Nonparametric Statistics, 2003, Vol. 00, pp. 1–35

Taylor & Francis
Taylor & Francis Group

LARGE SAMPLE THEORY FOR STATISTICS OF STABLE MOVING AVERAGES

TUCKER McELROY and DIMITRIS N. POLITIS*

Q1 *Department of Mathematics, University of California, San Diego, Lajolla, CA 92093-0112, USA*

(Received 21 March 2003; Revised 17 April 2003; In final form 25 April 2003)

We study the limit behavior of the partial sums, sample variance, and periodogram of the stable moving average process

$$x(t) = \int \psi(t+x)\mathbb{M}(\mathrm{d}x)$$

explored in Resnick, S., Samorodnitsky, G., and Xue, F. (1999). How misleading can sample ACF's of stable MA's be? (Very!). *Annals of Applied Probability*, **9**(3), 797–817. Each of these statistics has a rate of convergence involving the "characteristic exponent" α, which is an unknown parameter of the model. Through the employment of self-normalization, this number α can be removed from the limit distribution; the various limit distributions can then be approximated via subsampling. As a result, statistical inference for the mean can be conducted without knowledge (or explicit estimation) of α. New techniques, which are easily generalizable to a random field model, are presented to prove these results.

Q2 *Keywords*: Please supply

1 INTRODUCTION

Within the literature of dependent, heavy-tailed stationary time series, the discrete time stochastic process

$$X(t) = \int_{\mathbb{R}} \psi(t+x)\mathbb{M}(\mathrm{d}x) \tag{1}$$

has been studied in Resnick *et al*. (1999). Here, t is the integer index of the process, $\mathbb{M}$ is an α-stable random measure and ψ is a sufficiently regular real-valued function. In the above paper, the authors investigate the asymptotics of the sample autocovariances; they obtain a stochastic limit, and thereby conclude that the use of sample autocorrelation plots as a diagnostic for m-dependence is dubious.

Inherent in many statistics of such a process (1) is the fact that the characteristic exponent α is explicitly present in the rate of convergence to a nondegenerate distribution. Estimation of the characteristic exponent is a difficult practical problem – the ubiquitous Hill estimator for α has a bandwidth selection difficulty, namely the number of order statistics used in the

* Corresponding author. E-mail: politis@euclid.ucsd.edu

ISSN 1048-5252 print; ISSN 1029-0311 online
DOI: 10.1080/10485250310001622587

computation is determined by the practitioner. In McElroy and Politis (2002), this issue was addressed in the context of mean estimation via normalizing the sample mean with the sample standard deviation, under an infinite order moving average model with heavy-tailed inputs. The resulting *self-normalized* ratio's rate of convergence no longer depended upon the hidden α, which provided a fortuitous circumvention of the Hill estimator. Under the same infinite order moving average model, a similar method was used to get the asymptotics of the self-normalized periodogram in Klüppelberg and Mikosch (1993, 1994).

The paper at hand focuses on a different model (given by Eq. (1)) from that explored above, which is more general with respect to dependence structure but more specific distributionally, as the marginals are assumed to be α-stable, instead of merely being in an α domain of attraction. Since the above stochastic process can model a large family of dependent, heavy-tailed stationary time series, we have further studied the asymptotics of various statistics, such as the sample mean, sample variance and periodogram. Although some of the results in this direction were already known in Resnick *et al.* (1999), most of the material is new and all of the proof techniques are completely original; indeed, the analyses employed here reveal much of the structure of the stochastic process itself. These methods allow a fluid transition to higher dimensional index sets, so that the theorems are valid for random field models as well.

This paper is organized as follows: the second section develops the "Representation Lemma" and asymptotic results for the sample mean and sample variance, using new techniques. Moreover, these are demonstrated in Section 3 to be joint limit results using some intricate and delicate arguments. By contrast, in the paper by Resnick *et al.* (1999), a different approach is used to explore the sample covariance asymptotics. The fourth section gives an introduction into the analysis of the periodogram, which can be normalized by the sample variance, thus Q3 expanding and complementing the work of Klüppelberg and Mikoach (1992, 1993) with novel methodologies. Some applications, using subsampling methods, are given in the fifth section.

2 SAMPLE MEAN AND SAMPLE VARIANCE

2.1 The Model

Consider an α-stable random measure $\mathbb{M}$ with skewness intensity $\beta(\cdot)$ and Lebesgue control measure $\mathrm{d}x$ defined on the space $\mathbb{R}$; let ψ be a *filter function* in

$$\mathbb{L}_\delta := \left\{ f \colon \|f\|_\delta^\delta := \int_{\mathbb{R}} |f(x)|^\delta \, \mathrm{d}x < \infty \right\},$$

which is continuous and bounded for almost every x with respect to Lebesgue measure. Then we may construct the following stochastic integral with respect to an α-stable random measure $\mathbb{M}$ (see Samorodnitsky and Taqqu, 1994):

$$X(t) = \int_{\mathbb{R}} \psi(x+t)\mathbb{M}(\mathrm{d}x) \tag{2}$$

with $t \in \mathbb{Z}$. The number δ is in $(0,\alpha) \cap [0,1]$; α will be fixed throughout the discussion. When we speak of the sample mean, then $\alpha > 1$, but otherwise $0 < \alpha < 2$. Note that $\alpha = 2$ corresponds to a Gaussian stochastic process, and has been extensively studied; many results are similar, but for the sample variance (see below), there is a great difference between the $\alpha < 2$ and $\alpha = 2$ cases. We will assume that the skewness intensity $\beta(\cdot)$ of the random measure $\mathbb{M}$ has unit period, which is necessary for the stationarity of the process $X(\cdot)$.

Intuitively, we may think of X as the convolution of ψ and $\mathbb{M}$, in analogy with the infinite order moving average of classical time series analysis. The resulting time series is strictly stationary with dependence that extends over arbitrarily large lags (so long as ψ, does not have compact support; when ψ has compact support, the time series will only be m-dependent, where m is the diameter of the support region), and thus it makes for an interesting and relevant heavy-tailed model.

PROPOSITION 1 *The model defined by Eq.* (1) *is well-defined and stationary. That is, for each* t, *the random variable* $X(t)$ *is* α*-stable with location zero* (*unless* $a = 1$), *constant skewness and constant scale.*

Proof From Samorodnitsky and Taqqu (1994), it follows that the scale of $X(t)$ is

$$\left(\int_{\mathbb{R}} |\psi(t+x)|^{\alpha}\, \mathrm{d}x\right)^{1/\alpha} = \|\psi\|_{\alpha} < \infty$$

and the skewness is

$$\frac{(f_{\mathbb{R}}(\psi(t+x))^{\langle\alpha\rangle}\beta(x)\, \mathrm{d}x)}{\|\psi\|_{\alpha}^{\alpha}} = \frac{(\int_{\mathbb{R}}(\psi(x))^{\langle\alpha\rangle}\beta(x)\, \mathrm{d}x)}{\|\psi\|_{\alpha}^{\alpha}}.$$

We use the notation $a^{\langle\gamma\rangle} = \mathrm{sign}(a)\alpha^{\gamma}$. Since $|\beta(x)| \leq 1$, the skewness is bounded between -1 and 1. Note that stationarity follows from the translation invariance of Lebesgue measure, and the periodicity of the skewness intensity β. Moreover this shows that the process is non-degenerate unless ψ is zero almost everywhere (such filter functions we will not consider).

In the special case that $\alpha = 1$, the location is nonzero, and is equal to

$$\int_{\mathbb{R}} \psi(x+t) \log|\psi(x+t)\beta(x)|\mathrm{d}x$$

which is finite and independent of t so long as β has unit period.

Example – Infinite Order Moving Averages If one takes the filter function to have a step function form, *i.e.*,

$$\psi(\cdot) := \sum_{j\in\mathbb{Z}} \psi_j 1_{(j,j+1]} \tag{3}$$

for an appropriate sequence $\{\psi_j\}$, then a quick calculation yields, upon defining $Z(s) := \mathbb{M}(-s, -s+1]$, the familiar infinite order moving average with i.i.d (independent and identically distributed) α-stable inputs:

$$X(t) = \sum_{j\in\mathbb{Z}} \psi_j Z(t-j). \tag{4}$$

Since the filter function given in Eq. (3) is continuous almost-everywhere and appropriately summable (we require that $\sum_j |\psi_j|^{\delta} < \infty$), this example is subsumed by the model (1). Extensive work on examples of this type (with heavy-tailed input random variables) has been done in Davis and Resnick (1985, 1986); see also McElroy and Politis (2002).

2.2 The Representation Lemma

Since a step function of the form (3) can approximate an arbitrary continuous function fairly nicely (at least if the derivative is bounded), an obvious question is: to what extent does the previous example explain the more general model given by Eq. (1)? The following "Representation Lemma" takes a step towards answering that question.

LEMMA 1 *Define a new collection of random variables* $\{X(t, j);\ j \in \mathbb{Z}\}$ *by*

$$X(t, j) := \int_j^{j+1} \psi(x+t)\mathbb{M}(\mathrm{d}x) = \int_{\mathbb{R}} 1_{(j,j+1]}(x)\psi(x+t)\mathbb{M}(\mathrm{d}x) \tag{5}$$

for any $t \in \mathbb{Z}$ (*or any* $t \in \mathbb{R}$ *also works*). *Then* $X(t)$ *is almost surely equal to* $\sum_{j\in\mathbb{Z}} X(t, j)$.

Remark 1 This Representation Lemma will motivate and provide a whole new technique for analyzing statistics of processes following Eq. (1). If we consider $X(t, j-t)$, which is equal in distribution to

$$\left(\int_0^1 |\psi(x+j)|^\alpha\right)^{1/\alpha} Z(t-j),$$

we see that the extent to which $\int_0^1 |\psi(x+j)|^\alpha \,\mathrm{d}x$ differs from ψ_j^α is the extent to which our model truly differs from an infinite order moving average.

Remark 2 This lemma, and the preceding discussion, can be generalized to a random field scenario, where j and t belong to the d-dimensional integer lattice $\mathbb{Z}^d$.

Proof of the Representation Lemma We will use the following useful notation throughout this proof and the paper: let $B = (0, 1]$ denote the half-open unit interval (if we are in $\mathbb{R}^d$, then let $B = (0, 1]^d$), so that $B + j$ is just the interval $(j, j+1]$. Now observe that for any $m \in \mathbb{N}$,

$$\sum_{|j|\le m} X(t, j) = \int_{\mathbb{R}} 1_{(-m,m+1]}(x)\psi(t+x)\mathbb{M}(\mathrm{d}x) \tag{6}$$

due to the linearity of the stochastic integral. Now the idea is to take the limit as $m \to \infty$ on both sides of Eq. (6). First we show that the sum of the series in Eq. (5) is finite almost surely. If $1 < \alpha < 2$, then

$$\mathbb{E}\left|\sum_{j\in\mathbb{Z}} X(t, j)\right| \le \sum_{j\in\mathbb{Z}} \mathbb{E}|X(t, j)| = \sum_{j\in\mathbb{Z}} \left(\int_B |\psi(x+t+j)|^\alpha \,\mathrm{d}x\right)^{1/\alpha} c_{\alpha,\beta}(1),$$

where,

$$(c_{\alpha,\beta}(p))^p = \frac{2^{p-1}\Gamma(1-p/\alpha)}{p\int_0^\infty u^{-p-1}\sin^2 u \,\mathrm{d}u}\left(1+\beta^2\tan^2\frac{\alpha\pi}{2}\right)^{p/(2\alpha)} \cos\left(\frac{p}{\alpha}\arctan\left(\beta\tan\frac{\alpha\pi}{2}\right)\right)$$

as defined in, Samorodnitsky and Taqqu (1994, p. 18). This number is bounded above by $c_{\alpha,1}(1) < \infty$. It remains to show that the sum is finite. By the Mean Value Theorem, there exists a $y \in B$ (which depends on t and j) such that

$$\int_B |\psi(x+t+j)|^\alpha \,\mathrm{d}x = |\psi(y+t+j)|^\alpha$$

(in the d-dimensional case, this must be done iteratively for each coordinate). Therefore

$$\sum_{j\in\mathbb{Z}}\left(\int_B |\psi(x+t+j)|^\alpha \,\mathrm{d}x\right)^{1/\alpha} = \sum_{j\in\mathbb{Z}} |\psi(y+t+j)| \le \sup_{x\in B}\sum_{j\in\mathbb{Z}} |\psi(x+t+j)|,$$

which is finite, as demonstrated below. Now if $\alpha \le 1$, then find $\delta < \alpha$ stated in the assumptions on ψ, and compute:

$$\mathbb{E}\left(\left|\sum_{j\in\mathbb{Z}} X(t,j)\right|^\delta\right) \le \sum_{j\in\mathbb{Z}} \mathbb{E}|X(t,j)|^\delta \le \sum_{j\in\mathbb{Z}}\left(\int_B |\psi(x+t+j)|^\alpha \,\mathrm{d}x\right)^{\delta/\alpha}(c_{\alpha,1}(\delta))^\delta.$$

Again the expectation is finite, since $\delta < \alpha$; as for the sum, note that exponentiation by α is a concave function, so by an inverse application of Jensen's inequality we obtain the upper bound

$$\sum_{j\in\mathbb{Z}}\left(\int_B |\psi(x+t+j)\,\mathrm{d}x\right)^\delta = \sum_{j\in\mathbb{Z}} |\psi(y+t+j)|^\delta$$

for some $y \in B$, again by the Mean Value Theorem. We show that this is finite (and by letting $\delta = 1$, we obtain the proof for the case above). Let $A := \{x \in B\colon \sum_{j\in\mathbb{Z}} |\psi(x+j)|^\delta = \infty\}$. Then

$$\int_{\mathbb{R}} |\psi(x)|^\delta \,\mathrm{d}x = \sum_{j\in\mathbb{Z}} \int_B |\psi(x+j)|^\delta \,\mathrm{d}x \ge \int_A \sum_{j\in\mathbb{Z}} |\psi(x+j)|^\delta \,\mathrm{d}x = \infty\cdot\lambda(A),$$

where λ denotes Lebesgue measure. Since $\psi \in \mathbb{L}_\delta$, the Lebesgue measure of the set A must be zero. So the sum is finite almost everywhere, which can be modified to "everywhere" without loss of generality by taking a continuous version. This establishes the almost sure finiteness of the representation. From this it follows that

$$\sum_{|j|\le m} X(t,j) \xrightarrow{\text{a.s.}} \sum_{j\in\mathbb{Z}} X(t,j)$$

as $m \to \infty$. On the other hand,

$$1_{(-m,m+1]}(x)\psi(\cdot+t) \longrightarrow \psi(\cdot+t)$$

in the space $\mathbb{L}_\alpha$ for each t, so that

$$\int_{(-m,m+1]} \psi(x+t)\mathbb{M}(\mathrm{d}x) \xrightarrow{\text{P}} \int_{\mathbb{R}} \psi(x+t)\mathbb{M}(\mathrm{d}x)$$

as $m \to \infty$ (see Samorodnitsky and Taqqu, 1994). Finally, pick any $\varepsilon > 0$, and compute:

$$P\left[\left|\sum_{j\in\mathbb{Z}} X(t,j) - X(t)\right| > \varepsilon\right] \le P\left[\left|\sum_{j\in\mathbb{Z}} X(t,j) - \sum_{|j|\le m} X(t,j)\right| > \frac{\varepsilon}{2}\right]$$

$$+ P\left[\left|\sum_{|j|\le m} X(t,j) - X(t)\right| > \frac{\varepsilon}{2}\right] \longrightarrow 0+0$$

by the previous calculations, as $m \to \infty$. This is true for each ε, so we obtain

$$P\left[\left|\sum_{j\in\mathbb{Z}} X(t,j) - X(t)\right| > 0\right] = 0$$

which is almost sure equality. ∎

2.3 Convergence of the Sample Mean

Let us now proceed to the asymptotics of the scaled sample mean. The sample mean is $(1/n)\sum_{t=1}^{n} X(t)$, but we will multiply this by $n^{1-1/\alpha}$. Thus, we will consider the limiting behavior of the scaled sample mean

$$S_n = n^{-1/\alpha} \sum_{t=1}^{n} X(t).$$

When the characteristic exponent α is greater than one, the limit theorems developed here will later be put to a statistical use – since $\alpha > 1$, the mean exists, and an estimator for the mean will be constructed from S_n. Our result we believe to be novel, though perhaps unsurprising; the interesting component may lie in the method of proof, which performs a careful dissection of the underlying stochastic process. For ease of presentation, we restrict ourselves to a one-dimensional index set, though all the proofs have been carefully established in the higher-dimensional categories.

First, recall that $n^{-1/\alpha}$ is the correct normalization for a sum of n i.i.d α-stable random variables, in that S_n has a non-degenerate limit as $n \to \infty$. In fact, there is equality in distribution to an α-stable for each n; we have recourse to asymptotics only in the more general case that our data are drawn from the α-stable Domain of Normal Attraction, see McElroy and Politis (2002) for a discussion. The major difference in our situation is the lack of independence. However, due to the wonderful structure of the stochastic integral model, an α-stable limit is still obtained, whose scale parameter depends intimately upon the filter function ψ.

THEOREM 1 *Consider a time series generated from the model given by Eq.* (1), *where* $0 < \alpha \leq 2$. *Then the sample mean has an α-stable limit*:

$$n^{-1/\alpha} \sum_{t=1}^{n} X(t) \stackrel{\mathcal{L}}{\Longrightarrow} S_\infty(\alpha). \tag{7}$$

The limit variable $S_\infty(\alpha)$ is α-stable, with scale and skewness parameters

$$\left(\int_B |\Psi(x)|^\alpha \,\mathrm{d}x\right)^{1/\alpha}, \quad \frac{\int_B (\Psi(x))^{\langle\alpha\rangle}\beta(x)\,\mathrm{d}x}{\int_B |\Psi(x)|^\alpha \,\mathrm{d}x},$$

respectively, where $\Psi(x) := \sum_{j\in\mathbb{Z}} \psi(j+x)$. *If* $\alpha \neq 1$ *the location parameter is zero, but if* $\alpha = 1$ *we have the following value for the location*:

$$-\frac{2}{\pi}\int_B \sum_{j\in\mathbb{Z}} \psi(x+j)\log|\psi(x+j)|\beta(x)\,\mathrm{d}x.$$

If either $\alpha \neq 1$ or $\alpha = 1$ and $\beta = 0$, we can represent the limit variable as a stochastic integral:

$$S_\infty(\alpha) = \int_B \Psi(x)\mathbb{M}(\mathrm{d}x).$$

Remark 3 In the Gaussian $\alpha = 2$ case, this limiting scale is proportional to the standard deviation, and its square easily works out to be

$$\sigma^2 = \frac{1}{2}\sum_{h\in\mathbb{Z}} R(h),$$

where $R(h) = 2\int_{\mathbb{R}} \psi(x)\psi(x+h)\,\mathrm{d}x$ is the covariance function of the process (so $2\sigma^2$ is the limiting variance). Although the codifference (which for $\alpha = 2$ is just the covariance) for symmetric α-stable moving average processes tends to zero (see Theorem 4.7.3 of Samorodnitsky and Taqqu, 1994), this is not sufficient to guarantee the finiteness of σ^2 – we must have recourse to our $\mathbb{L}_1$ assumption on ψ.

Remark 4 When the model is symmetric, *i.e.* $\beta(x) = 0$ for all x, then the location parameters are also zero for $\alpha = 1$. It is only when asymmetry is present that we must distinguish the case that $\alpha = 1$.

Proof We begin with the representation lemma:

$$X(t) = \sum_{j\in\mathbb{Z}} \int_{B+j} \psi(x+t)\mathbb{M}(\mathrm{d}x) = \sum_{j\in\mathbb{Z}} \int_{B+j-t} \psi(x+t)\mathbb{M}(\mathrm{d}x) \tag{8}$$

Notice that at the end of Eq. (8), we have made the change of discrete variable $j \mapsto j - t$, with the result that each summand has scale

$$\sigma_j = \left(\int_B |\psi(x+j)|^\alpha\,\mathrm{d}x\right)^{1/\alpha}$$

and skewness

$$\beta_j = \frac{\int_B (\psi(x+j))^{\langle\alpha\rangle}\beta(x)\,\mathrm{d}x}{\sigma_j^\alpha},$$

respectively. This trick is instrumental to the general thrust of our method. Now consider a level m truncation of this sum, and take the scaled sample mean of this approximating model:

$$n^{-1/\alpha}\sum_{t=1}^{n}\sum_{|j|\leq m}\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x) = n^{-1/\alpha}\sum_{t=1}^{n}\int \sum_{|j|\leq m} 1_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x).$$

The αth power of the scale of this object is

$$\begin{aligned}\frac{1}{n}\int\left|\sum_{|j|\leq m}\sum_{t=1}^{n}1_{B+j-t}\psi(x+t)\right|^\alpha \mathrm{d}x &= \frac{1}{n}\sum_{s\in\mathbb{Z}}\int_{B+s}\left|\sum_{|j|\leq m}\sum_{t=1}^{n}1_{B+j-t}\psi(x+t)\right|^\alpha \mathrm{d}x\\ &= \frac{1}{n}\sum_{s\in\mathbb{Z}}\int_{B}\left|\sum_{|j|\leq m}\sum_{t=1}^{n}1_{B+j-s-t}\psi(x+s+t)\right|^\alpha \mathrm{d}x.\end{aligned} \tag{9}$$

Observe that we must have $j - s - t = 0$, or else the summand with that index will be zero itself. Also, the integral is nonzero only for values of s between $1 - n - m$ and m. Assuming, at this point, that n is significantly larger than m, we have $n - 2m$ terms of the form

$$\frac{1}{n} \int_B |\psi(x - m) + \cdots + \psi(x + m)|^\alpha \, \mathrm{d}x$$

and an additional $4m$ terms, all of which are different, but have a $1/n$ factor. These latter terms being $o(1)$, the former group adds up to

$$\frac{n - 2m}{n} \int_B |\psi(x - m) + \cdots + \psi(x + m)|^\alpha \, \mathrm{d}x$$

which tends to

$$\int_B \left| \sum_{|j| \le m} \psi(x + j) \right|^\alpha \mathrm{d}x \tag{10}$$

as $n \to \infty$. Similar arguments show that the skewness parameter tends to

$$\frac{\int_B (\sum_{|j| \le m} \psi(x + j))^{\langle \alpha \rangle} \beta(x) \, \mathrm{d}x}{\int_B | \sum_{|j| \le m} \psi(x + j)|^\alpha \, \mathrm{d}x}. \tag{11}$$

Moreover, when $\alpha = 1$, it can be shown that the location parameter converges to

$$-\frac{2}{\pi} \int_B \sum_{|j| \le m} \psi(x + j) \log |\psi(x + j)| \beta(x) \, \mathrm{d}x. \tag{12}$$

From this we may conclude the convergence in distribution

$$n^{-1/\alpha} \sum_{t=1}^{n} \sum_{|j| \le m} \int_{B+j-t} \psi(x + t) \mathbb{M}(\mathrm{d}x) \overset{\mathcal{L}}{\Longrightarrow} S_\infty^m(\alpha), \tag{13}$$

where the limit $S_\infty^m(\alpha)$ is α-stable with scale, skewness and location parameters given by Eqs. (10), (11) and (12), respectively. Note that, unfortunately, the mode of convergence cannot be strengthened to probability; this would require that

$$\int \left| n^{-1/\alpha} \sum_{t=1}^{n} \sum_{|j| \le m} 1_{B+j-t} \psi(x + t) - 1_B \sum_{|j| \le m} \psi(x + j) \right|^\alpha \mathrm{d}x \longrightarrow 0, \tag{14}$$

which is false (see Remark 5).

Now it remains to let m tend to infinity on both sides of the convergence (13). The right-hand side converges in probability to $S_\infty(\alpha)$, the desired limit, since the respective parameters converge; for the left-hand side, more subtle methods are necessary to interchange the limits (see Brockwell and Davis, 1991). Taking the difference of the left-hand side with the scaled sample mean yields

$$n^{-1/\alpha} \sum_{t=1}^{n} \sum_{|j| > m} \int_{B+j-t} \psi(x + t) \mathbb{M}\, (\mathrm{d}x)$$

for which we compute its scale's αth power (we measure the scale parameter, since a sufficient condition for a stable random variable to tend to zero in probability, is that its scale parameter tends to zero). First suppose that $\alpha > 1$:

$$\sigma^\alpha = \frac{1}{n}\int \left|\sum_{|j|>m}\sum_{t=1}^{n} 1_{B+j-t}\psi(x+t)\right|^\alpha \mathrm{d}x$$

$$\leq \frac{1}{n}\left\{\int\left(\sum_{|j|>m}\left|\sum_{t=1}^{n} 1_{B+j-t}\psi(x+t)\right|\right)^\alpha \mathrm{d}x\right\}^{(1/\alpha)\alpha}$$

$$\leq \frac{1}{n}\left\{\sum_{|j|>m}\left(\int\left|\sum_{t=1}^{n} 1_{B+j-t}\psi(x+t)\right|^\alpha \mathrm{d}x\right)^{(1/\alpha)}\right\}^\alpha$$

$$= \frac{1}{n}\left\{\sum_{|j|m}\left(\sum_{t=1}^{n}\int_B |\psi(x+j)|^\alpha \mathrm{d}x\right)^{1/\alpha}\right\}^\alpha$$

$$= \left\{\sum_{|j|>m}\left(\int_B |\psi(x+j)|^\alpha \mathrm{d}x\right)^{1/\alpha}\right\}^\alpha$$

by using the triangle inequality and Minkowski inequality for integrals. Therefore the scale is

$$\sigma \leq \sum_{|j|>m}\left(\int_B |\psi(x+j)|^\alpha \mathrm{d}x\right)^{1/\alpha}$$

which converges to zero as $m \to \infty$, independently of n, since the summation over all js is finite (this was established in the proof of the Representation Lemma). As for the case that $\alpha \leq 1$, the triangle inequality gives

$$\sigma^\alpha = \frac{1}{n}\int \left|\sum_{|j|>m}\sum_{t=1}^{n} 1_{B+j-t}\psi(x+t)\right|^\alpha \mathrm{d}x$$

$$\leq \frac{1}{n}\int \sum_{|j|>m}\left|\sum_{t=1}^{n} 1_{B+j-t}\psi(x+t)\right|^\alpha \mathrm{d}x$$

$$= \frac{1}{n}\sum_{|j|>m}\sum_{t=1}^{n}\int_B |\psi(x+j)|^\alpha \mathrm{d}x$$

$$= \sum_{|j|>m}\int_B |\psi(x+j)|^\alpha \mathrm{d}x.$$

Again, by reference to calculations in the argument for the Representation Lemma, we see this too tends to zero as m grows. Thus,

$$\lim_{m\longrightarrow\infty}\limsup_{n\longrightarrow\infty} P\left[\left|n^{-1/\alpha}\sum_{t=1}^{n}\sum_{|j|>m}\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x)\right| > \varepsilon\right] = 0$$

for any $\varepsilon > 0$. This calculation allows us to swap the limits in m and n in Eq. (13), and hence

$$n^{-1/\alpha} \sum_{t=1}^{n} \int \psi(x+t)\mathbb{M}(\mathrm{d}x) \stackrel{\mathcal{L}}{\Longrightarrow} S_\infty(\alpha).$$

Lastly, it is easy to see, so long as $\alpha \neq 1$, that

$$S_\infty(\alpha) \stackrel{\mathcal{L}}{=} \int_B \Psi(x)\mathbb{M}\,(\mathrm{d}x)$$

(where $\stackrel{\mathcal{L}}{=}$ denotes equality in distribution) since the scale, skewness and location parameters are identical. ■

Remark 5 The mode of convergence stated in Theorem 1 cannot be strengthened to convergence in probability. Consider the truncated model: convergence in probability requires that line (14) hold, by Proposition 3.5.1 of Samorodnitsky and Taqqu (1994). Since, by a simple calculuation,

$$\begin{aligned} &n^{-1/\alpha} \sum_{t=1}^{n} \sum_{|j|\le m} \int_{B+j-t} \psi(x+t)\mathrm{M}\,(\mathrm{d}x) \\ &\quad = op(1) + n^{-1/\alpha} \sum_{t=1}^{n} \int_{B-t} \sum_{|j|\le m} \psi(x+j+t)\mathbb{M}\,(\mathrm{d}x), \end{aligned}$$

we may conclude that convergence in probability is equivalent to the statement that

$$\int_{\mathbb{R}} \left| n^{-1/\alpha} \sum_{t=1}^{n} 1_{B-t}(x) \sum_{|j|\le m} \psi(x+j+t) - 1_B(x) \sum_{|j|\le m} \psi(x+j) \right|^{\alpha} \mathrm{d}x \qquad (15)$$

tends to zero as $n \to \infty$. Now, Eq. (15) is actually equal to

$$\int_{\mathbb{R}} \left| n^{-1/\alpha} \sum_{t=1}^{n} 1_{B-t}(x) \sum_{|j|\le m} \psi(x+j+t) \right|^{\alpha} \mathrm{d}x + \int_{\mathbb{R}} \left| 1_B(x) \sum_{|j|\le m} \psi(x+j) \right|^{\alpha} \mathrm{d}x$$

because of the disjoint supports of the two functions that compose the integrand. The latter term certainly does not tend to zero as n increases, and hence convergence in probability is not possible.

2.4 Convergence of the Sample Variance

We consider the appropriately scaled sample variance:

$$n^{-2/\alpha} \sum_{t=1}^{n} X^2(t).$$

As before, notice the unusual rate, which only agrees with the usual sample variance statistic in the case that $\alpha = 2$. Moreover, no centering by the sample mean is considered. In fact, the above statistic is asymptotically the same (in probability) as a centered version so long as $\alpha < 2$; these ideas are developed more fully at the end of this section.

We should also point out, that the limit theorem below is already well-known, and has been demonstrated in Resnick *et al.* (1999). However, the methods we use are somewhat different, and will allow for a joint sample mean – sample variance convergence result. Since, for a symmetric $X(t)$, the random variable $X^2(t)$ is the product of a positive $\frac{\alpha}{2}$ stable random variable and an independent χ^2 on one degree of freedom,[1] we may expect the limit, in light of Theorem 1, to be positive $\alpha/2$ stable as well, with a scale parameter depending upon the filter function ψ. This is indeed the case.

THEOREM 2 *Consider a time series generated from the model given by* (1), *where* $0 < \alpha < 2$. *Then the sample variance has a positively skewed* $\alpha/2$*-stable limit*:

$$n^{-2/\alpha} \sum_{t=1}^{n} X^2(t) \stackrel{\mathcal{L}}{\Longrightarrow} C \cdot \int_B \Psi_2(x)\tilde{\mathbb{M}}\,(\mathrm{d}x), \tag{16}$$

where $\Psi_2(x) := \sum_{j\in\mathbb{Z}} \psi^2(j+x)$*, and* C *is a constant only depending on* α*. The new random measure* $\tilde{\mathbb{M}}$ *is an* $\alpha/2$ *stable random measure with skewness intensity one. The limit random variable will have location zero and skewness one, and scale parameter*

$$\left(\int_B |\Psi_2(x)|^{\alpha/2}\,\mathrm{d}x\right)^{2/\alpha}.$$

Remark 6 Note that we have excluded the $\alpha = 2$ case from this theorem. In fact, this case operates in a completely different fashion, as the limit would be the deterministic variance (assuming that the process has zero mean). For $\alpha < 2$, the variance does not exist, and thus the sample variance's convergence to a random limit should not be surprising. Even under this non-classical situation, self-normalization can still be carried out.

Proof In light of the fact that Theorem 1 was proved in generality for $0 < \alpha < 1$, this theorem follows almost as a corollary. We begin with the following fact, which can be gleaned from Propositions 2.1 and 4.3 of Resnick *et al.* (1999):

$$n^{-2/\alpha} \sum_{t=1}^{n} X^2(t) = o_P(1) + n^{-2/\alpha} \left(\frac{C_\alpha}{C_{\alpha/2}}\right)^{2/\alpha} \sum_{t=1}^{n} \int_{\mathbb{R}} \psi^2(x+t)\tilde{\mathbb{M}}(\mathrm{d}x) \tag{17}$$

for constants C_α defined by

$$C_\alpha := \left(\int_{\mathbb{R}} x^{-\alpha} \sin x \,\mathrm{d}x\right)^{-1}.$$

[1] According to samorodnitsky and Taqqu (1994, p. 21), any symmetric α-stable random variable X can be written as

$$X = \sqrt{AG},$$

where A is a positive, $\alpha/2$ stable random variable and G is an independent standard Gaussian. Thus X^2 is equal in distribution to AG^2.

In fact, $(C_\alpha / C_{\alpha/2})^{2/\alpha}$ is the constant C appearing in line (16) of Theorem 2. Next, we apply the Representation Lemma 1 to the stochastic process on the right-hand side of Eq. (17), with filter function ψ^2 and $\alpha/2$ stable random measure $\tilde{\mathbb{M}}$. As a result, we have:

$$n^{-2/\alpha} \sum_{t=1}^{n} \int_{\mathbb{R}} \psi^2(x+t)\tilde{\mathbb{M}}(\mathrm{d}x) = n^{-2/\alpha} \sum_{t=1}^{n} \sum_{j \in \mathbb{Z}} \int_{B+j-t} \psi^2(x+t)\tilde{\mathbb{M}}(\mathrm{d}x).$$

So now considering the truncated version

$$n^{-2/\alpha} \sum_{t=1}^{n} \sum_{|j| \le m} \int_{B+j-t} \psi^2(x+t)\tilde{\mathbb{M}}(\mathrm{d}x),$$

we apply the methods of Theorem 1 to obtain its scale parameter's convergence to

$$\left(\int_B \left| \sum_{|j| \le m} \psi^2(x+j) \right|^{\alpha/2} \mathrm{d}x \right)^{2/\alpha}.$$

The skewness parameter is one, which is identical with the skewness parameter of the proposed limit, and the location parameters are all zero since $\alpha/2 \neq 1$. Thus we have the truncated convergence for each m:

$$n^{-2/\alpha} \sum_{t=1}^{n} \sum_{|j| \le m} \int_{B+j-t} \psi^2(x+t)\tilde{\mathbb{M}}(\mathrm{d}x) \overset{\mathcal{L}}{\Longrightarrow} \int_B \sum_{|j| \le m} \psi^2(x+j)\tilde{\mathbb{M}}(\mathrm{d}x)$$

so that we now need to let $m \to \infty$. But the proof of this is as in Theorem 2. Therefore we obtain the stated convergence in distribution. ∎

As a last note, we may wish to expand this theorem to include the $\alpha = 2$ case. In order to do this, we must "mean-correct" the scaled sample variance statistic; this has no effect on the limit if $\alpha < 2$, but if $\alpha = 2$ we have almost sure convergence to the variance of the Gaussian process. Actually, the mean-correction for $\alpha = 2$ is not neccessary for a mean zero process; therefore, let us consider a location-shifted model:

$$Z(t) := X(t) + \mu = \int_{\mathbb{R}} \psi(x+t)\mathbb{M}(\mathrm{d}x) + \mu. \tag{18}$$

This is the appropriate model for stable stationary dependent time series with nonzero location; in fact, these theorems have the statistical application of location estimation.

COROLLARY 1 *Consider a time series $Z(t)$ generated from the model given by Eq.* (18), *where $0 < \alpha \le 2$. Then the mean-corrected scaled sample variance either has a positively skewed $\alpha/2$-stable limit, or converges almost surely to the variance of $Z(t)$, that is,*

$$n^{-2/\alpha} \sum_{t=1}^{n} (Z(t) - \overline{Z})^2 \overset{\mathcal{L}}{\Longrightarrow} C \cdot \int_B \Psi_2(x)\tilde{\mathbb{M}}(\mathrm{d}x) \quad \text{if } 0 < \alpha < 2 \tag{19}$$

or

$$n^{-2/\alpha} \sum_{t=1}^{n} (Z(t) - \overline{Z})^2 \xrightarrow{a.s.} \mathrm{var}(Z(0)) \quad \text{if } \alpha = 2.$$

The limit random variable and the constant C in Eq. (19) *above are defined in Theorem* 2. *Also*, $\overline{Z} = (1/n)\sum_{t=1}^{n} Z(t)$ *denotes the sample mean.*

Proof First we establish that the sample second moments of the location zero and location μ time series have the same limit, as long as $\alpha < 2$:

$$\begin{aligned} n^{-2/\alpha}\sum_{t=1}^{n} Z^2(t) &= n^{-2/\alpha}\sum_{t=1}^{n}(X^2(t) + 2\mu X(t) + \mu^2) \\ &= n^{-2/\alpha}\left(\sum_{t=1}^{n} X^2(t) + 2\mu\sum_{t=1}^{n} X(t) + n\mu^2\right) \\ &= n^{-2/\alpha}\sum_{t=1}^{n} X^2(t) + 2\mu n^{-2/\alpha}\sum_{t=1}^{n} X(t) + \mu^2 n^{1-2/\alpha}. \end{aligned}$$

Notice that since $\alpha < 2$ the exponent $1 - 2/\alpha$ is negative. Now taking the difference of this expression with $n^{-2/\alpha}\sum_{t=1}^{n} X^2(t)$ yields

$$2\mu n^{-2/\alpha}\sum_{t=1}^{n} X(t) + \mu^2 n^{1-2/\alpha}$$

which tends to zero in probability, since $\sum_{t=1}^{n} X(t) = O_P(n^{1/\alpha})$. This shows that

$$n^{-2/\alpha}\sum_{t=1}^{n} X^2(t) = o_P(1) + n^{-2/\alpha}\sum_{t=1}^{n} Z^2(t)$$

as desired.

Next, we show that the sample variance for the location model is asymptotically the same as the sample second moment:

$$\begin{aligned} n^{-2/\alpha}\sum_{t=1}^{n}(Z(t) - \overline{Z})^2 &= n^{-2/\alpha}\sum_{t=1}^{n}(X(t) - \overline{X})^2 \\ &= n^{-2/\alpha}\left(\sum_{t=1}^{n} X^2(t) - n\overline{X}^2\right) \\ &= o_P(1) + n^{-2/\alpha}\left(\sum_{t=1}^{n} Z^2(t) - n\overline{X}^2\right) \end{aligned}$$

so that

$$\begin{aligned} n^{-2/\alpha}\sum_{t=1}^{n} Z^2(t) - n^{-2/\alpha}\sum_{t=1}^{n}(Z(t) - \overline{Z})^2 &= n^{-2/\alpha} n\overline{X}^2 + o_P(1) \\ &= \frac{1}{n}\left(n^{-1/\alpha}\sum_{t=1}^{n} X(t)\right)^2 + o_P(1) \\ &= O_P\left(\frac{1}{n}\right) \longrightarrow 0 \end{aligned}$$

as desired.

In the case that $\alpha = 2$, it is well known that

$$\frac{1}{n}\sum_{t=1}^{n}(Z(t)-\overline{Z})^2 \xrightarrow{\text{a.s.}} \mathrm{var}(Z(0)).$$

Now the variance is twice the squared scale for $\alpha = 2$ stable random variables – in this case the scale is simply

$$\left(\int_{\mathbb{R}} |\psi(x)|^{\alpha}\,\mathrm{d}x\right)^{1/\alpha} = \sqrt{\int_{\mathbb{R}} \psi^2(x)\,\mathrm{d}x};$$

therefore the variance $\mathrm{var}(Z(0))$ is

$$2\int_{\mathbb{R}} \psi^2(x)\,\mathrm{d}x. \qquad \blacksquare$$

3 SELF-NORMALIZED SAMPLE MEAN

We now move forward to a more ambitious goal – to describe the joint asymptotics of sample mean and sample variance. If Theorems 1 and 2 can be sensibly concatenated into a joint convergence result, then certain continuous functions – such as the quotient – of the sample mean and square root sample variance can be approximately determined. In terms of motivation, we remind the reader that classically, the division of sample mean minus the true mean by the square root of an estimate of variance allows us the freedom of not "pre-determining" the scale. In the case of heavy-tailed distributions, this self-normalization procedure cannot remove the limit distribution's dependence upon certain unknown model parameters; nevertheless, it is a beneficial procedure, since it effectively removes the characteristic exponent α from the rate of convergence, thereby providing us with a genuine $\sqrt{n}$-convergent statistic. To be more specific, the rates $n^{-1/\alpha}$ and $n^{-2/\alpha}$ of growth for scaled sample mean and scaled sample variance respectively will cancel in the quotient.

The paper of Logan *et al.* (1973) first obtained joint weak convergence results for the sample mean and sample second moment, for an i.i.d α-stable model. However, the quotient's limit was not a well-known random variable; even deriving a closed form expression for the probability density function of a generic stable random variable is still an open problem (for most characteristic exponents). Without knowledge of the limit quantiles, it is not possible to form confidence intervals for the mean. *Subsampling* methods offer a practical solution of these difficulties; see Politis *et al.* (1999). A full discussion will be given in the latter section on applications, but the basic idea is to use subsampling to approximate the limit distribution itself. When this procedure is valid (*e.g.*, it certainly works for *m*-dependent time series, and more generally holds for strong mixing processes), it provides a practical method for constructing approximate confidence intervals for the mean. This is the statistical context in which we wish to view the subsequent theorem.

To fix ideas, we consider the location model (18) introduced just before Corollary 1. Thus the centered scaled sample mean is

$$n^{1-1/\alpha}(\overline{Z}-\mu) = n^{-1/\alpha}\sum_{t=1}^{n}(Z(t)-\mu),$$

which is actually identical with the uncentered scaled sample mean S_n for the location zero model. In order to include the $\alpha = 2$ case, it is appropriate to center the scaled sample variance, as discussed earlier.

THEOREM 3 *Consider a time series generated from the model given by* (18), *where* $0 < \alpha \leq 2$; *if* $\alpha = 1$, *we assume in addition that* $\beta = 0$. *Then the centered sample mean and mean-corrected sample variance converge jointly in distribution to stable limits*:

$$\left(n^{-1/\alpha}\sum_{t=1}^{n}(Z(t)-\mu), n^{-2/\alpha}\sum_{t=1}^{n}(Z(t)-\overline{Z})^2(t)\right)$$
$$\stackrel{\mathcal{L}}{\Longrightarrow}\left(\int_B \Psi(x)\mathbb{M}(\mathrm{d}x), C\cdot\int_B \Psi_2(x)\tilde{\mathbb{M}}(\mathrm{d}x)\right). \tag{20}$$

For $\alpha = 2$ *we must replace the second component of the limit by* var(Z(0)). *The same notations as in Theorems* 1 *and* 2 *have been used.*

Remark 7 It may seem odd at first that the filter functions of the limit random variables are supported on the unit cube. The unit periodicity of Ψ and Ψ_2 is due to the fact that the integers index our process $X(t)$.

Proof Observe that it is sufficient to examine

$$\left(n^{-1/\alpha}\sum_{t=1}^{n}X(t), n^{-2/\alpha}\sum_{t=1}^{n}X^2(t)\right)$$

in view of Corollary 1. As mentioned in Remark 5, the modes of convergence cannot be strengthened from distribution to probability – the latter mode would have facilitated the immediate proof of this theorem, as a joint convergence in probability. The absence of such a firmer result has promulgated a deeper investigation of the stochastic structure.

As a preliminary remark, note that this theorem is well-known for independent random variables (*i.e.*, when ψ is supported in the unit interval) (see Logan *et al.*, 1973); and also for linear combinations of such (*i.e.*, when ψ is a step function) (see McElroy and Politis, 2002). The basic concept of this proof is to cut apart the stochastic process until a k-dependent sequence is revealed, for which joint convergence of sample mean and sample variance is valid, and then to paste the process back up again.

We begin with a valuable lemma, which is an easy corollary of some results of Davis and Hsing (1995). First, note that by definition a random vector $\mathbf{X}$ is α-stable if for any positive constants A, B and a constant vector $\mathbf{D}$

$$A\mathbf{X}^{(1)} + B\mathbf{X}^{(2)} \stackrel{\mathcal{L}}{=} (A^\alpha + B^\alpha)^{1/\alpha}\mathbf{X} + \mathbf{D},$$

where $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ are independent copies of $\mathbf{X}$. See, Samorodnitsky and Taqqu (1994, p. 58) for additional details.

LEMMA 2 *Let k be any positive integer, and suppose $X_1, \ldots, X_n$ is a k-dependent stationary time series, such that $\{X_1, \ldots, X_k\}$ forms an α-stable random vector for some characteristic exponent $\alpha \in (0, 2)$. If $\alpha = 1$, we also assume that the random variables are symmetric. Then*

$$\left(n^{-1/\alpha}\sum_{t=1}^{n}X_t, n^{-2/\alpha}\sum_{t=1}^{n}X_t^2\right)$$

converges jointly in distribution to $(S, \tilde{S})$, *where S is an* α*-stable random variable, and* $\tilde{S}$ *is a positively skewed* $\alpha/2$*-stable random variable.*

Proof Since $\{X_h, \ldots, X_{h+k-1}\}$ is a stable random vector for all h by stationarity, we have (letting $h = 1$)

$$\lim_{\lambda \longrightarrow \infty} \lambda^{\alpha} P\left[\|\mathbf{X}\| > \lambda, \frac{\mathbf{X}}{\|\mathbf{X}\|} \in A\right] = \tilde{C}_{\alpha} \Gamma_{\|\cdot\|}(A)$$

by Theorem 4.4.8 of Samorodnitsky and Taqqu (1994), whence the notation is taken: $\mathbf{X} = \{X_1, \ldots, X_k\}$, $\|\cdot\|$ is some norm on $\mathbb{R}^k$, and A is a Borel set in the space $\mathcal{S}_d^{\|\cdot\|}$ – the unit sphere under the topology generated by the norm $\|\cdot\|$. $\tilde{C}_{\alpha}$ is a positive constant, and the measure $\Gamma_{\|\cdot\|}$ is related to the spectral measure Γ of the random vector $\mathbf{X}$. Substituting $\lambda = tn^{1/\alpha}$ for some $t > 0$ and $x_n = n^{1/\alpha}$, we obtain

$$\lim_{n \longrightarrow \infty} nP\left[\|\mathbf{X}\| > tx_n, \frac{\mathbf{X}}{\|\mathbf{X}\|} \in A\right] = t^{-\alpha} \tilde{C}_{\alpha} \Gamma_{\|\cdot\|}(A).$$

Now let Θ be a random variable concentrated on $\mathcal{S}_d^{\|\cdot\|}$ with distribution $\tilde{C}_{\alpha}\Gamma_{\|\cdot\|}$; then we have

$$\lim_{n \longrightarrow \infty} nP\left[\|\mathbf{X}\| > tx_n, \frac{\mathbf{X}}{\|\mathbf{X}\|} \in A\right] = t^{-\alpha} P[\Theta \in A]$$

which implies that $\mathbf{X}$ is "jointly regulary varying" with index α, as defined in Davis and Hsing (1995). Now from example 5.4 of Davis and Hsing (1995), we see that the conditions of their Theorem 3.1 are satisfied; therefore, the assumptions of their example 5.3 are met, and we obtain the desired joint convergence with the choice $r = 2$, for values of α in $(0, 1) \cup (1, 2)$. The joint convergence result in example 5.3 will also hold true if the marginal distribution is symmetric, since that will ensure that the term $\mathbb{E}S_n(0, 1]$ in line (3.2) of Davis and Hsing (1995, p. 895) will be zero. ■

With this lemma in hand, we now proceed with the proof of our theorem. Consider the truncated model, as in Theorem 1; we define $W(t)$ as follows:

$$W(t) := \sum_{|j| \leq m} \int_{B+j-t} \psi(x + t)\mathbb{M}(\mathrm{d}x)$$

so that $W(1), \ldots, W(n)$ are stationary (they are identically distributed α-stable random variables with scale $\sigma_j = (\int_B |\psi(x + j)|^{\alpha}\, \mathrm{d}x)^{1/\alpha}$ and skewness $\int_B (\psi(x + j))^{<\alpha>} \beta(x)\, \mathrm{d}x / \sigma_j^{\alpha}$ and $(2m + 1)$-dependent. It is easy to see that $\mathbf{W} = \{W(1), \ldots, W(n)\}$ forms a strictly α-stable random vector, for any n. In fact, a sufficient condition for this is given by Theorem 2.1.5 of Samorodnitsky and Taqqu (1994) – for coefficients $b_1, \ldots, b_n$, the dot product $\sum_{t=1}^{n} b_t W(t)$ is strictly α-stable with scale

$$\left(\int \left\| \sum_{t=1}^{n} b_t \sum_{|j| \leq m} 1_{B+j-t}(x) \psi(x + t) \right\|^{\alpha} \mathrm{d}x \right)^{1/\alpha}$$

and skewness

$$\frac{\int (\sum_{t=1}^{n} b_t \sum_{|j| \leq m} 1_{B+j-t}(x) \psi(x + t))^{<\alpha>} \beta(x)\, \mathrm{d}x}{\int |\sum_{t=1}^{n} b_t \sum_{|j| \leq m} 1_{B+j-t}(x) \psi(x + t)|^{\alpha}\, \mathrm{d}x},$$

and location zero (either $\alpha \neq 1$, which implies location zero, or $\alpha = 1$ and $\beta = 0$, which likewise implies location zero). Applying Lemma 2 with $k = 2m + 1$, we obtain

$$\left(n^{-1/\alpha} \sum_{t=1}^{n} W(t), n^{-2/\alpha} \sum_{t=1}^{n} W^2(t) \right) \stackrel{\mathcal{L}}{\Longrightarrow} (S_\infty^m(\alpha), \tilde{S}_\infty^m(\alpha)). \tag{21}$$

The m subscript on the limit random variables reminds us that this result is for the truncated model; we have yet to let m tend to infinity. By restricting to the first component in Eq. (21) above, we see that in distribution

$$S_\infty^m(\alpha) = \int_B \sum_{|j| \leq m} \psi(x + j)\mathbb{M}(\mathrm{d}x)$$

which tends in probability to $\int_B \Psi(x)\mathbb{M}(\mathrm{d}x)$ as $m \to \infty$, as demonstrated in the proof of Theorem 1. The analysis of $W^2(t)$ requires considerably more scrutiny.

Asymptotics of the Second Component First we square out $W(t)$ and obtain

$$W^2(t) = \sum_{|j| \leq m} \left(\int_{B+j-t} \psi(x + t)\mathbb{M}(\mathrm{d}x) \right)^2 + \sum_{i \neq j} \left(\int_{B+i-t} \psi(x + t)\mathbb{M}(\mathrm{d}x) \right) \left(\int_{B+j-t} \psi(x + t)\mathbb{M}(\mathrm{d}x) \right);$$

here we introduce the notation $V(j, t) := \int_{B+j-t} \psi(x + t)\mathbb{M}(\mathrm{d}x)$ for brevity. Thus considering the whole normalized sum, we have:

$$n^{-2/\alpha} \sum_{t=1}^{n} W^2(t) = n^{-2/\alpha} \sum_{t=1}^{n} \sum_{|j| \leq m} V^2(j, t) + n^{-2/\alpha} \sum_{t=1}^{n} \sum_{i \neq j} V(i, t)V(j, t).$$

The first term is, up to terms that are $o_P(1)$, equal to

$$C \cdot n^{-2/\alpha} \sum_{t=1}^{n} \sum_{|j| \leq m} \int_{B+j-t} \psi^2(x + t)\tilde{\mathbb{M}}(\mathrm{d}x)$$

using Proposition 4.3 of Resnick *et al.* (1999) – it holds for each j, and therefore for the sum over the finite collection of js. C is the same constant that appeared in the proof of Theorem 2 – it is no way changed (or dependent on j), since it only depended on α and not on the underlying filter function. As seen in the proof of Theorem 2, the above term converges weakly to $C \int_B \sum_{|j| \leq m} \psi^2(x + j)\tilde{\mathbb{M}}(\mathrm{d}x)$. The second term tends to zero in probability – let us consider one choice of $i \neq j$. Since the indices are distinct, $V(i, t)$ and $V(j, t)$ are independent random variables, for each t, since the support of their filter functions are disjoint. A Markov Inequality argument can be used to show negligibility: choose a positive $\delta \leq 1$ such that

$\alpha \in (\delta, 2\delta)$. For instance, if $\alpha \in (1, 2)$ select $\delta = 1$; if $\alpha \in (0, 1]$, select $\delta = (2/3)\alpha$. If $\alpha = 2$, we use a separate argument. Now the δth absolute moment is

$$E\left|n^{-2/\alpha}\sum_{t=1}^{n} V(i,t)V(j,t)\right|^{\delta} = n^{-2(\delta/\alpha)} E\left|\sum_{t=1}^{n} V(i,t)V(j,t)\right|^{\delta}$$

$$\leq n^{-2(\delta/\alpha)} \sum_{t=1}^{n} E|V(i,t)|^{\delta} E|V(j,t)|^{\delta}$$

$$= n^{1-2\delta/\alpha} \cdot (E|V(0,0)|^{\delta})^2 \longrightarrow 0$$

by the Markov Inequality and choice of δ. But if $\alpha = 2$, we have

$$\frac{1}{n}\sum_{t=1}^{n} V(i,t)V(j,t) \xrightarrow{\text{a.s.}} E[V(i,t)V(j,t) = 0$$

by the Strong Law of Large Numbers, and our location zero assumption on our model. In summary, we have

$$n^{-2/\alpha}\sum_{t=1}^{n} W^2(t) = o_P(1) + n^{-2/\alpha}\sum_{t=1}^{n}\sum_{|j|\leq m}\int_{B+j-t}\psi^2(x+t)\tilde{\mathbb{M}}(\mathrm{d}x),$$

and we already know the weak limit of the right-hand side – namely

$$\tilde{S}_{\infty}^{m}(\alpha) = C\int_{B}\sum_{|j|\leq m}\psi^2(x+j)\tilde{\mathbb{M}}(\mathrm{d}x)$$

holds in the distributional sense.

As a final step, we must demonstrate the relation of convergence (21) to the asymptotics of $(n^{-1/\alpha}\sum_{t=1}^{n} X(t), n^{-2/\alpha}\sum_{t=1}^{n} X^2(t))$. In particular, we examine

$$n^{-1/\alpha}\sum_{t=1}^{n} X(t) - n^{-1/\alpha}\sum_{t=1}^{n} W(t)$$

and show that the limit as m tends to infinity of the limit superior as $n \to \infty$ in probability of the difference is zero; this was already verified in the proof of Theorem 1. Turning to the scaled sample second moment, we have

$$n^{-2/\alpha}\sum_{t=1}^{n} X^2(t) + o_P(1) = C \cdot n^{-2/\alpha}\sum_{t=1}^{n}\int_{\mathbb{R}}\psi^2(x+t)\tilde{\mathbb{M}}(\mathrm{d}x)$$

$$= C \cdot n^{-2/\alpha}\sum_{t=1}^{n}\sum_{j\in\mathbb{Z}}\int_{B+j-t}\psi^2(x+t)\tilde{\mathbb{M}}(\mathrm{d}x)$$

by Proposition 4.3 of Resnick *et al.* (1999) and the representation lemma. The difference of $C \cdot n^{-2/\alpha}\sum_{t=1}^{n}\sum_{j\in\mathbb{Z}}\int_{B+j-t}\psi^2(x+t)\tilde{\mathbb{M}}(\mathrm{d}x)$ with $n^{-2/\alpha}\sum_{t=1}^{n} W_t^2$ is (again, up to o_P (1) terms)

$$C \cdot n^{-2/\alpha}\sum_{|j|>m}\sum_{t=1}^{n}\int_{B+j-t}\psi^2(x+t)\tilde{\mathbb{M}}(\mathrm{d}x),$$

which is adequately controlled, as exhibited in the proof of Theorem 2. Hence

$$\left(n^{-1/\alpha}\sum_{t=1}^{n} X(t), n^{-2/\alpha}\sum_{t=1}^{n} X^2(t)\right) \overset{\mathcal{L}}{\Longrightarrow} \lim_{m\longrightarrow\infty}(S_\infty^m(\alpha), \tilde{S}_\infty^m(\alpha)).$$

Now it is easy to see that $(S_\infty^m(\alpha), \tilde{S}_\infty^m(\alpha))$ is jointly equal in distribution to

$$\left(\int_B \sum_{|j|\le m} \psi(x+j)\mathbb{M}(\mathrm{d}x), C\cdot\int_B \sum_{|j|\le m} \psi^2(x+j)\mathbb{M}(\mathrm{d}x)\right),$$

which converges in probability as $m \to \infty$ to

$$\left(\int_B \Psi(x)\mathbb{M}(\mathrm{d}x), C\cdot\int_B \Psi_2(x)\tilde{\mathbb{M}}(\mathrm{d}x)\right)$$

as desired. ■

Our next consideration, which closes the discussion of sample mean and sample variance, is a corollary that provides the application of location inference. Generally we are interested in the sampling distribution of a *root*, which is some function of a parameter and its estimator. Here we consider the root $\hat{\mu} - \mu$, where $\hat{\mu}$ is some estimate of the parameter μ; knowledge of the quantiles of the root's distribution facilitates the construction of confidence intervals. If we consider $\bar{Z}$ as our estimator $\hat{\mu}$ of the unknown location μ, then $\hat{\mu}$ is consistent as long as $\alpha > 1$, in which case the corollary below can be used to construct confidence intervals for μ.

COROLLARY 2 *Consider a time series generated from the model given by Eq.* (18), *where* $0 < \alpha \le 2$. *Let* $\hat{\sigma}^2$ *be the sample variance, and* $\hat{\sigma}$ *its square root, i.e.*

$$\hat{\sigma} = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n}(Z(t)-\bar{Z})^2}.$$

Then the following standardized root $T_n(\mu)$ *has a nondegenerate weak limit. When* $0 < \alpha < 2$ (*if* $\alpha = 1$, *we also assume that* $\beta = 0$), *we have*

$$T_n(\mu) := \frac{\bar{Z}-\mu}{\hat{\sigma}/\sqrt{n}} \overset{\mathcal{L}}{\Longrightarrow} \frac{\int_B \Psi(x)\mathbb{M}(\mathrm{d}x)}{\sqrt{C\cdot\int_B \Psi_2(x)\tilde{\mathbb{M}}(\mathrm{d}x)}}. \tag{22}$$

If $\alpha = 2$, *the whole denominator in Eq.* (22) *should be replaced by* $\sqrt{\mathrm{var}(Z(0))}$. *The same notations as in Theorems* 1 *and* 2 *have been used.*

Proof We begin with some algebra:

$$\begin{aligned}\frac{\bar{Z}-\mu}{\hat{\sigma}/\sqrt{n}} &= \frac{\sqrt{n}\sqrt{n-1}}{n}\frac{\sum_{t=1}^{n}(Z(t)-\mu)}{\sqrt{\sum_{t=1}^{n}(Z(t)-\bar{Z})^2}}\\ &= \sqrt{1-\frac{1}{n}}\frac{n^{-1/\alpha}\sum_{t=1}^{n}(Z(t)-\mu)}{\sqrt{n^{-2/\alpha}\sum_{t=1}^{n}(Z(t)-\bar{Z})^2}}.\end{aligned}$$

The $\sqrt{1 - 1/n}$ can be ignored in the face of n's approach to infinity; now applying the functional $f(x, y) = x/\sqrt{y}$ to the result (20) and invoking the continuous mapping theorem produces the desired conclusion. ■

This result is valid for $\alpha \leq 1$, but in this case it is of little use, as $\bar{Z}$ is not consistent for μ. Note finally that we cannot assume detailed knowledge of the limit distribution in Corollary 2, and thus computing its quantiles is only a theoretical possibility. In the final section of the paper, we explore how subsampling methods can estimate the limit distribution's quantiles.

4 SELF-NORMALIZED PERIODOGRAM

4.1 The Heavy-Tailed Periodogram

Let us now address the next major thrust of this paper: the periodogram. In classical time series analysis, when the random variables have finite variance, the spectral density is defined to be the Fourier Transform of the autocovariance sequence. For a sample of size n, this is estimated by the *periodogram* – the Discrete Fourier Transform of the sample autocovariance function – and evaluated at frequencies of the form $2\pi j/n$ for $j = 0, 1, \ldots, n-1$. The formula can be reduced (if we remove centerings) to

$$\left| n^{-1/2} \sum_{t=1}^{n} X(t) e^{-it\omega} \right|^2$$

for frequencies $\omega \in (0, 2\pi]$ and $i = \sqrt{-1}$. Typically this estimator of 2π times the spectral density is smoothed over a band of neighboring frequencies to obtain consistency.

When heavy-tails are present in the data, we can no longer define the spectral density, since the autocovariance sequence does not exist. In the case of linear processes, a theory for *heavy-tailed spectral densities* has been developed – see Klüppelberg and Mikosch (1992,
Q3 1993). Here we will be interested in the asymptotics of the *heavy-tailed periodogram*, which is defined to be

$$I(\omega) := n^{-2/\alpha} \left| \sum_{t=1}^{n} X(t) e^{-it\omega} \right|^2. \tag{23}$$

Notice that the rate n^{-1} has been replaced by $n^{-2/\alpha}$ – this is the appropriate rate of convergence to a non-degenerate limit distribution.

Given the results in the classical and linear heavy-tailed models, we do not expect a non-random limit. In fact, we cannot easily separate the limit into a product of a stochastic term and constants containing information about the data's periodicities:

THEOREM 4 *Consider the locationless model given by* (1), *and let* $0 < \alpha < 2$. *The heavy-tailed periodogram* $I(\omega)$, *for any* $\omega \in 2\pi\mathbb{Q}$, *converges in distribution to a non-degenerate limit. In particular, let* $S_j = \int_B \psi(x + j)\mathbb{M}(\mathrm{d}x)$; *then*

$$n^{-2/\alpha} \left| \sum_{t=1}^{n} X(t) e^{-it\omega} \right|^2 \stackrel{\mathcal{L}}{\Longrightarrow} (\gamma_c^2 + \gamma_s^2) \left| \sum_{j \in \mathbb{Z}} e^{-ij\omega} S_j \right|^2, \tag{24}$$

where γ_c and γ_s are constants defined as follows:

$$\gamma_c = g^{-1/\alpha}\left(\sum_{h=0}^{g-1} |\cos h\omega|^{\alpha}\right)^{1/\alpha}, \quad \gamma_s = g^{-1/\alpha}\left(\sum_{h=0}^{g-1} |\sin h\omega|^{\alpha}\right)^{1/\alpha},$$

where $\omega = 2\pi u/g$ for u and g relatively prime. In the above, $\mathbb{Q}$ denotes the set of rational numbers.

Q3 *Remark 8* A similar result has been established by Klüppelberg and Mikosch (1992); they look at an infinite order moving average model, and consider joint results for multiple periodicities of various types. The Klüppelberg – Mikosch setting corresponds to the case where ψ is a step function $\psi = \sum_{j\in\mathbb{Z}} 1_B +_j \Psi_j$; hence $S_j = \psi_j \mathbb{M}(B)$ for all j, and the limit in Eq. (24) reduces to

$$((\gamma_c \mathbb{M}(B))^2 + (\gamma_s \mathbb{M}(B))^2)\left|\sum_{j\in\mathbb{Z}} e^{-ij\omega}\psi_j\right|^2; \tag{25}$$

compare with Theorems 2.5 and 5.2 of the aforementioned paper. The second term in the above product (25) is $|\sum_{j\in\mathbb{Z}} e^{-ij\omega}\psi_j|^2$, which is actually related to a heavy-tailed analogue of the spectral density (for details, see Section 4.3).

Remark 9 To formulate this result in the context of random fields, we consider sampling the data from a rectangular subset K of the d-dimensional integer lattice $\mathbb{Z}^d$. Now the random measure, filter functions, and skewness intensity are defined on $\mathbb{R}^d$. If we let K be the cube $(0, n_1] \times \cdots \times (0, n_d] =$ intersected with the integer lattice $\mathbb{Z}^d$, the total number of observations is $N = n_1 \times \cdots \times n_d$. We use the following shorthand for sums:

$$\sum_{t=1}^{n} X(t) = \sum_{t_1=1}^{n_1}\cdots\sum_{t_d=1}^{n_d} X(t_1, \ldots, X_d).$$

Let us consider periodogram ordinates $\omega = (\omega_1, \ldots, \omega_d)$ – we need a scalar argument for the exponential function in the Fourier transform, so we take the dot product of t and ω, which will be denoted by $t'\omega$. So the heavy-tailed periodogram is defined as follows:

$$I(\omega) := N^{-2/\alpha}\left|\sum_{t=1}^{n} X(t)e^{-it'\omega}\right|^2$$

Then the result of Theorem 4 can be extended to random fields, so long as each frequency component ω_i is a rational multiple of 2π. The resulting limit will be a constant times

$$\left|\sum_{j\in\mathbb{Z}^d} e^{-ij'\omega} S_j\right|^2$$

with $S_j = \int_B \psi(x+j)\mathbb{M}(\mathrm{d}x)$.

Proof As indicated in the statement of the theorem, we will consider ω of the form $2\pi u/g$ where u and g are relatively prime integers, and g is positive. First, observe that

$$n^{-2/\alpha}\left|\sum_{t=1}^{n} X(t)e^{-it\omega}\right|^2 = \left(n^{-1/\alpha}\sum_{t=1}^{n} X(t)\cos t\omega\right)^2 + \left(n^{-1/\alpha}\sum_{t=1}^{n} X(t)\sin t\omega\right)^2.$$

Thus, it is sufficient to establish the joint convergence of the portions $n^{-1/\alpha}\sum_{t=1}^{n} X(t)\cos t\omega$ and $n^{-1/\alpha}\sum_{t=1}^{n} X(t)\sin t\omega$. Using the representation lemma, we have

$$n^{-1/\alpha}\sum_{t=1}^{n} X(t)\cos t\omega = n^{-1/\alpha}\sum_{t=1}^{n}\sum_{j\in\mathbb{Z}}\cos t\omega\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x), \tag{26}$$

$$n^{-1/\alpha}\sum_{t=1}^{n} X(t)\sin t\omega = n^{-1/\alpha}\sum_{t=1}^{n}\sum_{j\in\mathbb{Z}}\sin t\omega\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x). \tag{27}$$

By simple trigonometry, the first line is equal to

$$\begin{aligned} n^{-1/\alpha}\sum_{j\in\mathbb{Z}}\cos j\omega\sum_{t=1}^{n}\cos(t-j)\omega\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x) \\ -n^{-1/\alpha}\sum_{j\in\mathbb{Z}}\sin j\omega\sum_{t=1}^{n}\sin(t-j)\omega\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x) \end{aligned} \tag{28}$$

and the second line is

$$\begin{aligned} n^{-1/\alpha}\sum_{j\in\mathbb{Z}}\sin j\omega\sum_{t=1}^{n}\cos(t-j)\omega\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x) \\ -n^{-1/\alpha}\sum_{j\in\mathbb{Z}}\cos j\omega\sum_{t=1}^{n}\sin(t-j)\omega\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x). \end{aligned} \tag{29}$$

At this point we need a computational lemma:

LEMMA 3 *Fix m, and let f and h each be either cosine or sine. Then*

$$\begin{aligned} &n^{-1/\alpha}\sum_{|j|\le m} f(j\omega)\sum_{t=1}^{n} h((t-j)\omega)\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x) \\ &= o_P(1) + n^{-1/\alpha}\sum_{t=1}^{n} h(t\omega)\int_{B-t}\sum_{|j|\le m} f(j\omega)\psi(x+t+j)\mathbb{M}(\mathrm{d}x). \end{aligned}$$

Proof of Lemma The αth power of the scale of the difference is simply

$$\begin{aligned} &\frac{1}{n}\int\left|\sum_{|j|\le m} f(j\omega)\left(\sum_{t=1}^{n} h((t-j)\omega)1_{B+j-t}(x)\psi(x+t)\right.\right. \\ &\qquad\left.\left.-\sum_{t=1+j}^{n+j} h((t-j)\omega)1_{B+j-t}(x)\psi(x+t)\right)\right|^{\alpha}\mathrm{d}x \\ &\le \frac{1}{n}\int\left(\sum_{|j|\le m}\sum_{t\in K_{j,n}} 1_{B+j-t}(x)|\psi(x+t)|\right)^{\alpha}\mathrm{d}x, \end{aligned}$$

where $K_{j,n} = \{n+1, \ldots, n+j\} \cup \{1, \ldots, j\}$ for j positive, and $\{n+j+1, \ldots, n\} \cup \{1+j, \ldots, 0\}$ for j negative. Thus there are a finite number of terms of the form $1_{B+j-t}(x)|\psi(x+t)|$; hence the expression tends to zero as n tends to infinity. ∎

If we now apply Lemma 3 to the truncated versions of lines (28) and (29), we obtain, up to $o_P(1)$ terms,

$$n^{-1/\alpha} \sum_{t=1}^{n} \cos t\omega \int_{B-t} \sum_{|j|\le m} \cos j\omega \psi(x+t+j)\mathbb{M}(\mathrm{d}x)$$
$$- n^{-1/\alpha} \sum_{t=1}^{n} \sin t\omega \int_{B-t} \sum_{|j|\le m} \sin j\omega \psi(x+t+j)\mathbb{M}(\mathrm{d}x) \tag{30}$$

and

$$n^{-1/\alpha} \sum_{t=1}^{n} \cos t\omega \int_{B-t} \sum_{|j|\le m} \sin j\omega \psi(x+t+j)\mathbb{M}(\mathrm{d}x)$$
$$+ n^{-1/\alpha} \sum_{t=1}^{n} \sin t\omega \int_{B-t} \sum_{|j|\le m} \cos j\omega \psi(x+t+j)\mathbb{M}(\mathrm{d}x) \tag{31}$$

respectively. Now let us introduce the notations

$$W_c(t) := \int_{B-t} \sum_{|j|\le m} \cos j\omega \psi(x+t+j)\mathbb{M}(\mathrm{d}x)$$
$$W_s(t) := \int_{B-t} \sum_{|j|\le m} \sin j\omega \psi(x+t+j)\mathbb{M}(\mathrm{d}x)$$

which have scale parameters

$$\left(\int_B \left| \sum_{|j|\le m} \cos j\omega \psi(x+j) \right|^{\alpha} \mathrm{d}x \right)^{1/\alpha}$$
$$\left(\int_B \left| \sum_{|j|\le m} \sin j\omega \psi(x+j) \right|^{\alpha} \mathrm{d}x \right)^{1/\alpha};$$

hence $W_c(1), \ldots, W_c(n)$ and $W_s(1), \ldots, W_s(n)$ form iid collections! Due to assumptions on ω,

$$n^{-1/\alpha} \sum_{t=1}^{n} \cos t\omega W_c(t) = g^{-1/\alpha} \sum_{h=0}^{g-1} \cos h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t\in A_h} W_c(t),$$

where $A_h = \{t\colon 1 \le t \le n, t \bmod g = h\}$. A similar formula holds for W_s; employing these observations in Eqs. (30) and (31) yields

$$g^{-1/\alpha} \sum_{h=0}^{g-1} \left\{ \cos h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_c(t) - \sin h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_s(t) \right\}$$

$$g^{-1/\alpha} \sum_{h=0}^{g-1} \left\{ \cos h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_s(t) + \sin h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_c(t) \right\}$$

respectively. It is easy to demonstrate (use the Cramer–Wold device) that the vector

$$\left(n^{-1/\alpha} \sum_{t=1}^{n} W_c(t), n^{-1/\alpha} \sum_{t=1}^{n} W_s(t) \right)$$

converges in distribution, for any m, to the vector

$$\left(\int_B \sum_{|j| \le m} \cos j\omega\psi(x+j)\mathbb{M}(\mathrm{d}x), \int_B \sum_{|j| \le m} \sin j\omega\psi(x+j)\mathbb{M}(\mathrm{d}x) \right).$$

Since the size of A_h is approximately n/g, we can use the above observation with Slutsky's Theorem to establish the joint convergence

$$\begin{aligned} &\left(\cos h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_c(t), \sin h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_c(t), \right. \\ &\quad \left. \cos h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_s(t), \sin h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_s(t) \right) \\ &\stackrel{\mathcal{L}}{\Longrightarrow} \left(|\cos h\omega| \int_B \sum_{|j| \le m} \cos j\omega\psi(x+j)\mathbb{M}_h(\mathrm{d}x), \right. \\ &\quad |\sin h\omega| \int_B \sum_{|j| \le m} \cos j\omega\psi(x+j)\mathbb{M}_h(\mathrm{d}x), \\ &\quad |\cos h\omega| \int_B \sum_{|j| \le m} \sin j\omega\psi(x+j)\mathbb{M}_h(\mathrm{d}x), \\ &\quad \left. |\sin h\omega| \int_B \sum_{|j| \le m} \sin j\omega\psi(x+j)\mathbb{M}_h(\mathrm{d}x) \right) \end{aligned}$$

which holds jointly in $h = 0, \ldots, g-1$, due to the independence of the summands in the sets A_h. To clarify, this is a weak convergence of vectors with $4g$ components. The measures $\mathbb{M}_h$

are independent versions of the original measure $\mathbb{M}$. If we sum this convergence over h and divide by $g^{1/\alpha}$, we obtain

$$\left(g^{-1/\alpha} \sum_{h=0}^{g-1} \cos h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_c(t),\ g^{-1/\alpha} \sum_{h=0}^{g-1} \sin h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_c(t), \right.$$

$$\left. g^{-1/\alpha} \sum_{h=0}^{g-1} \cos h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_s(t),\ g^{-1/\alpha} \sum_{h=0}^{g-1} \sin h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_s(t) \right)$$

$$\stackrel{\mathcal{L}}{\Longrightarrow} \left(g^{-1/\alpha} \left(\sum_{h=0}^{g-1} |\cos h\omega|^{\alpha} \right)^{1/\alpha} \int_B \sum_{|j| \le m} \cos j\omega \psi(x+j) \mathbb{M}(\mathrm{d}x), \right.$$

$$g^{-1/\alpha} \left(\sum_{h=0}^{g-1} |\sin h\omega|^{\alpha} \right)^{1/\alpha} \int_B \sum_{|j| \le m} \cos j\omega \psi(x+j) \mathbb{M}(\mathrm{d}x),$$

$$g^{-1/\alpha} \left(\sum_{h=0}^{g-1} |\cos h\omega|^{\alpha} \right)^{1/\alpha} \int_B \sum_{|j| \le m} \sin j\omega \psi(x+j) \mathbb{M}(\mathrm{d}x),$$

$$\left. g^{-1/\alpha} \left(\sum_{h=0}^{g-1} |\sin h\omega|^{\alpha} \right)^{1/\alpha} \int_B \sum_{|j| \le m} \sin j\omega \psi(x+j) \mathbb{M}(\mathrm{d}x), \right)$$

using the fact that $\sum_{h=0}^{g-1} |f(h)| \int \phi(x) \mathbb{M}_h(\mathrm{d}x)$ equals $(\sum_{h=0}^{g-1} |f(h)|^{\alpha})^{1/\alpha} \int \phi(x) \mathbb{M}(\mathrm{d}x)$ in distribution. The constants in front of the limit vector's components are the same coefficients γ_c and γ_s defined in the statement of the theorem. If we apply the continuous mapping $(w, x, y, z) \mapsto (w - z, y + x)$ to the above convergence, we have

$$\left(g^{-1/\alpha} \sum_{h=0}^{g-1} \left\{ \cos h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_c(t) - \sin h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_s(t) \right\}, \right.$$

$$\left. g^{-1/\alpha} \sum_{h=0}^{g-1} \left\{ \cos h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_s(t) + \sin h\omega \left(\frac{n}{g}\right)^{-1/\alpha} \sum_{t \in A_h} W_c(t) \right\} \right)$$

$$\stackrel{\mathcal{L}}{\Longrightarrow} \left(\gamma_c \int_B \sum_{|j| \le m} \cos j\omega \psi(x+j) \mathbb{M}(\mathrm{d}x) - \gamma_s \int_B \sum_{|j| \le m} \sin j\omega \psi(x+j) \mathbb{M}(\mathrm{d}x), \right.$$

$$\left. \gamma_c \int_B \sum_{|j| \le m} \sin j\omega \psi(x+j) \mathbb{M}(\mathrm{d}x) + \gamma_s \int_B \sum_{|j| \le m} \cos j\omega \psi(x+j) \mathbb{M}(\mathrm{d}x) \right);$$

of course we recognize the left hand side as the truncation of

$$\left(n^{-1/\alpha} \sum_{t=1}^{n} X(t) \cos t\omega,\ n^{-1/\alpha} \sum_{t=1}^{n} X(t) \sin t\omega \right)$$

up to $o_P(1)$ terms. At this point we let m tend to ∞ as in the proof of Theorem 1; the only difference here is the presence of trigonometric terms, which have absolute value bounded by one, and hence are easily dealt with. Recalling that $S_j = \int_B \psi(x + j)\mathbb{M}(\mathrm{d}x)$, we have proved

$$\left(n^{-1/\alpha} \sum_{t=1}^{n} X(t) \cos t\omega, n^{-1/\alpha} \sum_{t=1}^{n} X(t) \sin t\omega\right)$$
$$\stackrel{\mathcal{L}}{\Longrightarrow} \left(\gamma_c \sum_{j\in\mathbb{Z}} \cos j\omega S_j - \gamma_s \sum_{j\in\mathbb{Z}} \sin j\omega S_j, \gamma_c \sum_{j\in\mathbb{Z}} \sin j\omega S_j + \gamma_s \sum_{j\in\mathbb{Z}} \cos j\omega S_j\right).$$

This is the desired goal; now we may apply the mapping $(x, y) \mapsto (x^2 + y^2)$ to obtain

$$n^{-2/\alpha} \left|\sum_{t=1}^{n} X(t)e^{-itw}\right|^2 \stackrel{\mathcal{L}}{\Longrightarrow} (\gamma_c^2 + \gamma_s^2) \left|\sum_{j\in\mathbb{Z}} e^{-ijw} S_j\right|^2$$

after some algebraic simplification of the right hand side. ∎

4.2 Adjustments for Location

Note that this theorem is proved for the "locationless model." If we wish to consider the *location model* $Z(t) = X(t) + \mu$, things are a bit different. The periodogram would then be

$$\left|n^{-1/\alpha} \sum_{t=1}^{n} Z(t)e^{-itw}\right|^2 = \left|n^{-1/\alpha} \sum_{t=1}^{n} X(t)e^{-itw} + \mu n^{-1/\alpha} \sum_{t=1}^{n} e^{-itw}\right|^2;$$

as long as $\omega \neq 0$, the second summand above tends to zero as $n \to \infty$, since $|\sum_{t=1}^{n} e^{-itw}| \leq 2/|1 - e^{-iw}|$. If, as is common in practice, we evaluate the periodogram at frequencies of the form $\omega = 2\pi j/n$, for $j \neq 0$, then the second term is actually equal to zero for every n. However, when $\omega = 0$ we obtain

$$n^{2-2/\alpha}|\bar{Z}|^2 = \left|n^{-1/\alpha} \sum_{t=1}^{n} X(t) + \mu n^{1-1/\alpha}\right|^2$$

which is asymptotically the same as $|n^{-1/\alpha} \sum_{t=1}^{n} X(t)e^{-itw}|^2$ if and only if $\alpha \leq 1$; otherwise the whole thing diverges – compare with Brockwell and Davis (1991, Proposition 10.3.1).

The other possibility is to define a *centered* periodogram:

$$\left|n^{-1/\alpha} \sum_{t=1}^{n} (Z(t) - \bar{Z})e^{-itw}\right|^2.$$

For $\omega = 0$, this *centered* periodogram is zero; otherwise we have

$$\left|n^{-1/\alpha} \sum_{t=1}^{n} X(t)e^{-itw} - n^{-1/\alpha} \sum_{t=1}^{n} (\bar{Z} - \mu)e^{-itw}\right|^2$$

and the second term is $n^{-1/\alpha}\bar{X}\sum_{t=1}^{n} e^{-itw}$ which is $O_P(n^{-1})$ for $\omega \neq 0$. Thus, for $\omega \neq 0$, this *centered* periodogram is asymptotically equivalent to the periodogram for the locationless model, and exactly equivalent if ω is chosen to be of form $2\pi j/n$. The *centered* periodogram seems to be the better choice, at least from the perspective of asymptotic theory.

4.3 Normalizing the Periodogram

We now develop some *α-stable spectral density* theory as an interpretation of the limit result in Theorem 4. Throughout this section, we will not concern ourselves with normalizations by 2π. Now, an alternative definition for the periodogram is the Fourier Transform of the sample autocovariance sequence. For stationary data, it is equivalent to transform the sample autocorrelation sequence; the formula is

$$\sum_{h\in\mathbb{Z}} e^{-ihw}\left(\frac{\sum_{t=1}^{n} X(t)X(t+h)}{\sum_{t=1}^{n} X^2(t)}\right). \tag{32}$$

For heavy-tailed infinite-order moving averages, it is well-known that the sample autocorrelation at lag h is actually consistent for the number

$$\frac{\sum_{j\in\mathbb{Z}} \psi_j \psi_{j+h}}{\sum_{j\in\mathbb{Z}} \psi_j^2},$$

even though the covariances are infinite – see Davis and Resnick (1986) for a derivation of the root's asymptotic distribution. For light-tailed infinite-order moving averages, the spectral density works out to be

$$\sum_{h\in\mathbb{Z}} e^{-ihw}\frac{\sum_{j\in\mathbb{Z}} \psi_j \psi_{j+h}}{\sum_{j\in\mathbb{Z}} \psi_j^2}, \tag{33}$$

i.e., the transform of the autocorrelation sequence. In analogy with this, the *heavy-tailed spectral density* can be defined by Eq. (33) when there is an infinite-order moving average model for the data. We observe that no theory for the Fourier transform of the sample autocovariance sequence can be developed in the heavy-tailed case, since those statistics will always have a random limit. In consequence of this, we are constrained to examine the Fourier transform of the sample autocorrelations. Now, rewriting Eq. (32), we obtain the *self-normalized periodogram*, which is the quotient of the usual periodogram and the sample variance:

$$I_N(\omega) := \frac{|\sum_{t=1}^{n} X(t)e^{-itw}|^2}{\sum_{t=1}^{n} X^2(t)}. \tag{34}$$

Klüppelberg and Mikosch (1993) first derived the asymptotic distribution of the above *self-normalized periodogram* under an infinite-order moving average model, and discovered a random limit. In a subsequent paper (Klüppelberg and Mikosch, 1994), it was demonstrated that smoothing over a nearby band of frequencies, as in the classical lightly-tailed case, produces a consistent estimate of the *heavy-tailed spectral density* given by Eq. (33). We now propose some extensions of this theory to our stable integral model.

According to Corollary 5.2 of Resnick, Samorodnitsky, and Xue (1999), the sample autocorrelations of the process $X(t)$ converge to a constant limit if and only if the filter function satisfies the following condition

$$\sum_{j\in\mathbb{Z}} \psi(x+j)\psi(x+j+h) = \rho(h)\sum_{j\in\mathbb{Z}} \psi^2(x+j) \tag{35}$$

for almost every $x \in B$ and some number $\rho(h)$. When condition (35) holds, then the sample autocorrelations are consistent for the number $\rho(h)$. As a remark, it is easily verified that ψs of step function form, as in line (3), satisfy this condition, and we obtain in that case

$$\rho(h) = \frac{\sum_{j \in \mathbb{Z}} \psi_j \psi_{j+h}}{\sum_{j \in \mathbb{Z}} \psi_j^2}.$$

Taking this as our starting point, we define the *α-stable spectral density* for an α stable moving average model given by Eq. (1) with filter function ψ satisfying condition (35) to be

$$f_\alpha(\omega) = \sum_{h \in \mathbb{Z}} e^{-ihw} \rho(h). \tag{36}$$

It is desirable to obtain the consistency of an appropriately smoothed self-normalized peridogram for the *α-stable spectral density*. In Klüppelberg and Mikosch (1994), these results were obtained for a heavy-talled moving average model; the theorem below delineates progress toward establishing a similar result for the general stable moving average case.

THEOREM 5 *Under the locationless model given by* (1) *with* $0 < \alpha < 2$, *the self-normalized periodogram* $I_N(\omega)$, *for any* $\omega \in 2\pi\mathbb{Q}$, *converges in distribution to a non-degenerate limit:*

$$\frac{|\sum_{t=1}^{n} X(t) e^{-itw}|^2}{\sum_{t=1}^{n} X^2(t)} \stackrel{\mathcal{L}}{\Longrightarrow} \frac{(\gamma_c^2 + \gamma_s^2) \left|\sum_{j \in \mathbb{Z}} e^{-ijw} S_j\right|^2}{C \int_B \Psi_2(x) \tilde{\mathbb{M}}(\mathrm{d}x)}. \tag{37}$$

The notation is the same as that defined in previous theorems.

Remark 10 Recall that for the scaled sample mean S_n, we constructed a normalized version $T_n(\mu)$ in Corollary 2, and thereby removed its explicit dependence on a from the root; Theorem 5 achieves the same objective for the periodogram $I(\omega)$ by producing $I_N(\omega)$. As mentioned in Section 4.2, we can replace the variables $X(t)$ by $Z(t)$ in the formula for the periodogram without affecting the asymptotic distribution (though with some difficulties at $\omega = 0$). Likewise, the sample variance is asymptotically the same when $Z(t)$ is substituted for $X(t)$ (so long as $\alpha < 2$), so it follows that the Theorem above also holds true for the locationless model (18).

PROPOSITION 2 *The limit random variable in Theorem* 4 *can be decomposed into the α-stable spectral density and an error term:*

$$\frac{\left|\sum_{j \in \mathbb{Z}} e^{-ijw} S_j\right|^2}{C \int_B \Psi_2(x) \tilde{\mathbb{M}}(\mathrm{d}x)} = f_\alpha(\omega) + \varepsilon(\omega).$$

The error term $\varepsilon(\omega)$ *is a complicated term, involving the Fourier transform of many random variables (see the proof of the proposition).*

Remark 11 Future work must focus on obtaining joint convergence results for the self-normalized periodogram over a band of frequencies, as in Klüppelberg and Mikosch (1993, 1994), and smoothing the result to remove the $\varepsilon(\omega)$ term. There is some evidence that the random variable $\varepsilon(\omega)$ is symmetric about zero. The hope is to construct a smoothed self-normalized periodogram which is consistent for $f_\alpha(\omega)$ when condition (35) is met.

Proof of the Theorem Returning to Eq. (26) in Theorem 4, we see that

$$n^{-1/\alpha}\sum_{t=1}^{n}\sum_{|j|\le m}\cos t\omega\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x)$$
$$=g^{-1/\alpha}\sum_{h=0}^{g-1}\cos h\omega\left(\frac{n}{g}\right)^{-1/\alpha}\sum_{t\in A_h}\sum_{|j|\le m}\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x).$$

As observed previously, the sum of the t's in A_h is over a collection of identically distributed $(2m+1)$-dependent stable random variables. Now we need a simple lemma.

LEMMA 4 *If the strictly stable random variables $Y(1),\ldots,Y(n)$ form a stationary and k-dependent sequence, then $(n^{-1/\alpha}\sum_{t=1}^{n}\cos t\omega Y(t), n^{-2/\alpha}\sum_{t=1}^{n}\cos^2 t\omega Y^2(t))$ converges to some limit, where $\omega\in 2\pi\mathbb{Q}$.*

Proof of Lemma If we let $\omega=2\pi u/g$ as before, the joint vector is equal to

$$\left(g^{-1/\alpha}\sum_{h=0}^{g-1}\cos h\omega\left(\frac{n}{g}\right)^{-1/\alpha}\sum_{t\in A_h}Y(t),\, g^{-2/\alpha}\sum_{h=0}^{g-1}\cos^2 h\omega\left(\frac{n}{g}\right)^{-2/\alpha}\sum_{t\in A_h}Y^2(t)\right)$$

which has characteristic function

$$E\exp\left\{i\theta g^{-1/\alpha}\sum_{h=0}^{g-1}\cos h\omega\left(\frac{n}{g}\right)^{-1/\alpha}\sum_{t\in A_h}Y(t)\right.$$
$$\left.+\,iv g^{-2/\alpha}\sum_{h=0}^{g-1}\cos^2 h\omega\left(\frac{n}{g}\right)^{-2/\alpha}\sum_{t\in A_h}Y^2(t)\right\}$$
$$=\prod_{h=0}^{g-1}E\exp\left\{i\theta g^{-1/\alpha}\cos h\omega\left(\frac{n}{g}\right)^{-1/\alpha}\sum_{t\in A_h}Y(t)\right.$$
$$\left.+\,iv g^{-2/\alpha}\cos^2 h\omega\left(\frac{n}{g}\right)^{-2/\alpha}\sum_{t\in A_h}Y^2(t)\right\}$$
$$\longrightarrow\prod_{h=0}^{g-1}E\exp\{i\theta g^{-1/\alpha}\cos h\omega S_h+iv g^{-2/\alpha}\cos^2 h\omega\tilde{S}_h\}$$
$$=E\exp\left\{i\theta g^{-1/\alpha}\sum_{h=0}^{g-1}\cos h\omega S_h+iv g^{-2/\alpha}\sum_{h=0}^{g-1}\cos^2 h\omega\tilde{S}_h\right\}$$

for some $(S_h,\tilde{S}_h)$, $0\le h<g$. The convergence follows from Lemma 2. ■

Of course, the same result will hold with *sine* instead of *cosine*; hence, considering the truncated model, we obtain the joint convergence of

$$\Bigg(n^{-1/\alpha}\sum_{t=1}^{n}\cos t\omega\sum_{|j|\le m}\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x),$$

$$n^{-2/\alpha}\sum_{t=1}^{n}\cos^2 t\omega\left(\sum_{|j|\le m}\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x)\right)^2,$$

$$n^{-1/\alpha}\sum_{t=1}^{n}\sin t\omega\sum_{|j|\le m}\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x),$$

$$n^{-2/\alpha}\sum_{t=1}^{n}\sin^2 t\omega\left(\sum_{|j|\le m}\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x)\right)^2\Bigg).$$

At this point we are not concerned with what the limit is – we will be able to deduce it later. Applying the mapping $(w, x, y, z) \mapsto (w, y, x+z)$ yields the joint convergence of

$$\Bigg(n^{-1/\alpha}\sum_{t=1}^{n}\cos t\omega\sum_{|j|\le m}\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x),$$

$$n^{-1/\alpha}\sum_{t=1}^{n}\sin t\omega\sum_{|j|\le m}\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x),$$

$$n^{-2/\alpha}\sum_{t=1}^{n}\left(\sum_{|j|\le m}\int_{B+j-t}\psi(x+t)\mathbb{M}(\mathrm{d}x)\right)^2\Bigg)$$

using the fact that $\cos^2 t\omega + \sin^2 t\omega = 1$. We have already dealt with the first and second components in the proof of Theorem 4, whereas the third component tends to

$$C\int_B\sum_{|j|\le m}\psi^2(x+j)\mathbb{M}(\mathrm{d}x)$$

as in the proof of Theorem 2. Next, let m tend to infinity, and we obtain the joint convergence of

$$\Bigg(n^{-1/\alpha}\sum_{t=1}^{n}\cos t\omega\int\psi(x+t)\mathbb{M}(\mathrm{d}x),$$

$$n^{-1/\alpha}\sum_{t=1}^{n}\sin t\omega\int\psi(x+t)\mathbb{M}(\mathrm{d}x), n^{-2/\alpha}\sum_{t=1}^{n}X^2(t)\Bigg).$$

Finally applying the mapping $(x, y, z) \mapsto (x^2 + y^2)/z$ concludes the proof by the continuous mapping theorem. ■

Proof of the Proposition The proof mainly consists in defining the error term $\varepsilon(\omega)$ and organizing the material appropriately. Define $\gamma(h)$ to be an analogue of the covariance of the process:

$$\gamma(h) = \sum_{j \in \mathbb{Z}} S_j S_{j+h}$$

for any $h \in \mathbb{Z}$. Then simple arithmetic yields

$$\left| \sum_{j \in \mathbb{Z}} e^{-ijw} S_j \right|^2 = \sum_{h \in \mathbb{Z}} e^{-ihw} \gamma(h).$$

Now we expand $\gamma(h)$, collecting the diagonal and off-diagonal terms of the resulting double-integral:

$$\begin{aligned}
\gamma(h) &= \int_B \int_B \sum_{j \in \mathbb{Z}} \psi(x+j)\psi(y+j+h)\mathbb{M}(\mathrm{d}x)\mathbb{M}(\mathrm{d}y) \\
&\overset{\mathcal{L}}{\Longrightarrow} C \cdot \int_B \sum_{j \in \mathbb{Z}} \psi(x+j)\psi(x+j+h)\tilde{\mathbb{M}}(\mathrm{d}x) \\
&\quad + \int \int_{x \neq y} \sum_{j \in \mathbb{Z}} \psi(x+j)\psi(y+j+h)\mathbb{M}(\mathrm{d}x)\mathbb{M}(\mathrm{d}y)
\end{aligned}$$

which we shall name $D(h)$ and $D'(h)$ respectively. It follows from condition (35) that

$$D(h) = \rho(h) \cdot D(0)$$

almost surely, since $D(h) - \rho(h)D(0)$ has scale parameter zero. Putting this together, the limit of the self-normalized periodogram is

$$\begin{aligned}
\frac{\left| \sum_{j \in \mathbb{Z}} e^{-ijw} S_j \right|^2}{C \cdot \int_B \Psi_2(x)\tilde{\mathbb{M}}(\mathrm{d}x)} &= \frac{\sum_{h \in \mathbb{Z}} e^{-ihw} \gamma(h)}{D(0)} = \sum_{h \in \mathbb{Z}} e^{-ihw} \frac{D(h)}{D(0)} + \sum_{h \in \mathbb{Z}} e^{-ihw} \frac{D'(h)}{D(0)} \\
&= \sum_{h \in \mathbb{Z}} e^{-ihw} \rho(h) + \varepsilon(\omega)
\end{aligned}$$

almost surely, where we define $\varepsilon(\omega)$ to be $\sum_{h \in \mathbb{Z}} e^{-ihw} D'(h)/D(0)$. This completes the proof of the proposition. ■

5 STATISTICAL APPLICATIONS

5.1 Mixing Properties

This final section of the paper describes how subsampling methods may be used for practical application of the results of the previous sections. The idea is to use the *subsampling distribution estimator* – which is calculated from the data– as an approximation of the limit distribution of our root; this yields approximate quantiles for the root's sampling distribution, and thus confidence intervals for the parameter of interest can be formed. For more details and background on these methods, see the book *Subsampling* (Politic *et al*., 1999).

Strong mixing is a condition on the dependence structure which is sufficient to insure the validity of the subsampling theorems. The strong mixing assumption requires that $\alpha_X(k) \to 0$ as $k \to \infty$; here $\alpha_X(k) := \sup_{A,B} |P[A \cap B] - P[A]P[B]|$, where A and B are events in the σ-fields generated by $\{X_t, t \le l\}$ and $\{X_t, t \ge l + k\}$, respectively, for any $l \ge 0$ (see Rosenblatt, 1956). Many time series models satisfy this assumption – for Gaussian processes, the summability of the autocovariance function implies the strong mixing property. When our process (1) is symmetric, the strong mixing condition is always satisfied, as the following Proposition demonstrates:

PROPOSITION 3 *The symmetric stochastic process $X(t)$ given by Eq.* (1) *– for any $\alpha \in (0, 2]$, filter function ψ satisfying the conditions specified above, and skewness intensity identically zero–is strong mixing.*

Proof Let $I(\theta_1, \theta_2; t)$ denote, for any real θ_1, θ_2, the following feature of the stochastic process:

$$I(\theta_1, \theta_2; t) = -\log Ee^{i(\theta_1 X(t) + \theta_2 X(0))} + \log Ee^{i\theta_1 X(t)} + \log Ee^{i\theta_2 X(0)}.$$

The asymptotic negligibility of $I(t)$ is a sufficient condition for the process to be strong mixing – (see Maruyama, 1970; Gross, 1993; Samorodnitsky and Taqqu, 1994). According to Gross (1993), so long as a stationary symmetric α-stable process satisfies *Condition S* (which is true – (see Samorodnitsky and Taqqu, 1994), a sufficient condition for strong mixing is

$$\lim_{t \longrightarrow \infty} I(1, -1; t) = 0$$

for $0 < \alpha \le 1$; for $\alpha > 1$, the additional assumption that

$$\lim_{t \longrightarrow \infty} I(1, 1; t) = 0$$

provides a sufficient condition. For moving average processes of type given by (1), both conditions are satisfied. In fact, $-I(1, -1; t)$ is the codifference of the process, which tends to zero as t grows, as demonstrated in Samorodnitsky and Taqqu (1994, Theorem 4.7.3). As for $I(1, 1; t)$, writing out the characteristic function for a symmetric α-stable random variable yields

$$I(1, 1; t) = \|\psi(\cdot + t) + \psi\|_\alpha^\alpha - 2\|\psi\|_\alpha^\alpha.$$

As in the proof of Theorem 4.7.3, we can approximate ψ by a compactly supported function with arbitrary precision, such that

$$\lim_{t \longrightarrow \infty} \|\psi(\cdot + t) + \psi\|_\alpha^\alpha = \lim_{t \longrightarrow \infty} \|\psi(\cdot + t)\|_\alpha^\alpha + \|\psi\|_\alpha^\alpha = 2\|\psi\|_\alpha^\alpha.$$

This concludes the proof for all cases of α. ■

5.2 The Sample Mean

Let us now focus on the mean estimation problem introduced in the second and third sections of this paper. Assume that the mean exists (so $\alpha > 1$ – see the discussion preceding Corollary 1), and define the *subsampling distribution estimator* of the standardized root (which will be denoted by $T_n(\mu)$ – see line (22)) to be the following empirical distribution function (edf):

$$K_b(x) := \frac{1}{n-b+1} \sum_{i=1}^{n-b+1} 1_{\{T_{b,i} \leq x\}}, \tag{38}$$

where $T_{b,i}$ is essentially the statistic $T_b(\mu)$ evaluated on the subseries $\{Z(i), \ldots, Z(b+i-1)\}$ (but with the unknown μ replaced by the estimate $\bar{Z}_n$); in other words,

$$T_{b,i} := \frac{\bar{Z}_{b,i} - \bar{Z}_n}{\hat{\sigma}_{b,i}/\sqrt{b}}.$$

The precise definitions of $\bar{Z}_{b,i}$ and $\hat{\sigma}_{b,i}$ are as follows:

$$\bar{Z}_{b,i} := \frac{1}{b} \sum_{t=i}^{b+i-1} Z(t),$$

$$\hat{\sigma}_{b,i} := \sqrt{\frac{1}{b} \sum_{t=i}^{b+i-1} (Z(t) - \bar{Z}_{b,i})^2}.$$

COROLLARY 3 *Restrict α to the interval $(1, 2]$ and let $\beta = 0$ in the location-shifted model (18), and let $J(x)$ denote the cdf of the limit random variable in Corollary 2, i.e.*

$$J(x) = P\left[\frac{\int_B \Psi(x)\mathbb{M}(\mathrm{d}x)}{\sqrt{C \cdot \int_B \Psi_2(x)\tilde{\mathbb{M}}(\mathrm{d}x)}} \leq x\right].$$

Then the subsampling distribution estimator K_b is consistent as an estimator of the true sampling distribution of $T_n(\mu)$, denoted by $J_n(x) = P\{T_n(\mu) \leq x\}$. In other words, if $b \to \infty$ as $n \to \infty$ but with $b/n \to 0$, we have

$$\sup_x |K_b(x) - J_n(x)| \xrightarrow{P} 0$$

and in addition

$$K_b^{-1}(t) \xrightarrow{P} J^{-1}(p)$$

for any $p \in (0, 1)$; here $G^{-1}(p)$ is the p-quantile of distribution G, i.e., $G^{-1}(p) := \inf\{x{:}G(x) \geq p\}$.

Proof The cdf j is a continuous function, since neither $\int_B \Psi(x)\mathbb{M}(\mathrm{d}x)$ nor $\int_B \Psi_2(x)\tilde{\mathbb{M}}(\mathrm{d}x)$ have point masses in their distributions. Now the proof of the corollary, using Proposition 3, follows at once from a simple extension of Proposition 11.4.3 of Politis *et al.* (1999). ■

From this result, we may use $K_b(x)$ as a cdf from which to draw quantiles and develop confidence intervals with asymptotic veracity. This is done as follows:

$$\begin{aligned} 1 - p &= P\left[J_n^{-1}\left(\frac{p}{2}\right) \le T_n(\mu) \le J_n^{-1}\left(1 - \frac{p}{2}\right)\right] \\ &\approx P\left[K_b^{-1}\left(\frac{p}{2}\right) \le T_n(\mu) \le K_b^{-1}\left(1 - \frac{p}{2}\right)\right] \\ &= P\left[\bar{Z} - K_b^{-1}\left(1 - \frac{p}{2}\right)\frac{\hat{\sigma}}{\sqrt{n}} \le \mu \le \bar{Z} - K_b^{-1}\left(\frac{p}{2}\right)\frac{\hat{\sigma}}{\sqrt{n}}\right] \end{aligned}$$

where $1 - p$ is the confidence level. Thus the approximate equal-tailed confidence interval for μ is

$$\left[\bar{Z} - K_b^{-1}\left(1 - \frac{p}{2}\right) \cdot \frac{\hat{\sigma}}{\sqrt{n}}, \bar{Z} - K_b^{-1}\left(\frac{p}{2}\right)\frac{\hat{\sigma}}{\sqrt{n}}\right]. \tag{39}$$

As we have alluded to, this procedure provides inferential information about μ without an *a priori* hypothesis on the value of α (except that $\alpha > 1$). Note that if the data is not actually heavy-tailed (*e.g.*, $\alpha = 2$), this procedure retains its validity. When we assume that $X(t)$ has a symmetric distribution (*i.e.*, the skewness intensity is identically zero), a partially symmetrized version of the subsampling distribution estimator K_b may have better accuracy properties; see McElroy and Politis (2002), Corollary 3 for more details. Finally, these subsampling techniques can be applied in the case of random fields. The mixing conditions become more complicated as the dimension d of the index set increases, and the details have been omitted for simplicity.

References

Bertail, P., Politis, D. and Romano, J. (1999). On subsampling estimators with unknown rate of convergence. *Journal of the American Statistical Association*, **94**(446), 569–579.

Billingsley, P. (1995). *Probability and Measure*, John Wiley and Sons.

Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods*, Springer.

Davis, R. and Hsing, T. (1995). Point process and partial sum convergence for weakly dependent random variable with infinite variance. *Annals of Probability*, **23**(2), 879–917.

Davis, R. and Mikosch, T. (1998). The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Annals of Statistics*, **26**(5), 2049–2080.

Davis, R. and Resnick, S. (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Annals of Probability*, **13**(1), 179–195.

Davis, R. and Resnick, S. (1986). Limit theory for the sample covariance and correlation functions of moving average. *Annals of Statistics*, **14**(2), 533–558.

Embrechts, P., Mikosch, T. and Klüppelberg, C. (1997). *Modeling Extremal Events for Insurance and Finance*. Springer-Verlag, Berlin.

Gross, A. (1993). Some mixing conditions for stationary symmetric stable stochastic processes, *Stochastic Processes and Their Applications*.

Klüppelberg, C. and Mikosch, T. (1993). Spectral estimates and stable processes. *Stochastic Processes and their Applications*, **47**, 323–344.

Klüppelberg, C. and Mikosch, T. (1994). Some limit theory for the self-normalized periodogram of stable processes. *Scandinavian Journal of Statistics*, **21**, 485–491.

Logan, B. F., Mallows, C. L., Rice, S. O. and Shepp, L. A. (1973). Limit distributions of self-normalized sums. *Annals of Probability*, **1**, 788–809.

Maruyama, G. (1970). Infinitely divisible processes. *Theory of Probability and its Applications*, **15**(1), 1–22.

McElroy, T. and Politis, D. (2002). Robust inference for the mean in the presence of serial correlation and heavy-tailed distributions. *Econometric Theory* (to appear).

Politis, D. and Romano, J. (1994). Large sample confidence regions based on subsamples under minimal assumptions. *Annals of Statistics*, **22**(4), 2031–2050.

Politis, D., romano, J. and Wolf, M. (1999). *Subsampling*. Springer, New York.

Politis, D., Romano, J. and Wolf, M. (1997). Subsampling for heteroskedastic time series. *Journal of Econometrics*, **81**(2), 281–317.

Resnick, S. (1997). Special Invited Paper: Heavy Tail Modeling and Teletraffic Data. *Annals of Statistics*, **25**(5), 1805–1849.

Resnick, S., Samorodnitsky, G. and Xue, F. (1999). How misleading can sample ACF's of Stable MA's be? (Very!). *Annals of Applied Probability*, **9**(3), 797–817.
Resnick, S. and Starica, C. (1998). Tail index estimation for dependent data. *Annals of applied Probability*, **8**(4), 1156–1183.
Romano, J. and Wolf, M. (1998). Inference for the mean in the heavy-tailed case. *Metrika*, **50**(1), 55–69.
Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences*, **42**, 43–47.
Samorodnitsky, G. and Taqqu, M. (1994). *Stable Non-Gaussian Random Processes*, Chapman and Hall.

Taylor & Francis
Taylor & Francis Group

Journal….**Non Parametric Statistics** Article ID… **GNST 031038**

TO: CORRESPONDING AUTHOR

AUTHOR QUERIES - TO BE ANSWERED BY THE AUTHOR

The following queries have arisen during the typesetting of your manuscript. Please answer the queries.

Q1	**Please check affiliation**	
Q2	**Please supply keywords**	
Q3	**Klüppelberg and Mikosch (1992) is cited in text, but not present in the reference list. Please check.**	
	Please check the shortened right handside running head.	

Production Editorial Department, Taylor & Francis Ltd.
4 Park Square, Milton Park, Abingdon OX14 4RN

Telephone: +44 (0) 1235 828600
Facsimile: +44 (0) 1235 829000