February 18, 2009

# WEAK CONVERGENCE OF THE FUNCTION-INDEXED INTEGRATED PERIODOGRAM FOR INFINITE VARIANCE PROCESSES

SAMI UMUT CAN, THOMAS MIKOSCH, AND GENNADY SAMORODNITSKY

Abstract. In this paper we study the weak convergence of the integrated periodogram indexed by classes of functions for linear and stochastic volatility processes with symmetric $\alpha$-stable noise. 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 representation as infinite Fourier series with iid $\alpha$-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 Töplitz coefficient matrices.

## 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, heavy-tailed time series whose marginal distributions exhibit power law behavior have attracted a lot of attention. The need for such models arises from application in areas as diverse as insurance, geophysics, finance and telecommunications. What these applications have in common is the fact that the observed time series are heavy-tailed in the sense that the fourth or even the second moments can be infinite. Infinite fourth moments are not un-typical for series of daily log-returns from exchange rates, stock indices, and other speculative prices. Infinite second moments can be observed in time series from insurance such as for windstorm, industrial fire and earthquake insurance. These aspects of financial and actuarial time series are described and studied in Embrechts et al. [15]; see also the references given therein and the recent surveys on the extremes of financial time series models by Mikosch [29], Davis and Mikosch [12, 13]. Infinite first moments are typical for the marginal distribution of the magnitudes of earthquakes; see for example Kagan [20]. Infinite variances are observed for the sizes of teletraffic data in the world wide web; see for example Crovella et al. [6, 7], Faÿ et al. [16], Leland et al. [27], Willinger et al. [43]. We also refer to the recent books Adler et al. [1] and Resnick [36] which give surveys on the modeling and statistical analyses of heavy-tailed phenomena.

Classical time series analysis deals with the second (or higher) 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 a 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 are not applicable anymore: they are defined through second order characteristics. However, various studies over the last 20 years have shown that the

1991 *Mathematics Subject Classification.* Primary 60F05; Secondary 60G52, 60G10, 62M15 .

*Key words and phrases.* Spectral analysis, infinite variance process, integrated periodogram, weighted integrated periodogram, stable process, linear process, empirical spectral distribution, asymptotic theory, random quadratic form, stochastic volatility process, entropy, time series.

Thomas Mikosch's research is partly supported by the Danish Research Council (FNU) Grant 272-06-0442. Gennady Samorodnitsky's research is partly supported by the ARO grant W911NF-07-1-0078 at Cornell University and a Villum Kann Rasmussen Visiting Professor Grant at the University of Copenhagen.

analysis of linear processes, i.e., processes of the form

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j \, \varepsilon_{t-j} \,, \quad t \in \mathbb{Z} \,, \tag{1.1}$$

with heavy-tailed iid *innovations* or *noise* $\varepsilon_j$, $j \in \mathbb{Z}$, and constant coefficients $\psi_j$, $j \in \mathbb{Z}$, is very similar to the classical (i.e., finite variance or even Gaussian) time series analysis. In classical time series analysis notions such as the autocovariances and the spectral density are defined only in terms of the coefficients $\psi_j$ and the variance $\sigma_\varepsilon^2$ of the noise. Most estimators and test statistics of 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. For surveys of this theory we refer to Chapter 7 in Embrechts et al. [15]; see also Klüppelberg and Mikosch [24], Mikosch [28].

One of the main goals of classical time series analysis is the study of the spectral properties of the underlying series under the assumption of finite variance of the marginal distributions. In this context, the *periodogram*

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

plays a prominent role as an estimator of the spectral density. Numerous estimation and test procedures are based on this statistic. In particular, integrated versions of the periodogram of the form

$$J_{n,X}(f) = \int_0^\pi I_{n,X}(\lambda) \, f(\lambda) \, d\lambda \,, \tag{1.2}$$

for appropriate classes of real-valued functions $f \in \mathcal{F}$ on $[0, \pi]$ are used for a multitude of applications. We mention a few of them.

We start with the class of the indicator functions

$$\mathcal{F}_I = \left\{ I_{[0,x]} : x \in [0, \pi] \right\} \,.$$

In this case, we consider the integrated periodogram

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

which is a process indexed by $x \in [0\,, \pi]$. Under general conditions, this type of process converges uniformly with probability 1 to the function

$$\sigma_\varepsilon^2 \int_0^x |\psi(\mathrm{e}^{-i\lambda})|^2 \, d\lambda \,, \quad x \in [0, \pi] \,,$$

where

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

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

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

In words, this is the Fourier series based on the *autocovariance function*

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

A necessary and sufficient condition for the existence of the autocovariance function is thus $\sum_j \psi_j^2 < \infty$ and $\sigma_\varepsilon^2 < \infty$. 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 as a basic tool for constructing goodness-of-fit tests for the underlying spectral distribution function. The theory is masterly presented in the classical book by Grenander and Rosenblatt [18]; see also Brockwell and Davis [5], Priestley [35].

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 [3] proposed to consider $(J_{n,X}(f))_{f\in\mathcal{F}_B}$, where

$$\mathcal{F}_B = \{I_{[0,x]}/f_X(x) : x \in [0,\pi]\}\,,$$

i.e., he considered the process

$$J_n(I_{[0,x]}/f_X) = \int_0^x \frac{I_{n,X}(\lambda)}{f_X(\lambda)}\,d\lambda\,, \quad x \in [0,\pi]\,.$$

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 of the form

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

are used to estimate the spectral density or to perform various tests about the spectrum of the underlying stationary sequence. A general reference on the integrated periodogram and its weighted versions as well as on statistical applications is Priestley [35], Chapter 6.

The weighted integrated periodogram is also the basis for one of the classical estimators for fitting ARMA and fractional ARIMA models. This method goes back to early work by Whittle [42]. In this context one considers the functional

$$J_{n,X}(1/f_X(\cdot;\theta)) = \int_0^\pi \frac{I_{n,X}(\lambda)}{f_X(\lambda;\theta)}\,d\lambda\,, \quad f_X(\cdot;\theta) \in \mathcal{F}_W\,,$$

where $\mathcal{F}_W$ is a class of spectral densities indexed by a parameter $\theta \in \Theta \subset \mathbb{R}^d$. The *Whittle estimator* $\widehat{\theta}_n$ of the true parameter $\theta_0 \subset \Theta$ is the minimizer of $J_{n,X}(1/f_X(\cdot;\theta))$ over the parameter set $\Theta$, or over a compact subset of it. This kind of estimation technique is one of the backbones of quasi-maximum likelihood estimation in parametric time series modeling. The so-defined estimator is known to be asymptotically equivalent to the corresponding least squares and Gaussian quasi-maximum likelihood estimators. Equivalence means that the estimator is consistent and asymptotically normal with the same $\sqrt{n}$-rate and asymptotic variance as in the other two cases. A general reference on parameter estimation in ARMA models is Chapter 8 in Brockwell and Davis [5]. When proving the asymptotic normality and consistency of $\widehat{\theta}_n$, one has to study the properties of the sequence $(J_{n,X}(1/f_X(\cdot;\widehat{\theta}_n)))$ which can be considered as 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 by Pollard [34], van der Vaart and Wellner [41]. Early on, this analogy was discovered by Dahlhaus [8] who gave some 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 Mikosch and Norvaiša [31]. A recent survey of non-parametric statistical methods related to the empirical spectral distribution indexed by classes of functions is Dahlhaus and Polonik [9].

It is the aim of this paper to give 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 infinite variance iid $\alpha$-stable sequences $(\varepsilon_t)$, linear processes $(X_t)$ with iid $\alpha$-stable noise and stochastic volatility processes with iid $\alpha$-stable multiplicative noise. Much of the theory on the integrated periodograms given below depends on tail estimates for random quadratic forms in iid infinite variance random variables. Such results are available for iid stable sequences. Although it seems feasible that the theory can be extended to the more general class of processes whose noise variables have regularly varying tail, we do not attempt to achieve this goal. The price would be more technicalities, the gain would be incremental.

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. As mentioned above, by now there exists quite a clear picture about the asymptotic theory of the sample autocovariances, the periodogram and its integrated versions when the innovation sequence in a linear process has infinite variance. For a detailed list of references in the case of iid and linear processes, see Embrechts et al. [15], Chapter 7. In addition, goodness-of-fit tests for heavy-tailed processes were considered in Klüppelberg and Mikosch [23] (this corresponds to the class of functions $\mathcal{F}_I$) in the case of short-memory linear processes, and for long-memory linear processes in Kokoszka and Mikosch [21]. (In contrast to common practice in time series analysis, in the infinite variance case one needs to define long memory of a linear process by suitable slow rates of convergence of the coefficients $\psi_j \to 0$ as $|j| \to \infty$.) Whittle estimation for infinite variance ARMA processes was studied in Mikosch et al. [30] and the corresponding fractional ARIMA case in Kokoszka and Taqqu [22].

We also consider the function-indexed integrated periodograms for stochastic volatility processes with infinite variance. The stochastic volatility process is one of the standard models in financial time series analysis; see Shephard [39, 40] or Andersen et al. [2]. A stochastic volatility model is defined by the relation

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

where the *volatility sequence* $(\sigma_t)$ is a strictly stationary non-negative process independent of the iid multiplicative noise sequence $(\varepsilon_t)$. For our purposes, $\varepsilon_t$ is S$\alpha$S for some $\alpha \in (0,2)$ as specified in Section 2.1. We also assume that $(\log \sigma_t)$ is a linear Gaussian process, i.e., there exist real values $c_j$, $j \in \mathbb{Z}$, and an iid standard normal sequence $(\eta_j)$ such that $\sum_j c_j^2 < \infty$ and

$$\log \sigma_t = \sum_{j=-\infty}^{\infty} c_j \, \eta_{t-j} \,, \quad t \in \mathbb{Z} \,.$$

Then, in particular, $E(\sigma_1^\alpha) < \infty$. This fact and the asymptotic relation (2.1) ensure that

$$P(X_1 > x) \sim E(\sigma_1^\alpha) \, P(\varepsilon_1 > x) \,;$$

see Jessen and Mikosch [19].

The asymptotic behavior of the sample autocovariance and autocorrelation functions of an infinite variance stochastic volatility model was studied in Davis and Mikosch [10, 11]. It turned out that this theory parallels the one for an iid sequence. The limit theory for the integrated periodograms of an iid sequence strongly depends on the weak limits of the sample autocovariance function, and this analogy remains valid for the infinite variance stochastic volatility models described above.

The study of infinite variance processes is a challenge insofar that most classical notions and methods of time series analysis seem to break down. This starts with the definition of the autocovariances and the spectral density. Nevertheless, for linear processes one can define these quantities in a straightforward way by a formal application of the corresponding finite variance expressions in terms of the coefficients $\psi_j$, $j \in \mathbb{Z}$. A particular difficulty consists of the fact that in the infinite variance case spectral analysis and Hilbert space techniques do not really agree with each other; we lose various tools of Fourier analysis, in particular, the isometry property provided by Parseval's formula. The latter tool was one of the main ingredients for the proofs in Mikosch and Norvaiša [31], in combination with techniques from empirical process theory.

The paper is organized as follows. In Section 2 we introduce stable distributions and the stochastic volatility process as well as 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 iid 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 Töplitz 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 iid sequence to linear processes and in Section 5 we derive the corresponding limit theory for stochastic volatility processes with stable innovations. The Appendix contains some auxiliary results about 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.

## 2. Some preliminaries

2.1. **Stable innovations.** In what follows, we study the weak limit theory for the weighted integrated periodogram indexed by a class of functions, where the *innovations* or *noise*, $\varepsilon_j$, $j \in \mathbb{Z}$, are supposed to be iid infinite variance *symmetric $\alpha$-stable* (S$\alpha$S) random variables. Recall that $Y_\alpha$ is said to have a *stable distribution* ($Y_\alpha \sim S_\alpha(\sigma,\beta,\mu)$) if there are parameters $0<\alpha\le 2$, $\sigma\ge 0$, $-1\le\beta\le 1$, and $\mu$ real such that its characteristic function has the form

$$E\mathrm{e}^{\,itY_\alpha} = \begin{cases} \exp\left\{i\mu t - \sigma^\alpha |t|^\alpha \left(1 - i\beta \mathrm{sign}(t)\,\tan(\pi\alpha/2)\right)\right\} & \text{if } \alpha \neq 1, \\ \exp\left\{i\mu t - \sigma |t| \left(1 + (2i\beta/\pi)\mathrm{sign}(t)\log|t|\right)\right\} & \text{if } \alpha = 1. \end{cases}$$

If $\beta=\mu=0$, then $Y_\alpha$ is symmetric $\alpha$–stable (S$\alpha$S).

For convenience we assume that $(\varepsilon_t)_{t\in\mathbb{Z}}$ is an iid sequence of $S_\alpha(1,0,0)$ random variables for some $\alpha\in(0,2)$. For $\alpha<2$ the random variable $\varepsilon_0$ has infinite variance in view of the relation

$$P(\varepsilon_0 > x) \sim c_\alpha\, x^{-\alpha}\,, \quad x \to \infty\,, \tag{2.1}$$

for some $c_\alpha > 0$; see Feller [17]. A general reference to the theory of stable distributions and processes is the monograph by Samorodnitsky and Taqqu [38].

2.2. **Preliminaries on the periodogram.** The following decomposition of the periodogram is fundamental:

$$I_{n,X}(\lambda) = \gamma_{n,X}(0) + 2\sum_{h=1}^{n-1}\cos(\lambda h)\,\gamma_{n,X}(h)\,, \tag{2.2}$$

where

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

denotes the *sample autocovariance function* of the sample $X_1,\dots,X_n$. Notice that the definition of $\gamma_{n,X}$ slightly deviates from the usual one where the $X_t$'s are centered by the sample mean. However, for the theory given below this centering is inessential. Centering with the sample mean $\overline{X}_n$ is not the most natural choice when dealing with infinite variance processes. Moreover, since the theory for the empirical spectral distribution provided in this paper could also be formulated in terms of the periodogram at the Fourier frequencies $\lambda_j = 2\pi j/n \in (0,\pi)$ (integrals then become Riemann sums evaluated at these frequencies), and $\sum_{t=1}^n \mathrm{e}^{-i\lambda_j t} = 0$, centering of the $X_t$'s in the definition of the periodogram is not necessary.

In what follows, we will frequently make use of the *self-normalized periodogram*

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

where

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

denotes the *sample autocorrelation function* of $X_1,\dots,X_n$.

In view of (2.2) we can rewrite $J_{n,\varepsilon}(f)$ as follows:

$$J_{n,\varepsilon}(f) = \gamma_{n,\varepsilon}(0)\,a_0(f) + 2\sum_{h=1}^{n-1} a_h(f)\,\gamma_{n,\varepsilon}(h)\,, \tag{2.3}$$

where

$$a_h(f) = \int_0^\pi \cos(\lambda h)\; f(\lambda)\,d\lambda\,,\quad h\in\mathbb{Z}\,, \tag{2.4}$$

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

$$\widetilde{J}_{n,\varepsilon}(f) = \rho_{n,\varepsilon}(0)\,a_0(f) + 2\sum_{h=1}^{n-1} a_h(f)\,\rho_{n,\varepsilon}(h)\,. \tag{2.5}$$

## 3. The iid case

In this section we study the limit behavior of the integrated periodograms $J_{n,\varepsilon}$ indexed by classes of functions for an iid 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 iid 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, i.e., we focus on random quadratic forms with Töplitz coefficient matrices given by the Fourier coefficients $a_h(f)$ defined in (2.4).

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

A glance at decomposition (2.3) convinces one 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 [14]; see also Brockwell and Davis [5], Section 13.3.

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

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

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

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

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

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

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

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

Here $Y_0, Y_1, Y_2, \ldots$ are independent stable random variables as described in Lemma 3.1. The 3-series theorem (see Petrov [33]) 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 $(\widetilde{X}_n)$ we need a slightly stronger assumption:

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

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; see, for example, Kwapień and Woyczyński [25]. They appear in a natural way in tail estimates for quadratic forms in iid stable random variables; see the Appendix.

Now we can formulate our result about the convergence of the finite–dimensional distributions:

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

$$(X_n(\mathbf{a}))_{\mathbf{a}\in\ell^\alpha\log\ell} \stackrel{fidi}{\longrightarrow} (Y(\mathbf{a}))_{\mathbf{a}\in\ell^\alpha\log\ell} \quad \text{and} \quad (\widetilde{X}_n(\mathbf{a}))_{\mathbf{a}\in\ell^\alpha\log\ell} \stackrel{fidi}{\longrightarrow} (\widetilde{Y}(\mathbf{a}))_{\mathbf{a}\in\ell^\alpha\log\ell}\,.$$

*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 $m \geq 1$,

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

where $\Longrightarrow$ denotes weak convergence. Since $\mathbf{a} \in \ell^\alpha$, $Y_m(\mathbf{a}) \Longrightarrow Y(\mathbf{a})$ as $m \to \infty$ follows from the 3-series theorem. According to Theorem 4.2 in Billingsley [4], it remains to show that

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

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 6.1 in the Appendix and the fact that $\mathbf{a} \in \ell^\alpha \log \ell$, we conclude that

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

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

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 (2.3) and its self-normalized version $\widetilde{J}_{n,\varepsilon}$ in (2.5).

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

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

*where $\mathbf{a}(f)$ is specified in* (2.4)*. Then*

$$n\,(n \log n)^{-1/\alpha}\,[J_{n,\varepsilon}(f) - a_0(f)\;\gamma_{n,\varepsilon}(0)]_{f\in\mathcal{F}} \quad \stackrel{fidi}{\longrightarrow} \quad 2\,[Y(\mathbf{a}(f))]_{f\in\mathcal{F}}\,,$$

$$(n/\log n)^{1/\alpha}\left[\widetilde{J}_{n,\varepsilon}(f) - a_0(f)\right]_{f\in\mathcal{F}} \quad \stackrel{fidi}{\longrightarrow} \quad 2\,[\widetilde{Y}(\mathbf{a}(f))]_{f\in\mathcal{F}}\,.$$

**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 $\gamma_X(h) \to 0$ as $h \to \infty$ is well known for numerous time series models. For example, if $f$ is the spectral density of an ARMA process, $\gamma_X(h) \to 0$ 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 Zygmund [44]. His Theorem (3.10) on p. 243 in Volume I 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) on p. 243 in 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 $\alpha = 2/(2+\beta)$.

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 Samorodnitsky and Taqqu [38], 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 $(0,\infty)$. Notice that $\ell^\alpha \log \ell$ can be characterized as follows:

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

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

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

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. This follows from the results in Samorodnitsky and Taqqu [38], Section 10.4.

The following is our main result on the weak convergence of the sequences $(X_n)$ and $(\widetilde{X}_n)$ 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 $\mathbb{C}(\mathcal{A})$:*

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

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

*Proof.* We restrict ourselves to show $X_n \Longrightarrow Y$. 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 $\delta$:

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

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

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

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 $\delta$ is small, where

$$c_k(\delta) = \sup_{d_{\mathcal{A}}(\mathbf{a},\mathbf{b})<\delta} |a_k - b_k|\,.$$

Now apply Lemma 6.2 to the $P_n'$'s:

$$\begin{aligned} P_n'(\epsilon,\delta) &\le \textit{const}\, \frac{1+\log^+|\epsilon|}{|\epsilon|^\alpha}\, \frac{1+\log n}{n\log n} \sum_{i=1}^{n-1}\sum_{k=i+1}^{n} |c_k(\delta)|^\alpha \left(1 + \log^+ \frac{1}{|c_k(\delta)|}\right) \\ &\le \textit{const} \sum_{k=1}^{\infty} h(c_k(\delta)) \to 0 \quad \text{as } \delta \to 0\,. \end{aligned} \tag{3.6}$$

The limit relation (3.6) is a consequence of condition (3.4). □

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 Töplitz matrices. The conditions on the parameter set $\mathcal{A}$ are nothing but restrictions on the infinite Töplitz matrices $(a_{i-j})_{i,j=1,2,\ldots}$. When specified to the particular case of Fourier coefficients as in (2.4), Theorem 3.5 yields the following.

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

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

*where* $\mathcal{A}$ *is a compact set of* $\ell^\alpha \log \ell$ *satisfying* (3.4) *and* $\mathbf{a}(f)$ *is specified in* (2.4)*. Then*

$$\begin{cases} n\,(n\log n)^{-1/\alpha}\,[J_{n,\varepsilon}(f) - a_0(f)\,\gamma_{n,\varepsilon}(0)]_{f\in\mathcal{F}} \implies 2\,[Y(\mathbf{a}(f))]_{f\in\mathcal{F}}, \\ (n/\log n)^{1/\alpha}\left[\widetilde{J}_{n,\varepsilon}(f) - a_0(f)\right]_{f\in\mathcal{F}} \implies 2\,[\widetilde{Y}(\mathbf{a}(f))]_{f\in\mathcal{F}}, \end{cases} \tag{3.7}$$

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

*Proof.* Let $T : \mathcal{F} \to \mathcal{A}$ 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

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

$n = 1, 2, \ldots$ converges in $L^1[0,\pi]$ to some function $f$ that has to be in $\mathcal{F}$. Therefore, $\mathbf{a} = Tf \in T\mathcal{F}$, and the latter set is compact.

The above argument shows that the $L^2[0,\pi]$ convergence in $\mathcal{F}$ is equivalent to the $\ell^\alpha \log \ell$ convergence in $T\mathcal{F}$. Since Theorem 3.5 implies weak convergence of the left hand side of (3.7) 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 (3.7) holds also in $\mathbb{C}(\mathcal{F})$. □

This result provides a solution to the problem of finding the weak limits of the specific random quadratic forms $J_{n,\varepsilon}$ in iid infinite S$\alpha$S random variables $\varepsilon_t$, uniformly over a whole class of functions $f \in \mathcal{F}$ satisfying some mild conditions.

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 present in proving tightness for the random quadratic forms $X_n$ and $\widetilde{X}_n$. The requirements prescribed by the geometry are usually given by entropy conditions; see Ledoux and Talagrand [26] 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 Chapter 12 of Samorodnitsky and Taqqu [38].

In this section we only consider vectors $\mathbf{a} \in \ell^\alpha \log \ell$ of the form (2.4), i.e., they are the Fourier coefficients of some functions $f$. Corollary 3.3 determines the structure of the limit process of the quadratic forms $J_{n,\varepsilon}$ via the convergence of their finite–dimensional distributions. Hence it suffices to show the tightness in $\mathbb{C}(\mathcal{F})$ for suitable classes $\mathcal{F}$. Klüppelberg and Mikosch [23] 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 some entropy condition.

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

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

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

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

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 $\alpha \in [1,2)$, define $\mathbf{a}(f)$ as in (2.4) 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*

$$N(\epsilon, \mathcal{F}, \rho_k) \le const \left[1 + \left(\frac{2^k}{\epsilon}\right)^\beta\right], \quad \epsilon > 0,\ k \ge 0\,. \tag{3.8}$$

*Then the weak convergence result* (3.7) *holds in* $\mathbb{C}(\mathcal{F})$.

In contrast to the finite variance case (see Dahlhaus [8], Mikosch and Norvaiša [31]) the entropy condition (3.8) 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 (3.8) are common in problems of continuity and boundedness for stable processes; see Chapter 10 in Samorodnitsky and Taqqu [38].

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

$$\lim_{m\to\infty} \limsup_{n\to\infty} P\left(\left\|\sum_{j=m}^n a_j \widehat{\gamma}_{n,\varepsilon}(j)\right\|_{\mathcal{F}} > \epsilon\right) = 0 \quad \text{for every } \epsilon > 0, \tag{3.9}$$

where

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

and

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

As in (6.4) on p. 1873 of Klüppelberg and Mikosch [23], one can argue that it suffices in (3.9) to consider $m$ and $n$ of some specific form. Let $a < b$ be two positive integers

$$m = 2^a \quad \text{and} \quad n = 2^{b+1} - 1\,,$$

and consider numbers

$$\epsilon_k = 2^{-k\theta}\,, \quad k \ge 1\,,$$

with $\theta > 0$. For $\theta$ sufficiently small and $a$ large enough we have

$$P\Big(\Big\|\sum_{j=m}^n a_j \widehat{\gamma}_{n,\varepsilon}(j)\Big\|_{\mathcal{F}} > \epsilon\Big) \le \sum_{k=a}^b P\Big(\Big\|\sum_{j=2^k}^{2^{k+1}-1} a_j \widehat{\gamma}_{n,\varepsilon}(j)\Big\|_{\mathcal{F}} > \epsilon_k\Big) = \sum_{k=a}^b p_k\,. \tag{3.10}$$

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

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

where

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

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

By virtue of Lemma 6.1, we have for all $f,g\in\mathcal{F}$,

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

where

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

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

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

Hence we are left to consider

$$\begin{aligned} &\sum_{k=a}^{b}\sum_{l=1}^{\infty} N(\epsilon_{k,l},\mathcal{F},\rho_k)\;\epsilon_k^{-\alpha}2^{-k(\alpha-1)+\alpha l}\epsilon_{k,l-1}^\alpha[1+k\;\log^+\epsilon_{k,l-1}^{-1}]\\ &=\sum_{k=a}^{b}2^{-k(\alpha-1-\alpha\theta)}\sum_{l=1}^{\infty}N(\epsilon_{k,l},\mathcal{F},\rho_k)\epsilon_{k,l-1}^\alpha[1+k\;\log^+\epsilon_{k,l-1}^{-1}]2^{\alpha l}\\ &\le const\sum_{k=a}^{b}2^{-k(\alpha-1-\alpha\theta)}\sum_{l=1}^{\infty}\left[1+\left(\frac{2^k}{\epsilon_{k,l}}\right)^\beta\right]\epsilon_{k,l-1}^\alpha\left[1+\log^+\epsilon_{k,l-1}^{-1}\right]2^{\alpha l}\,. \end{aligned} \tag{3.11}$$

Assume $\theta$ is so small that (3.10) holds. Define the numbers

$$\epsilon_{k,l} = 2^{-\gamma_1 l-\gamma_2 k}\,,\quad k,l\ge 0$$

with $\gamma_1,\gamma_2>0$ such that

$$1+\gamma_2 > \frac{1+\alpha\theta}{\alpha-\beta}\quad\text{and}\quad \gamma_1 > \frac{\alpha}{\alpha-\beta}\,.$$

For these parameter choices it is not difficult to see that (3.11) converges to zero by first letting $n\to\infty$ (i.e., $b\to\infty$) and then $m\to\infty$ (i.e., $a\to\infty$). This proves (3.9), hence the tightness of the considered processes in $\mathbb{C}(\mathcal{F})$. □

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 $K>0$, i.e.

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

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

$$N(\epsilon, \mathcal{G}_\Theta, \rho_k) \le const \left(\frac{2^k}{\epsilon}\right)^{a/b},$$

with $a/b \in (0, \alpha)$. This inequality follows from the following arguments. Let $\epsilon > 0, k \ge 0$. We can find $N \le c\big((K\pi 2^{k+1})/\epsilon\big)^{a/b}$ balls of radius at most $\big(\epsilon/(K\pi 2^{k+1})\big)^{1/b}$ 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

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

It follows that

$$N(\epsilon, \mathcal{G}_\Theta, \rho_k) \le N \le const \left(\frac{2^k}{\epsilon}\right)^{a/b}.$$

**Example 3.9.** Consider a Vapnik-Červonenkis (VC) class $\mathcal{G}$ of functions defined on $[0, \pi]$, with VC-index $V(\mathcal{G}) = 2$. (We refer to Section 2.6.2 of van der Vaart and Wellner [41] for more information on VC-classes of functions.) Given $\epsilon > 0$ and $k \ge 0$, we can find $N \le c(\pi 2^{k+1})/\epsilon$ 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, Theorem 2.6.7 of [41]. So there are $g_1, \dots, g_N \in \mathcal{G}$ such that for any $g \in \mathcal{G}$,

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

Then,

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

It follows that

$$N(\epsilon, \mathcal{G}, \rho_k) \le N \le const \left(\frac{2^k}{\epsilon}\right).$$

## 4. The linear process case

From (1.2) recall the definition of the integrated periodogram $J_{n,X}$ indexed by a class of functions $\mathcal{F}$. It is the aim of this section to show that the results for the case of an iid sequence $(\varepsilon_t)$ translate to the linear process case. The following decomposition will be crucial:

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

This decomposition is analogous to the decomposition of the spectral density $f_X$ of a linear process:

$$f_X(\lambda) = f_\varepsilon(\lambda)|\psi(\mathrm{e}^{-i\lambda})|^2\,.$$

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

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

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

$$J_{n,\varepsilon}(f)=\int_0^\pi I_{n,\varepsilon}(\lambda)\;f(\lambda)\;d\lambda\,,\quad f\in\mathcal{F}_\psi\,,$$

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*

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

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

$$\frac{n}{(n\log n)^{1/\alpha}}\sup_{f\in\mathcal{F}}\Big|\int_0^\pi f(x)R_n(x)\,dx\Big|\stackrel{P}{\longrightarrow}0.$$

*Proof.* From Proposition 5.1 in Mikosch et al. [30], substituting $n^{1/2}$ for $a_n$, we have the following decomposition for $R_n$:

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

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

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

We first show that

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

Note that

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

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 $n\to\infty$. 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 since, by assumption, $\sup_{f\in\mathcal{F}}\|f\|_2<\infty$,

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

where

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

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

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

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

$$P(|\varepsilon_m|>M\,m^{1/\alpha}(1+\log m)^{\frac{1}{\alpha}+\delta}\text{ for some }m\ge 1)\le\epsilon/2\,.$$

Write

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

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

$$\begin{aligned}&P(I_1(n)>k)-\epsilon/2\\ \le\;& P\Big(\sum_{j=1}^{n}|\psi_j|\Big(\int_0^{\pi}\Big|\sum_{m=1}^{j}\varepsilon_m e^{itx}\Big|^4dx\Big)^{1/2}\Big(\int_0^{\pi}\Big|\sum_{m=1}^{j}J_m\,e^{itx}\Big|^4dx\Big)^{1/2}>k\Big)\\ \le\;& k^{-1}\sum_{j=1}^{n}|\psi_j|\Big(\int_0^{\pi}E\Big|\sum_{m=1}^{j}J_m\,e^{itx}\Big|^4\,dx\Big)^{1/2}\\ =\;& k^{-1}\sum_{j=1}^{n}|\psi_j|\Big[\int_0^{\pi}\Big(\sum_{m=1}^{j}E(J_m^4)+6\sum_{m_1=1}^{j}\sum_{m_2=m_1+1}^{j}E(J_{m_1}^2)\,E(J_{m_2}^2)\,\cos((m_1-m_2)x)\,dx\Big)\Big]^{1/2}\end{aligned}$$

$$\leq \frac{c}{k}\sum_{j=1}^{n}|\psi_j|\Big[\Big(\sum_{m=1}^{j}\Big(m^{1/\alpha}(1+\log m)^{\frac{1}{\alpha}+\delta}\Big)^{4-\alpha}\Big)^{1/2}+\Big(\Big[\sum_{m=1}^{j}\Big(m^{1/\alpha}(1+\log m)^{\frac{1}{\alpha}+\delta}\Big)^{2-\alpha}\Big]^{2}\Big)^{1/2}\Big]$$

$$\leq \frac{c}{k}\sum_{j=1}^{\infty}|\psi_j|\, j^{2/\alpha}\,(1+\log j)^{\frac{(4-\alpha)}{2\alpha}+\tau}\,.$$

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

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

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

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

So we see that

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

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

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

follows from Mikosch et al. [32], 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. □

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

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

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

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

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

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

## 5. The stochastic volatility case

Recall the notion of a stochastic volatility process $(X_t)$, where $(\varepsilon_t)$ is iid S$\alpha$S for some $\alpha \in (0,2)$ and $(\sigma_t)$ a linear Gaussian process. The results for the empirical spectral distribution for this process are analogous to the iid case. In what follows, we will give the results and sketch some ideas of the proofs.

Our first observation is that Lemma 3.1 holds in a modified form for the sample autocovariance function $\gamma_{n,X}$ and autocorrelation function $\rho_{n,X}$ of the stochastic volatility process $(X_t)$; see Davis and Mikosch [10], Theorem 4.2 and Remark 4.2.

**Lemma 5.1.** *For every $m \geq 1$,*

$$(5.1)\left(\frac{n\gamma_{n,X}(0)}{n^{2/\alpha}}\,,\frac{n\gamma_{n,X}(h)}{(n\log n)^{1/\alpha}}\,,h=1,\ldots,m\right) \Longrightarrow (\|\sigma_0\|_\alpha^2\, Y_0, \|\sigma_0\sigma_1\|_\alpha Y_1,\ldots,\|\sigma_0\sigma_m\|_\alpha Y_m)\,,$$

*where the $Y_h$'s are independent, $Y_0$ is $S_{\alpha/2}(\sigma_1,1,0)$ and $(Y_h)_{h=1,\ldots,m}$ are iid $S_\alpha(\sigma_2,0,0)$ for some $\sigma_i=\sigma_i(\alpha)$, $i=1,2$, and for any random variable $Y$, $\|Y\|_\alpha = (E|Y|^\alpha)^{1/\alpha}$. In particular,*

$$(n/\log n)^{1/\alpha}(\rho_{n,\varepsilon}(h))_{h=1,\ldots,m} \Longrightarrow \left(\frac{\|\sigma_0\,\sigma_h\|_\alpha}{\|\sigma_0\|_\alpha^2}\frac{Y_h}{Y_0}\right)_{i=1,\ldots,m}. \tag{5.2}$$

Next, let $Y_0,Y_1,Y_2,\ldots$ be the limiting variables given in Lemma 5.1. For $\mathbf{a}\in\ell^\alpha$ we define the sequences of processes

$$(5.3)\qquad \begin{cases} X_n(\mathbf{a}) &=& (n\log n)^{-1/\alpha}\sum_{k=1}^{n-1} a_k\,[n\,\gamma_{n,X}(k)]\,, & Y(\mathbf{a}) &=& \sum_{k=1}^\infty a_k\,\|\sigma_0\sigma_k\|_\alpha\,Y_k\,,\\ \widetilde{X}_n(\mathbf{a}) &=& (n/\log n)^{1/\alpha}\sum_{k=1}^{n-1} a_k\,\rho_{n,X}(k)\,, & \widetilde{Y}(\mathbf{a}) &=& Y(\mathbf{a})/(\|\sigma_0\|_\alpha^2\,Y_0)\,.\end{cases}$$

Now we can formulate the following analog of Theorem 3.5.

**Theorem 5.2.** *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 $\mathbb{C}(\mathcal{A})$:*

$$(X_n(\mathbf{a}))_{\mathbf{a}\in\mathcal{A}} \Longrightarrow (Y(\mathbf{a}))_{\mathbf{a}\in\mathcal{A}} \quad \textit{and} \quad (\widetilde{X}_n(\mathbf{a}))_{\mathbf{a}\in\mathcal{A}} \Longrightarrow (\widetilde{Y}(\mathbf{a}))_{\mathbf{a}\in\mathcal{A}}$$

*where $X_n$, $\widetilde{X}_n$, $Y$ and $\widetilde{Y}$ are defined in (5.3), and the processes $Y$ and $\widetilde{Y}$ are sample–continuous.*

*Proof.* The proof of the convergence of the finite-dimensional distributions is analogous to the proof of Theorem 3.2; we sketch the main ideas. Using a Cramér–Wold argument, it suffices to prove the convergence of the one–dimensional distributions. From (5.1) and the continuous mapping theorem it follows that for every $m\geq 1$,

$$(n\log n)^{-1/\alpha}\sum_{k=1}^m a_k\,[n\,\gamma_{n,X}(k)] \Longrightarrow Y_m(\mathbf{a}) = \sum_{k=1}^m a_k\,\|\sigma_0\sigma_k\|_\alpha\,Y_k\,.$$

Since $\mathbf{a}\in\ell^\alpha$ and $\|\sigma_0\sigma_k\|_\alpha \leq \|\sigma_0\|_\alpha^2$, $Y_m(\mathbf{a}) \Longrightarrow Y(\mathbf{a})$ as $m\to\infty$ follows from the 3-series theorem. According to Theorem 4.2 in Billingsley [4], it remains to show that

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

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 6.1 in the Appendix conditional on $(\sigma_t)$, using the elementary inequality

$$E\Big(|c\,Y|^\alpha(1+\log^+(1/|c\,Y|)\Big) \;\leq\; 2\,c^\alpha\,(1+\log^+(1/c))\,E|Y|^\alpha(1+\log^+(1/|Y|))\,,\quad c>0\,, \tag{5.4}$$

and observing that $\sigma_1$ has finite moments of any order, we conclude that

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

In the last step we also used that $\mathbf{a} \in \ell^\alpha \log \ell$. This proves the theorem for $(X_n)$; the convergence of $(\widetilde{X}_n)$ can be shown analogously by utilizing (5.2).

For the proof of the tightness of $(X_n)$ and $(\widetilde{X}_n)$, one can follow the lines of the proof of Theorem 3.5. Again, an application of Lemma 6.2 conditional on $(\sigma_t)$ and (5.4) yield the same bounds for $P_n'$ (adapted to the stochastic volatility sequence). □

We obtain the following analog to Corollary 3.6.

**Corollary 5.3.** *Assume $\alpha \in (0,1)$ and let*

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

*where $\mathcal{A}$ is a compact set of $\ell^\alpha \log \ell$ satisfying* (3.4) *and $\mathbf{a}(f)$ is specified in* (2.4)*. Then*

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

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

Finally, we mention that the statement of Theorem 3.7 remains true for the stochastic volatility process $(X_t)$ in the case $\alpha \in [1,2)$, with the corresponding adaptations of the limiting processes $Y$ and $\widetilde{Y}$ given in (5.3). One can follow the lines of the proof of Theorem 3.7 but one needs to apply Lemma 6.1 conditional on $(\sigma_t)$ in combination with (5.4). We omit details.

## 6. Appendix

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

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

and

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

The following lemma is a consequence of Theorem 3.1 in Rosiński and Woyczyński [37]; see also Kwapień and Woyczyński [25].

**Lemma 6.1.** *For $\alpha \in (0,2)$, there exists a positive constant $D_\alpha$ such that for all $x > 0$:*

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

Let now $\mathbf{C} = (C_0, C_{s,t}, s,t = 1,2,\ldots)$ be a sequence of iid $S_1(1,0,0)$ random variables, independent of $(\varepsilon_t)$, and $\mathbf{b}$ be as above. The following lemma is a consequence of Lemma 6.1:

**Lemma 6.2.** *For $\alpha \in (0,1)$, there exists a positive constant $D_\alpha'$ such that for all $x > 0$:*

$$I_n(x) = P\left(\sum_{1\le s<t\le n} b_{s,t}\,C_{s,t}\,\varepsilon_s\varepsilon_t > x\right) \le D_\alpha' \frac{1 + \log^+ x}{x^\alpha}\,\Gamma_n(\mathbf{b})\,. \tag{6.1}$$

*Proof.* Apply Lemma 6.1 to $I_n(x)$, conditionally on $\mathbf{C}$:

$$
\begin{aligned}
I_n(x) &= P\left(\sum_{1\le s<t\le n} b_{s,t}\, C_{s,t}\, \varepsilon_s\varepsilon_t > x\right) \\
&= E_{\mathbf{C}} P\left(\sum_{1\le s<t\le n} b_{s,t}\, C_{s,t}\, \varepsilon_s\varepsilon_t > x \,\middle|\, \mathbf{C}\right) \qquad (6.2)\\
&\le const\, \frac{1+\log^+ x}{x^\alpha} \sum_{s=1}^n \sum_{t=i+1}^n |b_{s,t}|^\alpha\, E|C_0|^\alpha \left(1+\log^+ \frac{1}{|b_{s,t}C_0|}\right).
\end{aligned}
$$

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

$$
E|C_0|^\alpha \left(1+\log^+ \frac{1}{|xC_0|}\right) \le const \left(1+\log^+ \frac{1}{|x|}\right), \qquad (6.3)
$$

and combining (6.2) and (6.3), we thus obtain (6.1). □

**Acknowledgment.** 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.

## References


[1] R.J. Adler, R.E. Feldman and M.S. Taqqu (Eds.) (1998) *A Practical Guide to Heavy Tails.* Birkhäuser, Boston.

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

[3] M.S. Bartlett (1954) Problemes de l'analyse spectrale des séries temporelles stationnaires. *Publ. Inst. Statist. Uni. Paris* **III-3**, 119–134.

[4] P. Billingsley (1968) *Convergence of Probability Measures*. Wiley, New York.

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

[6] M. Crovella and A. Bestavros (1996) Self-similarity in world wide web traffic: evidence and possible causes. *Proceedings of the 1996 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems* **24**, 160–169.

[7] M. Crovella, A. Bestavros and M.S. Taqqu (1996) Heavy-tailed probability distributions in the world wide web. Preprint, 1996.

[8] R. Dahlhaus (1988) Empirical spectral processes and their applications to time series analysis. *Stochastic Processes and their Applications* **30**, 69–83.

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

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

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

[12] R.A. Davis and T. Mikosch (2009) Extreme value theory for GARCH processes. In: T.G. Andersen, R.A. Davis, J.-P. Kreiss, and T. Mikosch (Eds.) (2009) *The Handbook of Financial Time Series.* Springer, Heidelberg, to appear.

[13] R.A. Davis and T. Mikosch (2009) Extremes of stochastic volatility models. In: T.G. Andersen, R.A. Davis, J.-P. Kreiss and T. Mikosch (Eds.) *The Handbook of Financial Time Series.* Springer, Heidelberg, to appear.

[14] R.A. Davis and S.I. Resnick (1986) Limit theory for the sample covariance and correlation functions of moving averages. *Annals of Statistics* **14** 533–558.

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

[16] G. Faÿ, B. Gonzalez-Arevalo, T. Mikosch, and G. Samorodnitsky (2006) Modeling teletraffic arrivals by a Poisson cluster process. *QESTA* **54**, 121–140

[17] W. Feller (1971) *An Introduction to Probability Theory and its Applications*, volume 2, 2nd edition. Wiley, New York.

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

[19] A.H. Jessen and T. Mikosch (2006) Regularly varying functions. *Publ. Inst. Math. (Beograd) (N.S.)* **80(94)**, 171-192.

[20] Y.Y. Kagan (1997) Earthquake size distribution and earthquake insurance. *Communications in Statistics: Stochastic Models* **13**, 775–797.

[21] P. Kokoszka and T. Mikosch (1997) The integrated periodogram for long–memory processes with finite or infinite variance. *Stochastic Processes and their Applications* **66**, 55–78.

[22] P. Kokoszka and M.S. Taqqu (1996) Parameter estimation for infinite variance fractional ARIMA. *Annals of Statistics* **24**, 1880–1913.

[23] C. Klüppelberg and T. Mikosch (1996) The integrated periodogram for stable processes. *Annals of Statistics* **24**, 1855–1879.

[24] C. Klüppelberg and T. Mikosch (1996) Self–normalized and randomly centred spectral estimates. In: C.C. Heyde, Yu.V. Prokhorov, R. Pyke and S.T. Rachev (Eds.) *Proceedings of the Athens International Conference on Applied Probability and Time Series,* vol. 2: Time Series. Springer, Berlin, pp. 259–271.

[25] S. Kwapień and W.A. Woyczyński (1992) *Random Series and Stochastic Integrals: Single and Multiple.* Birkhäuser, Basel.

[26] M. Ledoux and M. Talagrand (1991) *Probability in Banach Spaces: Isoperimetry and Processes.* Springer, Berlin.

[27] W.E. Leland, W. Willinger, M.S. Taqqu and D.V. Wilson (1993) On the self-similar nature of Ethernet traffic. *Computer Communications Review* **23**, 183–193.

[28] T. Mikosch (1998) Periodogram estimates from heavy–tailed data. In: R.A. Adler, R. Feldman and M.S. Taqqu (Eds.) *A Practical Guide to Heavy Tails: Statistical Techniques for Analysing Heavy–Tailed Distributions.* Birkhäuser, Boston, pp. 241–258.

[29] T. Mikosch (2003) Modelling dependence and tails of financial time series. In: B. Finkenstädt and H. Rootzén (Eds.) *Extreme Values in Finance, Telecommunications and the Environment.* Chapman & Hall, London, pp. 185–286.

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

[31] T. Mikosch and R. Norvaiša (1997) Uniform convergence of the empirical spectral distribution function. *Stochastic Processes and their Applications* **70**, 85–114.

[32] T. Mikosch, S. Resnick and G. Samorodnitsky (2000) The maximum of the periodogram for a heavy-tailed sequence. *Annals of Applied Probability* **28**, 885–908.

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

[34] D. Pollard (1984) *Convergence of Stochastic Processes.* Springer, Berlin.

[35] M. Priestley (1981) *Spectral Analysis and Time Series*, volume I and II. Academic Press, New York.

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

[37] J. Rosiński and W.A. Woyczyński (1987) Multilinear forms in Pareto-like random variables and product random measures. *Colloqia Mathematica* **51**, 303–313.

[38] G. Samorodnitsky and M. Taqqu (1994) *Stable Non-Gaussian Random Processes*. Chapman & Hall, New York.

[39] N. Shephard (1996) Statistical aspects of ARCH and stochastic volatility. In: D.R. Cox, D.V. Hinkley, and O.E. Barndorff-Nielsen (Eds.) *Time Series Models in Econometrics, Finance and Other Fields.* Chapman & Hall, London, pp. 1-67.

[40] N. Shephard (Ed.) (2005) *Stochastic Volatility: Selected Readings.* Oxford University Press, Oxford (UK).

[41] A.W. van der Vaart and J.A. Wellner (1996) *Weak Convergence and Empirical Processes.* Springer, New York.

[42] P. Whittle (1951) *Hypothesis Testing in Time Series Analysis.* Almqvist och Wicksel, Uppsala.

[43] W. Willinger, M.S. Taqqu, R. Sherman and D. Wilson (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.

[44] A. Zygmund (2002) *Trigonometric Series,* volumes I and II, 3rd edition Cambridge University Press, Cambridge (UK).

Center for Applied Mathematics, Cornell University, 657 Rhodes Hall, Ithaca, NY 14853, U.S.A.
*E-mail address*: suc2@cornell.edu

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

School of Operations Research and Information Engineering, Cornell University, 220 Rhodes Hall, Ithaca, NY 14853, U.S.A.
*E-mail address*: gennady@orie.cornell.edu