

# Weak convergence of the function-indexed integrated periodogram for infinite variance processes

SAMI UMUT CAN<sup>1</sup>, THOMAS MIKOSCH<sup>2</sup> and  
GENNADY SAMORODNITSKY<sup>3</sup>

<sup>1</sup>*Center for Applied Mathematics, Cornell University, 657 Rhodes Hall, Ithaca, NY 14853, USA.*

*E-mail: [suc2@cornell.edu](mailto:suc2@cornell.edu)*

<sup>2</sup>*Department of Mathematics, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark. E-mail: [mikosch@math.ku.dk](mailto:mikosch@math.ku.dk)*

<sup>3</sup>*School of Operations Research and Information Engineering, Cornell University, 220 Rhodes Hall, Ithaca, NY 14853, USA. E-mail: [gennady@orie.cornell.edu](mailto:gennady@orie.cornell.edu)*

In this paper, we study the weak convergence of the integrated periodogram indexed by classes of functions for linear processes with symmetric  $\alpha$ -stable innovations. Under suitable summability conditions on the series of the Fourier coefficients of the index functions, we show that the weak limits constitute  $\alpha$ -stable processes which have representations as infinite Fourier series with i.i.d.  $\alpha$ -stable coefficients. The cases  $\alpha \in (0, 1)$  and  $\alpha \in [1, 2)$  are dealt with by rather different methods and under different assumptions on the classes of functions. For example, in contrast to the case  $\alpha \in (0, 1)$ , entropy conditions are needed for  $\alpha \in [1, 2)$  to ensure the tightness of the sequence of integrated periodograms indexed by functions. The results of this paper are of additional interest since they provide limit results for infinite mean random quadratic forms with particular Toeplitz coefficient matrices.

**Keywords:** asymptotic theory; empirical spectral distribution; entropy; infinite variance process; integrated periodogram; linear process; random quadratic form; spectral analysis; stable process; time series; weighted integrated periodogram

## 1. Introduction

Over the last decades, efforts have been made to get a better understanding of non-Gaussian time series in the time and frequency domains. In particular, time series whose marginal distributions exhibit power law behavior have attracted a lot of attention. The need for such models arises from applications in areas as diverse as insurance, geophysics, finance and telecommunications. Infinite fourth moments are not untypical for series of daily log-returns from exchange rates, stock indices, and other speculative prices, whereas infinite second moments can be observed in time series from insurance such as for windstorm, industrial fire and earthquake insurance [10,11,13,26]. Infinite first moments are typical for the marginal distribution of the magnitudes of earthquakes [17]. Infinite variances are observed for the sizes of teletraffic data in the World Wide Web [6,7,14,24,38]; see also the recent books [1,33].

Classical time series analysis deals with the second (or higher) order moment structure of a stationary sequence. Heavy-tailed modeling requires, in addition, that one takes into account the

interplay between the dependence structure and the tails of the series. An important task is to understand the classical statistical estimators and test procedures when big shocks to the underlying system are present. When the marginal distributions have infinite variance, the notions of autocovariance, autocorrelation and spectral distribution lose their meaning. However, various studies over the last twenty years have shown that the analysis of linear processes  $X_t = \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j}$ ,  $t \in \mathbb{Z}$ , with heavy-tailed i.i.d. innovations  $(\varepsilon_j)_{j \in \mathbb{Z}}$  and constant coefficients  $(\psi_j)_{j \in \mathbb{Z}}$  is very similar to classical (finite variance) time series analysis, where notions such as autocovariances and spectral density are defined only in terms of the  $\psi_j$ 's and the innovation variance  $\sigma_\varepsilon^2$ . Most estimators and test statistics from classical time series analysis can be modified insofar that one considers self-normalized (or studentized) versions of them and for these versions, an asymptotic theory exists which parallels the classical theory with Gaussian limit processes. In contrast to the latter theory, the limits involve infinite variance stable distributions and processes [13, 21, 25].

One of the main goals of classical (finite variance) time series analysis is to study the spectral properties of the linear process  $(X_t)$ . In this context, the *periodogram*

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

plays a prominent role as an estimator of the spectral density. Numerous estimation and test procedures are based on this statistic and integrated versions of the form  $J_{n,X}(f) = \int_0^\pi I_{n,X}(\lambda) f(\lambda) d\lambda$  for appropriate classes of real-valued functions  $f \in \mathcal{F}$  on  $[0, \pi]$ . In applications, the class  $\mathcal{F}_I = \{I_{[0,x]}: x \in [0, \pi]\}$  is most important. The resulting integrated periodogram is a process indexed by  $x \in [0, \pi]$ . Under general conditions,  $J_{n,X}(I_{[0,\cdot]})$  converges uniformly with probability 1 to the function  $\sigma_\varepsilon^2 \int_0^\pi |\psi(e^{-i\lambda})|^2 d\lambda$ , where

$$\psi(e^{-i\lambda}) = \sum_{j=-\infty}^{\infty} \psi_j e^{-i\lambda j}, \quad \lambda \in [0, \pi],$$

and  $|\psi(e^{-i\lambda})|^2$  is the corresponding *power transfer function*. The latter is one of the essential building blocks of the *spectral density* of the stationary process  $(X_t)$ :

$$f_X(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} |\psi(e^{-i\lambda})|^2 = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma_X(h), \quad \lambda \in [0, \pi].$$

This is the Fourier series based on the *autocovariance function*

$$\gamma_X(h) = \text{cov}(X_0, X_h) = \sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+|h|}, \quad h \in \mathbb{Z}.$$

Since  $J_{n,X}(I_{[0,\cdot]})$  estimates the spectral distribution function of the stationary process  $(X_t)$ , it has been used for a long time as the *empirical spectral distribution function*, both as an estimator and a basic tool for constructing goodness-of-fit tests for the underlying spectral distribution function

[4,16,32]. Since the limit process of the properly centered and normalized process  $J_{n,X}(I_{[0,\cdot]})$  depends on the (in general unknown) spectral density  $f_X$ , Bartlett [2] proposed to consider the integrated periodogram based on  $\mathcal{F}_B = \{I_{[0,x]}/f_X(x) : x \in [0, \pi]\}$ . Under general conditions, this process converges uniformly, with probability 1, to the function  $f(x) \equiv x$  and the limit process can be shown to be independent of the coefficients of the linear process, but depends on the fourth moment of  $\varepsilon_1$ . More generally, weighted integrated periodograms based on the classes  $\mathcal{F}_g = \{I_{[0,x]}g(x) : x \in [0, \pi]\}$  for suitable functions  $g$  are used to estimate the spectral density or to perform various tests concerning the spectrum of the underlying stationary sequence [32].

The weighted integrated periodogram is also the basis of one of the classical estimators for fitting ARMA and fractional ARIMA models. This method goes back to early work by Whittle [37]. In this context, one considers the functional  $J_{n,X}(1/f_X(\cdot; \theta))$ ,  $f_X(\cdot; \theta) \in \mathcal{F}_W$ , where  $\mathcal{F}_W$  is a class of spectral densities indexed by a parameter  $\theta \in \Theta \subset \mathbb{R}^d$ . The Whittle estimator  $\hat{\theta}_n$  of the true parameter  $\theta_0 \in \Theta$  is the minimizer of  $J_{n,X}(1/f_X(\cdot; \theta))$  over  $\Theta$ . This estimation technique is one of the backbones of quasi-maximum likelihood estimation in parametric time series modeling. The Whittle estimator is known to be asymptotically equivalent to the corresponding least-squares and Gaussian quasi-maximum likelihood estimators [4]. When proving the asymptotic normality and consistency of  $\hat{\theta}_n$ , one has to study the properties of the sequence  $(J_{n,X}(1/f_X(\cdot; \hat{\theta}_n)))$  which, again, can be considered as a weighted integrated periodogram indexed by a class of functions.

The above examples have in common that one always considers a stochastic process  $(J_{n,X}(f))_{f \in \mathcal{F}}$  for some class of functions. In all cases, one is interested in the asymptotic behavior of the process  $J_{n,X}$ , uniformly over the class  $\mathcal{F}$ . This is analogous to the case of the empirical distribution function indexed by classes of functions. General references in this context are the monographs [31,36]. Early on, this analogy was discovered by Dahlhaus [8], who gave the uniform convergence theory for  $J_{n,X}$  under entropy and exponential moment conditions. The almost sure and weak convergence theory under entropy and power moment conditions was given in [28]. A recent survey of nonparametric statistical methods related to the empirical spectral distribution indexed by classes of functions is [9].

It is the aim of this paper to develop an analogous weak convergence theory for heavy-tailed stationary processes. We will understand ‘heavy-tailedness’ in the sense of infinite variance of the marginal distributions. Our focus will be on linear processes  $(X_t)$  with i.i.d. symmetric  $\alpha$ -stable ( $S\alpha S$ ) innovations  $(\varepsilon_t)$  for some  $\alpha \in (0, 2)$ . Recall that a random variable  $Y_\alpha$  has a *symmetric stable distribution* ( $Y_\alpha \sim S_\alpha(\sigma, 0, 0)$ ) if there are parameters  $0 < \alpha \leq 2$ , and  $\sigma \geq 0$  such that its characteristic function has the form  $E e^{itY_\alpha} = e^{-\sigma^\alpha |t|^\alpha}$ . For convenience, we also assume that  $\sigma = 1$  for the distribution of  $\varepsilon_1$ . For  $\alpha < 2$ , the random variable  $\varepsilon_1$  is known to have infinite variance [15,35]. Much of the theory given below depends on tail estimates for random quadratic forms in i.i.d. infinite variance random variables. Such results are available for i.i.d. stable sequences. Although it seems feasible that the theory can be extended to the more general class of processes whose innovations have regularly varying tails, we do not attempt to achieve this goal. The price would be more technicalities and the gain would be negligible.

We intend to show how the classical (finite variance) tools and methods have to be modified in the infinite variance stable situation which can be considered as a boundary case of the classical one when some of the innovations assume extremely large values. By now, there exists quite a clear picture concerning the asymptotic theory of the sample autocovariances, the periodogram

and its integrated versions when the innovation sequence in a linear process has infinite variance; see [13], Chapter 7. In addition to the latter reference, goodness-of-fit tests for heavy-tailed processes (corresponding to the class  $\mathcal{F}_I$ ) were considered for short- and long-memory linear processes [18,20], and Whittle estimation for infinite variance ARMA and FARIMA processes was also studied [19,27].

The paper is organized as follows. In Section 2, we introduce some useful notation for the integrated periodogram. Our main goal is to prove the weak convergence of the integrated periodograms indexed by suitable classes of functions. We achieve this goal for an i.i.d. sequence in Section 3, first by showing the *convergence of the finite-dimensional distributions* (Section 3.1), then the *tightness*. The conditions and methods are rather different in the cases  $\alpha \in (0, 1)$  (Section 3.2) and  $\alpha \in [1, 2)$  (Section 3.3). The case  $\alpha \in (0, 1)$  is treated in the more general context of random quadratic forms with Toeplitz coefficient matrices satisfying some summability condition. The case  $\alpha \in [1, 2)$  requires entropy conditions and the corresponding techniques. In Section 4, we extend the limit theory for the integrated periodograms from an i.i.d. sequence to linear processes. The Appendix contains some auxiliary results concerning tail estimates of random quadratic forms in stable random variables. The weak convergence results of this paper might also be of separate interest in the context of infinite variance random quadratic forms. The theory for such quadratic forms is not well studied. We also refer to an extended version of this paper [5] which covers the class of stochastic volatility processes with regularly varying marginal distributions.

## 2. Preliminaries on the periodogram

The following decomposition of the periodogram is fundamental:

$$I_{n,X}(\lambda) = \gamma_{n,X}(0) + 2 \sum_{h=1}^{n-1} \cos(\lambda h) \gamma_{n,X}(h), \quad (2.1)$$

where  $\gamma_{n,X}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} X_t X_{t+h}$ ,  $h \in \mathbb{Z}$ , denotes the *sample autocovariance function* of the sample  $X_1, \dots, X_n$ . Note that the definition of  $\gamma_{n,X}$  deviates slightly from the usual one where the  $X_t$ 's are centered by the sample mean. However, for the theory given below, this centering is not essential. Centering with the sample mean  $\bar{X}_n$  is not the most natural choice when dealing with infinite variance processes. In what follows, we will frequently make use of the *self-normalized periodogram*

$$\tilde{I}_{n,X}(\lambda) = \frac{I_{n,X}(\lambda)}{\gamma_{n,X}(0)} = \rho_{n,X}(0) + 2 \sum_{h=1}^{n-1} \cos(\lambda h) \rho_{n,X}(h),$$

where  $\rho_{n,X}(h) = \gamma_{n,X}(h)/\gamma_{n,X}(0)$ ,  $h \in \mathbb{Z}$ , denotes the *sample autocorrelation function* of  $X_1, \dots, X_n$ . In view of (2.1), we can rewrite  $J_{n,\varepsilon}(f)$  as follows:

$$J_{n,\varepsilon}(f) = \gamma_{n,\varepsilon}(0) a_0(f) + 2 \sum_{h=1}^{n-1} a_h(f) \gamma_{n,\varepsilon}(h), \quad (2.2)$$

where

$$a_h(f) = \int_0^\pi \cos(\lambda h) f(\lambda) d\lambda, \quad h \in \mathbb{Z}, \quad (2.3)$$

are the *Fourier coefficients* of  $f$ . We also introduce the self-normalized version of  $J_{n,\varepsilon}$ :

$$\tilde{J}_{n,\varepsilon}(f) = \rho_{n,\varepsilon}(0)a_0(f) + 2 \sum_{h=1}^{n-1} a_h(f)\rho_{n,\varepsilon}(h). \quad (2.4)$$

### 3. The i.i.d. case

In this section, we study the limit behavior of the integrated periodograms  $J_{n,\varepsilon}$  indexed by classes of functions for an i.i.d.  $S\alpha S$  sequence with  $\alpha \in (0, 2)$ . In Section 3.1, we consider the convergence of the finite-dimensional distributions. In Sections 3.2 and 3.3, we prove the tightness of the processes in the cases  $\alpha \in (0, 1)$  and  $\alpha \in [1, 2)$ , respectively. In the case  $\alpha \in (0, 1)$ , we solve a more general weak convergence problem for random quadratic forms in the i.i.d. sequence  $(\varepsilon_t)$ ; the convergence of the integrated periodograms indexed by classes of functions is only a special case. The case  $\alpha \in [1, 2)$  is more involved. Among others, entropy conditions will be needed and we only prove results on the weak convergence of the empirical spectral distribution, that is, we focus on random quadratic forms with Toeplitz coefficient matrices given by the Fourier coefficients  $a_h(f)$  defined in (2.3).

#### 3.1. Convergence of the finite-dimensional distributions

A glance at decomposition (2.2) is enough to see that the convergence of the finite-dimensional distributions of  $J_{n,\varepsilon}$  is essentially determined by the weak limit behavior of the sample autocovariances  $\gamma_{n,\varepsilon}(h)$ . For this reason, we recall a well-known result due to Davis and Resnick [12]; see also [4], Section 13.3.

**Lemma 3.1.** *For every  $m \geq 1$ ,*

$$\left( \frac{n\gamma_{n,\varepsilon}(0)}{n^{2/\alpha}}, \frac{n\gamma_{n,\varepsilon}(h)}{(n \log n)^{1/\alpha}}, h = 1, \dots, m \right) \Longrightarrow (Y_0, Y_1, \dots, Y_m), \quad (3.1)$$

where  $\Longrightarrow$  denotes weak convergence, the  $Y_h$ 's are independent,  $Y_0$  is  $S_{\alpha/2}(\sigma_1, 1, 0)$  and  $(Y_h)_{h=1,\dots,m}$  are i.i.d.  $S_\alpha(\sigma_2, 0, 0)$  for some  $\sigma_i = \sigma_i(\alpha)$ ,  $i = 1, 2$ . In particular,

$$(n/\log n)^{1/\alpha}(\rho_{n,\varepsilon}(h))_{h=1,\dots,m} \Longrightarrow (Y_h/Y_0)_{h=1,\dots,m}. \quad (3.2)$$

The latter result is an immediate consequence of (3.1) and the continuous mapping theorem. Lemma 3.1 yields the weak convergence for any finite linear combination of the sample autocovariances and autocorrelations. It also suggests that the weak limit of the standardized process

$J_{n,\varepsilon}(f)$  will be determined by the infinite series  $\sum_{h=1}^{\infty} a_h(f)Y_h$ . But this also means that we need to require additional assumptions on the sequence  $(a_h(f))$ .

We will treat this problem in a more general context. Consider a sequence

$$\mathbf{a} = (a_1, a_2, \dots) \in \ell^\alpha,$$

that is,  $\mathbf{a}$  satisfies the summability condition  $\sum_h |a_h|^\alpha < \infty$ . For such an  $\mathbf{a}$ , we define the sequences of processes

$$\begin{cases} X_n(\mathbf{a}) = (n \log n)^{-1/\alpha} \sum_{k=1}^{n-1} a_k [n\gamma_{n,\varepsilon}(k)], & Y(\mathbf{a}) = \sum_{k=1}^{\infty} a_k Y_k, \\ \tilde{X}_n(\mathbf{a}) = (n/\log n)^{1/\alpha} \sum_{k=1}^{n-1} a_k \rho_{n,\varepsilon}(k), & \tilde{Y}(\mathbf{a}) = Y(\mathbf{a})/Y_0. \end{cases} \quad (3.3)$$

Here,  $Y_0, Y_1, Y_2, \dots$  are independent stable random variables, as described in Lemma 3.1. The 3-series theorem [30] implies that  $\mathbf{a} \in \ell^\alpha$  is equivalent to the a.s. convergence of the infinite series  $Y(\mathbf{a})$  in (3.3). However, for the weak convergence of  $(X_n)$  and  $(\tilde{X}_n)$ , we need a slightly stronger assumption:

$$\mathbf{a} \in \ell^\alpha \log \ell = \left\{ \mathbf{a} = (a_1, a_2, \dots) \in \ell^\alpha : \sum_{k=1}^{\infty} |a_k|^\alpha \log^+ \frac{1}{|a_k|} < \infty \right\}.$$

This assumption ensures the weak convergence of the random quadratic forms in (3.3); see the proof of Theorem 3.2 below. Assumptions of this type frequently occur in the literature on infinite variance quadratic forms (e.g., [22]). They appear in a natural way in tail estimates for quadratic forms in i.i.d. stable random variables; see the Appendix.

We can now formulate our result concerning the convergence of the finite-dimensional distributions.

**Theorem 3.2.** *For any  $\alpha \in (0, 2)$ ,*

$$(X_n(\mathbf{a}))_{\mathbf{a} \in \ell^\alpha \log \ell} \xrightarrow{\text{fidi}} (Y(\mathbf{a}))_{\mathbf{a} \in \ell^\alpha \log \ell} \quad \text{and} \quad (\tilde{X}_n(\mathbf{a}))_{\mathbf{a} \in \ell^\alpha \log \ell} \xrightarrow{\text{fidi}} (\tilde{Y}(\mathbf{a}))_{\mathbf{a} \in \ell^\alpha \log \ell}.$$

**Proof.** Using a Cramér–Wold argument, it suffices to prove the convergence of the one-dimensional distributions. From (3.1) and the continuous mapping theorem, it immediately follows that for every  $m \geq 1$ ,

$$(n \log n)^{-1/\alpha} \sum_{k=1}^m a_k [n\gamma_{n,\varepsilon}(k)] \implies Y_m(\mathbf{a}) = \sum_{k=1}^m a_k Y_k,$$

where  $\implies$  denotes weak convergence. Since  $\mathbf{a} \in \ell^\alpha$ ,  $Y_m(\mathbf{a}) \implies Y(\mathbf{a})$  as  $m \rightarrow \infty$  follows from the 3-series theorem. According to [3], Theorem 4.2, it remains to show that

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left( (n \log n)^{-1/\alpha} \left| \sum_{k=m+1}^{n-1} a_k [n \gamma_{n,\varepsilon}(k)] \right| > \epsilon \right) = 0$$

for every  $\epsilon > 0$  and  $\mathbf{a} \in \ell^\alpha \log \ell$ . We write  $p_{n,m}(\mathbf{a}; \epsilon)$  for the above probabilities. Applying Lemma A.1 in the Appendix and the fact that  $\mathbf{a} \in \ell^\alpha \log \ell$ , we conclude that

$$p_{n,m}(\mathbf{a}; \epsilon) \leq \text{const} \sum_{k=m+1}^{\infty} |a_k|^\alpha \left[ 1 + \log^+ \frac{1}{|a_k|} \right] \rightarrow 0 \quad \text{as } m \rightarrow \infty.$$

(The constant on the right-hand side depends on  $\epsilon$ .) This proves the theorem for  $(X_n)$ ; the convergence of  $(\tilde{X}_n)$  can be shown analogously by utilizing (3.2).  $\square$

As an immediate corollary of Theorem 3.2, we obtain the following result which solves the problem of finding the limits of the finite-dimensional distributions for the integrated periodogram  $J_{n,\varepsilon}$  in (2.2) and its self-normalized version  $\tilde{J}_{n,\varepsilon}$  in (2.4).

**Corollary 3.3.** *Let  $\alpha \in (0, 2)$  and*

$$\mathcal{F} = \{f \in L^2[0, \pi]: \mathbf{a}(f) = (a_1(f), a_2(f), \dots) \in \ell^\alpha \log \ell\},$$

where  $\mathbf{a}(f)$  is specified in (2.3). We then have

$$\begin{aligned} n(n \log n)^{-1/\alpha} [J_{n,\varepsilon}(f) - a_0(f) \gamma_{n,\varepsilon}(0)]_{f \in \mathcal{F}} &\xrightarrow{\text{fidi}} 2[Y(\mathbf{a}(f))]_{f \in \mathcal{F}}, \\ (n/\log n)^{1/\alpha} [\tilde{J}_{n,\varepsilon}(f) - a_0(f)]_{f \in \mathcal{F}} &\xrightarrow{\text{fidi}} 2[\tilde{Y}(\mathbf{a}(f))]_{f \in \mathcal{F}}. \end{aligned}$$

**Remark 3.4.** The condition  $\mathbf{a}(f) \in \ell^\alpha \log \ell$  is, in general, not easily verified. However, if  $f$  represents the spectral density of a stationary process  $(X_n)$  with absolutely summable autocovariance function  $\gamma_X$ , then, up to a constant multiple,  $f$  is represented by the Fourier series of  $\gamma_X$ , and the rate of decay of  $\gamma_X(h) \rightarrow 0$  as  $h \rightarrow \infty$  is well known for numerous time series models. For example, if  $f$  is the spectral density of an ARMA process,  $\gamma_X(h) \rightarrow 0$  at an exponential rate and then  $\mathbf{a}(f) \in \ell^\alpha \log \ell$  is satisfied for every  $\alpha > 0$ .

Conditions ensuring that  $\mathbf{a}(f) \in \ell^\alpha$  can be found in the literature on Fourier series, for example, in [39]. Theorem (3.10) on page 243 in Volume I of that reference yields, for Lipschitz continuous functions  $f$  with exponent  $\beta \in (0, 1]$ , that  $\mathbf{a}(f) \in \ell^\alpha$  for  $\alpha > 2/(2\beta + 1)$ , but not necessarily for  $\alpha = 2/(2\beta + 1)$ . This means, in particular, that Lipschitz continuous functions do not necessarily satisfy  $\mathbf{a}(f) \in \ell^\alpha$  for small values  $\alpha < 1$ . Zygmund's Theorem (3.13), [39], page 243, Volume I, states that  $\mathbf{a}(f) \in \ell^\alpha$  if  $f$  is of bounded variation and Lipschitz continuous with exponent  $\beta \in (0, 1]$  such that  $\alpha > 2/(2 + \beta)$ , but this statement is not necessarily valid for  $\alpha = 2/(2 + \beta)$ .

We also note that  $\mathbf{a}(f) \notin \ell^\alpha$  for  $f(x) = I_{[0,x]}$ ,  $x \in (0, \pi]$  and  $\alpha \leq 1$ . Indeed, then  $a_k(f) = k^{-1} \sin(xk)$ ,  $k = 1, 2, \dots$  and  $\sum_k |a_k(f)|^\alpha = \infty$ . The latter condition implies that the series  $Y(\mathbf{a}(f))$  diverges a.s. by the 3-series theorem and the 0–1 law. Hence, Corollary 3.3 does not apply to the important class of indicator functions when  $\alpha < 1$ . Moreover,  $(J_{n,\varepsilon}(f))$  is not tight. Indeed, it follows from the argument above and from [22], Theorem 6.2.1 that for some  $\delta > 0$ , for every  $K > 0$ ,

$$\begin{aligned} \delta &\leq \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} P \left( (n \log n)^{-1/\alpha} \left| \sum_{k=1}^m a_k(f) [n \gamma_{n,\varepsilon}(k)] \right| > K \right) \\ &\leq \text{const} \liminf_{n \rightarrow \infty} P \left( (n \log n)^{-1/\alpha} |J_{n,\varepsilon}(f) - a_0(f) \gamma_{n,\varepsilon}(0)| > K \right). \end{aligned}$$

### 3.2. Tightness and weak convergence in the case $\alpha \in (0, 1)$

In order to derive a full weak convergence counterpart of the convergence in terms of the finite-dimensional distributions in Corollary 3.3, it remains to establish tightness of the corresponding family of laws. We start, once again, in the more general context of random fields indexed by sequences in  $\ell^\alpha \log \ell$ . Since we are dealing with the weak convergence of infinite-dimensional objects, we may expect difficulties which are due to the geometric properties of the underlying path spaces. It is also not completely surprising that the case  $\alpha \in (0, 1)$  is the ‘better one’ in comparison with  $\alpha \in [1, 2)$ ; see, for example, the results on boundedness, continuity and oscillations of  $\alpha$ -stable processes in [35], Chapter 10. Note, however, that the constraint  $\mathbf{a}(f) \in \ell^\alpha \log \ell$  is harder to satisfy for smaller  $\alpha$  than for larger  $\alpha$ ; see also Remark 3.4.

In the present case  $\alpha \in (0, 1)$ , we introduce the function

$$h(x) = \begin{cases} |x|^\alpha \log(b + |x|^{-1}), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

where  $b$  is chosen so large that  $h$  is concave on  $(0, \infty)$ . Note that

$$\ell^\alpha \log \ell = \left\{ \mathbf{a}: \sum_{k=1}^{\infty} h(a_k) < \infty \right\}$$

and this set is a linear metric space when endowed with the metric  $d(\mathbf{a}, \mathbf{b}) = \sum_{k=1}^{\infty} h(a_k - b_k)$ .

Assume that  $\mathcal{A}$  is a compact set of  $\ell^\alpha \log \ell$  with the additional property that

$$\sum_{k=1}^{\infty} \sup_{\mathbf{a} \in \mathcal{A}} h(a_k) < \infty. \quad (3.4)$$

Observe that  $\mathcal{A}$  is then also a compact subset of  $\ell^\alpha$  and  $(Y(\mathbf{a}))_{\mathbf{a} \in \mathcal{A}}$  is sample-continuous as a random element with values in  $\mathbb{C}(\mathcal{A})$ , the space of continuous functions on  $\mathcal{A}$  equipped with the uniform topology; see [35], Section 10.4. The following is our main result on the weak convergence of the sequences  $(X_n)$  and  $(\tilde{X}_n)$  of infinite variance random quadratic forms in the case  $\alpha \in (0, 1)$ .



**Theorem 3.5.** Assume  $\alpha \in (0, 1)$ . For a compact set  $\mathcal{A}$  of  $\ell^\alpha \log \ell$  satisfying (3.4), the following weak convergence result holds in  $\mathbb{C}(\mathcal{A})$ :

$$(X_n(\mathbf{a}))_{\mathbf{a} \in \mathcal{A}} \Longrightarrow (Y(\mathbf{a}))_{\mathbf{a} \in \mathcal{A}} \quad \text{and} \quad (\tilde{X}_n(\mathbf{a}))_{\mathbf{a} \in \mathcal{A}} \Longrightarrow (\tilde{Y}(\mathbf{a}))_{\mathbf{a} \in \mathcal{A}},$$

where  $X_n$ ,  $\tilde{X}_n$ ,  $Y$  and  $\tilde{Y}$  are defined in (3.3) and the processes  $Y$  and  $\tilde{Y}$  are sample-continuous.

**Proof.** We restrict ourselves to showing that  $X_n \Longrightarrow Y$ . In view of Theorem 3.2, it suffices to prove the tightness of the processes  $X_n$  in  $(\mathbb{C}(\mathcal{A}), d_{\mathcal{A}})$ , where  $d_{\mathcal{A}}$  is the restriction of  $d$  to  $\mathcal{A}$ . We have, for positive  $\epsilon$  and  $\delta$ ,

$$\begin{aligned} & P\left(\sup_{d_{\mathcal{A}}(\mathbf{a}, \mathbf{b}) < \delta} |X_n(\mathbf{a}) - X_n(\mathbf{b})| > \epsilon\right) \\ & \leq P\left(\sum_{k=1}^{n-1} \sup_{d_{\mathcal{A}}(\mathbf{a}, \mathbf{b}) < \delta} |a_k - b_k| [n\gamma_{n, |\epsilon|}(k)] > \epsilon(n \log n)^{1/\alpha}\right) = P_n(\epsilon, \delta). \end{aligned} \quad (3.5)$$

We want to show that  $P_n(\epsilon, \delta)$  can be made arbitrarily small for all  $n$ , provided that  $\delta$  is small. We solve this problem in a modified form: let  $\mathbf{C} = (C_0, C_{s,t}, s, t = 1, 2, \dots)$  be a sequence of i.i.d.  $S_1(1, 0, 0)$  random variables, independent of  $(\varepsilon_t)$ , and  $(b_{s,t})$  a double array of real numbers. We then have

$$C_0 \sum_{1 \leq s < t \leq n} |b_{s,t}| |\varepsilon_s \varepsilon_t| \stackrel{d}{=} \sum_{1 \leq s < t \leq n} b_{s,t} C_{s,t} |\varepsilon_s \varepsilon_t| \stackrel{d}{=} \sum_{1 \leq s < t \leq n} b_{s,t} C_{s,t} \varepsilon_s \varepsilon_t.$$

By virtue of this argument, it suffices to replace the products  $|\varepsilon_t \varepsilon_s|$  in the quadratic form in (3.5) with the products  $C_{s,t} \varepsilon_t \varepsilon_s$ . This means that it suffices to show that

$$P'_n(\epsilon, \delta) = P\left(\sum_{k=1}^{n-1} c_k(\delta) \sum_{j=1}^{n-k} C_{j, j+k} \varepsilon_j \varepsilon_{j+k} > \epsilon(n \log n)^{1/\alpha}\right)$$

can be made arbitrarily small for all  $n$ , provided that  $\delta$  is small, where  $c_k(\delta) = \sup_{d_{\mathcal{A}}(\mathbf{a}, \mathbf{b}) < \delta} |a_k - b_k|$ . Now, first apply Lemma A.2 to the  $P'_n$ 's and then condition (3.4):

$$\begin{aligned} P'_n(\epsilon, \delta) & \leq \text{const} \frac{1 + \log^+ |\epsilon|}{|\epsilon|^\alpha} \frac{1 + \log n}{n \log n} \sum_{i=1}^{n-1} \sum_{k=i+1}^n |c_k(\delta)|^\alpha \left(1 + \log^+ \frac{1}{|c_k(\delta)|}\right) \\ & \leq \text{const} \sum_{k=1}^{\infty} h(c_k(\delta)) \rightarrow 0 \quad \text{as } \delta \rightarrow 0. \end{aligned} \quad \square$$

Theorem 3.5 provides the limit process for a very general class of random quadratic forms with infinite first moments. The coefficient matrices of these quadratic forms are given by Toeplitz matrices. The conditions on the parameter set  $\mathcal{A}$  are nothing but restrictions on the infinite Toeplitz matrices  $(a_{i-j})_{i,j=1,2,\dots}$ . When specified to the particular case of Fourier coefficients, as in (2.3), Theorem 3.5 yields the following.

**Corollary 3.6.** Assume that  $\alpha \in (0, 1)$  and let

$$\mathcal{F} = \{f \in L^2[0, \pi]: \mathbf{a}(f) = (a_1(f), a_2(f), \dots) \in \mathcal{A}\},$$

where  $\mathcal{A}$  is a compact set of  $\ell^\alpha \log \ell$  satisfying (3.4) and  $\mathbf{a}(f)$  is specified in (2.3). We then have

$$\begin{cases} n(n \log n)^{-1/\alpha} [J_{n,\varepsilon}(f) - a_0(f)\gamma_{n,\varepsilon}(0)]_{f \in \mathcal{F}} \implies 2[Y(\mathbf{a}(f))]_{f \in \mathcal{F}}, \\ (n/\log n)^{1/\alpha} [\tilde{J}_{n,\varepsilon}(f) - a_0(f)]_{f \in \mathcal{F}} \implies 2[\tilde{Y}(\mathbf{a}(f))]_{f \in \mathcal{F}}, \end{cases} \quad (3.6)$$

where the convergence holds in  $\mathbb{C}(\mathcal{F})$ .

**Proof.** Let  $T: \mathcal{F} \rightarrow \mathcal{A}$  be defined by  $Tf = \mathbf{a}(f)$ . We claim that  $T\mathcal{F} \subset \mathcal{A}$  is closed, hence compact. Indeed, if  $(f_n) \subset \mathcal{F}$  is such that  $Tf_n$  converges in  $\ell^\alpha \log \ell$  to some point  $\mathbf{a} \in \mathcal{A}$ , then (as  $0 < \alpha < 1$ ), the sequence

$$f_n(\lambda) = \frac{1}{\pi} \sum_{j=-\infty}^{\infty} a_{|j|}(f_n) \cos j\lambda, \quad \lambda \in [0, \pi],$$

$n = 1, 2, \dots$  converges in  $L^1[0, \pi]$  to some function  $f$  that must necessarily be in  $\mathcal{F}$ . Therefore,  $\mathbf{a} = Tf \in T\mathcal{F}$ , and the latter set is compact. The above argument shows that  $L^2[0, \pi]$  convergence in  $\mathcal{F}$  is equivalent to  $\ell^\alpha \log \ell$  convergence in  $T\mathcal{F}$ . Since Theorem 3.5 implies weak convergence of the left-hand side of (3.6) to its right-hand side in  $\mathbb{C}(\mathcal{A})$  (when each function  $f \in \mathcal{F}$  is identified with  $Tf \in \mathcal{A}$ ), we conclude that weak convergence in (3.6) also holds in  $\mathbb{C}(\mathcal{F})$ .  $\square$

### 3.3. Tightness and weak convergence in the case $\alpha \in [1, 2)$

Establishing full weak convergence in the case  $\alpha \in [1, 2)$  is more difficult than in the case  $\alpha \in (0, 1)$ . Indeed, for  $\alpha \in (0, 1)$ , we were allowed to switch from the random variables  $\varepsilon_t$  to their absolute values, due to the specific geometry of the spaces  $\ell^\alpha$  and, in particular,  $\ell^\alpha \log \ell$ . The spaces  $\ell^\alpha$ ,  $\alpha \in [1, 2)$ , have a much more complicated structure and, therefore, the particular geometry of these spaces will be need to be invoked in proving tightness for the random quadratic forms  $X_n$  and  $\tilde{X}_n$ . The requirements prescribed by the geometry are usually given by entropy conditions; see [23] for a general treatment of random elements with values in Banach spaces. Entropy conditions are typically needed when  $\alpha$ -stable processes with  $\alpha \in [1, 2)$  appear; see the discussion in [35], Chapter 12.

In this section, we only consider vectors  $\mathbf{a} \in \ell^\alpha \log \ell$  of the form (2.3), that is, they are the Fourier coefficients of some functions  $f$ . Corollary 3.3 determines the structures of the limit processes of the quadratic forms  $J_{n,\varepsilon}$  via the convergence of their finite-dimensional distributions. It hence suffices to show the tightness in  $\mathbb{C}(\mathcal{F})$  for suitable classes  $\mathcal{F}$ . [20] considered the special case of the one-dimensional class  $\mathcal{F}_I$  of indicator functions on  $[0, \pi]$ . We extend their approach to more general classes of functions, using an entropy condition.

For  $f, g \in \mathcal{F}$ , let

$$d_j(f, g) = j|a_j(f) - a_j(g)|, \quad j \geq 1.$$

Each  $d_j$  defines a pseudo-metric on  $\mathcal{F}$ . Let

$$\rho_k(f, g) = \max_{2^k \leq j < 2^{k+1}} d_j(f, g), \quad k \geq 0.$$

Recall that the  $\epsilon$ -covering number  $N(\epsilon, \mathcal{F}, \rho_k)$  of  $(\mathcal{F}, \rho_k)$  is the minimal integer  $m$  for which we can find functions  $f_1, \dots, f_m \in \mathcal{F}$  such that  $\sup_{f \in \mathcal{F}} \min_{i=1, \dots, m} \rho_k(f, f_i) < \epsilon$ .

**Theorem 3.7.** Assume that  $\alpha \in [1, 2)$ , define  $\mathbf{a}(f)$  as in (2.3) and let  $\mathcal{F}$  be a subset of  $L^2[0, \pi]$  satisfying:

- (i)  $\mathbf{a}(f) \in \ell^\alpha \log \ell$  for all  $f \in \mathcal{F}$ ;
- (ii)  $\exists \beta \in (0, \alpha)$  such that

$$N(\epsilon, \mathcal{F}, \rho_k) \leq \text{const} [1 + (2^k / \epsilon)^\beta], \quad \epsilon > 0, k \geq 0. \quad (3.7)$$

The weak convergence result (3.6) then holds in  $\mathbb{C}(\mathcal{F})$ .

In contrast to the finite variance case [8, 28], the entropy condition (3.7) is a rather strong one. Indeed, in the papers mentioned, integrability of some power of  $\log N(\epsilon)$  in a neighborhood of the origin suffices. However, conditions such as (3.7) are common in problems of continuity and boundedness for stable processes; see [35], Chapter 10.

**Proof of Theorem 3.7.** The convergence of the finite-dimensional distributions follows from Theorem 3.2. We restrict ourselves to proving tightness for  $J_{n, \epsilon}$ , which follows by proving that

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left( \left\| \sum_{j=m}^n a_j \widehat{\gamma}_{n, \epsilon}(j) \right\|_{\mathcal{F}} > \epsilon \right) = 0 \quad \text{for every } \epsilon > 0, \quad (3.8)$$

where  $\|g\|_{\mathcal{F}} = \sup_{f \in \mathcal{F}} |g(f)|$  and

$$\widehat{\gamma}_{n, \epsilon}(j) = (n \log n)^{-1/\alpha} [n \gamma_{n, \epsilon}(j)], \quad j = 1, 2, \dots$$

As in [20], (6.4), one can argue that it suffices in (3.8) to consider  $m$  and  $n$  of some specific form. Let  $a < b$  be two positive integers,  $m = 2^a$  and  $n = 2^{b+1} - 1$ , and consider numbers  $\epsilon_k = 2^{-k\theta}$ ,  $k \geq 1$ , with  $\theta > 0$ . For  $\theta$  sufficiently small and  $a$  large enough, we have

$$P \left( \left\| \sum_{j=m}^n a_j \widehat{\gamma}_{n, \epsilon}(j) \right\|_{\mathcal{F}} > \epsilon \right) \leq \sum_{k=a}^b P \left( \left\| \sum_{j=2^k}^{2^{k+1}-1} a_j \widehat{\gamma}_{n, \epsilon}(j) \right\|_{\mathcal{F}} > \epsilon_k \right) = \sum_{k=a}^b p_k. \quad (3.9)$$

Consider an array  $(\epsilon_{k,l})$  of positive numbers such that  $\epsilon_{k,l} \rightarrow 0$  as  $l \rightarrow \infty$  for each  $k \geq 0$ . Then,

$$p_k \leq N(\epsilon_{k,0}, \mathcal{F}, \rho_k) p_{k,0} + \sum_{l=1}^{\infty} N(\epsilon_{k,l}, \mathcal{F}, \rho_k) p_{k,l},$$

where

$$p_{k,0} = \sup_{f \in \mathcal{F}} P \left( \left| \sum_{j=2^k}^{2^{k+1}-1} a_j(f) \widehat{\gamma}_{n,\varepsilon}(j) \right| > \epsilon_k/2 \right),$$

$$p_{k,l} = \sup_{f, g \in \mathcal{F}, \rho_k(f, g) \leq \epsilon_{k,l-1}} P \left( \left| \sum_{j=2^k}^{2^{k+1}-1} [a_j(f) - a_j(g)] \widehat{\gamma}_{n,\varepsilon}(j) \right| > 2^{-(l+1)} \epsilon_k \right).$$

By virtue of Lemma A.1, we have, for all  $f, g \in \mathcal{F}$ ,

$$P \left( \left| \sum_{j=2^k}^{2^{k+1}-1} [a_j(f) - a_j(g)] \widehat{\gamma}_{n,\varepsilon}(j) \right| > 2^{-(l+1)} \epsilon_k \right) \leq \text{const } b_{k,l},$$

where

$$b_{k,l} = \epsilon_k^{-\alpha} 2^{\alpha l} \sum_{j=2^k}^{2^{k+1}-1} |a_j(f) - a_j(g)|^\alpha [1 + \log^+(1/|a_j(f) - a_j(g)|)].$$

Assuming that  $\rho_k(f, g) \leq \epsilon_{k,l-1}$ , we have

$$\begin{aligned} b_{k,l} &\leq \text{const } \epsilon_k^{-\alpha} 2^{\alpha l} \epsilon_{k,l-1}^\alpha \sum_{j=2^k}^{2^{k+1}-1} j^{-\alpha} [1 + \log j \log^+ \epsilon_{k,l-1}^{-1}] \\ &\leq \text{const } \epsilon_k^{-\alpha} 2^{\alpha l} \epsilon_{k,l-1}^\alpha 2^{-k(\alpha-1)} [1 + k \log^+ \epsilon_{k,l-1}^{-1}]. \end{aligned}$$

Hence, we are left to consider

$$\begin{aligned} &\sum_{k=a}^b \sum_{l=1}^{\infty} N(\epsilon_{k,l}, \mathcal{F}, \rho_k) \epsilon_k^{-\alpha} 2^{-k(\alpha-1)+\alpha l} \epsilon_{k,l-1}^\alpha [1 + k \log^+ \epsilon_{k,l-1}^{-1}] \\ &= \sum_{k=a}^b 2^{-k(\alpha-1-\alpha\theta)} \sum_{l=1}^{\infty} N(\epsilon_{k,l}, \mathcal{F}, \rho_k) \epsilon_{k,l-1}^\alpha [1 + k \log^+ \epsilon_{k,l-1}^{-1}] 2^{\alpha l} \\ &\leq \text{const } \sum_{k=a}^b 2^{-k(\alpha-1-\alpha\theta)} \sum_{l=1}^{\infty} \left[ 1 + \left( \frac{2^k}{\epsilon_{k,l}} \right)^\beta \right] \epsilon_{k,l-1}^\alpha [1 + \log^+ \epsilon_{k,l-1}^{-1}] 2^{\alpha l}. \end{aligned} \quad (3.10)$$

Assume that  $\theta$  is so small that (3.9) holds. Define the numbers  $\epsilon_{k,l} = 2^{-\gamma_1 l - \gamma_2 k}$ ,  $k, l \geq 0$  with  $\gamma_1, \gamma_2 > 0$  such that  $1 + \gamma_2 > (1 + \alpha\theta)/(\alpha - \beta)$  and  $\gamma_1 > \alpha/(\alpha - \beta)$ . For these parameter choices, it is not difficult to see that (3.10) converges to zero by first letting  $n \rightarrow \infty$  (i.e.,  $b \rightarrow \infty$ ) and then  $m \rightarrow \infty$  (i.e.,  $a \rightarrow \infty$ ). This proves (3.8), hence the tightness of the processes considered in  $\mathbb{C}(\mathcal{F})$ .  $\square$

In what follows, we give examples of function spaces  $\mathcal{F}$  satisfying condition (ii) of Theorem 3.7.

**Example 3.8.** Consider a space of indexed functions  $\mathcal{G}_\Theta = \{g_\theta: \theta \in \Theta\}$  that are defined on  $[0, \pi]$  such that  $(\Theta, \tau)$  is a compact metric space, the mapping  $\theta \mapsto g_\theta$  is Hölder continuous with exponent  $b > 0$  and constant  $K > 0$ , that is,

$$\sup_{0 \leq x \leq \pi} |g_{\theta_1}(x) - g_{\theta_2}(x)| \leq K(\tau(\theta_1, \theta_2))^b \quad \text{for all } \theta_1, \theta_2 \in \Theta,$$

and the number of balls (in metric  $\tau$ ) of radius at most  $\epsilon$  necessary to cover  $\Theta$  is of the order  $\epsilon^{-a}$  for some  $0 < a < b\alpha$ . Then,  $\mathcal{G}_\Theta$  satisfies  $N(\epsilon, \mathcal{G}_\Theta, \rho_k) \leq \text{const}(2^k/\epsilon)^{a/b}$  with  $a/b \in (0, \alpha)$ . Indeed, let  $\epsilon > 0, k \geq 0$ . We can find  $N \leq c((K\pi 2^{k+1})/\epsilon)^{a/b}$  balls of radius at most  $(\epsilon/(K\pi 2^{k+1}))^{1/b}$  covering  $\Theta$ . Call them  $B_1, \dots, B_N$ , with centers  $\theta_1, \dots, \theta_N$ . Now, given  $\theta \in \Theta$ , we have  $\theta \in B_i$  for some  $i \in \{1, \dots, N\}$  and

$$\begin{aligned} \rho_k(g_\theta, g_{\theta_i}) &= \max_{2^k \leq j < 2^{k+1}} j \left| \int_0^\pi \cos(jx) (g_\theta(x) - g_{\theta_i}(x)) dx \right| \\ &\leq 2^{k+1} \pi \sup_{0 \leq x \leq \pi} |g_\theta(x) - g_{\theta_i}(x)| \leq 2^{k+1} \pi K \tau(\theta, \theta_i)^b \leq \epsilon. \end{aligned}$$

The desired bound now follows since  $N(\epsilon, \mathcal{G}_\Theta, \rho_k) \leq N \leq \text{const}(2^k/\epsilon)^{a/b}$ .

**Example 3.9.** Consider a Vapnik–Červonenkis (VC) class  $\mathcal{G}$  of functions defined on  $[0, \pi]$  with VC index  $V(\mathcal{G}) = 2$ ; see [36], Section 2.6.2 for more information on VC classes of functions. Given  $\epsilon > 0$  and  $k \geq 0$ , we can find  $N \leq c(\pi 2^{k+1})/\epsilon$  balls of radius at most  $\epsilon/(\pi 2^{k+1})$  that cover  $\mathcal{G}$  in the norm  $\frac{1}{\pi} \int_0^\pi |\cdot| dx$ ; see, for example, [36], Theorem 2.6.7. Therefore, there exist  $g_1, \dots, g_N \in \mathcal{G}$  such that for any  $g \in \mathcal{G}$ ,

$$\min_{1 \leq i \leq N} \frac{1}{\pi} \int_0^\pi |g(x) - g_i(x)| dx < \frac{\epsilon}{\pi 2^{k+1}}.$$

We then have

$$\begin{aligned} \min_{1 \leq i \leq N} \rho_k(g, g_i) &= \min_{1 \leq i \leq N} \max_{2^k \leq j < 2^{k+1}} j \left| \int_0^\pi \cos(jx) (g(x) - g_i(x)) dx \right| \\ &\leq \min_{1 \leq i \leq N} 2^{k+1} \int_0^\pi |g(x) - g_i(x)| dx \leq \epsilon. \end{aligned}$$

It follows that  $N(\epsilon, \mathcal{G}, \rho_k) \leq N \leq \text{const} 2^k/\epsilon$ .

## 4. The linear process case

It is the aim of this section to show that the results for the case of an i.i.d. sequence  $(\varepsilon_t)$  translate to the linear process case. The following decomposition will be crucial:

$$I_{n,X}(\lambda) = I_{n,\varepsilon}(\lambda) |\psi(e^{-i\lambda})|^2 + R_n(\lambda). \quad (4.1)$$

This decomposition is analogous to the decomposition  $f_X(\lambda) = f_\varepsilon(\lambda) |\psi(e^{-i\lambda})|^2$  of the spectral density  $f_X$  of a linear process. We will show that the normalized integrated remainder term  $\int_0^\pi R_n(\lambda) f(\lambda) d\lambda$  is negligible, uniformly over the class of functions  $\mathcal{F}$ , in comparison to the normalized main part

$$\int_0^\pi I_{n,\varepsilon}(\lambda) |\psi(e^{-i\lambda})|^2 f(\lambda) d\lambda, \quad f \in \mathcal{F},$$

which can be treated by the methods of the previous section. Note that, for a given sequence of coefficients  $(\psi_j)_{j \in \mathbb{Z}}$ , the functions  $|\psi(e^{-i\cdot})|^2 f$  constitute just another class of functions on  $[0, \pi]$ ,  $\mathcal{F}_\psi$  say, and therefore we will study the process  $J_{n,\varepsilon}(f)$ ,  $f \in \mathcal{F}_\psi$ , for suitable classes  $\mathcal{F}_\psi$ .

**Lemma 4.1.** *Let  $R_n$  be the remainder term appearing in the decomposition (4.1) of the periodogram  $I_{n,X}$ . Suppose that the linear filter  $(\psi_j)$  of the process  $X$  satisfies*

$$\sum_{j=-\infty}^{\infty} |\psi_j| |j|^{2/\alpha} (1 + \log^+ |j|)^{(4-\alpha)/(2\alpha)+\tau} < \infty \quad (4.2)$$

for some  $\tau > 0$  and  $\mathcal{F}$  is a collection of real-valued functions defined on  $[0, \pi]$  such that  $\sup_{f \in \mathcal{F}} \|f\|_2 < \infty$ . We then have

$$\frac{n}{(n \log n)^{1/\alpha}} \sup_{f \in \mathcal{F}} \left| \int_0^\pi f(x) R_n(x) dx \right| \xrightarrow{P} 0.$$

**Proof.** From [27], Proposition 5.1, substituting  $n^{1/2}$  for  $a_n$ , we have the following decomposition for  $R_n$ :

$$R_n(x) = n^{-1} (\psi(e^{ix}) L_n(x) K_n(-x) + \psi(e^{-ix}) L_n(-x) K_n(x) + |K_n(x)|^2), \quad (4.3)$$

where  $\psi$  is the transfer function as defined before and

$$\begin{aligned} L_n(x) &= \sum_{t=1}^n \varepsilon_t e^{-ixt}, & K_n(x) &= \sum_{j=-\infty}^{\infty} \psi_j e^{-ixj} U_{nj}(x), \\ U_{nj}(x) &= \left( \sum_{t=1-j}^{n-j} - \sum_{t=1}^n \right) \varepsilon_t e^{-ixt}. \end{aligned}$$

We first show that

$$\frac{1}{(n \log n)^{1/\alpha}} \sup_{f \in \mathcal{F}} \left| \int_0^\pi f(x) |K_n(x)|^2 dx \right| \xrightarrow{P} 0. \quad (4.4)$$

Note that

$$\begin{aligned} \left| \int_0^\pi f(x) |K_n(x)|^2 dx \right| &\leq \int_0^\pi |f(x)| \left( \sum_{j=-\infty}^\infty |\psi_j| |U_{nj}(x)| \right)^2 dx \\ &\leq \text{const} \left( \sum_{j=-\infty}^{-1} + \sum_{j=1}^\infty \right) |\psi_j| \int_0^\pi |f(x)| |U_{nj}(x)|^2 dx. \end{aligned}$$

The convergence in (4.4) will follow if we can show that the suprema over  $f \in \mathcal{F}$  of the two infinite sums in the last expression are bounded in probability as  $n \rightarrow \infty$ . We will prove this for the second sum; the first one can be handled analogously.

We have, by definition of the terms  $U_{nj}(x)$ , the Cauchy–Schwarz inequality and the fact that, by assumption,  $\sup_{f \in \mathcal{F}} \|f\|_2 < \infty$ ,

$$\begin{aligned} &\sup_{f \in \mathcal{F}} \sum_{j=1}^\infty |\psi_j| \int_0^\pi |f(x)| |U_{nj}(x)|^2 dx \\ &\leq \sup_{f \in \mathcal{F}} \sum_{j=1}^n |\psi_j| \int_0^\pi |f(x)| \left| \sum_{t=1-j}^0 \varepsilon_t e^{-ixt} - \sum_{t=n-j+1}^n \varepsilon_t e^{-ixt} \right|^2 dx \\ &\quad + \sup_{f \in \mathcal{F}} \sum_{j=n+1}^\infty |\psi_j| \int_0^\pi |f(x)| \left| \sum_{t=1-j}^{n-j} \varepsilon_t e^{-ixt} - \sum_{t=1}^n \varepsilon_t e^{-ixt} \right|^2 dx \\ &\leq c[I_1(n) + I_2(n) + I_3(n) + I_4(n)], \end{aligned}$$

where

$$\begin{aligned} I_1(n) &= \sum_{j=1}^n |\psi_j| \left( \int_0^\pi \left| \sum_{t=1-j}^0 \varepsilon_t e^{-ixt} \right|^4 dx \right)^{1/2}, \\ I_2(n) &= \sum_{j=1}^n |\psi_j| \left( \int_0^\pi \left| \sum_{t=n-j+1}^n \varepsilon_t e^{-ixt} \right|^4 dx \right)^{1/2}, \\ I_3(n) &= \sum_{j=n+1}^\infty |\psi_j| \left( \int_0^\pi \left| \sum_{t=1-j}^{n-j} \varepsilon_t e^{-ixt} \right|^4 dx \right)^{1/2}, \\ I_4(n) &= \sum_{j=n+1}^\infty |\psi_j| \left( \int_0^\pi \left| \sum_{t=1}^n \varepsilon_t e^{-ixt} \right|^4 dx \right)^{1/2}. \end{aligned}$$

It remains to show that each sequence  $I_k(n)$ ,  $k = 1, 2, 3, 4$ , is tight. Now,

$$I_1(n) \stackrel{d}{=} \sum_{j=1}^n |\psi_j| \left( \int_0^\pi \left| \sum_{m=1}^j \varepsilon_m e^{ixm} \right|^4 dx \right)^{1/2}.$$

Let  $\epsilon > 0$ . Choose  $M > 0$  so large that the following holds, for  $\delta = \frac{2\alpha}{4-\alpha}\tau$ :

$$P(|\varepsilon_m| > Mm^{1/\alpha}(1 + \log m)^{1/\alpha+\delta} \text{ for some } m \geq 1) \leq \epsilon/2.$$

Write

$$J_m = \varepsilon_m I_{\{|\varepsilon_m| \leq Mm^{1/\alpha}(1+\log m)^{1/\alpha+\delta}\}}.$$

Then, for  $k > 0$ , we have, for  $\delta$  chosen as above,

$$\begin{aligned} & P(I_1(n) > k) - \epsilon/2 \\ & \leq P\left(\sum_{j=1}^n |\psi_j| \left( \int_0^\pi \left| \sum_{m=1}^j \varepsilon_m e^{ix} \right|^4 dx \right)^{1/2} \left( \int_0^\pi \left| \sum_{m=1}^j J_m e^{ix} \right|^4 dx \right)^{1/2} > k\right) \\ & \leq k^{-1} \sum_{j=1}^n |\psi_j| \left( \int_0^\pi E \left| \sum_{m=1}^j J_m e^{ix} \right|^4 dx \right)^{1/2} \\ & = k^{-1} \sum_{j=1}^n |\psi_j| \left[ \int_0^\pi \left( \sum_{m=1}^j E(J_m^4) \right. \right. \\ & \quad \left. \left. + 6 \sum_{m_1=1}^j \sum_{m_2=m_1+1}^j E(J_{m_1}^2) E(J_{m_2}^2) \cos((m_1 - m_2)x) dx \right) \right]^{1/2} \\ & \leq \frac{c}{k} \sum_{j=1}^n |\psi_j| \left[ \left( \sum_{m=1}^j (m^{1/\alpha}(1 + \log m)^{1/\alpha+\delta})^{4-\alpha} \right)^{1/2} \right. \\ & \quad \left. + \left( \left[ \sum_{m=1}^j (m^{1/\alpha}(1 + \log m)^{1/\alpha+\delta})^{2-\alpha} \right]^2 \right)^{1/2} \right] \\ & \leq \frac{c}{k} \sum_{j=1}^\infty |\psi_j| j^{2/\alpha} (1 + \log j)^{(4-\alpha)/(2\alpha)+\tau}. \end{aligned}$$

By virtue of (4.2), the last expression can be made smaller than  $\epsilon/2$  by choosing  $k$  large enough, which proves the tightness of  $I_1(n)$ . Similar arguments show that  $I_j(n)$ ,  $j = 2, 3, 4$ , are tight sequences as well. The convergence in (4.4) follows.



By the decomposition (4.3), the proof will be finished if we can also establish that

$$\frac{1}{(n \log n)^{1/\alpha}} \sup_{f \in \mathcal{F}} \left| \int_0^\pi f(x) \psi(e^{ix}) L_n(x) K_n(-x) dx \right| \xrightarrow{P} 0. \quad (4.5)$$

We have, by the Cauchy–Schwarz inequality and the identity  $|L_n(x)|^2 = n I_{n,\varepsilon}(x)$ ,

$$\begin{aligned} & \left| \int_0^\pi f(x) \psi(e^{ix}) L_n(x) K_n(-x) dx \right| \\ & \leq c \|f\|_2 \left( \int_0^\pi |L_n(x) K_n(-x)|^2 dx \right)^{1/2} \\ & \leq c \|f\|_2 n^{1/2} \left( \sup_{0 \leq x \leq \pi} I_{n,\varepsilon}(x) \right)^{1/2} \left( \int_0^\pi |K_n(-x)|^2 dx \right)^{1/2}. \end{aligned}$$

We therefore see that

$$\begin{aligned} & \frac{1}{(n \log n)^{1/\alpha}} \sup_{f \in \mathcal{F}} \left| \int_0^\pi f(x) \psi(e^{ix}) L_n(x) K_n(-x) dx \right| \\ & \leq \frac{c}{n^{1/\alpha-1/2}} \frac{(\sup_{0 \leq x \leq \pi} I_{n,\varepsilon}(x))^{1/2}}{(\log n)^{1/\alpha}} \left( \int_0^\pi |K_n(-x)|^2 dx \right)^{1/2}. \end{aligned}$$

Similar arguments as for (4.4) ensure the tightness of the sequence  $\int_0^\pi |K_n(-x)|^2 dx$ . The tightness of the term

$$\frac{(\sup_{0 \leq x \leq \pi} I_{n,\varepsilon}(x))^{1/2}}{(\log n)^{1/\alpha}}$$

follows from [29] Theorem 2.1 (for  $0 < \alpha < 1$ ) and Proposition 3.1 (for  $1 \leq \alpha < 2$ ). Thus, we conclude that (4.5) holds, and Lemma 4.1 is proved.  $\square$

**Remark 4.2.** A referee kindly pointed out that Lemma 4.1 remains valid under the following condition, which is weaker than (4.2): assume that there exists a sequence  $(\omega_n)$  of positive numbers such that

$$(n \log n)^{-1/\alpha} \left( \sum_{k=1}^n \omega_k^{-\alpha} \right)^{2/\alpha-1/2} \sum_{|j| \leq n} |\psi_j| \left( \sum_{l=1}^j \omega_l^{4-\alpha} \right)^{1/2} \rightarrow 0.$$

Condition (4.2) follows by taking  $\omega_n = n^{1/\alpha} (1 + \log n)^{1/\alpha+\delta}$  for some positive  $\delta$ .

By (4.1), we may write, for each  $f$ ,

$$\begin{aligned} J_{n,X}(f) - a_0(f|\psi|^2) \gamma_{n,\varepsilon}(0) &= J_{n,\varepsilon}(f|\psi|^2) - a_0(f|\psi|^2) \gamma_{n,\varepsilon}(0) \\ &\quad + \int_0^\pi f(x) R_n(x) dx, \end{aligned}$$

where  $|\psi|^2$  stands for  $|\psi(e^{-i\cdot})|^2$ . Combining this decomposition with Lemma 4.1, we can now state the following analogs to Corollary 3.6 and Theorem 3.7.

**Corollary 4.3.** *Assume that  $\alpha \in (0, 1)$  or  $\alpha \in [1, 2)$  and let  $\mathcal{F}$  be as defined as in Corollary 3.6 or Theorem 3.7, respectively. Suppose that the set  $\{f: [0, \pi] \rightarrow \mathbb{R}: f|\psi|^2 \in \mathcal{F}\} = \mathcal{F}_\psi$  satisfies  $\sup_{f \in \mathcal{F}_\psi} \|f\|_2 < \infty$  and (4.2) holds for some  $\tau > 0$ . We then have*

$$\begin{cases} n(n \log n)^{-1/\alpha} [J_{n,X}(f) - a_0(f|\psi|^2)\gamma_{n,\varepsilon}(0)]_{f \in \mathcal{F}_\psi} \Longrightarrow 2[Y(\mathbf{a}(f|\psi|^2))]_{f \in \mathcal{F}_\psi}, \\ (n/\log n)^{1/\alpha} [\tilde{J}_{n,X}(f) - a_0(f|\psi|^2)]_{f \in \mathcal{F}_\psi} \Longrightarrow 2[\tilde{Y}(\mathbf{a}(f|\psi|^2))]_{f \in \mathcal{F}_\psi}, \end{cases} \quad (4.6)$$

where the convergence holds in  $\mathbb{C}(\mathcal{F}_\psi)$ .

## Appendix

For an array  $\mathbf{b} = (b_{s,t})$  of real numbers, define the quadratic forms

$$Q_{n,\varepsilon}(\mathbf{b}) = \sum_{1 \leq s \neq t \leq n} b_{s,t} \varepsilon_s \varepsilon_t$$

and

$$\Gamma_n(\mathbf{b}) = \sum_{1 \leq s \neq t \leq n} |b_{s,t}|^\alpha \left( 1 + \log^+ \frac{1}{|b_{s,t}|} \right).$$

The following lemma is a consequence of [34], Theorem 3.1; see also [22].

**Lemma A.1.** *For  $\alpha \in (0, 2)$ , there exists a positive constant  $D_\alpha$  such that for all  $x > 0$ ,*

$$P(Q_{n,\varepsilon}(\mathbf{b}) > x) \leq D_\alpha \frac{1 + \log^+ x}{x^\alpha} \Gamma_n(\mathbf{b}).$$

Now, let  $\mathbf{C} = (C_0, C_{s,t}, s, t = 1, 2, \dots)$  be a sequence of i.i.d.  $S_1(1, 0, 0)$  random variables, independent of  $(\varepsilon_t)$ , and let  $\mathbf{b}$  be as above. The following lemma is a consequence of Lemma A.1.

**Lemma A.2.** *For  $\alpha \in (0, 1)$ , there exists a positive constant  $D'_\alpha$  such that for all  $x > 0$ ,*

$$I_n(x) = P\left(\sum_{1 \leq s < t \leq n} b_{s,t} C_{s,t} \varepsilon_s \varepsilon_t > x\right) \leq D'_\alpha \frac{1 + \log^+ x}{x^\alpha} \Gamma_n(\mathbf{b}). \quad (\text{A.1})$$

**Proof.** Apply Lemma A.1 to  $I_n(x)$ , conditionally on  $\mathbf{C}$ :

$$\begin{aligned}
 I_n(x) &= P\left(\sum_{1 \leq s < t \leq n} b_{s,t} C_{s,t} \varepsilon_s \varepsilon_t > x\right) \\
 &= E_{\mathbf{C}} P\left(\sum_{1 \leq s < t \leq n} b_{s,t} C_{s,t} \varepsilon_s \varepsilon_t > x \mid \mathbf{C}\right) \\
 &\leq \text{const} \frac{1 + \log^+ x}{x^\alpha} \sum_{s=1}^n \sum_{t=i+1}^n |b_{s,t}|^\alpha E|C_0|^\alpha \left(1 + \log^+ \frac{1}{|b_{s,t} C_0|}\right),
 \end{aligned} \tag{A.2}$$

Because  $\alpha \in (0, 1)$ , we also have, for  $x > 0$ ,

$$E|C_0|^\alpha \left(1 + \log^+ \frac{1}{|x C_0|}\right) \leq \text{const} \left(1 + \log^+ \frac{1}{|x|}\right), \tag{A.3}$$

and combining (A.2) and (A.3), we thus obtain (A.1).  $\square$

## Acknowledgements

Thomas Mikosch's research was supported in part by the Danish Research Council (FNU) Grant 272-06-0442. Gennady Samorodnitsky's research was supported in part by the ARO Grant W911NF-07-1-0078 at Cornell University and a Villum Kann Rasmussen Visiting Professor Grant at the University of Copenhagen. The paper was written when Sami Umut Can and Gennady Samorodnitsky visited the Department of Mathematical Sciences at the University of Copenhagen in 2008/2009. They take pleasure in thanking the Department for hospitality and financial support. The authors also thank the anonymous referees for suggestions that helped to improve the paper.

## References

- [1] Adler, R.J., Feldman, R.E. and Taqqu, M.S. (eds.) (1998). *A Practical Guide to Heavy Tails*. Boston: Birkhäuser. [MR1652283](#)
- [2] Bartlett, M.S. (1954). Problemes de l'analyse spectrale des séries temporelles stationnaires. *Publ. Inst. Statist. Uni. Paris* **III-3** 119–134. [MR0067453](#)
- [3] Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley. [MR0233396](#)
- [4] Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods*, 2nd ed. New York: Springer. [MR1093459](#)
- [5] Can, S.U., Mikosch, T. and Samorodnitsky, G. (2009). Weak convergence of the function-indexed integrated periodogram for infinite variance processes. Extended technical report. Available at [www.math.ku.dk/~mikosch](http://www.math.ku.dk/~mikosch).
- [6] Crovella, M. and Bestavros, A. (1996). Self-similarity in world wide web traffic: Evidence and possible causes. In *Proceedings of the 1996 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems* **24** 160–169.

- [7] Crovella, M., Bestavros, A. and Taqqu, M.S. (1996). Heavy-tailed probability distributions in the world wide web. Preprint.
- [8] Dahlhaus, R. (1988). Empirical spectral processes and their applications to time series analysis. *Stochastic Process. Appl.* **30** 69–83. [MR0968166](#)
- [9] Dahlhaus, R. and Polonik, W. (2002). Empirical processes and nonparametric maximum likelihood estimation for time series. In *Empirical Process Techniques for Dependent Data* (H.G. Dehling, T. Mikosch and M. Sørensen, eds.) 275–298. Boston: Birkhäuser. [MR1958786](#)
- [10] Davis, R.A. and Mikosch, T. (2009). Extreme value theory for GARCH processes. In *The Handbook of Financial Time Series* (T.G. Andersen, R.A. Davis, J.-P. Kreiss and T. Mikosch, eds.) 187–200. Heidelberg: Springer.
- [11] Davis, R.A. and Mikosch, T. (2009). Extremes of stochastic volatility models. In *The Handbook of Financial Time Series* (T.G. Andersen, R.A. Davis, J.-P. Kreiss and T. Mikosch, eds.) 355–364. Heidelberg: Springer.
- [12] Davis, R.A. and Resnick, S.I. (1986). Limit theory for the sample covariance and correlation functions of moving averages. *Ann. Statist.* **14** 533–558. [MR0840513](#)
- [13] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Finance and Insurance*. Berlin: Springer. [MR1458613](#)
- [14] Faÿ, G., Gonzalez-Arevalo, B., Mikosch, T. and Samorodnitsky, G. (2006). Modeling teletraffic arrivals by a Poisson cluster process. *Qesta* **54** 121–140. [MR2268057](#)
- [15] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications* **2**, 2nd ed. New York: Wiley.
- [16] Grenander, U. and Rosenblatt, M. (1984). *Statistical Analysis of Stationary Time Series*, 2nd ed. New York: Chelsea. [MR0890514](#)
- [17] Kagan, Y.Y. (1997). Earthquake size distribution and earthquake insurance. *Comm. Statist. Stochastic Models* **13** 775–797.
- [18] Kokoszka, P. and Mikosch, T. (1997). The integrated periodogram for long-memory processes with finite or infinite variance. *Stochastic Process. Appl.* **66** 55–78. [MR1431870](#)
- [19] Kokoszka, P. and Taqqu, M.S. (1996). Parameter estimation for infinite variance fractional ARIMA. *Ann. Statist.* **24** 1880–1913. [MR1421153](#)
- [20] Klüppelberg, C. and Mikosch, T. (1996). The integrated periodogram for stable processes. *Ann. Statist.* **24** 1855–1879. [MR1421152](#)
- [21] Klüppelberg, C. and Mikosch, T. (1996). Self-normalized and randomly centred spectral estimates. In *Proceedings of the Athens International Conference on Applied Probability and Time Series*, vol. 2: *Time Series* (C.C. Heyde, Yu.V. Prokhorov, R. Pyke and S.T. Rachev, eds.) 259–271. Berlin: Springer. [MR1466751](#)
- [22] Kwapien, S. and Woyczyński, W.A. (1992). *Random Series and Stochastic Integrals: Single and Multiple*. Basel: Birkhäuser. [MR1167198](#)
- [23] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces: Isoperimetry and Processes*. Berlin: Springer. [MR1102015](#)
- [24] Leland, W.E., Willinger, W., Taqqu, M.S. and Wilson, D.V. (1993). On the self-similar nature of Ethernet traffic. *Computer Communications Review* **23** 183–193.
- [25] Mikosch, T. (1998). Periodogram estimates from heavy-tailed data. In *A Practical Guide to Heavy Tails: Statistical Techniques for Analysing Heavy-Tailed Distributions* (R.A. Adler, R. Feldman and M.S. Taqqu, eds.) 241–258. Boston: Birkhäuser. [MR1652289](#)
- [26] Mikosch, T. (2003). Modelling dependence and tails of financial time series. In *Extreme Values in Finance, Telecommunications and the Environment* (B. Finkenstädt and H. Rootzén, eds.) 185–286. London: Chapman & Hall.

- [27] Mikosch, T., Gadrich, T., Klüppelberg, C. and Adler, R.J. (1995). Parameter estimation for ARMA models with infinite variance innovations. *Ann. Statist.* **23** 305–326. [MR1331670](#)
- [28] Mikosch, T. and Norvaiša, R. (1997). Uniform convergence of the empirical spectral distribution function. *Stochastic Process. Appl.* **70** 85–114. [MR1472960](#)
- [29] Mikosch, T., Resnick, S. and Samorodnitsky, G. (2000). The maximum of the periodogram for a heavy-tailed sequence. *Ann. Appl. Probab.* **28** 885–908. [MR1782277](#)
- [30] Petrov, V.V. (1995). *Limit Theorems of Probability Theory*. Oxford: Oxford Univ. Press. [MR1353441](#)
- [31] Pollard, D. (1984). *Convergence of Stochastic Processes*. Berlin: Springer. [MR0762984](#)
- [32] Priestley, M. (1981). *Spectral Analysis and Time Series*, **I** and **II**. New York: Academic Press.
- [33] Resnick, S.I. (2006). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. New York: Springer. [MR2271424](#)
- [34] Rosiński, J. and Woyczyński, W.A. (1987). Multilinear forms in Pareto-like random variables and product random measures. *Colloq. Math.* **51** 303–313. [MR0891300](#)
- [35] Samorodnitsky, G. and Taqqu, M. (1994). *Stable Non-Gaussian Random Processes*. New York: Chapman & Hall. [MR1280932](#)
- [36] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes*. New York: Springer. [MR1385671](#)
- [37] Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis*. Uppsala: Almqvist & Wicksel. [MR0040634](#)
- [38] Willinger, W., Taqqu, M.S., Sherman, R. and Wilson, D. (1995). Self-similarity through high variability: Statistical analysis of ethernet lan traffic at the source level. Proceedings of the ACM/SIGCOMM'95, Cambridge, MA. *Computer Communications Review* **25** 100–113.
- [39] Zygmund, A. (2002). *Trigonometric Series*, **I** and **II**, 3rd ed. Cambridge, UK: Cambridge Univ. Press. [MR1963498](#)

Received March 2009 and revised September 2009