Limit Theorems for Heavy-Tailed Random Fields With Subsampling Applications

Tucker McElroy
University of California, San Diego

Dimitris N. Politis
University of California, San Diego

Abstract

We examine random fields defined by a linear filter of heavy-tailed input random variables, and establish limit theorems for the sample mean and sample variance, as well as for their joint laws; in addition we establish limit theorems for the "heavy-tailed linear periodogram." Lastly, a discussion of subsampling methodology is given, and its utility in producing valid inference for the mean is demonstrated.

1 Introduction

1.1 The Statistical Problems

Consider a strictly stationary random field $\{X(t), t \in \mathbb{Z}^d\}$, observed over some region $K \subset \mathbb{Z}^d$. In this paper we are concerned with estimating the mean and heavy-tailed spectral density of this random field under the assumptions that the marginal distributions are heavy-tailed and that the dependence structure is linear. We focus on the sample mean $|K|^{-1} \sum_{t \in K} X(t)$ as an estimator of the mean $\theta := \mathbb{E}X(t)$, which is assumed to be finite, and the periodogram $|K|^{-1} |\sum_{t \in K} X(t) \exp\{-it'\omega\}|^2$ as an estimator of the heavy-tailed spectral density, which is defined subsequently.

The Sample Mean Problem Our point of view is one of generality: the sample mean is a ubiquitous estimate of location; in particular, it is generally consistent for θ even if the regularity condition of finite variance breaks down. Also, since we allow for asymmetric data, the sample mean is typically preferable to the sample median. Now inference (i.e., confidence intervals and hypothesis tests) for θ are based on the distribution of the sample mean, which is crucially affected by dependence and/or heavy tails. Self-normalization and subsampling were used in the context of a time series to estimate the limit distribution *without* knowledge (or explicit estimation) of either the dependence or the heavy-tailed index, in McElroy (2001). Here we consider the case of random fields, and investigate the statistical behavior of the normalized sample mean in higher dimensions.

The Spectral Density Problem In classical time series analysis, the spectral density gives a measure of the oscillatory character and dependence structure of the stochastic process. However, when the marginal distributions have infinite variance, it is unclear how to even define a spectral density. Fortunately, there is an intuitive way to do this when the process has a linear form: autocorrelation coefficients are well-defined for heavy-tailed moving averages (see Davis and Resnick (1986)), and thus one considers the Fourier Transform of this sequence. Consistency of a "heavy-tailed" periodogram was first considered by Klüppelberg and Mikosch (1993) for time series; this paper extends their results to random fields using slightly different techniques.

Subsampling The normalized sum of *iid* (independent and identically distributed) heavy-tailed random variables converges weakly to a non-normal limit (a stable law); thus it satisfies a non-central limit theorem. In order to develop confidence intervals for θ, we need the quantiles of this stable law, which, unfortunately, are generally unknown, because both the scale and the index of stability (the heavy-tailed index) will generally be unknown. The recourse is to use subsampling methodology to estimate the limit quantiles; cf. Politis et al (1999). A second practical problem is that the rate of convergence of the sum is generally unknown (it is not the common $\sqrt{n}$ which occurs in the Central Limit Theorem), which prevents us from forming the correct statistic. This is solved by self-normalization, i.e. by dividing by some appropriate measure of scale, such as the square root of the sample variance. If this is done, the limit is no longer a stable random variable, but has a well-defined continuous cdf (cumulative distribution function), so that subsampling theory can still be applied. The *iid* case has been extensively studied; viz. Logan et al (1973), Arcones and Giné (1989), Romano and Wolf (1999), and Politis et al (1999, Chapter 11). Similar difficulties exist with the periodogram, which are also resolved by self-normalization – see Klüppelberg and Mikosch (1994).

The paper at hand endeavors to generalize these results to dependent data defined by a linear random *field*. In particular, we establish new limit theorems (of the stable type) for the self-normalized sample mean and self-normalized periodogram of random fields. In addition we show how subsampling can be used for practical statistical inference.

Other literature on this topic includes: Davis and Resnick (1985, 1986), Resnick (1986, 1987), Davis and Hsing (1995). The first two papers are primarily concerned with the limit behavior of sample autocorrelations for this linear model, while Resnick (1986, 1987) considers point process techniques used to prove many of these results. Davis and Hsing (1995) examine models with long range dependence which are not linear. Resnick (1997) discusses an estimator for the heavy-tail index, and the attendant weaknesses, such as large volatility. Many of the proofs in Sections two and three are based on techniques from Davis and Resnick (1986).

This paper is organized in the following manner: Subsection 1.2 discusses the theoretical background for the model considered. Section 2 deals entirely with the random fields results, and is centered around three propositions: the partial sums, the sample variance, and their joint convergence respectively. Section 3 is concerned with the mathematics of the heavy-tailed self-normalized periodogram. Next, Section 4 discusses subsampling and its applications to the asymptotic results of Sections 2 and 3. Finally, Section 5 is an appendix with some of the more technical proofs. Due to the length of the paper, much of the background material on random fields, heavy-tailed random variables, and subsampling will be assumed (but references are given).

1.2 Background: The Model

Let $\mathbb{Z}^d$ denote the integer lattice in d-dimensional Euclidean space, and let K be a subset of $\mathbb{Z}^d$ which is the "observation region" of the data, i.e. the locations at which the data is collected. We consider a random field $\{X(t)\}$ which has a linear dependence structure:

$$X(t) = \sum_{j \in \mathbb{Z}^d} \psi(j) Z(t-j) \tag{1}$$

The random variables $\{Z(t),\, t \in \mathbb{Z}^d\}$ are independent and identically distributed (hereafter abbreviated as *iid*). This model is a generalization of infinite order moving average time series to random fields. Throughout this paper we will use the term "linear" to denote this infinite order moving average with *iid* residuals. The filter coefficients $\{\psi(j)\}$ need to satisfy a summability condition, which is discussed below. We use Z without an index to denote a common version of the Z_t's which is equal in distribution. We also assume that the Z_t's are

Heavy-tailed random variables of parameter α, for some $\alpha \in (0,2)$. We define $HT(\alpha)$, the collection of Heavy-Tailed random variables of parameter α, as follows: $Z \in HT(\alpha)$ if and only if

$$\mathbb{P}[|Z| > x] = x^{-\alpha} L(x) \tag{2}$$

$$\frac{\mathbb{P}[Z > x]}{\mathbb{P}[|Z| > x]} \to p, \qquad \frac{\mathbb{P}[Z \leqslant -x]}{\mathbb{P}[|Z| > x]} \to q \tag{3}$$

as $x \to \infty$. Here p and q are between 0 and 1 and add up to 1. $L(x)$ is a "slowly varying" function, i.e. $L(ux)/L(x) \to 1$ as $x \to \infty$ for any fixed u ; an example of a slowly varying function is the logarithmic function. Note that it easily follows that the right and left tails of Z behave like

$$\mathbb{P}[Z > x] \sim p x^{-\alpha} L(x), \qquad \mathbb{P}[Z \leqslant -x] \sim q x^{-\alpha} L(x)$$

where "$\sim$" denotes that the ratio tends to unity as $x \to \infty$. We require the filter coefficients $\{\psi(j)\}$ to be in l_δ for some $\delta < \alpha$ – see Brockwell and Davis (1991, Chapter 13) – in order to ensure that the sum on the right hand side of (1) converges almost surely.

The terminology is descriptive and fairly standard: "heavy tails" refers to the slow (polynomial) rate of decay of $\mathbb{P}[|Z| > x]$. Let $DOM(\alpha)$ denote all random variables that obey an α-stable limit theorem, i.e. if $Z_i, i = 1, 2, \cdots$ is an *iid* sequence from $DOM(\alpha)$, then there exist real constants $a_n > 0$ and b_n such that

$$a_n^{-1} \left(\sum_{i=1}^{n} (Z_i - b_n) \right) \overset{\mathcal{L}}{\Longrightarrow} S \tag{4}$$

where S is an α-stable random variable, and the convergence is weak.

Note that for $\alpha \in (0,2)$, $HT(\alpha) = DOM(\alpha)$; for $\alpha = 2$, $DOM(2)$ contains all square integrable random variables, as well as the random variables in $HT(2)$. We will think of the data as being in $HT(\alpha)$, but will extensively use the fact that for $0 < \alpha < 2$ this is the same as $DOM(\alpha)$ when deriving results.

Of course if Z is itself an α stable random variable, then $Z \in DOM(\alpha)$. If in addition it is symmetric (written Z is sαs), then X (the common version of $X(t)$) has the law of a sαs as well, but scaled by $(\sum_j |\psi(j)|^\alpha)^{\frac{1}{\alpha}}$.

There are a few facts about the choice of a_n in equation (4): first, the sequence should satisfy

$$n\mathbb{P}[|Z| > a_n x] \to x^{-\alpha}$$

as $n \to \infty$ for every positive x. (In particular, if we take a_n that satisfies this, then we can prove the limit result for the domain of attraction.) It is easy to check that $a_n := \inf\{x : \mathbb{P}[|Z| > x] \leqslant n^{-1}\}$ satisfies this condition. It is well-known – see Bingham, Goldie, and Teugels (1987) – that this sequence can be expressed as $a_n = n^{\frac{1}{\alpha}} L(n)$ (where L is slowly varying, but is not necessarily the same slowly varying function in (2)). Given this, a suitable choice for b_n is

$$b_n = \mathbb{E}[Z; |Z| \leqslant a_n]. \tag{5}$$

Notice that since $\{\psi(j)\} \in l_p$, i.e. $(\sum_j |\psi(j)|^p)^{\frac{1}{p}} < \infty$, they are also in l_α since $p < \alpha$, so $(\sum_j |\psi(j)|^\alpha)^{\frac{1}{\alpha}} < \infty$. The following notation will be used: Ψ will denote the whole sequence of $\{\psi(j), j \in \mathbb{Z}^d\}$, and Ψ_p will denote its l_p norm. It is true that $\{X(t), t \in \mathbb{Z}^d\}$ forms a strictly stationary random field, since applying a shift operator to the law for the Z-series does not effect the distribution. Now if we take $\alpha > 1$ (this assumption is made, for obvious reasons, in the sample mean problem), the mean does exist, and we shall call it $\eta := \mathbb{E}(Z)$. Thus $\mathbb{E}X(t) = \psi_\infty \eta =: \theta$, where $\psi_\infty := \sum_{j\in\mathbb{Z}} \psi(j)$. For the first half of the paper, when dealing with the sample mean problem, we make the assumption that $\alpha \in (1, 2)$. For the second portion of the paper that deals with the periodogram, this assumption is relaxed to $\alpha \in (0, 2)$.

2 Sample Mean Results

For notation, let $\mathbf{n}$ be the d-dimensional vector with components $n_1, n_2, \cdots, n_d$, and let $N = \Pi_{i=1}^d n_i$. Also, let $\mathbf{1}$ be the vector $(1, 1, \cdots, 1)$ in $\mathbb{Z}^d$. By $o_P(1)$ we denote a random variable that tends to zero in probability as $\min_i n_i \to \infty$ (so that all components grow, though not necessarily at the same rate). The observation region K mentioned in the previous section will be the d-dimensional cube $(0, n_1] \times (0, n_2] \times \cdots \times (0, n_d]$ intersected with the integer lattice $\mathbb{Z}^d$. This choice of K is for simplicity; more general shapes for K could be considered, but the mathematics gets extremely complicated.

This section will treat the convergence of the partial sums of our random field model (1). Since we are interested in the estimation of the mean, we will always assume that the heavy-tailed parameter α is strictly greater than one. This section is broken down into the following subsections: first there is a treatment of the convergence of the partial sums, and then a discussion of the partial sums of squares (the sample variance statistic); finally these results are combined into the desired joint limit theorem.

2.1 Partial Sums

Let $\sum_{t=\mathbf{1}}^{\mathbf{n}} = \sum_{t_1=1}^{n_1} \sum_{t_2=1}^{n_2} \cdots \sum_{t_d=1}^{n_d}$, and let a_n be the rate which satisfies (4) for the given random field $\{Z(t), t \in \mathbb{Z}^d\}$. The size of our observation region is N, so we will use a_N as the appropriate rate. We begin with the following basic lemma:

Lemma 2.1 *For any $j \in \mathbb{Z}^d$,*

$$\frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} Z(t) = o_P(1) + \frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} Z(t-j) \tag{6}$$

$$\frac{1}{a_N^2} \sum_{t=\mathbf{1}}^{\mathbf{n}} Z^2(t) = o_P(1) + \frac{1}{a_N^2} \sum_{t=\mathbf{1}}^{\mathbf{n}} Z^2(t-j) \tag{7}$$

Proof Consider the first line above – equation (6). We examine the difference (so without loss of generality we assume that Z has mean zero)

$$\frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} Z(t) - \frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} Z(t-j) = \frac{1}{a_N} \sum_{t \in K\Delta(K-j)} Z(t)$$

for any fixed vector $j = (j_1, j_2, \cdots, j_d)$, where Δ denotes the symmetric difference of two sets, and the set $K-j$ denotes K shifted by the vector $-j$. Upon examination of the set $K\Delta(K-j)$, we see that we can chop it up into (overlapping) blocks of various sizes: there are two blocks of size $j_1 \times n_2 \times \cdots \times n_d$, and two blocks of size $n_1 \times j_2 \times \cdots \times n_d$, and so forth. Thus in the ith pair of blocks, there are $n' = N \cdot \frac{j_i}{n_i}$ terms present in the sum; these terms are *iid*, and converge to an α-stable law at rate $a_{n'}^{-1}$. So if we denote this block by A_i, then we have

$$\frac{1}{a_{n'}} \sum_{t \in A_i} Z(t) = O_P(1).$$

Hence

$$\frac{1}{a_N} \sum_{t \in A_i} Z(t) = \frac{1}{a_{n'}} \left(\frac{j_i^{\frac{1}{\alpha}} L(n')}{n_i^{\frac{1}{\alpha}} L(N)} \right) \sum_{t \in A_i} Z(t) = O_P \left(n_i^{-\frac{1}{\alpha}} \frac{L(n')}{L(N)} \right).$$

The term $\frac{L(n')}{L(N)}$ cannot tend to infinity (if it diverges at all) faster than $n_i^{\frac{1}{\alpha}}$, since L is slowly varying. Therefore the whole expression above tends to zero. Since this can be easily established for each of the d block pairs, the first part of the Lemma is proved. For the second line (7), we observe that the *iid* random variables $Z^2(t)$ are in $DOM(\frac{\alpha}{2})$; thus, using the same notations,

$$\frac{1}{a_N^2}\sum_{t\in B} Z^2(t) = \frac{1}{a_{n'}^2}\left(\frac{j_1^{\frac{2}{\alpha}}L^2(n')}{n_1^{\frac{2}{\alpha}}L^2(N)}\right)\sum_{t\in B} Z^2(t) = O_P\left(n_1^{-\frac{2}{\alpha}}\frac{L^2(n')}{L^2(N)}\right).$$

This completes the second part of the Lemma. †

The following result is elementary:

Lemma 2.2 *Assume that the random variables $Z(t)$ have been centered to have mean zero. Then*

$$\frac{1}{a_N}\sum_{t=\mathbf{1}}^{\mathbf{n}} Z(t) \stackrel{\mathcal{L}}{\Longrightarrow} S$$

as $\min_i n_i \to \infty$, where S is an α-stable law with some scale $\sigma > 0$, skewness β, and location μ. Even if the $Z(t)$'s do not have mean zero, the following convergence holds:

$$\frac{1}{a_N^2}\sum_{t=\mathbf{1}}^{\mathbf{n}} Z^2(t) \stackrel{\mathcal{L}}{\Longrightarrow} \tilde{S}$$

as $\min_i n_i \to \infty$, where $\tilde{S}$ is an $\frac{\alpha}{2}$-stable law with some scale σ, skewness $\beta = 1$, and location μ, i.e. it is a totally right skewed stable random variable. The symbol $\stackrel{\mathcal{L}}{\Longrightarrow}$ is used to denote convergence in law.

Proof The random variables $Z(t)$ are summed over the region K; so the left hand side is a sum of $N = |K|$ *iid* random variables. The first result then follows from the fact that $Z \in DOM(\alpha)$. For the latter result, the normalized sum is asymptotically the same as a mean zero version; the convergence follows from the fact that $Z^2 \in DOM(\frac{\alpha}{2})$. †

Theorem 2.1

$$\frac{1}{a_N}\sum_{t=\mathbf{1}}^{\mathbf{n}} (X(t) - \theta) \stackrel{\mathcal{L}}{\Longrightarrow} \psi_\infty \cdot S$$

as $\min_i n_i \to \infty$, where $\psi_\infty = \sum_{j\in\mathbb{Z}^d}\psi(j)$.

Proof The proof of this Theorem will be broken into several parts, due to the intricacy of the calculations. For notational convenience, we introduce the centered versions

$$Y(t) = X(t) - \theta, \qquad W(t) = Z(t) - \eta.$$

Let B_m be the cube in $\mathbb{Z}^d$ of width $2m+1$ centered at the origin, so that the coordinates of each side run between $-m$ and m. We consider the field $W(t-j)$ for $j \in B_m$; the result of Lemma 2.1 holds true for each $j \in B_m$, and hence it will also hold true when we apply the continuous mapping

$$(A(j) \quad j \in B_m) \mapsto \sum_{j \in B_m} \psi(j) A(j)$$

for any field $\{A(j)\}$. If we arrange the field as a long vector, using some arbitrary choice of ordering (such as lexicographical), this mapping amounts to a dot product with the same ordering of the coefficient field $\{\psi(j), j \in B_m\}$. Thus we obtain

$$\frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} \sum_{j \in B_m} \psi(j) W(t-j) = o_P(1) + \frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} \sum_{j \in B_m} \psi(j) W(t). \tag{8}$$

Let us abbreviate the sum on the left hand side by defining

$$Y^{(m)}(t) = \sum_{j \in B_m} \psi(j) W(t-j).$$

Then it follows from Lemma 2.2 – since $W(t)$ has mean zero – that for fixed m

$$\frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} Y^m(t) \overset{\mathcal{L}}{\Longrightarrow} \sum_{j \in B_m} \psi(j) \cdot S,$$

where S is an α-stable random variable (the same as that occurring in the first part of Lemma 2.2). We wish to now let $m \to \infty$ on both sides of this convergence; the right hand side converges almost surely to $\psi_\infty \cdot S$. For the left hand side we have the following Lemma:

Lemma 2.3 *Consider the difference*

$$\frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} Y(t) - \frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} Y^{(m)}(t); \tag{9}$$

the limit as $m \to \infty$ *of the* $\limsup_{\mathbf{n} \to \infty}$ *in probability of this expression is zero.*

Proof The difference easily decomposes into three terms:

$$\begin{aligned}
&\frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} \sum_{j \in B_m^c} \psi(j) \left(W(t-j) 1_{\{|W(t-j)| \leqslant a_N\}} - b_N \right) \\
&+ \frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} \sum_{j \in B_m^c} \psi(j) W(t-j) 1_{\{|W(t-j)| > a_N\}} \\
&+ \frac{N b_N}{a_N} \sum_{j \in B_m^c} \psi(j),
\end{aligned}$$

where b_N was defined in (5). We divide each of these terms up into $2d$ terms, according to a division of B_m^c into (overlapping) chunks. Each piece is defined by fixing one index j_i to range between either $m+1$ and ∞ or $-(m+1)$ and $-\infty$; all other indices may take on any integer value. This produces $2d$ blocks, and the sum over each individual block will be shown to tend to zero in probability. The proof for each block is quite similar, so we prove only the first case:

$$D_1 := \{j \in \mathbb{Z}^d : j_1 > m\}. \tag{10}$$

Thus, we must show that

$$\frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} \sum_{j \in D_1} \psi(j) \left(W(t-j) 1_{\{|W(t-j)| \leqslant a_N\}} - b_N \right) \tag{11}$$

$$+ \frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} \sum_{j \in D_1} \psi(j) W(t-j) 1_{\{|W(t-j)| > a_N\}} \tag{12}$$

$$+ \frac{N b_N}{a_N} \sum_{j \in D_1} \psi(j) \tag{13}$$

has the desired limit behavior described (9) (see Billingsley (1995)).

The Third Term (13) First note that

$$b_N := \mathbb{E}\left[W 1_{\{|W| \leqslant a_N\}}\right] = \mathbb{E}[W] - \mathbb{E}\left[W 1_{\{|W| > a_N\}}\right] = -\mathbb{E}\left[W 1_{\{|W| > a_N\}}\right],$$

so that the absolute value of the third term is bounded by

$$\frac{N}{a_N} \sum_{j \in D_1} \psi(j) |b_N| \leqslant \frac{N}{a_N} \sum_{j \in D_1} |\psi(j)| \mathbb{E}\left[|W| 1_{\{|W| > a_N\}}\right] \to \frac{\alpha}{\alpha - 1} \sum_{j \in D_1} |\psi(j)|$$

by Karamata's Theorem – see Feller (1971) – where the limit is taken as $N \to \infty$ (which is implied by $\min_i n_i \to \infty$); thus the limit of this as $m \to \infty$ is zero, due to the summability of the filter coefficients.

The Second Term (12) If we write out the second term in full vector form, we consider the following probability, and use Markov's Inequality for any $\gamma > 0$ with the $\mathbb{L}^1$ norm:

$$\begin{aligned}
&\mathbb{P}\left[a_N^{-1}\left|\sum_{t\in K}\sum_{j\in D_1}\psi(j)W(t-j)1_{\{|W(t-j)|>a_N\}}\right|>\gamma\right]\\
&\leqslant \frac{1}{\gamma}\frac{1}{a_N}\mathbb{E}\left[\sum_{t\in K}\sum_{j\in D_1}|\psi(j)||W(t-j)|1_{\{|W(t-j)|>a_N\}}\right]\\
&=\frac{1}{\gamma}\frac{1}{a_N}\sum_{t\in K}\sum_{j\in D_1}|\psi(j)|\mathbb{E}\left[|W|1_{\{|W|>a_N\}}\right]\\
&=\frac{1}{\gamma}\frac{N}{a_N}\sum_{j\in D_1}|\psi(j)|\mathbb{E}\left[|W|1_{\{|W|>a_N\}}\right]\\
&\to\frac{1}{\gamma}\frac{\alpha}{\alpha-1}\sum_{j\in D_1}|\psi(j)|
\end{aligned}$$

as $N\to\infty$, due again to Karamata's Theorem. Recall that the set D_1 was defined in equation (10). Finally, we let m go to ∞ and obtain zero, due to the summability of the filter coefficients.

The First Term (11) First we introduce the notation $\tilde{D}_1 := \{k\in\mathbb{Z}^d : k_1 > m - n_1\}$ and $C_1 := (0,n_2]\times\cdots\times(0,n_d]\bigcap\mathbb{Z}^{d-1}$, which is a subset of the hyperplane on axes 2 through d. The first term has the following form:

$$\begin{aligned}
&a_N^{-1}\sum_{t\in K}\sum_{j\in D_1}\psi(j)\left(W(t-j)1_{\{|W(t-j)|\leqslant a_N\}}-b_N\right)\\
&=a_N^{-1}\sum_{k\in\tilde{D}_1}\left\{\sum_{t\in C_1}\Xi_{k_1,n_1}(k_2+t_2,\cdots,k_d+t_d)\right\}\left(W(-k)1_{\{|W(-k)|\leqslant a_N\}}-b_N\right)
\end{aligned}$$

where Ξ is defined as follows:

$$\Xi_{k_1,n_1}(s) := \begin{cases}\psi(m+1,s)+\cdots+\psi(n_1+k_1,s) & -n_1+m+1\leqslant k_1\leqslant m\\ \psi(k_1+1,s)+\cdots+\psi(k_1+1,s) & k_1>m\end{cases}$$

for any $s\in\mathbb{Z}^{d-1}$. Now we apply the Chebyshev Inequality to the following probability:

$$\begin{aligned}
&\mathbb{P}\left[a_N^{-1}\left|\sum_{k\in\tilde{D}_1}\left\{\sum_{t\in C_1}\Xi_{k_1,n_1}(k_2+t_2,\cdots,k_d+t_d)\right\}\left(W(-k)1_{\{|W(-k)|\leqslant a_N\}}-b_N\right)\right|>\gamma\right]\\
&\leqslant\frac{1}{\gamma^2}\frac{1}{N}\sum_{k\in\tilde{D}_1}\left\{\sum_{t\in C_1}\Xi_{k_1,n_1}(k_2+t_2,\cdots,k_d+t_d)\right\}^2\frac{N}{a_N^2}Var\left(W1_{\{|W|\leqslant a_N\}}\right).
\end{aligned}$$

In squaring out the expression in the first line, we note that any "off-diagonal" terms are independent, and thus the expectation of those terms is zero (since b_N is the centering of the random variables). So this leaves only the "diagonal" terms in the squaring, which are written in the second line. The last term has finite $\limsup_{N\to\infty}$, due again to Karamata's Theorem. As for the sum of coefficients, the following technical claim holds:

Claim 2.1

$$\lim_{m\to\infty}\limsup_{N\to\infty}\frac{1}{N}\sum_{k\in\tilde{D}_1}\left\{\sum_{t\in C_1}\Xi_{k_1,n_1}(k_2+t_2,\cdots,k_d+t_d)\right\}^2=0$$

Together, the three parts of the difference over block D_1 tend to zero, and the Lemma is established. †

The proof of Theorem 2.1 now follows immediately from Lemma 2.3. †

2.2 Sample Variance

The proofs for the sample variance are extremely similar to those for the partial sums, so some of the more laborious details are omitted.

Theorem 2.2

$$\frac{1}{a_N^2}\sum_{t=\mathbf{1}}^{\mathbf{n}}X^2(t)\overset{\mathcal{L}}{\Longrightarrow}\Psi_2^2\tilde{S}$$

as $\min_i n_i\to\infty$, *where* $\Psi_2=\left(\sum_{j\in\mathbb{Z}^d}\psi^2(j)\right)^{\frac{1}{2}}$ *and* $\tilde{S}$ *is the* $\frac{\alpha}{2}$ *totally right skewed stable random variable from Lemma 2.2.*

Proof Because the random variable $X(t)$ is squared, this proof is a bit more complicated than Theorem 2.1. Thus, we first establish the following preliminary Lemma:

Lemma 2.4

$$\frac{1}{a_N^2}\sum_{t=\mathbf{1}}^{\mathbf{n}}X^2(t)=o_P(1)+\frac{1}{a_N^2}\sum_{t=\mathbf{1}}^{\mathbf{n}}\sum_{j\in\mathbb{Z}^d}\psi^2(j)Z^2(t-j)\tag{14}$$

Proof of Lemma The difference between the right and left hand sides of (14) is

$$\frac{1}{a_N^2}\sum_{t=\mathbf{1}}^{\mathbf{n}}\sum_{i\neq j\in\mathbb{Z}^d}\psi(i)\psi(j)Z(t-i)Z(t-j)$$

which in the $\mathbb{L}^1$ norm is bounded by

$$\frac{1}{a_N^2}\sum_{t=\mathbf{1}}^{\mathbf{n}}\sum_{i\neq j\in\mathbb{Z}^d}|\psi(i)||\psi(j)|(\mathbb{E}\,|Z(t)|)^2\leqslant\frac{N}{a_N^2}(\mathbb{E}\,|Z|)^2\cdot\left(\sum_{i\in\mathbb{Z}^d}|\psi(i)|\right)^2,$$

and this tends to zero as $N\to\infty$. This proves the Lemma. †

Now we return to the proof of Theorem 2.2 , which follows similar lines to that of Theorem 2.1. By the previous lemma, it suffices to examine the convergence of

$$\frac{1}{a_N^2}\sum_{t=\mathbf{1}}^{\mathbf{n}}\sum_{j\in\mathbb{Z}^d}\psi^2(j)Z^2(t-j).$$

Again we consider this sum on the d-dimensional cube B_m, and by Lemma 2.1 we have

$$a_N^{-2}\sum_{t=\mathbf{1}}^{\mathbf{n}}\sum_{j\in B_m}\psi^2(j)Z^2(t-j)=o_P(1)+a_N^{-2}\sum_{t=\mathbf{1}}^{\mathbf{n}}\sum_{j\in B_m}\psi^2(j)Z^2(t) \tag{15}$$

so that

$$a_N^{-2}\sum_{t=\mathbf{1}}^{\mathbf{n}}\sum_{j\in B_m}\psi^2(j)Z^2(t-j)\overset{\mathcal{L}}{\Longrightarrow}\sum_{j\in B_m}\psi^2(j)\cdot\tilde{S} \tag{16}$$

by Lemma 2.2. The idea is now to let m increase to infinity on both sides of this convergence. On the right side this is clearly valid, and almost sure convergence to $\Psi_2^2\cdot\tilde{S}$ is obtained. As for the left hand side, we must demonstate that the limit as $m\to\infty$, for any choice of $\gamma>0$, of

$$\limsup_{\min_i n_i\to\infty}\mathbb{P}\left[\left|a_N^{-2}\sum_{t=\mathbf{1}}^{\mathbf{n}}\sum_{j\in B_m^c}\psi^2(j)Z^2(t-j)\right|>\gamma\right]$$

is zero, just as in (9). We decompose this sum into two terms:

$$\frac{1}{a_N^2}\sum_{t=\mathbf{1}}^{\mathbf{n}}\sum_{j\in B_m^c}\psi^2(j)Z^2(t-j)1_{\{|Z(t-j)|\leqslant a_N\}}+\frac{1}{a_N^2}\sum_{t=\mathbf{1}}^{\mathbf{n}}\sum_{j\in B_m^c}\psi^2(j)Z^2(t-j)1_{\{|Z(t-j)|>a_N\}}$$

and each term is further divided into $2d$ overlapping blocks as in Theorem 2.1. Considering only the sum over the first block D_1, we have

$$\frac{1}{a_N^2}\sum_{t=\mathbf{1}}^{\mathbf{n}}\sum_{j\in D_1}\psi^2(j)Z^2(t-j)1_{\{|Z(t-j)|\leqslant a_N\}}+\frac{1}{a_N^2}\sum_{t=\mathbf{1}}^{\mathbf{n}}\sum_{j\in D_1}\psi^2(j)Z^2(t-j)1_{\{|Z(t-j)|>a_N\}}. \tag{17}$$

The Second Term of (17) Choose any $\gamma > 0$, then by the use of Chebyshev's inequality with $\mathbb{E}|\cdot|^{\frac{1}{2}}$, we have

$$\begin{aligned}
&\mathbb{P}\left[a_N^{-2}\left|\sum_{t=\mathbf{1}}^{\mathbf{n}}\sum_{j\in D_1}\psi^2(j)Z^2(t-j)1_{\{|Z(t-j)|>a_N\}}\right|>\gamma\right] \\
&\leqslant \frac{1}{\sqrt{\gamma}}a_N^{-1}\sum_{t=\mathbf{1}}^{\mathbf{n}}\sum_{j\in D_1}|\psi(j)|\mathbb{E}\left[|Z(t-j)|1_{\{|Z(t-j)|>a_N\}}\right] \\
&\leqslant \frac{1}{\sqrt{\gamma}}a_N^{-1}\sum_{t=\mathbf{1}}^{\mathbf{n}}\sum_{j\in D_1}|\psi(j)|\mathbb{E}\left[|Z|1_{\{|Z|>a_N\}}\right] \\
&\leqslant \frac{1}{\sqrt{\gamma}}\frac{N}{a_N}\sum_{j\in D_1}|\psi(j)|\mathbb{E}\left[|Z|1_{\{|Z|>a_N\}}\right] \\
&\to \frac{1}{\sqrt{\gamma}}\sum_{j\in D_1}|\psi(j)|\frac{\alpha}{\alpha-1}
\end{aligned}$$

where the limit is as $\min_i n_i \to \infty$, and we have used Karamata's Theorem. The sum of the coefficients now tends to zero as $m \to \infty$, and thus the second term is accounted for.

First Term of (17) Now the first term can be rewritten as

$$a_N^{-2}\sum_{k\in\tilde{D}_1}\left\{\sum_{t\in C_1}\Omega_{k_1,n_1}(k_2+t_2,\cdots,k_d+t_d)\right\}Z^2(-k)1_{\{|Z(-k)|\leqslant a_N\}}$$

with Ω defined for any $s \in \mathbb{Z}^{d-1}$ by

$$\Omega_{k_1,n_1}(s) := \begin{cases}\psi^2(m+1,s)+\cdots+\psi^2(n_1+k_1,s) & -n_1+m+1\leqslant k_1\leqslant m \\ \psi^2(k_1+1,s)+\cdots+\psi^2(k_1+1,s) & k_1>m\end{cases}$$

We next apply the Markov Inequality to get the $\mathbb{L}^1$ norm of the previous quantity, for any $\gamma > 0$:

$$\begin{aligned}
&\mathbb{P}\left[a_N^{-2}\left|\sum_{k\in\tilde{D}_1}\left\{\sum_{t\in C_1}\Omega_{k_1,n_1}(k_2+t_2,\cdots,k_d+t_d)\right\}Z^2(-k)1_{\{|Z(-k)|\leqslant a_N\}}\right|>\gamma\right] \\
&\leqslant \frac{1}{\gamma}\frac{1}{a_N^2}\sum_{k\in\tilde{D}_1}\left\{\sum_{t\in C_1}|\Omega_{k_1,n_1}(k_2+t_2,\cdots,k_d+t_d)|\right\}\mathbb{E}\left[Z^2(-k)1_{\{|Z(-k)|\leqslant a_N\}}\right] \\
&= \frac{1}{\gamma}\frac{N}{a_N^2}\mathbb{E}\left[Z^2 1_{\{|Z|\leqslant a_N\}}\right]\cdot\frac{1}{N}\sum_{k\in\tilde{D}_1}\left\{\sum_{t\in C_1}|\Omega_{k_1,n_1}(k_2+t_2,\cdots,k_d+t_d)|\right\}
\end{aligned}$$

The first term in the product has finite limit superior as $\min_i n_i \to \infty$. The following claim finishes the proof of the theorem:

Claim 2.2 *The sum of the filter coefficients are bounded as* $\mathbf{n} \to \infty$, *and the limit of this as* $m \to \infty$ *is zero.* †

2.3 Joint Convergence

In this next part we demonstrate the joint convergence of the random variables previously studied, i.e. sample mean and sample variance. As a consequence, a limit theorem for the self-normalized quantity

$$\frac{\sum_{t=\mathbf{1}}^{\mathbf{n}} (X(t)-\theta)}{\sqrt{\sum_{t=\mathbf{1}}^{\mathbf{n}} \left(X(t)-\overline{X}\right)^2}}$$

is obtained. As usual, $\overline{X} = N^{-1}\sum_{t=\mathbf{1}}^{\mathbf{n}} X(t)$ denotes the sample mean.

Theorem 2.3 *The scaled first and second sample moments converge jointly to a nondegenerate bivariate distribution:*

$$\left(\frac{1}{a_N}\sum_{t=\mathbf{1}}^{\mathbf{n}} (X(t)-\theta)\,, \frac{1}{a_N^2}\sum_{t=\mathbf{1}}^{\mathbf{n}} X^2(t)\right) \stackrel{\mathcal{L}}{\Longrightarrow} \left(\psi_\infty S, \Psi_2^2 \tilde{S}\right)$$

and hence

$$\frac{\sum_{t=\mathbf{1}}^{\mathbf{n}} (X(t)-\theta)}{\sqrt{\sum_{t=\mathbf{1}}^{\mathbf{n}} \left(X(t)-\overline{X}\right)^2}} \stackrel{\mathcal{L}}{\Longrightarrow} \frac{\psi_\infty S}{\Psi_2\sqrt{\tilde{S}}}$$

as $\min_i n_i \to \infty$*. The joint characteristic function of* S *and* $\tilde{S}$ *is given by*

$$\mathbb{E}\left[\exp\{i\phi S + i\tau\tilde{S}\}\right] = \exp\{\int_{\mathbb{R}} \left(\exp(iy\phi + iy^2\tau) - 1 - iy\phi\right) |y|^{-(1+\alpha)} K(y)\, dy\}$$

where K *is equal to* αp *or* αq*, depending on whether* y *is positive or negative respectively, with* p *and* q *given by equation (3). The limit variable* $S/\sqrt{\tilde{S}}$ *is nondegenerate.*

Remark If the random field $Z(t)$ actually has a finite variance, then the above ratio will converge to a normal distribution, as is well-known. Thus, centering the denominator by the sample mean costs us nothing when $\alpha < 2$, but also has the advantage of giving a nondegenerate limit when the variance is finite.

Proof We first observe that the limiting distribution of the sample variance and the sample second moment are the same, when $\alpha < 2$:

$$a_N^{-2}\sum_{t=\mathbf{1}}^{\mathbf{n}} \left(X(t)-\overline{X}\right)^2 - a_N^{-1}\sum_{t=\mathbf{1}}^{\mathbf{n}} X^2(t) = \frac{N}{a_N^2}\overline{X}^2 = O_P(1/N)$$

since $\overline{X}^2 = O_P(a_N^2/N^2)$. Next, we introduce some notation:

$$T_N^{(m)} := \frac{1}{a_N}\sum_{t=\mathbf{1}}^{\mathbf{n}} \left(X^{(m)}(t) - \theta^{(m)}\right), \quad W_N^{(m)} := \frac{1}{a_N^2}\sum_{t=\mathbf{1}}^{\mathbf{n}} \left(X^{(m)}(t)\right)^2$$

where

$$X^{(m)}(t) := \sum_{j \in B_m} \psi(j) Z(t-j), \quad \theta^{(m)} := \mathbb{E}\left[X^{(m)}(t)\right],$$

with B_m defined in the proof of Theorem 2.1. From that proof we also know that

$$T_N^{(m)} \xrightarrow{\mathcal{L}} T^{(m)}, \qquad W_N^{(m)} \xrightarrow{\mathcal{L}} W^{(m)}$$

with

$$T^{(m)} := \sum_{j \in B_m} \psi(j) S, \qquad W^{(m)} := \sum_{j \in B_m} \psi^2(j) \tilde{S}.$$

More precisely, we can write, from (8) and (15),

$$\begin{aligned} T_N^{(m)} &= o_P(1) + \tilde{T}_N^{(m)}, & \tilde{T}_N^{(m)} &\xrightarrow{\mathcal{L}} T^{(m)} \\ W_N^{(m)} &= o_P(1) + \tilde{W}_N^{(m)}, & \tilde{W}_N^{(m)} &\xrightarrow{\mathcal{L}} W^{(m)} \end{aligned}$$

where

$$\tilde{T}_N^{(m)} := \sum_{j \in B_m} \psi(j) \frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} (Z(t) - \eta), \qquad \tilde{W}_N^{(m)} := \sum_{j \in B_m} \psi^2(j) \frac{1}{a_N^2} \sum_{t=\mathbf{1}}^{\mathbf{n}} Z^2(t).$$

We may concatenate these statements to produce the joint convergence

$$\left(T_N^{(m)}, W_N^{(m)}\right) = o_P(1) + \left(\tilde{T}_N^{(m)}, \tilde{W}_N^{(m)}\right) \tag{18}$$

$$\left(\tilde{T}_N^{(m)}, \tilde{W}_N^{(m)}\right) \xrightarrow{\mathcal{L}} \left(T^{(m)}, W^{(m)}\right) \tag{19}$$

This second line (19) holds true because it holds true for *iid* sequences (see Logan et al (1973)) for the first demonstration of this in the case that the inputs Z's are actually stable random variables; Resnick (1986, page 95) handles the case of *iid* inputs in DOM (α)) and therefore also for finite linear combinations of such. The characteristic function of the limiting variables S and $\tilde{S}$ will be, according to Logan et al (1973),

$$\mathbb{E}\left[\exp\{i\phi S + i\tau\tilde{S}\}\right] = \exp\{\int_{\mathbb{R}} \left(\exp(iy\phi + iy^2\tau) - 1 - iy\phi\right) |y|^{-(1+\alpha)} K(y)\, dy\}$$

This function K, as mentioned in the theorem, is either αp or αq depending on whether y is positive or negative respectively, and is connected to the skewness of the original variables. Hence, there is significant dependence between S and $\tilde{S}$. If $\phi = 0$ or $\tau = 0$ it is easy to see that we recover the $\alpha/2$ stable and α stable characteristic functions for $\tilde{S}$ and S respectively. Thus putting (18) and (19) together, we find that

$$\left(T_N^{(m)}, W_N^{(m)}\right) \xrightarrow{\mathcal{L}} \left(T^{(m)}, W^{(m)}\right). \tag{20}$$

All that remains at this point is to take the limit in probability of these expressions as m tends to ∞, as in the Propositions. Now we also know that

$$\lim_{m\to\infty} T_N^{(m)} = T_N, \qquad \lim_{m\to\infty} W_N^{(m)} = W_N$$

(the limits are in probability) from Lemma 2.3 and (16), where

$$T_N := \frac{1}{a_N}\sum_{t=\mathbf{1}}^{\mathbf{n}} (X(t)-\theta), \qquad W_N := \frac{1}{a_N^2}\sum_{t=\mathbf{1}}^{\mathbf{n}} X^2(t).$$

On the right side of (20), we also know that

$$T^{(m)} \xrightarrow{a.s.} T := \psi_\infty S, \qquad W^{(m)} \xrightarrow{a.s.} W := \Psi_2^2 \tilde{S}$$

which gives the joint weak convergence

$$\left(\frac{1}{a_N}\sum_{t=\mathbf{1}}^{\mathbf{n}} (X(t)-\theta), \frac{1}{a_N^2}\sum_{t=\mathbf{1}}^{\mathbf{n}} X^2(t)\right) \stackrel{\mathcal{L}}{\Longrightarrow} \left(\psi_\infty S, \Psi_2^2 \tilde{S}\right)$$

Finally, if we apply the continuous function $f(x,y) = \frac{x}{\sqrt{y}}$ to the above convergence, the proof is complete. The ratio $S/\sqrt{\tilde{S}}$ is not constant, because if it were one would deduce that the square of the α-stable variable S has a positively skewed $\alpha/2$ stable distribution, which is never true. $\dagger$

3 Self-Normalized Periodogram

Classical Spectral Density In classical time series, where the data has finite variance, the spectral density is defined to be the Discrete Fourier Transform of the autocovariance sequence:

$$f_C(\omega) = \sum_{h\in\mathbb{Z}} \exp\{-ih\omega\}\gamma(h)$$

for $\omega \in (-\pi,\pi]$, where $\gamma(h) := \mathbb{E}X(t)X(t+h) - \mathbb{E}X(t)\mathbb{E}X(t+h)$. Some may notice that the spectral density is typically defined by $f_C/2\pi$; we have omitted the constant 2π for ease of presentation. In the case of a linear model

$$X(t) = \sum_{j\in\mathbb{Z}} \psi(j) Z(t-j)$$

for finite variance, mean zero *iid* inputs $Z(t)$, we know that the autocovariance is given by

$$\gamma(h) = \sum_{j\in\mathbb{Z}} \psi(j)\psi(j+h)\ Var(Z)$$

so that the autocorrelation is simply

$$\rho(h) = \frac{\sum_{j\in\mathbb{Z}} \psi(j)\psi(j+h)}{\sum_{j\in\mathbb{Z}} \psi^2(j)}; \tag{21}$$

notice that this does not depend on $Var(Z)$. Therefore, we may write the spectral density as the Fourier Transform of the autocorrelation sequence, multiplied by $Var(X)$:

$$f_C(\omega) = \sum_{h\in\mathbb{Z}} \exp\{-ih\omega\}\rho(h) \cdot Var(X). \tag{22}$$

The classic estimator of $f_C(\omega)$ is called the "periodogram":

$$I(\omega) := \left| n^{-\frac{1}{2}} \sum_{t=1}^{n} X(t) \exp\{-it\omega\} \right|^2$$

which happens to be inconsistent; typically it is smoothed over a band of frequencies to obtain consistency.

All of this theory can easily be generalized to d-dimensions; the model is given by equation (1), and we replace the exponents by inner products.

Heavy-Tailed Spectral Density Now suppose that the data is heavy-tailed as in Section 2 with $\alpha \in (0, 2)$, so that covariances do not exist. However, in the linear model (1), we can define "autocorrelations" by

$$\rho(h) = \frac{\sum_{j\in\mathbb{Z}^d} \psi(j)\psi(j+h)}{\sum_{j\in\mathbb{Z}^d} \psi^2(j)}$$

in analogy with (21) – see Davis and Resnick (1986). Carrying the analogy forward, we define the "heavy-tailed linear spectral density" by

$$f_{HT}(\omega) := \sum_{h\in\mathbb{Z}^d} \exp\{-ih'\omega\}\rho(h)$$

for any vector of frequencies $\{\omega_1, \omega_2, \cdots, \omega_d\}$ each in $(-\pi, \pi]$. Note that this formula differs from equation (22) only by $Var(X)$. The periodogram is now

$$I(\omega) := \left| \frac{1}{a_N} \sum_{t=1}^{n} X(t) \exp\{-it'\omega\} \right|^2$$

where the sum is taken over the observation rectangle K, and $a'b$ denotes the dot product of two vectors a and b.

Normalizing the Periodogram Unfortunately, $I(\omega)$ grows at rate a_N, so we are in the same difficulties as with the sample mean. By normalizing the periodogram, we remove the rate problem, and obtain f_{HT} times a random variable for the limit. Thus we introduce the "self-normalized periodogram" as follows:

$$I_N(\omega) := \frac{\left|\sum_{t=\mathbf{1}}^{\mathbf{n}} X(t)\exp\{-it'\omega\}\right|^2}{\sum_{t=\mathbf{1}}^{\mathbf{n}} X^2(t)}$$

which is $I(\omega)$ divided by the sample variance. This normalization is suggested to us by the classical case, since the $Var(X)$ term occuring in equation (22) will be exactly accounted for by the limit of the sample variance.

Now by Theorem 2.2

$$S_N := \frac{1}{a_N^2}\sum_{t=1}^{n} X^2(t) \overset{\mathcal{L}}{\Longrightarrow} \Psi_2^2 \tilde{S}$$

for an $\frac{\alpha}{2}$-stable positive random variable $\tilde{S}$. Thus we might expect that $I_N(\omega)$ has a well-defined limit. As long as the random variables are appropriately centered, this is indeed true, as the following theorem to demonstrates:

Theorem 3.1 *Let $\alpha \in (0,2)$, and a vector of frequencies ω with each component a rational multiple of 2π. Also assume that the sequence $b_N = 0$ for N sufficiently large. Then*

$$I_N(\omega) = \frac{I(\omega)}{S_N} \overset{\mathcal{L}}{\Longrightarrow} \frac{\left|\sum_{j\in\mathbb{Z}^d}\psi(j)\exp\{-ij'\omega\}\right|^2}{\Psi_2^2}\cdot\frac{(U^2+V^2)}{\tilde{S}} = f_{HT}(\omega)\frac{(U^2+V^2)}{\tilde{S}} \tag{23}$$

as $\min_i n_i \to \infty$. The random variables U and V are α-stable random variables, which have the following joint characteristic function with $\tilde{S}$:

$$\begin{aligned}&\mathbb{E}\left[\exp\{i\eta_1 U + i\eta_2 V + i\eta_3\tilde{S}\}\right]\\ &= \Pi_{l=1}^{Q}\exp\{\alpha\int_0^\infty \left(\cos(yf_lQ^{-1/\alpha})\exp(iy^2\eta_3 Q^{-2/\alpha}) - 1\right) y^{-(1+\alpha)}\,dy\}\end{aligned}$$

which is valid for all values of $\alpha \in (0,2)$. The constants Q and f_l depend on the frequencies ω, and are described in the proof of Proposition 3.1*.*

The assumption that $b_N = 0$ is easily satisfied by $X(t)$ with a distribution symmetric about zero. The Theorem is important, as it in fact suggests that after some smoothing, our normalized periodogram will be a consistent estimator for the heavy-tailed linear spectral density function. For random fields this is still an open problem, though it has been dealt with in the $d = 1$ case by Klüppelberg and Mikosch (1993, 1994).

Remark We have restricted the components of ω to be rational multiples of 2π. The proof for irrational multiples of 2π is far more complicated (see Klüppelberg and Mikosch (1993, 1994)); but for most applications, one typically evaluates the periodogram only at frequencies of the form $\frac{2\pi k}{n}$ for $k = 0, 1, \cdots, n-1$. In signal processing, for example, one takes n to be a power of 2 and plots the periodogram at all points $\frac{2\pi k}{n}$.

Theorem 3.1 will follow immediately from Theorem 3.2, which is stated in what follows. Throughout, we employ the same random field notations introduced at the beginning of section 2. We begin with the following proposition; the joint convergence below is the first building block of the proof's architecture.

Proposition 3.1 *Fix ω such that each component is a rational multiple of 2π. Let $\alpha \in (0,2)$. Define centered variables $\tilde{Z}(t) = Z(t) - b_N$. Then the following joint convergence result holds:*

$$\left(\frac{1}{a_N^2} \sum_{t=\mathbf{1}}^{\mathbf{n}} \tilde{Z}^2(t), \frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} \tilde{Z}(t) \cos t'\omega, \frac{1}{a_N} \sum_{t=\mathbf{1}}^{\mathbf{n}} \tilde{Z}(t) \sin t'\omega \right) \stackrel{\mathcal{L}}{\Longrightarrow} \left(\tilde{S}, U, V \right) \tag{24}$$

where $\tilde{S}$ is a totally right skewed $\frac{\alpha}{2}$-stable random variable, and U and V are α-stable random variables. Their joint characteristic function is given in Theorem 3.1.

Proof This proof is deferred to the appendix.

We will now develop this result to investigate the joint asymptotic properties of $I(\omega)$ and the sample variance $\hat{\sigma^2}$. Suppose again that $b_N = 0$, so that $\tilde{Z}(t) = Z(t)$. First observe that

$$\begin{aligned} &\sum_{t=1}^{n} \sum_{j \in B_m} \psi(j) Z(t-j) \cos t'\omega \qquad (25) \\ &= \sum_{j \in B_m} \psi(j) \cos j'\omega \sum_{s=1-j}^{n-j} Z(s) \cos s'\omega - \sum_{j \in B_m} \psi(j) \sin j'\omega \sum_{s=1-j}^{n-j} Z(s) \sin s'\omega \\ &= o_P(a_N) + \sum_{j \in B_m} \psi(j) \cos j'\omega \sum_{s=1}^{n} Z(s) \cos s'\omega - \sum_{j \in B_m} \psi(j) \sin j'\omega \sum_{s=1}^{n} Z(s) \sin s'\omega \end{aligned}$$

by the law of cosines and application of Lemma 2.1. In a similar fashion, we obtain

$$\begin{aligned} &\sum_{t=1}^{n} \sum_{j \in B_m} \psi(j) Z(t-j) \sin t'\omega \qquad (26) \\ &= o_P(a_N) + \sum_{j \in B_m} \psi(j) \cos j'\omega \sum_{s=1}^{n} Z(s) \sin s'\omega + \sum_{j \in B_m} \psi(j) \sin j'\omega \sum_{s=1}^{n} Z(s) \cos s'\omega \end{aligned}$$

by using the law of sines. These statements (25) and (26), together with (14), produce the joint statement

$$\left(a_N^{-2}\sum_{t=1}^{n}\sum_{j\in B_m}\psi^2(j)Z_{t-j}^2, a_N^{-1}\sum_{t=1}^{n}\sum_{j\in B_m}\psi(j)Z_{t-j}\cos t'\omega, a_N^{-1}\sum_{t=1}^{n}\sum_{j\in B_m}\psi(j)Z_{t-j}\sin t'\omega\right)$$

$$\begin{aligned}
&= o_P(1) + \Bigg(\sum_{j\in B_m}\psi^2(j)\ a_N^{-2}\sum_{t=1}^{n}\psi^2(j)Z^2(t),\\
&\sum_{j\in B_m}\psi(j)\cos j'\omega\ a_N^{-1}\sum_{s=1}^{n}Z(s)\cos s'\omega - \sum_{j\in B_m}\psi(j)\sin j'\omega\ a_N^{-1}\sum_{s=1}^{n}Z(s)\sin s'\omega,\\
&\sum_{j\in B_m}\psi(j)\cos j'\omega\ a_N^{-1}\sum_{s=1}^{n}Z(s)\sin s'\omega + \sum_{j\in B_m}\psi(j)\sin j'\omega\ a_N^{-1}\sum_{s=1}^{n}Z(s)\cos s'\omega\Bigg)\\
&\stackrel{\mathcal{L}}{\Longrightarrow}\left(\sum_{j\in B_m}\psi^2(j)\tilde{S}, \psi_c^m U - \psi_s^m V, \psi_c^m V + \psi_s^m U\right)
\end{aligned} \tag{27}$$

by using Proposition 3.1. The constants ψ_c^m and ψ_s^m are defined by the formulas

$$\psi_c^m := \sum_{j\in B_m}\psi(j)\cos j'\omega \qquad \psi_s^m := \sum_{j\in B_m}\psi(j)\sin j'\omega.$$

Next apply the continuous mapping $(x,y,z)\mapsto(x,y^2+z^2)$ to the weak convergence in (27), and we obtain

$$\begin{aligned}
&\left(a_N^{-2}\sum_{t=1}^{n}\sum_{j\in B_m}\psi^2(j)Z^2(t-j), \left|a_N^{-1}\sum_{t=1}^{n}\sum_{j\in B_m}\psi(j)Z(t-j)e^{it'\omega}\right|^2\right)\\
&\stackrel{\mathcal{L}}{\Longrightarrow}\left(\sum_{j\in B_m}\psi^2(j)\tilde{S}, ((\psi_c^m)^2+(\psi_s^m)^2)(U^2+V^2)\right)
\end{aligned} \tag{28}$$

after using the simple identity $(aU-bV)^2+(aV+bU)^2=(a^2+b^2)(U^2+V^2)$. So we are finally in the situation of Lemma 2.3 , so that we should take the limit as $m\to\infty$ in the convergence (28). The right hand side clearly converges almost surely to

$$(\Psi_2^2\tilde{S}, |\sum_{j\in\mathbb{Z}^d}\psi(j)e^{ij'\omega}|^2(U^2+V^2)).$$

So if we can handle the left hand side of (28), we have proved the following theorem:

Theorem 3.2 *Let $\alpha\in(0,2)$, and consider a vector of frequencies ω such that each component is a rational multiple of 2π. Also assume that the sequence $b_N=0$ for N sufficiently large. Then the periodogram and sample variance converge jointly*

$$\left(I(\omega), \hat{\sigma^2}_n\right)\stackrel{\mathcal{L}}{\Longrightarrow}\left(\left|\sum_{j\in\mathbb{Z}^d}\psi(j)e^{ij'\omega}\right|^2(U^2+V^2), \Psi_2^2\tilde{S}\right)$$

as $min_i n_i \to \infty$*, and the self-normalized periodogram therefore obeys*

$$I_N(\omega) \stackrel{\mathcal{L}}{\Longrightarrow} \frac{\left|\sum_{j\in\mathbb{Z}^d}\psi(j)e^{ij'\omega}\right|^2 (U^2+V^2)}{\Psi_2^2\tilde{S}}. \tag{29}$$

The random variable $(U^2+V^2)/\tilde{S}$ *is nondegenerate, so the heavy-tailed periodogram is not consistent. The joint characteristic function of* U*,* V*, and* $\tilde{S}$ *is given in Theorem* 3.1*.*

Proof The previous discussion leading up to (28) is the bulk of the proof. We must show that the periodogram for the truncated series is asymptotically the same as the periodogram; as for the sample variance, this was already established in (14) and (15). But by applying the same techniques used to prove (9), we can establish

$$\begin{aligned} a_N^{-1}\sum_{t=1}^n X(t)\cos t'\omega &= o_P(1) + a_N^{-1}\sum_{t=1}^n\sum_{j\in B_m}\psi(j)Z(t-j)\cos t'\omega \\ a_N^{-1}\sum_{t=1}^n X(t)\sin t'\omega &= o_P(1) + a_N^{-1}\sum_{t=1}^n\sum_{j\in B_m}\psi(j)Z(t-j)\sin t'\omega \end{aligned}$$

with some minor adjustments (since $Z_t\cos t'\omega$ and $Z_t\sin t'\omega$ are not identically distributed; however, by partitioning them into orbits, as in the proof of Proposition 3.1, after much labor we get the same result). Put in a vector format, we have

$$\begin{aligned} &\left(a_N^{-1}\sum_{t=1}^n X(t)\cos t'\omega, a_N^{-1}\sum_{t=1}^n X(t)\sin t'\omega\right) \\ &= o_P(1) + \left(a_N^{-1}\sum_{t=1}^n\sum_{j\in B_m}\psi(j)Z(t-j)\cos t'\omega, a_N^{-1}\sum_{t=1}^n\sum_{j\in B_m}\psi(j)Z(t-j)\sin t'\omega\right) \end{aligned}$$

where $o_P(1)$ here is a short hand for the limit superior of the probability of the difference over $min_i n_i \to \infty$ tending to zero as m increases to infinity. Now applying the continuous functional $(x,y)\mapsto x^2+y^2$, which preserves the $o_P(1)$ relation, we have

$$I(\omega) = o_P(1) + \left|a_N^{-1}\sum_{t=1}^n\sum_{j\in B_m}\psi_j Z(t-j)e^{it'\omega}\right|^2.$$

The nondegeneracy of $(U^2+V^2)/\tilde{S}$ follows from the fact that U^2+V^2 can be expressed as a sum of squared α-stable variables, which never has an $\alpha/2$ stable distribution. All of this argument goes smoothly for $\alpha \leqslant 1$; in the case that $\alpha > 1$, we should replace Z by $Z - \mathbb{E}Z$ to make Proposition 3.1 work out correctly. However, it is easy to check that this makes no difference asymptotically to $I_N(\omega)$, because both its numerator and denominator grow at rate a_N^2 . †

Remark As mentioned in Klüppelberg and Mikosch (1994), the self-normalized periodogram has the nice property of being independent of the possibly unknown parameter α. Just as the self-normalized sample mean enjoys robustness under $\alpha \in (1,2)$ as discussed in Section 2, the self-normalized periodogram is robust for $\alpha \in (0,2)$.

4 Subsampling Applications

The objective of the previous limit results for the sample mean in Section 2 is to establish confidence intervals for the mean via the quantiles of the limiting distribution. Self-normalization by the sample variance was used to remove the unknown rate a_N of convergence (see Theorem 2.3), so that the ratio of partial sums and sample variance could be formed by the practitioner. The second ingredient we need is a way of estimating the quantiles of the limit, which is the complicated random variable

$$\frac{\psi_\infty \cdot S}{\Psi_2 \sqrt{\tilde{S}}}; \tag{30}$$

this can be accomplished by subsampling.

The concept of subsampling is developed in the book by Politis, Romano, and Wolf (1999). Subsets of the observation region K are chosen, for each set K, and the statistic is calculated over the random variables in that subest. This is done for all the subsets that can fit into K, and then an empirical distribution function is calculated from those values. The result is an estimate of the limit cdf, and its quantiles may be used as approximations.

Let us denote the ratio in Theorem 2.3 by

$$T_K(\theta) := \sqrt{N}\frac{\left(\hat{\theta}_K - \theta\right)}{\hat{\sigma}_K} = \frac{\sum_{t\in K}\left(X(t)-\theta\right)}{\sqrt{\frac{1}{N}\sum_{t\in K}\left(X(t)-\overline{X}\right)^2}},$$

where $\hat{\theta}_K := \frac{1}{N}\sum_{t\in K} X(t)$ and $\hat{\sigma}_K := \sqrt{\sum_{t\in K}\left(X(t)-\overline{X}\right)^2}$. From here on we utilize the notation of Chapter 5 of Politis, Romano, and Wolf (1999), so let $\mathbf{b}$ be a vector with components $(b_1, b_2, \cdots, b_d)$, which give the various dimensions of a "rectangular" subset of K, i.e. the subset is b_1 by b_2 by b_3 , etc. B will denote this set, and we let $b = \Pi_{i=1}^d b_i$ be the volume that it encloses. The vector $\mathbf{q}$ gives the positions of the various subsampling blocks within the larger blocks, so $q_i = n_i - b_i + 1$ for $i = 1, 2, \cdots, d$. Thus $q = \Pi_{i=1}^d q_i$ gives the total number of those blocks. Next we define the "subsampling distribution estimator" of $T_K(\theta)$ to be the following empirical distribution function (edf):

$$L_{\mathbf{n},\mathbf{b}}(x) := \frac{1}{q} \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{q}} 1_{\{T_{\mathbf{b},\mathbf{i}} \leq x\}} \tag{31}$$

where $T_{\mathbf{b},\mathbf{i}}$ is essentially the sum $T_K(\theta)$ evaluated on the subseries $\{X(t)\}$ with t in a scaled version of K with side lengths given by the vector $\mathbf{b}$ (but with the unknown θ replaced by the estimate $\hat{\theta}_K$) . Thus

$$T_{\mathbf{b},\mathbf{i}} := \sqrt{b}\, \frac{\hat{\theta}_{B+\mathbf{i}} - \hat{\theta}_K}{\hat{\sigma}_{B+\mathbf{i}}}.$$

Now we must briefly discuss mixing conditions – see Bulinskii (1981) or Bulinskii (1986), p. 311 for a discussion of numerous related mixing conditions. Let $\hat{\alpha}_X(k; l_1)$ be the mixing coefficients discussed in Politis et al (1999, p.122), i.e.

$$\hat{\alpha}_X(k; l_1) := \sup_{E_2 = E_1 + t} \left| \mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1)\, \mathbb{P}(A_2) \right|$$

with $A_1 \in \mathcal{F}(E_1)$, $A_2 \in \mathcal{F}(E_2)$, $|E_1| \leqslant l_1$, and $\rho(E_1, E_2) \geqslant k$. Here, E_1 and E_2 are subsets of $\mathbb{Z}^d$, and $|E|$ denotes the cardinality of E while ρ is the Euclidean distance metric. Also, $\mathcal{F}(E_i)$ is the σ-algebra generated by random variables X_t with $t \in E_i$, and $i = 1$ or 2. These coefficients are actually dominated by the strong mixing coefficients introduced by Rosenblatt (1956) – general conditions for a linear series (the $d = 1$ case) to be strong mixing are given by Withers (1981); they require that the $\psi(j)$ tend to zero fast enough (with j), and that the Zs have an absolutely continuous distribution. We make the following assumption on the mixing coefficients of the random field:

$$N^{-1} \sum_{k=1}^{\hat{n}} k^{d-1} \hat{\alpha}_X(k; b) \to 0 \tag{32}$$

where $\hat{n} := \max_i n_i$. This mixing condition (32) is easily seen to be satisfied if the random field has a compactly supported filter function ψ, for example. Now we can state the desired corollary:

Corollary 4.1 *Let $J(\cdot)$ be the cdf of the limit random variable given in (30), and choose the vector $\mathbf{b} = \mathbf{b}_K$ such that $b_i \to \infty$ and $b_i/n_i \to 0$ as $n_i \to \infty$, for $i = 1, 2, \cdots d$; also assume that the mixing condition (32) holds. Then*

$$L_{\mathbf{n},\mathbf{b}}(x) \xrightarrow{P} J(x)$$

for every continuity point x of $J(\cdot)$.

Proof This result follows immediately from Theorem 2.3 and Corollary 5.3.1 of Politis, Romano, and Wolf (1999) (notice that $\tau_u = \sqrt{u}$, so $\tau_{\mathbf{b}}/\tau_{\mathbf{n}} \to 0$, as required). †

Remark Since the limit random variable in Theorem 2.3 is absolutely continuous, we may form the asymptotically correct $(1-t)100$ percent equal-tailed confidence intervals for θ :

$$\left[\hat{\theta}_K - L^{-1}_{\mathbf{n},\mathbf{b}}(1-t/2)\cdot\frac{\hat{\sigma}_K}{N}, \hat{\theta}_K - L^{-1}_{\mathbf{n},\mathbf{b}}(t/2)\cdot\frac{\hat{\sigma}_K}{N}\right]$$

for a $1-t$ confidence level (here, $L^{-1}(\cdot)$ denotes the quantile function of a cdf $L(\cdot)$). Notice that nowhere in our procedure or in the interval construction do we need explicit knowledge of the value of α; herein lies the advantage of our method.

As with the sample mean, subsampling can also be used to approximate the limit distribution of the periodogram $I(\omega)$. Corollary 4.1 above will hold for the self-normalized periodogram if we just let $J(\cdot)$ be the cdf of the limit random variable of $I_N(\omega)$, which is given by (29). Then the subsampling distribution estimator $L_{\mathbf{n},\mathbf{b}}(x)$ must be alterred slightly: we use the same equation (31), but now $T_{\mathbf{b},\mathbf{i}}$ is defined by

$$T_{\mathbf{b},\mathbf{i}} := \frac{\tilde{I}_{B+\mathbf{i}}(\omega)}{\hat{\sigma^2}_{B+\mathbf{i}}};$$

then the corollary still holds as stated.

5 Appendix

This appendix contains proof of Proposition 3.1.

Proof of Proposition 3.1 We center all variables $Z(t)$ by b_N, in view of equation (4); thus we consider the centered variables $\tilde{Z}(t) = Z(t) - b_N$. Fix ω, and let $U(t) := \tilde{Z}(t)\cos t'\omega$ and $V(t) := \tilde{Z}(t)\sin t'\omega$. Then choose any real numbers η_1, η_2, η_3, so that the characteristic function of the left hand side of (24) is

$$\begin{aligned} &\mathbb{E}\exp\left\{i\eta_1 a_N^{-1}\sum_{t=\mathbf{1}}^{\mathbf{n}} U(t) + i\eta_2 a_N^{-1}\sum_{t=\mathbf{1}}^{\mathbf{n}} V(t) + i\eta_3 a_N^{-2}\sum_{t=\mathbf{1}}^{\mathbf{n}}\tilde{Z}^2(t)\right\} \qquad (33)\\ &= \mathbb{E}\exp\left\{ia_N^{-1}\sum_{t=\mathbf{1}}^{\mathbf{n}}\tilde{Z}(t)\left(\eta_1\cos t'\omega + \eta_2\sin t'\omega\right) + i\eta_3 a_N^{-2}\sum_{t=\mathbf{1}}^{\mathbf{n}}\tilde{Z}^2(t)\right\}. \end{aligned}$$

Now let $f(t) := \eta_1\cos t'\omega + \eta_2\sin t'\omega$, and notice that this function is periodic in each component t_i if we fix the other components, due to the choice of the ω_i's. This makes f into a periodic function on $\mathbb{Z}^d$ with a finite orbit, say of length Q. Thus we can partition the observation region K into regions G_l^N, each of which consists of all points $t \in K$ such that $f(t)$ is constant. Thus we set f_l to be the value of $f(t)$ on the set G_l^N, and we know the size of G_l^N is $G = \frac{N}{Q}$. Note that f_l does not depend on N, even though it is the value of $f(t)$ on the subset G_l^N. Now we use this in (33) along with independence of the inputs to obtain

$$
\begin{aligned}
&\mathbb{E}\exp\left\{ia_N^{-1}\sum_{l=1}^{Q}\sum_{t\in G_l^N}\tilde{Z}(t)f(t)+i\eta_3 a_N^{-2}\sum_{l=1}^{Q}\sum_{t\in G_l^N}\tilde{Z}^2(t)\right\}\\
&=\Pi_{l=1}^{Q}\mathbb{E}\exp\left\{ia_N^{-1}\sum_{t\in G_l^N}\tilde{Z}(t)f_l+i\eta_3 a_N^{-2}\sum_{t\in G_l^N}\tilde{Z}^2(t)\right\}\\
&=\Pi_{l=1}^{Q}\mathbb{E}\exp\left\{if_l\left(Q^{\frac{1}{\alpha}}\frac{L(N)}{L(G)}\right)^{-1}a_G^{-1}\sum_{t\in G_l^N}\tilde{Z}(t)+i\eta_3\left(Q^{\frac{1}{\alpha}}\frac{L(N)}{L(G)}\right)^{-2}a_G^{-2}\sum_{t\in G_l^N}\tilde{Z}^2(t)\right\}\\
&\to\Pi_{l=1}^{Q}\mathbb{E}\exp\left\{if_lQ^{-\frac{1}{\alpha}}S_l+i\eta_3Q^{-\frac{2}{\alpha}}\tilde{S}_l\right\}\\
&=\mathbb{E}\exp\left\{iQ^{-\frac{1}{\alpha}}\sum_{l=1}^{Q}f_lS_l+iQ^{-\frac{2}{\alpha}}\eta_3\sum_{l=1}^{Q}\tilde{S}_l\right\}\\
&=\mathbb{E}\exp\left\{i\eta_1Q^{-\frac{1}{\alpha}}\sum_{l=1}^{Q}c_l^1S_l+i\eta_2Q^{-\frac{1}{\alpha}}\sum_{l=1}^{Q}c_l^2S_l+i\eta_3Q^{1-\frac{2}{\alpha}}\sum_{t=1}^{Q}\tilde{S}_l\right\}.
\end{aligned}
$$

The limit in the middle lines was taken as $\min_i n_i \to \infty$, which forced $N \to \infty$ and $G \to \infty$. Also we used the simple identity $a_N = Q^{\frac{1}{\alpha}}\frac{L(N)}{L(G)}a_G$, and $\frac{L(N)}{L(G)} = \frac{L(QG)}{L(G)} \to 1$ as $G \to \infty$ since the function $L(\cdot)$ is slowly varying. Finally, c_l^1 and c_l^2 are the constants obtained when we decompose $f_l = \eta_1 c_l^1 + \eta_2 c_l^1$. Now $\{S_l; l = 1, 2, \cdots, Q\}$ are independent α-stable random variables, and $\{\tilde{S}_l; l = 1, 2, \cdots, Q\}$ are independent totally right skewed $\frac{\alpha}{2}$-stable random variables. The third equality is valid due to the joint convergence of the terms for $l = 1, 2, \cdots, Q$.

In the previous calculation we have used the fact that

$$
\left(a_G^{-1}\sum_{t\in G_l^N}\tilde{Z}(t),a_G^{-2}\sum_{t\in G_l^N}\tilde{Z}^2(t)\right)\overset{\mathcal{L}}{\Longrightarrow}\left(S,\tilde{S}\right)
$$

jointly, which is the generalization of Theorem 2.3 to the case that $\alpha \in (0, 2)$. The proof is similar, the main difference being the centering by b_N when $\alpha \leq 1$.

Thus we may conclude that

$$
\left(a_N^{-2}\sum_{t=\mathbf{1}}^{\mathbf{n}}\tilde{Z}^2(t),a_N^{-1}\sum_{t=\mathbf{1}}^{\mathbf{n}}U(t),\sum_{t=\mathbf{1}}^{\mathbf{n}}V(t)\right)\overset{\mathcal{L}}{\Longrightarrow}\left(Q^{-\frac{2}{\alpha}}\sum_{l=1}^{Q}\tilde{S}_l,Q^{-\frac{1}{\alpha}}\sum_{l=1}^{Q}c_l^1S_l,Q^{-\frac{1}{\alpha}}\sum_{l=1}^{Q}c_l^2S_l\right),
$$

which is the right hand side of (24) when we make the following associations: let $\tilde{S} := Q^{-\frac{2}{\alpha}} \sum_{l=1}^{Q} \tilde{S}_Q$, $U := Q^{-\frac{1}{\alpha}} \sum_{l=1}^{Q} c_l^1 S_l$, and $V := Q^{-\frac{1}{\alpha}} \sum_{l=1}^{Q} c_l^2 S_l$. The joint characteristic function for U, V, and $\tilde{S}$ may be deduced from the expression

$$\Pi_{l=1}^{Q} \mathbb{E} \exp \left\{ i f_l Q^{-\frac{1}{\alpha}} S_l + i\eta_3 Q^{-\frac{2}{\alpha}} \tilde{S}_l \right\}$$

Now for each l, the pair $(S_l, \tilde{S}_l)$ has the characteristic function given in Theorem 2.3; putting these facts together produces the characteristic function given in Theorem 3.1. †

Acknowledgements The authors would like to thank the editor, Dmitri Chibisov, for comments regarding the joint characteristic function in Theorems 2.3, 3.1, and 3.2, which led to substantial improvements in the paper. We also thank an anonymous referee for useful comments on the text.

References

[1] Arcones, M. and Giné, E. (1989). The bootstrap of the mean with arbitrary bootstrap sample size. *Annals of the Institute Henri Poincaré* **25**, 457–481.

[2] Beran, J., Sherman R., Taqqu, M., and Willinger W. (1995) Long-range dependence in Variable-bit rate video traffic. *IEEE Trans. Comm.* **43** 1566–1579.

[3] Billingsley, P. (1995) *Probability and Measure.* John Wiley and Sons, New York.

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

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

[6] Bulinskii, A.V. (1981) *Limit Theorems for Random Processes and Fields*, Izdat. MGU, Moscow (in Russian).

[7] Bulinskii, A.V. (1986) Limit Theorems Under Weak Dependence Conditions. *Probability Theory and Mathematical Statistics* **1**, 307–326.

[8] Cline, D. (1983) Infinite Series of Random Variables with Regularly Varying Tails. Technical Report 83-24, Institute of Applied Mathematics and Statistics, University of British Columbia.

[9] Davis, R. and Hsing, T. (1995) Point Process and Partial Sum Convergence for Weakly Dependent Random Variables with Infinite Variance. *Annals of Probability* **23**, No. 2, 879–917.

[10] Davis, R., and Mikosch, T. (1998) The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Annals of Statistics* **26**, No. 5, 2049–2080.

[11] Davis, R., and Resnick, S. (1985) Limit theory for moving averages of random variables with regularly varying tail probabilities. *Annals of Probability* **13**, No. 1, 179–195.

[12] Davis, R., and Resnick, S. (1986) Limit theory for the sample covariance and correlation functions of moving averages. *Annals of Statistics* **14**, No. 2, 533–558.

[13] Durrett, R. (1996) *Probability: Theory and Examples.* Duxbury Press.

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

[15] Ivanov, I., and Leonenko, N. (1989). *Statistical Analysis of Random Fields.* Kluwer Academic Publishers, The Netherlands.

[16] Klüppelberg, C., and Mikosch, T. (1993) Spectral Estimates and Stable Processes. *Stochastic Processes and their Applications* **47** 323–344.

[17] Klüppelberg, C., and Mikosch, T. (1994) Some Limit Theory for the Self-Normalized Periodogram of Stable Processes. *Scandinavian Journal of Statistics* **21** 485–491.

[18] 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.

[19] McElroy, T. (2001) *Statistical Inference for Parameters of Time Series Exhibiting the Noah and Joseph Effects.* Ph.D. Thesis, Dept. of Mathematics, University of California, San Diego.

[20] Politis, D. and Romano, J. (1994) Large Sample Confidence Regions Based on Subsamples Under Minimal Assumptions. *Annals of Statistics* **22**, No. 4, 2031–2050.

[21] Politis, D., Romano, J., and Wolf, M. (1999) *Subsampling.* Springer, New York.

[22] Resnick, S. (1986) Point processes, regular variation, and weak convergence. *Advances in Applied Probability* **18**, 66–138.

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

[24] Resnick, S. (1997) Special Invited Paper: Heavy Tail Modeling and Teletraffic Data. *Annals of Statistics* **25**, No. 5, 1805–1849.

[25] Romano, J., and Wolf, M. (1999) Subsampling inference for the mean in the heavy-tailed case. *Metrika* **50** (1999) 1, 55-69

[26] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences* **42**, 43–47.

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

[28] Withers, C.S. (1981) Conditions for Linear Processes to be Strong Mixing. *Probability Theory and Related Fields* **57**, 477–480