infty$  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)

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

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

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

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

$$\begin{aligned}
 |J_{23}| &= 8 \left| \frac{m_n^2}{n^2} \sum_{h_1=1}^{r_n-1} \sum_{h_3=1}^{r_n-h_1} \left( \sum_{h_2=r_n+1-h_3}^{r_n} + \sum_{h_2=r_n+1}^{n-h_1-h_3-1} \right) (n - h_1 - h_2 - h_3) p_{h_1 h_3 h_2} f_{h_1+h_3, h_2+h_3} \right| \\
 &\leq c \frac{m_n r_n^3}{n} (m_n p_0) + \left[ c \frac{m_n^2}{n} \sum_{h_1=1}^{r_n-1} \sum_{h_3=1}^{r_n-h_1} \sum_{h_2=r_n+1}^{n-h_1-h_3-1} |p_{h_1 h_3 h_2} - p_{h_1 h_3} p_0| \right]
 \end{aligned}$$

$$\begin{aligned}
 & + c \frac{m_n^2}{n^2} p_0 \sum_{h_1=1}^{r_n-1} \sum_{h_3=1}^{r_n-h_1} p_{h_1 h_3} \cos(\lambda_{j_1}(h_1 + h_3)) \\
 & \quad \times \left[ \sum_{h_2=r_n+1}^{n-h_1-h_3-1} (n - h_1 - h_2 - h_3) \cos(\lambda_{j_1}(h_1 + h_2)) \right] \\
 \leq & c \frac{m_n r_n^3}{n} + c \frac{m_n^2 r_n^2}{n} \sum_{h_2=r_n+1}^{\infty} \xi_{h_2} + c \frac{r_n^2}{n} (m_n p_0)^2 (\sin(\lambda_j/2))^{-2}.
 \end{aligned}$$

In the last step we used Lemma A.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  $k, l > r_n, s \leq r_n$  or  $k > r_n, l, s \leq r_n$  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 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*

$$m_n^2 n \sum_{h=r_n+1}^n \xi_h \rightarrow 0, \quad n \rightarrow \infty, \tag{5.11}$$

where  $(\xi_t)$  is the mixing rate function. Then

$$Q_0 = \frac{m_n^2}{n^2} \sum_{h_1, h_2, h_3 > r_n} (n - h_1 - h_2 - h_3)_+ p_{h_1 h_2 h_3} \cos(\lambda h_1 + x_1) \cos(\omega h_3 + x_2) \rightarrow 0, \tag{5.12}$$

$$\frac{m_n^2}{n^2} \sum_{h_1, h_2, h_3 > r_n} (n - h_1 - h_2 - h_3)_+ p_{h_1 h_2 h_3} \sin(\lambda h_1 + x_1) \sin(\omega h_3 + x_2) \rightarrow 0. \tag{5.13}$$

**Proof.** Recall (5.10). Write  $g_{h_1 h_3} = \cos(\lambda h_1 + x_1) \cos(\omega h_3 + x_2)$ . Then we have

$$\begin{aligned}
 |Q_0| \leq & \frac{m_n^2}{n^2} \sum_{h_1, h_2, h_3 > r_n} (n - h_1 - h_2 - h_3)_+ |p_{h_1 h_2 h_3} - p_{h_1} p_{h_3}| \\
 & + \left| \frac{m_n^2}{n^2} \sum_{h_3=r_n+1}^{n-2r_n-3} \sum_{h_1=r_n+1}^{n-h_3-r_n-2} \sum_{h_2=r_n+1}^{n-h_1-h_3-1} (n - h_1 - h_2 - h_3) (p_{h_1} - p_0^2) (p_{h_3} - p_0^2) g_{h_1 h_3} \right| \\
 & + \left| \frac{m_n^2}{n^2} \sum_{h_2=r_n+1}^{n-2r_n-3} \sum_{h_3=r_n+1}^{n-h_2-r_n-2} \sum_{h_1=r_n+1}^{n-h_2-h_3-1} (n - h_1 - h_2 - h_3) p_0^2 (p_{h_3} - p_0^2) g_{h_1 h_3} \right|
 \end{aligned}$$

$$\begin{aligned}
 & + \left| \frac{m_n^2}{n^2} \sum_{h_2=r_n+1}^{n-2r_n-3} \sum_{h_1=r_n+1}^{n-h_2-r_n-2} \sum_{h_3=r_n+1}^{n-h_1-h_2-1} (n-h_1-h_2-h_3) p_0^2 (p_{h_1} - p_0^2) g_{h_1 h_3} \right| \\
 & + \left| \frac{m_n^2}{n^2} \sum_{h_2=r_n+1}^{n-2r_n-3} \sum_{h_1=r_n+1}^{n-h_2-r_n-2} \sum_{h_3=r_n+1}^{n-h_1-h_2-1} (n-h_1-h_2-h_3) p_0^4 g_{h_1 h_3} \right| \\
 & = \sum_{i=1}^5 Q_i.
 \end{aligned}$$

By virtue of (5.11),  $Q_1$  is negligible. Similarly,  $Q_2 \leq m_n^2 (\sum_{h=r_n+1}^n \xi_h)^2 \rightarrow 0$ . As to  $Q_3$ , Lemma A.1 and mixing imply that

$$Q_3 \leq c \frac{(m_n p_0)^2}{n^2} \sum_{h_2=r_n+1}^{n-2r_n-3} \sum_{h_3=r_n+1}^{n-h_2-r_n-2} (n-h_2-h_3) |p_{h_3} - p_0^2| (\sin(\lambda/2))^{-2} \leq cn \sum_{h_3=r_n+1}^n \xi_{h_3} \rightarrow 0.$$

A similar bound applies to  $Q_4$ . A double application of Lemma A.1 yields

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

Collecting these bounds, we proved (5.12). Similar arguments apply to (5.13). □

Thus, we showed that  $J_1$  and  $J_2$  are negligible as  $n \rightarrow \infty$ . Hence, (5.6) holds.

**Proof of (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)

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

Then we have

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

*Proof that  $G_3$  and  $G_6$  are negligible*

Using mixing and Lemma A.1, we have as  $n \rightarrow \infty$ ,

$$\begin{aligned}
 |G_3| &= 8 \frac{m_n^2}{n^2} \left| \sum_{B_3} (n-h_1-h_2)((p_{h_1 h_2} - p_0 p_{h_2}) + p_0(p_{h_2} - p_0^2) + p_0^3) f_{h_1 h_2} \right| \\
 &\leq c m_n^2 \sum_{h_1=r_n+1}^n \xi_{h_1} + c \frac{m_n}{n} \sum_{h_2=r_n+1}^n \xi_{h_2} + c \frac{(m_n p_0)^3}{m_n (\sin(\lambda_j/2))^2} = G'_3 \rightarrow 0.
 \end{aligned}$$

We also have

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

By (5.12),  $G_{62}$  is negligible and the same arguments as for  $G_3$  show that  $G_{61} \leq r_n G'_3 \rightarrow 0$ . Thus,  $G_6$  is negligible as  $n \rightarrow \infty$ .

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

First, observe that

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

and we also know that (5.5) holds. Thus, (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 A.1,

$$\begin{aligned} |G_1 - P_1| &= 4 \frac{m_n^2}{n^2} \left| \sum_{h=r_n+1}^{n-1} (n-h) ((p_h - p_0^2) + p_0^2) 0.5(1 + \cos(2\lambda_j h)) - (m_n p_0)^2 \right| \\ &\leq c \frac{m_n}{n} m_n \sum_{h=r_n+1}^{n-1} \xi_h + c(m_n p_0)^2 \left| \frac{1}{2n^2} \sum_{h=r_n+1}^{n-1} (n-h) - 1 \right| + c \frac{1}{n} (m_n p_0)^2 \rightarrow 0. \end{aligned}$$

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

$$\begin{aligned} |G_2 - P_2| &\leq cG'_3 + \left| 2 \frac{m_n^2}{n^2} \sum_{h_2=1}^{r_n} \sum_{h_1=r_n+1}^{n-h_2-1} (n-h_1-h_2) (p_{h_1 h_2} + p_{h_2 h_1}) \right. \\ &\quad \times (\cos(\lambda_j h_1) + \cos(\lambda_j (2h_2 + h_1))) \\ &\quad \left. - 4m_n^2 p_0 \sum_{h=1}^{r_n} \frac{n-h}{n} p_h \cos(\lambda_j h) \right| \\ &\leq cG'_3 + c \frac{r_n^2}{n} (m_n p_0)^2 + c \frac{m_n r_n}{n} m_n \sum_{h_2=r_n+1}^{n-1} \xi_{h_2} + c(m_n p_0)^2 \frac{r_n}{n(\sin(\lambda_j))^2} \rightarrow 0. \end{aligned}$$

Here we used (5.9) to rewrite  $p_{h_1 h_2}$  such that the mixing condition and Lemma A.1 can be applied.

Finally, we turn to  $G_4$  and  $G_5$ . By virtue of (5.12) and (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  $\{1 \leq h_1, h_3 \leq r_n, r_n < h_2 < n - h_1 - h_3\}$ . We write  $G_{4+5}$  for the remaining sum. Recall that

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

Then we have

$$\begin{aligned} |G_{4+5} - P_3| &= \left| 4 \frac{m_n^2}{n^2} \sum_{h_1=1}^{r_n} \sum_{h_3=1}^{r_n} \sum_{h_2=r_n+1}^{n-h_1-h_3-1} (n - h_1 - h_2 - h_3) ((p_{h_1 h_2 h_3} - p_{h_1} p_{h_3}) + p_{h_1} p_{h_3}) \right. \\ &\quad \times (2f_{h_1 h_3} + 2\cos(\lambda_{j_1}(h_1 + 2h_2 + h_3))) \\ &\quad \left. - 4 \frac{m_n^2}{n^2} \sum_{h_1=r_n+1}^{n-1} \sum_{h_2=r_n+1}^{n-1} (n - h_1)(n - h_2) p_{h_1} p_{h_2} f_{h_1 h_2} \right| \\ &\leq c \frac{r_n^3}{n} (m_n p_0)^2 + c \frac{m_n r_n^2}{n} m_n \sum_{h_3=r_n+1}^{n-1} \xi_{h_3} + c \frac{r_n^2}{n(\sin(\lambda_{j_1}))^2} (m_n p_0)^2. \end{aligned}$$

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

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

## 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  $P(X_0 > x, X_t > x)/[P(X > x)]^2 \rightarrow 1$  as  $x \rightarrow \infty$  and also suggested diagnostic conditions in this situation. Hill [28] proposed the quantities  $\lim_{x \rightarrow \infty} [P(X_0 > x, X_t > x)/[P(X > x)]^2 - 1]$  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 of 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, for example, 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), goodness-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.

### Appendix A: Some trigonometric sum formulas

Equations (A.1) and (A.2) are given in Gradshteyn and Ryzhik [23], 1.341 on page 29; (A.3) and (A.4) are 1.352 on page 31; and (A.5) and (A.6) are listed as 1.353 on page 31. For any  $\lambda, x$  and  $n \geq 1$ , the following identities hold

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

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

$$\sum_{k=1}^{n-1} k \cos(k\lambda) = \frac{n \sin((2n - 1)\lambda/2)}{2 \sin(\lambda/2)} - \frac{1 - \cos n\lambda}{4(\sin(\lambda/2))^2}, \tag{A.3}$$

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

$$\sum_{k=1}^{n-1} p^k \sin(k\lambda) = \frac{p \sin(\lambda) - p^n \sin(n\lambda) + p^{n+1} \sin((n - 1)\lambda)}{1 - 2p \cos(\lambda) + p^2}, \tag{A.5}$$

$$\sum_{k=0}^{n-1} p^k \cos(k\lambda) = \frac{1 - p \cos(\lambda) - p^n \cos(n\lambda) + p^{n+1} \cos((n - 1)\lambda)}{1 - 2p \cos(\lambda) + p^2}. \tag{A.6}$$

Using these formulas, direct calculation yields for any frequency  $\lambda$ ,

$$\begin{aligned} & \sum_{s=1}^{n-h} [\cos(\lambda s) \sin(\lambda(s+h)) + \cos(\lambda(s+h)) \sin(\lambda s)] \\ &= \sum_{s=1}^{n-h} \sin(2\lambda s + \lambda h) = \frac{\sin(\lambda n) \sin(\lambda(n-h+1))}{\sin \lambda} - \sin(\lambda h). \end{aligned} \tag{A.7}$$

For distinct frequencies  $\lambda, \omega$ , we then obtain

$$\begin{aligned} & \sum_{s=1}^{n-h} [\cos(\lambda s) \sin(\omega(s+h)) + \cos(\lambda(s+h)) \sin(\omega s)] \\ &= 0.5 \sum_{s=1}^{n-h} [\sin((\lambda + \omega)s + \omega h) - \sin((\lambda - \omega)s - \omega h)] \\ & \quad + 0.5 \sum_{s=1}^{n-h} [\sin((\lambda + \omega)s + \lambda h) - \sin((\lambda - \omega)s + \lambda h)] \\ &= -\sin(\omega h) \\ & \quad + 0.5 \frac{\sin((n-h+1)(\lambda + \omega)/2)}{\sin((\lambda + \omega)/2)} \\ & \quad \times [\sin(\omega h + (n-h)(\lambda + \omega)/2) + \sin(\lambda h + (n-h)(\lambda + \omega)/2)] \\ & \quad - 0.5 \frac{\sin((n-h+1)(\lambda - \omega)/2)}{\sin((\lambda - \omega)/2)} \\ & \quad \times [\sin(-\omega h + (n-h)(\lambda - \omega)/2) + \sin(\lambda h + (n-h)(\lambda - \omega)/2)], \end{aligned} \tag{A.8}$$

$$\begin{aligned} & \sum_{s=1}^{n-h} [\cos(\lambda s) \cos(\omega(s+h)) + \cos(\lambda(s+h)) \cos(\omega s)] \\ &= 0.5 \sum_{s=1}^{n-h} [\cos((\lambda + \omega)s + \omega h) + \cos((\lambda - \omega)s - \omega h) + \cos((\lambda + \omega)s + \lambda h) \\ & \quad + \cos((\lambda - \omega)s + \lambda h)] \\ &= -\cos(\omega h) - \cos(\lambda h) \\ & \quad + 0.5 \frac{\sin((n-h+1)(\lambda + \omega)/2)}{\sin((\lambda + \omega)/2)} \\ & \quad \times [\cos(\omega h + (n-h)(\lambda + \omega)/2) + \cos(\lambda h + (n-h)(\lambda + \omega)/2)] \\ & \quad + 0.5 \frac{\sin((n-h+1)(\lambda - \omega)/2)}{\sin((\lambda - \omega)/2)} \\ & \quad \times [\cos(-\omega h + (n-h)(\lambda - \omega)/2) + \cos(\lambda h + (n-h)(\lambda - \omega)/2)], \end{aligned} \tag{A.9}$$

$$\begin{aligned}
 & \sum_{s=1}^{n-h} [\sin(\lambda s) \sin(\omega(s+h)) + \sin(\lambda(s+h)) \sin(\omega s)] \\
 &= 0.5 \sum_{s=1}^{n-h} [\cos((\lambda + \omega)s + \omega h) - \cos((\lambda - \omega)s - \omega h) + \cos((\lambda + \omega)s + \lambda h) \\
 &\quad - \cos((\lambda - \omega)s + \lambda h)] \\
 &= 0.5 \frac{\sin((n-h+1)(\lambda - \omega)/2)}{\sin((\lambda - \omega)/2)} \\
 &\quad \times [\cos(-\omega h + (n-h)(\lambda - \omega)/2) + \cos(\lambda h + (n-h)(\lambda - \omega)/2)] \\
 &\quad - 0.5 \frac{\sin((n-h+1)(\lambda + \omega)/2)}{\sin((\lambda + \omega)/2)} \\
 &\quad \times [\cos(\omega h + (n-h)(\lambda + \omega)/2) + \cos(\lambda h + (n-h)(\lambda + \omega)/2)].
 \end{aligned} \tag{A.10}$$

Next, assume the conditions of Theorem 5.1. Then a direct application of (A.1)–(A.4) yields for  $\lambda \in (0, \pi)$  the following relations:

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

$$\begin{aligned}
 & -\cos(x) \left( \frac{n \sin((2n-1)\lambda/2)}{2 \sin(\lambda/2)} - \frac{(r_n+1) \sin((2r_n-1)\lambda/2)}{2 \sin(\lambda/2)} \right. \\
 & \quad \left. - \frac{\cos((r_n+1)\lambda) - \cos(n\lambda)}{4(\sin(\lambda/2))^2} \right) \\
 & + \sin(x) \left( \frac{n \cos((2n-1)\lambda/2)}{2 \sin(\lambda/2)} - \frac{(r_n+1) \cos((2r_n-1)\lambda/2)}{2 \sin(\lambda/2)} \right. \\
 & \quad \left. - \frac{\sin(n\lambda) - \sin((r_n+1)\lambda)}{4(\sin(\lambda/2))^2} \right).
 \end{aligned}$$

**Lemma A.1.** *Under the assumptions of Theorem 5.1 the following relations hold uniformly for  $\lambda \in (0, 2\pi)$ , as  $n \rightarrow \infty$ ,*

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

## Appendix B: 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  $(\psi_j)$  of the linear process representation of  $(X_t)$  (cf. (3.2)):

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

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

The case  $\phi \in (0, 1)$ ,  $\theta + \phi > 0$ ,  $p > 0$ . A direct application of (3.4) yields that

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

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

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

$$L^{(2)}(n, x, \alpha, \lambda)$$

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

Then

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

where

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

The corresponding spectral density is given by

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

The case  $\phi \in (0, 1)$ ,  $\theta + \phi < 0$ ,  $q > 0$ . In view of (3.4), we have

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

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 (3.4) vanishes for  $i \geq 1$ . Thus,

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

For  $h = 2k > 0$ ,

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

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

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

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

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

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 (3.4) vanishes for  $i \geq 0$ . Thus,

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

For  $h = 2k > 0$ ,

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

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

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

The corresponding spectral density is

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



## Acknowledgments

We would like to thank the reviewers of our paper for careful reading and comments, in particular for pointing out several useful references. Thomas Mikosch's research is partly supported by the Danish Research Council (FNU) Grants 09-072331 "Point process modelling and statistical inference" and 10-084172 "Heavy tail phenomena: Modeling and estimation". The research of Yuwei Zhao is supported by the Danish Research Council Grant 10-084172.

## References

- [1] Andersen, T.G., Davis, R.A., Kreiss, J.P. and Mikosch, T. (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. Theory Related Fields* **150** 337–372. [MR2824860](#)
- [3] Basrak, B., Davis, R.A. and Mikosch, T. (2002). Regular variation of GARCH processes. *Stochastic Process. Appl.* **99** 95–115. [MR1894253](#)
- [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. *Stochastic Process. Appl.* **119** 1055–1080. [MR2508565](#)
- [6] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation. Encyclopedia of Mathematics and Its Applications* **27**. Cambridge: Cambridge Univ. Press. [MR0898871](#)
- [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 ed. Oakland, CA: Holden-Day Inc. [MR0595684](#)
- [9] Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*, 2nd ed. *Springer Series in Statistics*. New York: Springer. [MR1093459](#)
- [10] Brockwell, P.J. and Davis, R.A. (2002). *Introduction to Time Series and Forecasting*, 2nd ed. *Springer Texts in Statistics*. New York: Springer. [MR1894099](#)
- [11] Davis, R.A. and Hsing, T. (1995). Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Probab.* **23** 879–917. [MR1334176](#)
- [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. [MR1915537](#)
- [13] Davis, R.A. and Mikosch, T. (2009). Extreme value theory for GARCH processes. In *Handbook of Financial Time Series* (T.G. Andersen, R.A. Davis, J.P. Kreiss and T. Mikosch, eds.) 187–200. Berlin: Springer.
- [14] Davis, R.A. and Mikosch, T. (2009). Extremes of stochastic volatility models. In *Handbook of Financial Time Series* (T.G. Andersen, R.A. Davis, J.P. Kreiss and T. Mikosch, eds.) 355–364. Berlin: Springer.
- [15] Davis, R.A. and Mikosch, T. (2009). The extremogram: A correlogram for extreme events. *Bernoulli* **15** 977–1009. [MR2597580](#)
- [16] Davis, R.A., Mikosch, T. and Cribben, I. (2012). Towards estimating extremal serial dependence via the bootstrapped extremogram. *J. Econometrics* **170** 142–152. [MR2955945](#)
- [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. *Stochastic Process. Appl.* **122** 3790–3811. [MR2965926](#)
- [19] Doukhan, P. (1994). *Mixing: Properties and Examples. Lecture Notes in Statistics* **85**. New York: Springer. [MR1312160](#)
- [20] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events: For Insurance and Finance. Applications of Mathematics (New York)* **33**. Berlin: Springer. [MR1458613](#)
- [21] Fasen, V., Klüppelberg, C. and Schlather, M. (2010). High-level dependence in time series models. *Extremes* **13** 1–33. [MR2593949](#)
- [22] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications. Vol. II*, 2nd ed. New York: Wiley. [MR0270403](#)
- [23] Gradshteyn, I.S. and Ryzhik, I.M. (1980). *Table of Integrals, Series, and Products*. New York: Academic Press. Corrected and enlarged edition edited by Alan Jeffrey, Incorporating the fourth edition edited by Yu. V. Geronimus [Yu. V. Geronimus] and M. Yu. Tseytlin [M. Yu. Tseitlin], Translated from the Russian. [MR0582453](#)
- [24] Grenander, U. and Rosenblatt, M. (1984). *Statistical Analysis of Stationary Time Series*, 2nd ed. New York: Chelsea. [MR0890514](#)
- [25] Hagemann, A. (2011). Robust spectral analysis. Working paper, UICU.
- [26] Hannan, E.J. (1960). *Time Series Analysis. Methuen's Monographs on Applied Probability and Statistics*. London: Methuen. [MR0114281](#)
- [27] Hill, J.B. (2009). On functional central limit theorems for dependent, heterogeneous arrays with applications to tail index and tail dependence estimation. *J. Statist. Plann. Inference* **139** 2091–2110. [MR2497563](#)
- [28] Hill, J.B. (2011). Extremal memory of stochastic volatility with an application to tail shape inference. *J. Statist. Plann. Inference* **141** 663–676. [MR2732937](#)
- [29] Ibragimov, I.A. and Linnik, Y.V. (1971). *Independent and Stationary Sequences of Random Variables*. Groningen: Wolters-Noordhoff. With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, Translation from the Russian edited by J. F. C. Kingman. [MR0322926](#)
- [30] Jakubowski, A. (1993). Minimal conditions in  $p$ -stable limit theorems. *Stochastic Process. Appl.* **44** 291–327. [MR1200412](#)
- [31] Jakubowski, A. (1997). Minimal conditions in  $p$ -stable limit theorems. II. *Stochastic Process. Appl.* **68** 1–20. [MR1454576](#)
- [32] Jessen, A.H. and Mikosch, T. (2006). Regularly varying functions. *Publ. Inst. Math. (Beograd) (N.S.)* **80** 171–192. [MR2281913](#)
- [33] Kallenberg, O. (1983). *Random Measures*, 3rd ed. Berlin: Akademie-Verlag. [MR0818219](#)
- [34] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes. Springer Series in Statistics*. New York: Springer. [MR0691492](#)
- [35] Ledford, A.W. and Tawn, J.A. (2003). Diagnostics for dependence within time series extremes. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **65** 521–543. [MR1983762](#)
- [36] Lee, J. and Subba Rao, S. (2012). The quantile spectral density and comparison based tests for non-linear time series. Working paper.
- [37] Mikosch, T. and Rezapur, M. (2013). Stochastic volatility models with possible extremal clustering. *Bernoulli* **19** 1688–1713.
- [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. [MR1789987](#)
- [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. [MR1805791](#)
- [40] Petrov, V.V. (1995). *Limit Theorems of Probability Theory: Sequences of Independent Random Variables. Oxford Studies in Probability* **4**. New York: The Clarendon Press Oxford Univ. Press. [MR1353441](#)

- [41] Pham, T.D. and Tran, L.T. (1985). Some mixing properties of time series models. *Stochastic Process. Appl.* **19** 297–303. [MR0787587](#)
- [42] Priestley, M.B. (1981). *Spectral Analysis and Time Series*. London, New York: Academic Press.
- [43] Resnick, S.I. (1986). Point processes, regular variation and weak convergence. *Adv. in Appl. Probab.* **18** 66–138. [MR0827332](#)
- [44] Resnick, S.I. (1987). *Extreme Values, Regular Variation, and Point Processes. Applied Probability. A Series of the Applied Probability Trust* **4**. New York: Springer. [MR0900810](#)
- [45] Resnick, S.I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer Series in Operations Research and Financial Engineering*. New York: Springer. [MR2271424](#)

*Received April 2012 and revised November 2012*