LIMIT THEORY FOR BILINEAR PROCESSES WITH HEAVY TAIRED NOISE

RICHARD A. DAVIS AND SIDNEY I. RESNICK

Colorado State University and Cornell University

December 1995

Abstract. We consider a simple stationary bilinear model

\[ X_t = \xi X_{t-1} Z_{t-1} + Z_t, \quad t = 0, \pm 1, \pm 2, \ldots \]

generated by heavy tailed noise variables \( \{Z_t\} \). A complete analysis of weak limit behavior is given by means of a point process analysis. A striking feature of this analysis is that the sample correlation converges in distribution to a non-degenerate limit. A warning is sounded about trying to detect non-linearities in heavy tailed models by means of the sample correlation function.

1. Introduction.

Current efforts in time series analysis attempt to deal with data which exhibit features such as long range dependence, non-linearity and heavy tails. There are numerous data sets from the fields of telecommunications, finance and economics which appear to be compatible with the assumption of heavy-tailed marginal distributions. Examples include file lengths, cpu time to complete a job, call holding times, inter-arrival times between packets in a network and lengths of on/off cycles (Duffy, et al 1993, 1994; Meier–Hellstern et al, 1991; Willinger, Taqqui, Sherman and Wilson, 1995).

A key question of course is how to fit models to data which require heavy tailed marginal distributions. In the traditional setting of a stationary time series with finite variance, every purely non-deterministic process can be expressed as a linear process driven by an uncorrelated input sequence. For such time series, the autocorrelation function can be well approximated by that of an finite order ARMA\((p,q)\) model. In particular, one can choose an autoregressive model of order \(p\) \(AR(p)\) such that the ACF of the two models agree for lags \(1 = \ldots, p\) (see Brockwell and Davis (1991), p. 240). So from a second order point of view, linear models are sufficient for data analysis. In the infinite variance case, we have no such confidence that linear models are sufficiently flexible and rich enough for modelling purposes.

For a stationary time series \(\{X_t\}\) with infinite variance, there is no analogue of a linear process representation or approximation to it. If \(\{X_t\}\) is the linear process,

\[ X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \]

Key words and phrases. extreme value theory, Poisson processes, bilinear and non-linear processes, stable laws, point processes, stationary processes, sample correlation.

Richard Davis was partially supported by NSF Grant DMS-9504596 at Colorado State University. Sidney Resnick was partially supported by NSF Grant DMS-9400835 Cornell University and some support was also received from NSA Grant MDA904-95-H-1096

Typeset by \hbox{AMSTex}
where \( \{Z_t\} \) is an iid sequence of random variables with infinite variance, then one can still define an analogue of the ACF in terms of the coefficients \( \{\psi_j\} \) of the linear filter. Namely, \( \rho(h) = \sum_{j=0}^{\infty} \psi_j \psi_{j+h} / \sum_{j=0}^{\infty} \psi_j^2 \). Somewhat surprisingly, the sample ACF defined for heavy tailed data as

\[
\hat{\rho}_H(h) = \frac{\sum_{t=1}^{n-h} X_t X_{t+h}}{\sum_{t=1}^{n} X_t^2}, \quad h = 1, 2, \ldots,
\]

has a number of desirable properties such as consistency (\( \hat{\rho}_H(h) \stackrel{P}{\longrightarrow} \rho(h) \)) and a reasonably fast rate of convergence (see Davis and Resnick (1985b, 1986)). On the other hand, if the model is non-linear, then it is not clear what, if anything, \( \hat{\rho}(h) \) converges to. One of the principal objectives in this paper is to show that for a class of bilinear models, \( \hat{\rho}_H(h) \) converges in distribution to a non-degenerate random variable depending on \( h \). This means that other model fitting and diagnostic tools which rest on the sample ACF, such as the AIC for identifying the order of an AR model and the Yule-Walker estimates for fitting an AR model will not converge to constants either, but will converge in distribution to non-degenerate random variables.

Failure to account for non-linearities can have dramatic consequences in the analysis and be quite misleading. A full discussion is contained in the forthcoming paper by Feigin and Resnick (1996). Here we briefly illustrate the effect of non-linearities on estimation procedures for autoregressive processes. We simulated three independent samples (\( \text{test}_i, \ i = 1, 2, 3 \)) of size 5000 from the bilinear process

\[(1.1) \quad X_t = Z_{t-1} X_{t-1} + Z_t, \quad t = 0, \pm 1, \pm 2, \ldots, \]

where \( \{Z_t\} \) are iid Pareto random variables,

\[P[Z_1 > x] = 1/x, \quad x > 1.\]

A stationary solution for (1.1) is of the form

\[(1.2) \quad X_t = Z_t + \sum_{k=1}^{\infty} (1.1)^k \left( \prod_{j=1}^{k-1} Z_{t-j} \right) Z_{t-k}^2.\]

The erratic nature of the behavior of \( \hat{\rho}_H \) is illustrated in Figure 1.1 which graphs the heavy tail ACF for \( \text{test}_i, \ i = 1, 2, 3 \). The graphs look rather different reflecting the fact that we are basically sampling independently three times from the non-degenerate limit distribution of the heavy tailed ACF. If one were not aware of the non-linearity in the data, one would be tempted to model with a low order moving average based for example on the left hand plot. Furthermore, partial autocorrelation plots and plots of the AIC statistic as a function of the order of the model all show similar erratic behavior as one moves from independent sample to independent sample. So failure to account for non-linearity means there is great potential to be misled in the sorts of models one tries to fit.

![Figure 1.1. Heavy tailed ACF for 3 bilinear samples.](image-url)
In contrast, we present in Figure 1.2 comparable heavy tailed ACF plots for three independent samples of size 1500 of AR(2) data. The AR(2) is

\[ X_t = 1.3X_{t-1} - 0.7X_{t-2} + Z_t, t = 0, \pm 1, \pm 2, \ldots \]

and the innovations have a Pareto distribution as for the bilinear example. Here, the pictures look identical reflecting the fact that we are sampling from degenerate distributions.

![Figure 1.2. Heavy tailed ACF for 3 autoregressive samples.](image)

The potential for encountering such problems in modeling real data is illustrated in Section 5 of Resnick (1995) where 3802 interarrival times of ISDN D-channel packets are analyzed. From the point of view of the AIC criterion, the best fitting autoregression model is found and the autoregressive coefficients are estimated by the LP estimators of Feigin and Resnick (1994). The residuals of the autoregressive model are analyzed and pass a test for independence (Feigin, Resnick and Starica (1995)). However, when the residuals are split into three subsamples and the sample ACF is computed for each of the subsamples, we obtain three different looking functions (see Figure 1.3). One explanation could be the presence of non-linearity in the data.

![Figure 1.3. ACF of partitioned data.](image)

Section 2 of this paper deals with some mathematical preliminaries about tail properties of variables of the type appearing as the summands in (1.2). Section 3 provides a detailed point process analysis of asymptotic properties of a simple bilinear process. In Section 4 we consider some corollaries of the limit results of Section 3 with emphasis on the limiting behavior of the extremes, partial sums and sample autocorrelations from observations on a bilinear model. Unlike the linear process case, the sample autocorrelations of the bilinear process have non-degenerate limit laws.

The principle thrust of this paper is to point out that second order methods depending on the sample autocorrelation function for identification and estimation of models involving non-linearities can misguide the analyst and result in an inappropriate model being selected. In future work, we hope to discuss the weak limit behavior of higher order non-linear processes, develop an estimation theory for a broad class of non-linear models and to develop methods for the detection of non-linearities in heavy tailed phenomena.
2. Analytic results on tail weights.

We assume throughout that \( \{Z_n, -\infty < n < \infty\} \) are iid non-negative random variables with common distribution \( F \) whose tail satisfies

\[
1 - F(x) = x^{-\alpha} L(x), \quad \alpha > 0, \; x > 0
\]

where \( L \) is slowly varying at infinity. Let \( c > 0 \) be a positive constant satisfying

\[
c^{\alpha/2} E Z_1^{\alpha/2} < 1.
\]

Then it is easy to see, for instance using Holder's inequality, that

\[
X_t = \sum_{j=1}^{\infty} c^j \left( \prod_{i=1}^{j-1} Z_{i-t} \right) Z_{t-j}^2, \quad t = 0, \pm 1, \pm 2, \ldots
\]

is a well defined stationary process since the infinite series converges. Furthermore, \( \{X_t\} \) satisfies the bilinear recursion

\[
X_t = c Z_{t-1} X_{t-1} + Z_t, \quad t = 0, \pm 1, \pm 2, \ldots
\]

Set

\[
Y_t^{(0)} = Z_t,
\]

\[
Y_t^{(j)} = \left( \prod_{i=1}^{j-1} Z_{t-i} \right) Z_{t-j}^2, \quad j \geq 1,
\]

so that

\[
X_t = \sum_{j=0}^{\infty} c^j Y_t^{(j)}.
\]

We use the convention that \( \prod_{i=1}^{0} Z_i = 1 \). The condition (2.2) is stronger than Liu's (1989) condition for convergence of the infinite series in (2.3), but is required for the regular variation analysis of the tail of the distribution of \( X_t \).

We now begin with a series of lemmas designed to understand the tail behavior of \( Y_t^{(j)} \) as well as sums of these variables.

**Lemma 2.1.** Suppose \( Y_1, \ldots, Y_k \) are non-negative random variables (but not necessarily independent or identically distributed). If \( Y_1 \) has distribution \( F \) satisfying (2.1) and if as \( x \to \infty \)

\[
\frac{P[Y_i > x]}{1 - F(x)} \to c_i, \quad i = 1, \ldots, k
\]

and

\[
\frac{P[Y_i > x, Y_j > x]}{1 - F(x)} \to 0, \quad i \neq j,
\]

then

\[
\frac{P[\sum_{i=1}^{k} Y_i > x]}{1 - F(x)} \to c_1 + \cdots + c_k.
\]
Proof. Let $k = 2$. Define $b_n$ to satisfy
\[ n(1 - F(b_n)) \to 1, \quad n \to \infty \]
and on $[0, \infty]$ define the measure $\nu$ by $\nu(x, \infty] = x^{-\alpha}$. The definition of $b_n$ yields vague convergence
\[ nP[Y_t/b_n \in \cdot] \overset{\nu}{\to} c_t \nu \]
in the space of measures on $(0, \infty)$ and (2.6) implies
\[ nP[b_n^{-\lambda}(Y_1, Y_2) \in (dx, dy)] \overset{\nu}{\to} c_1 \nu(dx) \epsilon_0(dy) + c_2 \epsilon_0(dx) \nu(dy) \]
in the space of measures on $[0, \infty]^2 \setminus \{0\}$. The proof is completed as on p. 227 of Resnick (1987).

The case for general $k$ follows by induction. \[ \square \]

We now verify (2.6) and (2.7) for the variables defined in (2.5).

**Lemma 2.2.** For the variables $\{Y_t^{(j)}, j \geq 1\}$ we have, as $x \to \infty$, for all $k > j \geq 1$

\[ (1) \quad \frac{P[Y_t^{(j)} > x]}{P[Y_t^{(k)} > x]} \to c_{jk} := (EZ_1^{\alpha/2})^{j-k}, \]

and

\[ (2) \quad \frac{P[Y_t^{(j)} > x, Y_t^{(k)} > x]}{P[Y_t^{(k)} > x]} \to 0. \]

Proof. A result of Breiman (1965) (see also Resnick, 1986) says that if $\xi$ is a non-negative random variable satisfying (2.1) and if $\eta$ another non-negative random variable independent of $\xi$ satisfying $E\eta^\gamma < \infty$ for some $\gamma > \alpha$ then
\[ P[\eta \xi > x] \sim E\eta^\alpha P[\xi > x], \quad x \to \infty. \]

Since $Y_t^{(j)}$ satisfies (2.5) and $EZ_1^\gamma < \infty$ for $\alpha/2 < \gamma < \alpha$, this Breiman result applies to give for $j \geq 1$

\[ P[Y_t^{(j)} > x] \sim E\left(\prod_{i=1}^{j-1} Z_{t-i}\right)^{\alpha/2} P[Z_1^2 > x] = (EZ_1^{\alpha/2})^{j-1} x^{-\alpha/2} L(\sqrt{x}). \]

The result (1) now easily follows.

For (2), observe that
\[ P[Y_t^{(j)} > x, Y_t^{(k)} > x] = P[\left(\prod_{i=1}^{j-1} Z_{t-i}\right) Z_{t-j}^2 > x, \left(\prod_{i=1}^{k-1} Z_{t-i}\right) Z_{t-k}^2 > x] \]
\[ = P[AZ_{t-j}^2 > x, ABZ_{t-k}^2 > x] \]
\[ \leq P[A \leq \epsilon, AZ_{t-j}^2 > x] + P[A > \epsilon, AZ_{t-j}^2 > x, ABZ_{t-k}^2 > x] \]
\[ \leq P[A \leq \frac{x}{\epsilon}, Z_{t-j}^2 > x] + P[Z_{t-j}^2 > \frac{x}{\epsilon}, BZ_{t-k}^2 > \frac{x}{\epsilon}] \]
\[ = I + II. \]
Now \( I \) is handled by the result of Breiman's quoted above:

\[
\limsup_{x \to \infty} \frac{P[A_1 \cap \ldots \cap A_k, Z_{i-j} > x]}{P[Z_i^2 > x]} = E(A_1 \cap \ldots \cap A_k)^{\alpha/2} \to 0, \quad \epsilon \to 0.
\]

For \( II \) we have

\[
\frac{II}{P[Z_i^2 > x]} \sim \frac{\epsilon^{\alpha/2} P[Z_i^2 > x] E(B)^{\alpha/2} \epsilon^{\alpha/2} P[Z_i^2 > x]}{P[Z_i^2 > x]} \to 0.
\]

This completes the proof of (2).

A similar proof works if either \( j \) or \( k \) is 0. \( \square \)

Combining Lemmas 2.1 and 2.2 yield

**Corollary 2.3.** If \( \{Z_i\} \) satisfies (2.1), then

\[
P[\sum_{j=1}^{k} c^j Y_{i-j} > x] \sim \sum_{j=1}^{k} c^{j\alpha/2} (EZ_1^{\alpha/2})^{j-1} P[Z_1^2 > x], \quad x \to \infty.
\]

We now extend Corollary 2.3 so that the number of summands can be infinite.

**Corollary 2.4.** If \( \{Z_i\} \) satisfies (2.1) and \( c \) satisfies (2.2), then

\[
\lim_{x \to \infty} \frac{P[\sum_{j=1}^{\infty} c^j Y_{i-j} > x]}{P[Z_i^2 > x]} = \sum_{j=1}^{\infty} c^{j\alpha/2} (EZ_1^{\alpha/2})^{j-1} = \frac{c^{\alpha/2}}{(1 - c^{\alpha/2} EZ^{\alpha/2})}.
\]

**Proof.** The proof follows closely the argument of Cline (1983) outlined in Resnick (1987), p. 228. Clearly for any \( k \geq 1 \)

\[
P[\sum_{j=1}^{k} c^j Y_{i-j} > x] \leq P[\sum_{j=1}^{\infty} c^j Y_{i-j} > x]
\]

so that applying Corollary 2.3,

\[
\liminf_{x \to \infty} \frac{P[\sum_{j=1}^{\infty} c^j Y_{i-j} > x]}{P[Z_i^2 > x]} \geq \frac{P[\sum_{j=1}^{k} c^j Y_{i-j} > x]}{P[Z_i^2 > x]} = \sum_{j=1}^{k} c^{j\alpha/2} (EZ_1^{\alpha/2})^{j-1}.
\]

Letting \( k \to \infty \) yields a lower bound for (2.9).

The upper bound which allows Breiman's (1965) result to work also allows us to pass a limit inside an infinite summation which results in

\[
\lim_{x \to \infty} \frac{\sum_{j=1}^{\infty} P[c^j Y_{i-j} > x]}{P[Z_i^2 > x]} = \frac{c^{\alpha/2}}{(1 - c^{\alpha/2} EZ^{\alpha/2})}.
\]

To get the upper bound for (2.9) we proceed as on p. 229 of Resnick (1987). Assuming for convenience that \( 0 < \alpha < 1 \) (with a similar Holder argument when this assumption is not true) we must show

\[
\limsup_{x \to \infty} \sum_{j=1}^{\infty} \frac{E[c^j W_j Z_{i-j}^2 I[W_j Z_{i-j}^2 \leq x]]}{x P[Z_i^2 > x]^{1/2}} \leq (\text{const}) \sum_{j=1}^{\infty} c^j EW_j^\frac{\alpha}{2}
\]
for some \( \delta < \alpha \) where

\[
W_j = \prod_{i=1}^{j-1} Z_{t-i}.
\]

But

\[
\sum_{j=1}^{\infty} c^j \frac{\sum_{j=1}^{\infty} E(c^j W_j Z_{1}^j 1_{(Z_{1}^j, Z_{t-j}^j \leq x)}))}{x P[Z_{1}^j > x]} = \sum_{j=1}^{\infty} c^j \int_0^\infty w \frac{E(Z_{t-j}^j 1_{[Z_{t-j}^j \leq x/w]}))}{E(Z_{1}^j 1_{[Z_{1}^j \leq x]}))} f_{W_j}(dw) \frac{E(Z_{1}^j 1_{[Z_{1}^j \leq x]}))}{x P[Z_{1}^j > x]}
\]

and applying Potter’s inequalities (Bingham et al, 1987; Resnick, 1987; Geluk and de Haan, 1987) this is bounded by

\[
= \sum_{j=1}^{\infty} c^j \left[ \int_0^1 w(w^{-1})^{1-\alpha+\alpha-\delta} f_{W_j}(dw) \right] \text{(const)} + \int_1^{\infty} u f_{W_j}(dw)
\]

\[
\leq \sum_{j=1}^{\infty} c^j 2 \text{(const)} E W_j^\delta < \infty
\]

and the desired result follows. The rest of the proof mimics the material on p. 229-230 of Resnick (1987). \( \square \)

3. Point process convergence.

In this section, we investigate the limit behavior of a sequence of point processes associated with a bilinear time series model. Let \( \{X_t\} \) be the simple first order bilinear time series defined as a stationary solution to the equations

\[
X_t = c X_{t-1} Z_{t-1} + Z_t,
\]

where \( \{Z_t\} \) is an iid sequence of random variables with regularly varying tail probabilities. Specifically, we assume

\[
P(|Z_1| > x) = x^{-\alpha} L(x), \ \alpha > 0, \ L(x) \text{ is regularly varying,}
\]

and

\[
P(Z_1 > x) \rightarrow p \quad \text{and} \quad P(Z_1 < -x) \rightarrow q,
\]

as \( x \rightarrow \infty, 0 \leq p \leq 1 \) and \( q = 1 - p \). Similar to the condition imposed on \( c \) in Section 2, we assume

\[
|c|^{\alpha/2} E|Z_1|^{\alpha/2} < 1.
\]

Under this condition (see Liu (1989)), there exists a unique stationary solution to the equations (3.1) given by

\[
X_t = \sum_{j=0}^{\infty} c^j Y_t^{(j)},
\]

where

\[
Y_t^{(j)} = \begin{cases} 
Z_t, & \text{if } j = 0, \\
\left(\prod_{i=1}^{j-1} Z_{t-i}\right) Z_{t-j}^2, & \text{if } j \geq 1.
\end{cases}
\]
The object of interest in this section is the sequence of point processes based on the points \( \{b_n^{-2}X_t, t = 1, \ldots, n\} \), where \( b_n \) is the \( 1 - n^{-1} \) quantile of \( |Z_1| \), i.e.

\[
(3.6) \quad b_n = \inf\{x : P(|Z_1| > x) < n^{-1}\}.
\]

Before discussing the relevant limit theory, we quickly review the salient facts of point process theory.

For a locally compact Hausdorff topological space \( \mathbb{E} \), we let \( M_p(\mathbb{E}) \) be the space of Radon point measures on \( \mathbb{E} \). This means \( m \in M_p(\mathbb{E}) \) is of the form

\[
m = \sum_{i=1}^{\infty} \epsilon_{x_i},
\]

where \( x_i \in \mathbb{E} \) are the locations of the point masses of \( m \) and \( \epsilon_{x_i} \) denotes the point measure defined by

\[
\epsilon_{x_i}(A) = \begin{cases} 
1, & \text{if } x \in A, \\
0, & \text{if } x \notin A.
\end{cases}
\]

We emphasize that we assume that all measures in \( M_p(\mathbb{E}) \) are Radon which means that for any \( m \in M_p(\mathbb{E}) \) and any compact \( K \subset \mathbb{E} \), \( m(K) < \infty. \) On the space \( M_p(\mathbb{E}) \) we use the vague metric \( \rho(\cdot, \cdot) \). Its properties are discussed for example in Resnick (1987, Section 3.4) and Kallenberg (1983). Note that a sequence of measures \( m_n \in M_p(\mathbb{E}) \) converge vaguely to \( m_0 \in M_p(\mathbb{E}) \) if for any continuous function \( f: \mathbb{E} \mapsto [0, \infty) \) with compact support we have \( m_n(f) \to m_0(f) \) where \( m_n(f) = \int_\mathbb{E} f \, dm_n. \) The non-negative continuous functions with compact support with be denoted by \( C^+_K(\mathbb{E}) \).

A Poisson process on \( \mathbb{E} \) with mean measure \( \mu \) will be denoted by \( \text{PRM}(\mu) \). The primary example of interest in our applications is the case when \( \mathbb{E}_m = [-\infty, \infty]^m \setminus \{0\} \), where compact sets are closed subsets of \( [-\infty, \infty]^m \) which are bounded away from \( 0 \).

We begin with the following point process convergence result which underpins the main results of this section. This result is a slight generalization of Proposition 3.2 in Feigin, Kratz and Resnick (1994) to the case when \( Z_t \) may have negative values.

**Proposition 3.1.** Suppose the marginal distribution \( F \) of the iid sequence \( \{Z_t\} \) satisfies (3.2)--(3.3) and \( m \) is a fixed positive integer. Suppose further that \( \sum_{s=1}^{\infty} \epsilon_{x_s} \) is \( \text{PRM}(\mu) \) where \( \mu(dx) = \alpha(x^{-\alpha-1}_{[x>0]} + q(-x)^{-\alpha-1}_{[x<0]}dx \) and \( \{U_{kl}, U_{kl}'; k \geq 1, l \geq 1\} \) are iid with distribution \( F \). If \( \epsilon_t \in [-\infty, \infty]^m \) denotes the basis element with \( t \)th component equal to one and the rest zero and \( \mathbb{E}_m = [-\infty, \infty]^m \setminus \{0\} \), then

\[
\sum_{t=1}^{n} \epsilon(t^{-1}(Z_{t-1}, i=1, \ldots, m), Z_{t-1}, j=1, \ldots, m) \Rightarrow \sum_{s=1}^{\infty} \epsilon(j, \epsilon_{1}, sgn(j_1)_{[0]} U_{11}, \ldots, U_{s, m-1})
\]

\[
+ \sum_{s=1}^{\infty} \epsilon(j, \epsilon_{2}, U_{11}, sgn(j_1)_{[0]} U_{11}', \ldots, U_{s, m-2})
\]

\[
+ \cdots + \sum_{s=1}^{\infty} \epsilon(j, \epsilon_{m}, U_{11}, \ldots, U_{s, m-1}, sgn(j_1)_{[0]})
\]

in \( M_p(\mathbb{E} \times [-\infty, \infty]^m) \).

Now for \( k = 1, \ldots, m, \) consider the point processes defined on space \( \mathbb{E}_k := [-\infty, \infty] \setminus 0 \) given by

\[
I^k_n = \sum_{t=1}^{n} \epsilon_{-2Y_t(k)},
\]

where \( Y_t(k) \) is as defined in (3.5). We first establish the joint convergence of \( (I^{(1)}_n, \ldots, I^{(m)}_n) \) on \( M_p^m(\mathbb{E}_k) \).
Proposition 3.2. Under the assumptions of Proposition 3.1, we have
\[ (I_n^{(1)}, \ldots, I_n^{(m)}) \Rightarrow (I^{(1)}, \ldots, I^{(m)}) \]
on $M^m_p(E_1)$, where $I^k = \sum_{s=1}^{\infty} \epsilon_{I^s(U_u \ldots U_{us-1})}$.

Proof. For $k \in \{1, \ldots, m\}$, the restriction
\[ g(x_1, \ldots, x_m, u_1, \ldots, u_m) = (x_k, u_1, \ldots, u_{k-1}) \]
is a continuous mapping from $E_m \times [-\infty, \infty]^m$ into $E_1 \times [-\infty, \infty]^{k-1}$ with the property that $g^{-1}(K)$ is compact for every compact $K \subseteq E_1 \times [-\infty, \infty]^{k-1}$. This mapping, therefore, induces a continuous mapping (see Resnick (1987)) from $M_p(E_m \times [-\infty, \infty]^m)$ into $M_p(E_1 \times [-\infty, \infty]^{k-1})$ and hence
\[ \bar{I}_n^k := \sum_{t=1}^{n} \epsilon_{i_{t-k+1}^i \ldots i_{t-1}^i, i_{t-k}^i \ldots i_{t-1}^i} \Rightarrow \bar{I}_k^k := \sum_{s=1}^{\infty} \epsilon_{i_{s-k}^i \ldots i_{s-1}^i, i_{s}^i \ldots i_{s-1}^i} \]
where the convergence is joint in $k = 1, \ldots, m$. If $M$ and $-M$ are continuity points of $F$, then this convergence also holds for these point processes when restricted to the set $E_1 \times [-M, M]^{k-1}$. That is,
\[ I_n^{k, M} := \bar{I}_n^k \cap (E_1 \times [-M, M]^{k-1}) \Rightarrow I^{k, M} := \bar{I}_k^k \cap (E_1 \times [-M, M]^{k-1}) \]
jointly for $k = 1, \ldots, m$ in $M_p^m(E_1 \times [-\infty, \infty]^{k-1})$.

Now consider the mapping $f_k : E_1 \times [-M, M]^{k-1} \mapsto E_1 \times [-M, M]^{k-1}$ defined by
\[ f_k(x, u_1, \ldots, u_{k-1}) = \begin{cases} x^2 u_1 \cdots u_k, & \text{if } \sum_{i=1}^{k-1} |u_i| < \infty, \\ 17, & \text{otherwise}. \end{cases} \]
Observe that if $K$ is the compact set in $E_1$ given by $\{x : |x| > b\}$, then $f_k^{-1}(K) \cap (E_1 \times [-M, M]^{k-1}) \subseteq \{x : |x| > (b/M^{k-1})^{1/2} \times [-M, M]^{k-1}\}$ which is compact in $E_1 \times [-\infty, \infty]^{k-1}$. It follows that $f_k^{-1}(K)$ restricted to $E_1 \times [-M, M]^{k-1}$ is compact for any compact subset $K$ of $E_1$, and since $f_k$ is continuous on the support of $\bar{I}_k^k$ a.s., we have by Corollary 1.2 of Resnick (1986) and (3.8) that
\[ (\bar{I}_n^k \circ f_1^{-1}, \ldots, \bar{I}_n^k \circ f_m^{-1}) \Rightarrow (\bar{I}_1^k \circ f_1^{-1}, \ldots, \bar{I}_m^k \circ f_m^{-1}) \]
in $M^m_p(E_1)$. Since the point processes $I^k = \bar{I}_k^k \circ f_k^{-1}$ are well defined Poisson processes, we have as $M \to \infty$
\[ (\bar{I}_1^k \circ f_1^{-1}, \ldots, \bar{I}_m^k \circ f_m^{-1}) \Rightarrow (I^1, \ldots, I^m) \]
pointwise in the vague metric a.s. Noting that $I_n^k = \bar{I}_n^k \circ f_k^{-1}$, the conclusion of the proposition will follow (see Theorem 4.2 in Billingsley (1968)), once we show that for each $k$ and any $\eta > 0$,
\[ \lim_{M \to \infty} \limsup_{n \to \infty} P[\rho(\bar{I}_n^k \circ f_k^{-1}, \bar{I}_n^k \circ f_k^{-1}) > \eta] = 0. \]
By the form of the metric $\rho$, it is enough to show for any $h \in C^k_F(E_1)$ that
\[ \lim_{M \to \infty} \limsup_{n \to \infty} P[\bar{I}_n^k \circ f_k^{-1}(h) - I_n^k \circ f_k^{-1}(h)] > \eta = 0. \]
If the support of $h$ is contained in the set $G_\delta = \{x : |x| > \delta\}$, then the probability in (3.9) is bounded by $P[\bar{I}_n^k(G_\delta \times K_M^c) \geq 1]$ where $\delta' = (\delta/M^{k-1})^{1/2}$ and $K_M^c$ is the complement of $K_M := [-M, M]^{k-1}$. Using (3.7), this probability converges as $n \to \infty$ to,$
\[ P[\bar{I}_k^k(G_\delta \times K_M^c) \geq 1] \to 0 \text{ as } M \to \infty, \]
which completes the proof. $\square$
Proposition 3.3. On the space \( \mathcal{M}_p(\mathbb{E}_m) \) with vague metric \( \rho_m \),

\[
\rho_m \left( \sum_{t=1}^{n} \varepsilon_{b_n^2}^{(1)}(Y_t^{(1)}, \ldots, Y_t^{(m)}), \sum_{t=1}^{n} \sum_{i=1}^{m} \varepsilon_{b_n^2}^{(i)} y_{t,i} \right) \rightarrow 0.
\]

Proof. Let \( B = [b_1, c_1] \times \cdots \times [b_m, c_m] \) be a bounded rectangle in \( \mathbb{E}_m \). Then either \( B \) is bounded away from each of the coordinate axes or intersects exactly one in an interval (see p. 181 of Davis and Resnick (1985)). If \( B \) is bounded away from each of the coordinate axes, then

\[
E \left[ \sum_{t=1}^{n} \varepsilon_{b_n^2}^{(1)}(Y_t^{(1)}, \ldots, Y_t^{(m)})(B) \right] = nP[b_n^2(Y_1^{(1)}, \ldots, Y_1^{(m)}) \in B] \rightarrow 0
\]

by Lemma 2.2. The remainder of the proof of the proposition follows the same lines of reasoning given for Proposition 2.1 of Davis and Resnick (1985) and hence is omitted. \( \square \)

Theorem 3.4. Suppose \( \{X_t\} \) is the bilinear process (3.1) where the marginal distribution \( F \) of the iid noise \( \{Z_t\} \) satisfies (3.2)–(3.3), the constant \( c \) satisfies (3.4) and \( b_n \) is given by (3.6). If \( \sum_{t=1}^{\infty} \varepsilon_{s,t} \) is PRM(\( \mu \)) with \( \mu \) given in Proposition 3.1 and \( \{U_{s,k}, s \geq 1, k \geq 1\} \) are iid with distribution \( F \), then

(i) in \( \mathcal{M}_p(\mathbb{E}_2) \),

\[
\sum_{t=1}^{n} \varepsilon_{b_n^2}^{(1)}X_t = \sum_{k=1}^{\infty} \sum_{s=1}^{\infty} \varepsilon_{x_k^2}^{b} W_{s,k},
\]

where

\[
W_{s,k} = \begin{cases} 
\prod_{i=1}^{k-1} U_{s,i}, & \text{if } k > 1, \\
1, & \text{if } k = 1, \\
0, & \text{if } k < 1.
\end{cases}
\]

(ii) In \( \mathcal{M}_p(\mathbb{E}_{n+1}) \),

\[
\sum_{t=1}^{n} \varepsilon_{b_n^2}(X_t, X_{t-1}, \ldots, X_{t-k}) \Rightarrow \sum_{k=1}^{\infty} \sum_{s=1}^{\infty} \varepsilon_{x_k^2}^{b}(W_{s,k} W_{s,k-1, \ldots, W_{s,k-1, \ldots, W_{s,k}, \ldots, W_{s,k}, \ldots, W_{s,k}, \ldots, W_{s,k}, \ldots})).
\]

Proof. (i) Propositions 3.2 and 3.3 imply that

\[
\sum_{t=1}^{n} \varepsilon_{b_n^2}(Y_t^{(1)}, \ldots, Y_t^{(m)}) \Rightarrow \sum_{k=1}^{m} \sum_{s=1}^{\infty} \varepsilon_{x_k^2}^{b} (U_{s,1} \cdots U_{s,k-1}) a_k = \sum_{k=1}^{m} \sum_{s=1}^{\infty} \varepsilon_{x_k^2}^{b} W_{s,k} a_k.
\]

on \( \mathcal{M}_p(\mathbb{E}_m) \). Now the map

\[
(y_1, \ldots, y_m) \mapsto \sum_{k=1}^{m} c_k y_k
\]

induces a continuous map from \( \mathcal{M}_p(\mathbb{E}_m) \rightarrow \mathcal{M}_p(\mathbb{E}_2) \) and so by the continuous mapping theorem applied to the convergence in (3.10) we obtain

\[
\sum_{t=1}^{n} \varepsilon_{b_n^2}^{(1)}(Y_t^{(1)}, \ldots, Y_t^{(m)}) \Rightarrow \sum_{k=1}^{m} \sum_{s=1}^{\infty} \varepsilon_{x_k^2}^{b} W_{s,k}.
\]
in \( M_p(\mathbb{E}_d) \). As \( m \to \infty \),

\[
\sum_{k=1}^{m} \sum_{s=1}^{\infty} \varepsilon_{j}^{2} c_{s} W_{s-k} \to \sum_{k=1}^{\infty} \sum_{s=1}^{\infty} \varepsilon_{j}^{2} c_{s} W_{s-k}
\]

pointwise in the vague metric and so by Theorem 4.2 in Billingsley (1968), it suffices to show that for any \( \eta > 0 \) and \( f \in C_{b}^{\infty}(\mathbb{E}_1) \),

\[
\lim_{m \to \infty} \limsup_{n \to \infty} P[|\sum_{t=1}^{n} f(b_n^{-2} \sum_{k=1}^{m} c^k Y_{t}^{(k)}) - \sum_{t=1}^{n} f(b_n^{-2} X_t)| > \eta] = 0.
\]

To prove (3.11) first note that

\[
P[b_n^{-2} \sqrt{n} \sum_{t=1}^{n} c^k Y_{t}^{(k)} - X_t > \eta] \leq P[b_n^{-2} \sqrt{n} |Y_t^{(0)}| + \sum_{k=m+1}^{\infty} |c|^k |Y_t^{(k)}| > \eta]
\]

\[
\leq n P[b_n^{-2} |Y_t^{(0)}| > \eta/2] + n P[b_n^{-2} \sum_{k=m+1}^{\infty} |c|^k |Y_t^{(k)}| > \eta/2]
\]

which, by Corollary 2.4,

\[
\to 0 + (\eta/2)^{-\alpha/2} \sum_{j=m+1}^{\infty} |c|^j (E|Z_1|^\alpha/2)^{j-1}
\]

\[
\to 0, \text{ as } m \to \infty.
\]

The remainder of the proof of (3.11) is now identical to the argument given for (2.11) in Davis and Resnick (1985) with this last result substituting for Lemma 2.3 of the Davis and Resnick paper.

(ii) We shall only sketch out the proof in the case \( h = 1 \), the general case being a straightforward adaptation of this argument. First observe that \( Y_{t}^{(k)} = Z_{t-1} Y_{t-1}^{(k-1)} \) so that

\[
(Y_{t}^{(1)}, \ldots, Y_{t}^{(m)}, Y_{t-1}^{(1)}, \ldots, Y_{t-1}^{(m-1)}) = (Y_{t}^{(1)}, Z_{t-1}(Y_{t-1}^{(1)}, \ldots, Y_{t-1}^{(m-1)}), Y_{t-1}^{(1)}, \ldots, Y_{t-1}^{(m-1)}).
\]

Using a slight modification to the arguments given in Propositions 3.2 and 3.3 we obtain the point process convergence result

\[
\sum_{t=1}^{n} \varepsilon_{t} b_n^{-2} (Y_{t}^{(1)}, \ldots, Y_{t}^{(m)}, Y_{t-1}^{(1)}, \ldots, Y_{t-1}^{(m-1)}) \Rightarrow \sum_{k=1}^{m} \sum_{s=1}^{\infty} \varepsilon_{j}^{2} (W_{s-k} c_{k} + W_{s-k-1} c_{k+1})
\]

in \( M_p(\mathbb{R}^{2m-1}) \) where the \( \varepsilon_{t} \) are the unit basis elements in \( \mathbb{R}^{2m-1} \). Then, using the continuous mapping of \( M_p(\mathbb{R}^{2m-1}) \to M_p(\mathbb{E}_2) \) induced by the function

\[
(s_1, \ldots, s_m, u_1, \ldots, u_{m-1}) \mapsto \left( \sum_{k=1}^{m} c^k s_k, \sum_{k=1}^{m-1} c^k u_k \right),
\]

we obtain

\[
\sum_{t=1}^{n} \varepsilon_{t} b_n^{-2} (\sum_{k=1}^{m} c^k Y_{t}^{(k)}, \sum_{k=1}^{m-1} c^k Y_{t-1}^{(k)}) \Rightarrow \sum_{k=1}^{m} \sum_{s=1}^{\infty} \varepsilon_{j}^{2} (c_{k} W_{s-k} c_{k-1} + W_{s-k-1} c_{k+1})
\]

The rest of the proof of (ii) is the same as that given in (i). \( \square \)
Remark 3.1. While it was not required in the proofs of the results in this section, it can be shown that $X_t$ has regularly varying tail probabilities with index $\alpha/2$. This assertion extends Lemma 2.2 to non-positive $Z_t$ and/or negative coefficient $c$. A direct proof of this property can be fashioned after the argument used in Lemma 2.2 as in Cline (1983) for linear processes.

4. Applications.

By applying continuous functionals to the basic convergence result of Theorem 3.4, the limiting behavior for a number of statistics can be easily derived. We now explore some of these applications.

(A) Extremes. The point process convergence in (i) of Theorem 3.4 allows one to compute the joint limiting distribution of any collection of upper and lower extreme order statistics. To illustrate these computations in a simple case, let $M_n = \max\{X_1, \ldots, X_n\}$ and note that $\{b_n^{-2}M_n \leq x\} \equiv \{N_n(x, \infty) = 0\}$ where, $N_n$ is the point process $N_n = \sum_{t=1}^n \delta_{b_n^{-2}X_t}$. It follows that

$$P[b_n^{-2}M_n \leq x] = P[N_n(x, \infty) = 0] \rightarrow P[N(x, \infty) = 0],$$

where $N = \sum_{k=1}^\infty \sum_{i=1}^\infty \delta_{j_{x,k}c^iW_{s,k}}$. Now the event $\{N(x, \infty) = 0\}$ is equivalent to the event that none of the points $\{j_{x,k}c^iW_{s,k}, s \geq 1, k \geq 1\}$ exceeds $x$. The latter can be expressed as the set $\cap_{s=1}^\infty \{j_{x,k}c^iW_{s,k} \leq x\}$, where $V_s = \max_{k=1}^\infty (c^kW_{s,k})$ and since $\{j_{x,k}c^iW_{s,k}\}$ are the points of a PRM on $[0, \infty]$ with mean measure $\nu(x, \infty) = EV_1^{\alpha/2}x^{-\alpha/2}$ (see equation (4.4) of Resnick (1986)), we have

$$P[N(x, \infty) = 0] = \begin{cases} 0, & \text{if } x \leq 0, \\ \exp\{E(V_1^{\alpha/2})x^{-\alpha}\}, & \text{if } x > 0. \end{cases}$$

(B) Partial sums. If the exponent $\alpha$ of regular variation is less than four, then the partial sum $S_n = \sum_{t=1}^n X_t$ of the bilinear process $\{X_t\}$ is asymptotically stable with exponent $\alpha/2$. These results are essentially special cases of Theorem 3.1 in Davis and Hsing (1995). For the case $\alpha \in (0,2)$, $X_t$ has regularly varying tails with exponent $\alpha/2 < 1$ (see Remark 3.1) and hence a direct application of Theorem 3.1 (i) of Davis and Hsing (1995) yields,

$$b_n^{-2}S_n \Rightarrow S := \sum_{k=1}^\infty \sum_{i=1}^\infty j_{x,k}c^iW_{s,k} = \sum_{i=1}^\infty j_{x,k}c^iA_s,$$

where $A_s = \sum_{k=1}^\infty c^kW_{s,k}$. (The characteristic function of $S$ is given in Theorem 3.2 of Davis and Hsing (1995).)

For the case $\alpha \in [2,4)$, a direct application of Theorem 3.2 in Davis and Hsing (1995) is more difficult since condition (3.2) of the theorem must be checked. Instead, we present a different approach under the simplified assumptions that $\alpha \in (2,4)$ and the distribution of $Z_t$ is symmetric about zero. The condition $\alpha > 2$ implies that $\Var(Z_t) < \infty$ and the symmetry of $Z_t$ allows for a simpler expression of the limit random variable in terms of the points of the truncated sequence $\{X_t^{(m)} = \sum_{k=0}^m c^kY_t^{(k)}\}$, we obtain

$$b_n^{-2}\sum_{i=1}^n (X_t^{(m)} - \mu_n) \Rightarrow S_1 + \cdots + S_m := S^{(m)},$$

where $\mu_n = E(Z_1^{(m)}I_{Z_1^{(m)} \leq \mu_n})$ and

$$S_k = \begin{cases} \lim_{c \to 0} \sum_{x=0}^\infty j_{x,k}^21_{|z|>\gamma} - \alpha(\alpha - 2)^{-1}(\epsilon^{1-\alpha/2} - 1), & \text{if } k = 1, \\ \sum_{x=1}^\infty j_{x,k}^2c^kW_{s,k}, & \text{if } k > 1. \end{cases}$$
(It is easy to check that \( b_n^{-2} \sum_{t=1}^{n} Y_t(0) = o_p(1) \). Using characteristic functions, one can show that \( S^{(m)} \Rightarrow S := \sum_{k=1}^{\infty} S_k \). Next we show that for any \( \eta > 0 \),

\[
(4.1) \quad \lim_{m \to \infty} \limsup_{n \to \infty} P[b_n^{-2}|S_n - \sum_{t=1}^{n} X_t^{(m)}| > \eta] = 0
\]

from which the limit,

\[
b_n^{-2}(S_n - n\mu_n) \Rightarrow S,
\]

will follow immediately from Theorem 4.2 of Billingsley (1968). For \( \delta > 0 \) fixed, write

\[
b_n^{-2}(S_n - \sum_{t=1}^{n} X_t^{(m)}) = b_n^{-2} \sum_{t=1}^{n} \sum_{k=m+1}^{\infty} c^k Y_t^{(k)} = b_n^{-2} \sum_{t=1}^{n} \sum_{k=m+1}^{\infty} c^k Y_t^{(k)} \mathbb{1}_{[Z_{t-k} \leq b_n^2 \delta]} + b_n^{-2} \sum_{t=1}^{n} \sum_{k=m+1}^{\infty} c^k Y_t^{(k)} \mathbb{1}_{[Z_{t-k} > b_n^2 \delta]}
\]

(4.2)

The absolute value of the second term in (4.2) has expectation bounded by

\[
nb_n^{-2} \sum_{k=m+1}^{\infty} \mathbb{E}[|c^k (E[Z_1])^{k-1} E(Z_1^2 \mathbb{1}_{[Z_1 \leq b_n^2 \delta]} - (\text{const}) \sum_{k=m+1}^{\infty} r^k nb_n^{-2} E(Z_1^2 \mathbb{1}_{[Z_1 \leq b_n^2 \delta]}),
\]

where, by assumption (3.4) \( r := \mathbb{E}[cZ_1] < (\mathbb{E}[cZ_1]^{1/2})^{2/\alpha} < 1 \) for \( \alpha > 2 \). Using Karamata’s theorem, the right hand side of (4.3) is asymptotic to

\[
(\text{const}) r^m nb_n^{-2} (b_n^2 \delta)^{-\alpha/2} \to (\text{const}) r^m \delta^{-\alpha/2}, \quad \text{as } n \to \infty,
\]

(4.4)

On the other hand, the mean zero assumption of \( Z_t \) implies that the \( Y_t^{(k)} \mathbb{1}_{[Z_{t-k} \leq b_n^2 \delta]} \)'s are uncorrelated for all \( t \) and \( k \) so that the variance of the first term in (4.2) is

\[
b_n^{-4} \sum_{t=1}^{n} \sum_{k=m+1}^{\infty} c^{2k} \text{Var}(Y_t^{(k)} \mathbb{1}_{[Z_t \leq b_n^2 \delta]}) \leq nb_n^{-4} \sum_{k=m+1}^{\infty} c^{2k} (E[Z_1^2])^{-k-1} E\left(Z_1^2 \mathbb{1}_{[Z_1 \leq b_n^2 \delta]}\right).
\]

Using Karamata’s theorem again, the right hand side is

\[
\leq (\text{const}) \gamma^m \delta^2 nP[Z_1^2 > b_n^2 \delta] \to (\text{const}) \gamma \delta^{2-\alpha/2}, \quad \text{as } n \to \infty,
\]

(4.5)

where \( \gamma = \mathbb{E}(cZ_1)^2 < 1 \). This, combined with (4.4) proves (4.1) as asserted.

(C) The sample correlation function. We now consider the behavior of the vector of heavy tailed sample correlations \( \{\rho_H(l), l = 1, \ldots, h\} \) for integers \( h = 1, 2, \ldots \). Recall that \( \rho_H(l) \) was defined in Section 1 to be

\[
\rho_H(l) = \frac{\sum_{t=1}^{n-l} X_tX_{t+l}}{\sum_{t=1}^{n} X_t^2}.
\]

In Davis and Resnick (1985, 1986) we showed that for a heavy tailed MA(\( \infty \)) process, the sample ACF was a consistent estimate of the model ACF expressed in terms of the coefficients of the linear filter. This is of course also the case in the classical setting where the innovation variables have finite second moment (see Brockwell and Davis, 1991). In contrast to this phenomena of constant limits, we find for the non-linear process that sample correlations converge in distribution to non-degenerate limit random variables depending on the lag.
Theorem 4.1. Suppose \( \{X_t\} \) is the bilinear process (3.1) where the marginal distribution \( F \) of the iid noise \( \{Z_t\} \) satisfies (3.2)-(3.3), and the constant \( c \) satisfies (3.4). If \( 0 < \alpha < 4 \) we have for any \( h = 1, 2, \ldots \) that
\[
(\hat{\beta}_H(l), l = 1, \ldots, h) \Rightarrow (L_i, i = 1, \ldots, h)
\]
in \( \mathbb{R}^h \), where in the notation of Theorem 3.4
\[
L_i = \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} j_s^4 c^{2k-i} W_{s,k} W_{s,k-i}, \quad i = 1, \ldots, h.
\]

Proof. Theorem 3.4 (ii) implies
\[
\left( \sum_{t=1}^{n} \epsilon_{n}^{2}(X_t, X_{t-1}), l = 1, \ldots, h \right) \Rightarrow \left( \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} \epsilon_{f}^{2}(c^k W_{s,k}, c^{k-1} W_{s,k-1}), l = 1, \ldots, h \right)
\]
in \( \mathcal{L}^p(\mathbb{E}_2) \). In order to simplify the exposition, we focus on convergence of a single component in (4.5) but at the end of the discussion it should be obvious that joint convergence ensues.

For convenience we focus on the first component convergence in (4.5):
\[
\sum_{t=1}^{n} \epsilon_{n}^{2}(X_t, X_{t-1}) \Rightarrow \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} \epsilon_{f}^{2}(c^k W_{s,k}, c^{k-1} W_{s,k-1}).
\]

Pick \( \delta > 0 \) and apply a restriction of the state space to
\[
\mathbb{E}_\delta = \{(x_1, x_2) \in \mathbb{E}_2 : |x_1| \vee |x_2| > \delta \}
\]
to obtain
\[
\sum_{t=1}^{n} \epsilon_{n}^{2}(X_t, X_{t-1})(\cdot \cap \mathbb{E}_\delta) \Rightarrow \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} \epsilon_{f}^{2}(c^k W_{s,k}, c^{k-1} W_{s,k-1})(\cdot \cap \mathbb{E}_\delta).
\]
As in the discussion after (3.8), because the state space has been compactified by restriction, we may apply the functional which multiplies components to obtain
\[
\sum_{t=2}^{n} 1_{|X_t| \vee |X_{t-1}| > \delta} \epsilon_{n}^{2}(X_t, X_{t-1}) \Rightarrow \sum_{k=1}^{\infty} \sum_{s=1}^{\infty} 1_{|f(c^k W_{s,k}) \vee f(c^{k-1} W_{s,k-1}) > \delta} \epsilon_{f}^{2}(c^{2k-1} W_{s,k} W_{s,k-1}).
\]

Each point process on the previous line has only a finite number of points and applying the summation functional we get
\[
\gamma_n, \delta(1) := \sum_{t=2}^{n} 1_{|X_t| \vee |X_{t-1}| > \delta} b_n^{-4}(X_t, X_{t-1})
\]
\[
\Rightarrow \gamma_{\infty, \delta}(1) := \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} 1_{|f(c^k W_{s,k}) \vee f(c^{k-1} W_{s,k-1}) > \delta} j_s^4 c^{2k-1} W_{s,k} W_{s,k-1}.
\]

We claim
\[
\gamma_n, 0(1) \Rightarrow \gamma_{\infty, 0}(1)
\]
in $\mathbb{R}$. To prove this we check (Billingsley, 1968, Theorem 4.2)

\[(4.7) \quad \gamma_{\infty, \delta}(1) \Rightarrow \gamma_{\infty, 0}(1), \quad \delta \downarrow 0,\]

and

\[(4.8) \quad \lim_{\delta \downarrow 0} \lim_{n \to \infty} \sup P[|\gamma_{n, \delta}(1) - \gamma_{n, 0}(1)| > \eta] = 0.\]

To verify (4.7), it will be sufficient to check that the series

\[(4.9) \quad \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} j_s^2 c^{2k-1} W_{s,k} W_{s,k-1} = \sum_{s=1}^{\infty} j_s^3 B_s,\]

where $B_s = \sum_{k=1}^{\infty} c^{2k-1} W_{s,k} W_{s,k-1}$, is absolutely convergent. Since $\alpha/4 < 1$ we have by the triangle inequality

\[(4.10) \quad E|B_s|^{\alpha/4} \leq \sum_{k=1}^{\infty} (c(c^{2k-1})^{\alpha/4} E|W_{s,k-1}|^{\alpha/4} E|Z_l|^{\alpha/4}) \leq (\text{const}) \sum_{k=1}^{\infty} (E|c Z_l|^{\alpha/2})^{k-1} < \infty,\]

The last inequality follows from (3.4). The independence of the $B_s$ together with (4.10) implies that $\sum_{s=1}^{\infty} \epsilon_s |B_s|$ is PRM with intensity measure $\mu(x, \infty) = (E|B_1|^{\alpha/4}) x^{-\alpha/4}$ and hence has absolutely summable points a.s. (see Resnick (1986) and Davis and Resnick (1985), p. 192).

It remains to check (4.8). This is a standard argument mimicking the one given in Davis and Resnick (1985), p. 193. The probability in (4.8) is bounded by

\[P\left[\sum_{i=1}^{n \mathbb{1}_{X_i 1}^{\leq \varepsilon_n \eta}} \frac{1}{\eta}\left|\sum_{i=1}^{n} \frac{1}{\varepsilon_n \eta}\right| X_i 1_{\left|X_i 1\right| < \varepsilon_n \eta} > \eta\right] \leq \frac{n}{b_n^2} E\left|X_1 1_{\left|X_1 1\right| \leq \varepsilon_n \eta}\right| / \eta\]

which by Cauchy’s inequality is dominated by

\[\frac{n}{b_n^2} E\left|X_1 1_{\left|X_1 1\right| \leq \varepsilon_n \eta}\right| / \eta\]

and since $P[|X_1| > x]$ is regularly varying with index $-\alpha/2 \in (-2, 0)$ we get by Karamata’s theorem that

\[\lim_{\delta \downarrow 0} \lim_{n \to \infty} \frac{n}{b_n^2} E\left|X_1 1_{\left|X_1 1\right| \leq \varepsilon_n \eta}\right| / \eta = 0\]

which proves (4.8).

We have now checked $\gamma_{n, 0}(1) \Rightarrow \gamma_{\infty, 0}(1)$ and in fact, examining the proof of this fact shows that

\[(4.11) \quad (\gamma_{n, 0}(0), \gamma_{n, 0}(1)) \Rightarrow (\gamma_{\infty, 0}(0), \gamma_{\infty, 0}(1))\]

where $\gamma_{n, 0}(0) = \sum_{i=2}^{n} X_i X_{i-1} / b_n^2$. Dividing the first component into the second in (4.11) yields the first component convergence given in the statement of the Theorem. This finishes our discussion of the proof.
REFERENCES


RICHARD A. DAVIS, DEPARTMENT OF STATISTICS, COLORADO STATE UNIVERSITY, Ft. COLLINS, CO 80523 USA

E-mail address: rda@stat.colostate.edu

SIDNEY I. RESNICK, CORNELL UNIVERSITY, SCHOOL OF OPERATIONS RESEARCH AND INDUSTRIAL ENGINEERING, ETC BUILDING, ITHACA, NY 14853 USA

E-mail address: sid@orie.cornell.edu