

LIMIT THEORY FOR BILINEAR PROCESSES WITH HEAVY-TAILED NOISE

BY RICHARD A. DAVIS¹ AND SIDNEY I. RESNICK²

Colorado State University and Cornell University

We consider a simple stationary bilinear model $X_t = cX_{t-1}Z_{t-1} + Z_t$, $t = 0, \pm 1, \pm 2, \dots$, 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 nondegenerate limit. A warning is sounded about trying to detect nonlinearities 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, nonlinearity 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, interarrival times between packets in a network and lengths of on/off cycles [Duffy, McIntosh, Rosenstein and Willinger (1993, 1994); Meier-Hellstern, Wirth, Yan and Hoefflin (1991); Willinger, Taqqu, 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 nondeterministic process can be expressed as a linear process driven by an uncorrelated input sequence. For such time series, the autocorrelation function (ACF) can be well approximated by that of a finite order ARMA(p, q) model. In particular, one can choose an autoregressive (AR) model of order p [AR(p)] such that the ACF of the two models agree for lags $1, \dots, p$ [see Brockwell and Davis (1991), page 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 modeling 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

Received December 1995; revised July 1996.

¹Research partially supported by NSF Grant DMS-95-04596 at Colorado State University.

²Research partially supported by NSF Grant DMS-94-00535 at Cornell University. Additional support was received from NSA Grant MDA904-95-H-1036.

AMS 1991 subject classifications. 60E07, 60F17, 60G55, 60G70, 62M10.

Key words and phrases. Extreme value theory, Poisson processes, bilinear and nonlinear processes, stable laws, point processes, stationary processes, sample correlation.

linear process,

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

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

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

has a number of desirable properties such as consistency [$\hat{\rho}_H(h) \rightarrow_P \rho(h)$] and a reasonably fast rate of convergence [see Davis and Resnick (1985b, 1986)]. On the other hand, if the model is nonlinear, 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 nondegenerate random variable depending on h . This means that other model fitting and diagnostic tools which rest on the sample ACF, such as the Akaike Information Criterion (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 nondegenerate random variables.

Failure to account for nonlinearities can have dramatic consequences in the analysis and be quite misleading. Additional discussion is contained in Feigin and Resnick (1996). Here we briefly illustrate the effect of nonlinearities on estimation procedures for autoregressive processes. We simulated three independent samples (test_i , $i = 1, 2, 3$) of size 5000 from the bilinear process

$$(1.1) \quad X_t = 0.1Z_{t-1}X_{t-1} + Z_t, \quad t = 0, \pm 1, \pm 2, \dots,$$

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

$$(1.2) \quad X_t = Z_t + \sum_{k=1}^{\infty} (0.1)^k \left(\prod_{j=1}^{k-1} Z_{t-j} \right) Z_{t-k}^2.$$

The erratic nature of the behavior of $\hat{\rho}_H$ is illustrated in Figure 1, which graphs the heavy-tailed ACF for test_i , $i = 1, 2, 3$. The graphs look rather different, reflecting the fact that we are basically sampling independently three times from the nondegenerate limit distribution of the heavy-tailed ACF. If one were not aware of the nonlinearity 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

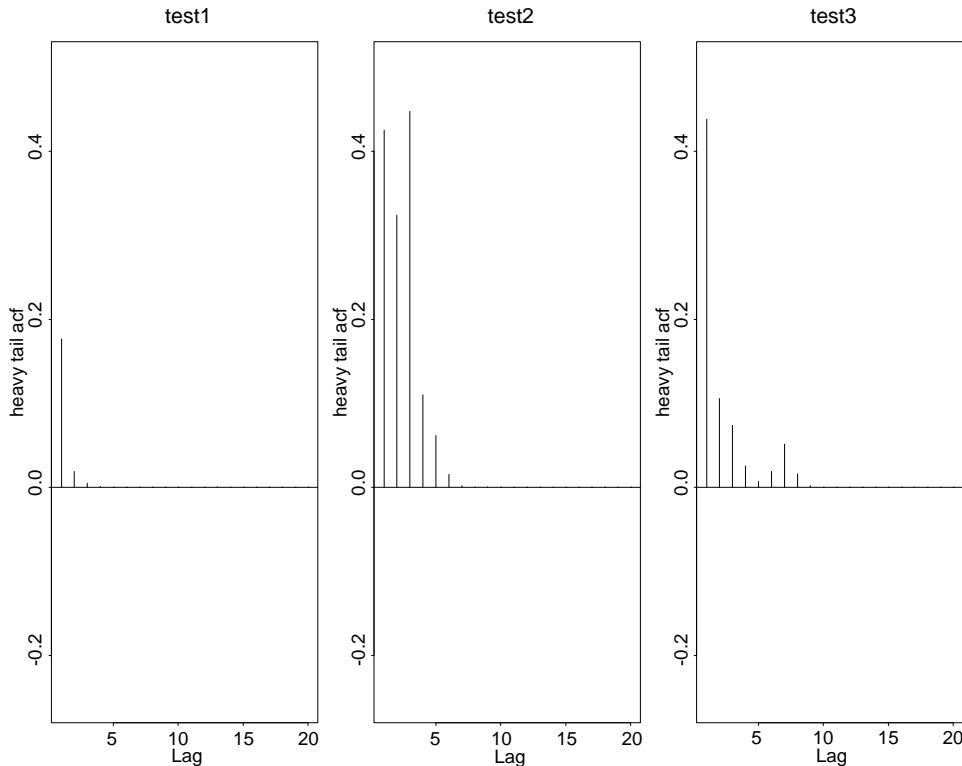


FIG. 1. Heavy-tailed ACF for three bilinear samples.

behavior as one moves from independent sample to independent sample. So failure to account for nonlinearity means there is great potential to be misled in the sorts of models one tries to fit.

In contrast, we present in Figure 2 comparable heavy-tailed ACF plots for three independent samples of size 1500 of AR(2) data. The AR(2) is

$$X_t = 1.3X_{t-1} - 0.7X_{t-2} + Z_t, \quad t = 0, \pm 1, \pm 2, \dots,$$

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.

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 linear programming (LP) estimators of Feigin and Resnick (1994). The residuals of the autoregressive model are analyzed and pass a test for independence [Feigin, Resnick and Stărică (1995)]. However, when the residuals are split into three subsamples and the

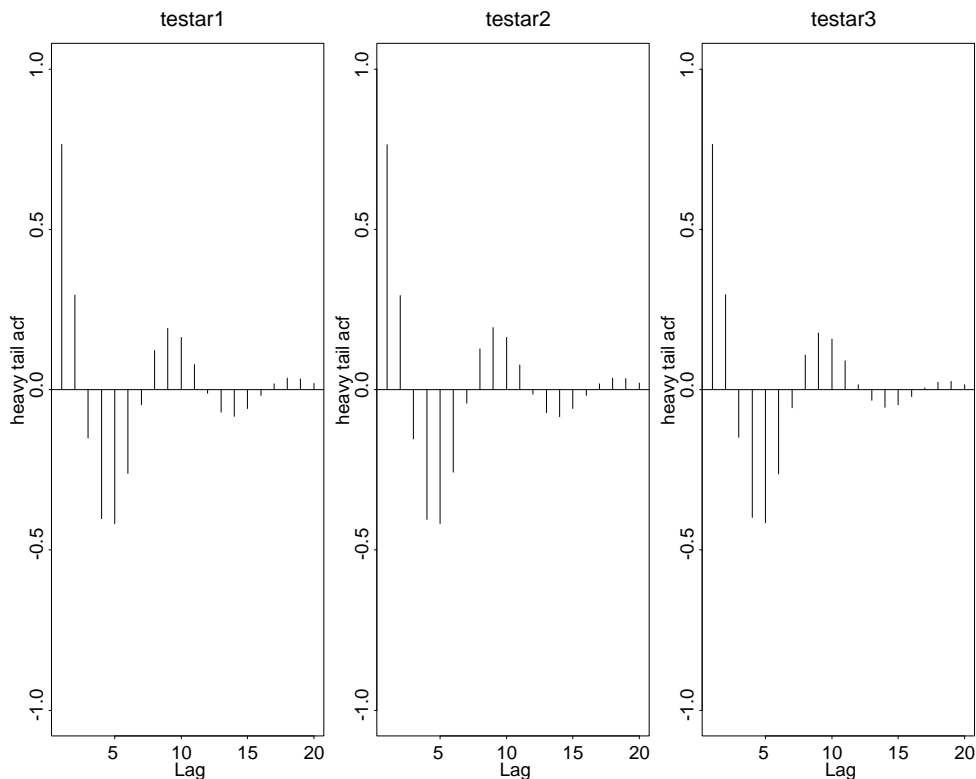


FIG. 2. *Heavy-tailed ACF for three autoregressive samples.*

sample ACF is computed for each of the subsamples, we obtain three different looking functions (see Figure 3). One explanation could be the presence of nonlinearity in the data.

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 nondegenerate limit laws.

The principal 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 nonlinearities 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 nonlinear processes, develop an estimation theory for a broad class of nonlinear models and develop methods for the detection of nonlinearities in heavy-tailed phenomena.

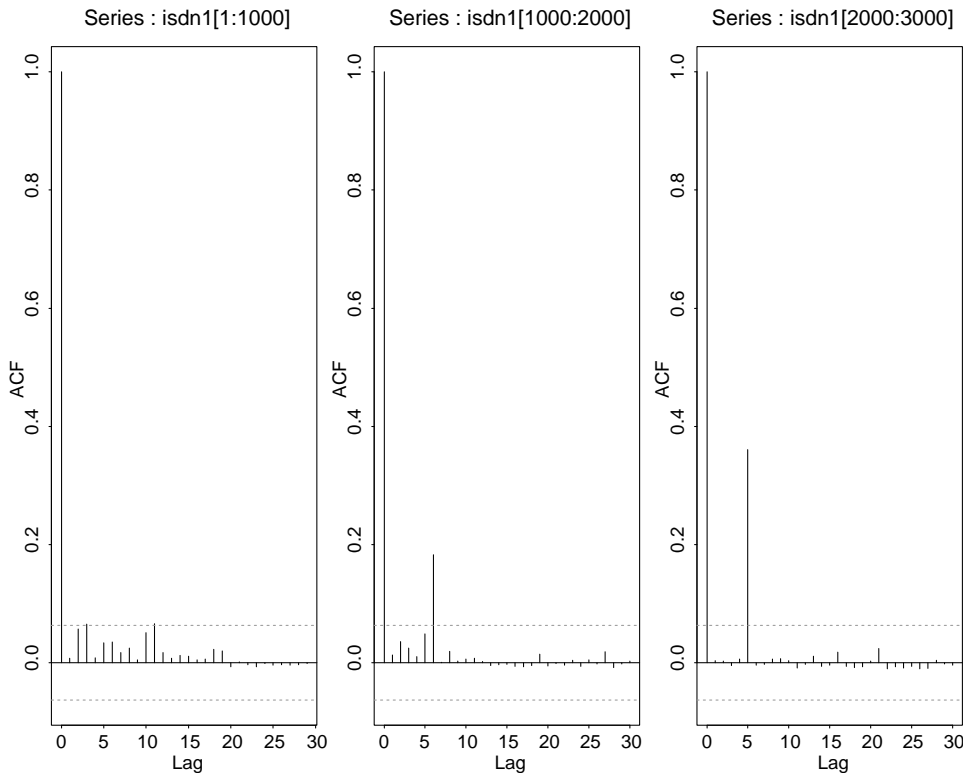


FIG. 3. ACF of partitioned data.

2. Analytic results on tail weights. We assume throughout that $\{Z_n, -\infty < n < \infty\}$ are iid nonnegative random variables with common distribution F whose tail satisfies

$$(2.1) \quad 1 - F(x) = x^{-\alpha}L(x), \quad \alpha > 0, x > 0,$$

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

$$(2.2) \quad c^{\alpha/2}EZ_1^{\alpha/2} < 1.$$

Then it is easy to see, for instance using Hölder’s inequality, that

$$(2.3) \quad X_t = Z_t + \sum_{j=1}^{\infty} c^j \left(\prod_{i=1}^{j-1} Z_{t-i} \right) Z_{t-j}^2, \quad t = 0, \pm 1, \pm 2, \dots,$$

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

$$(2.4) \quad X_t = cZ_{t-1}X_{t-1} + Z_t, \quad t = 0, \pm 1, \pm 2, \dots$$

Set

$$(2.5) \quad \begin{aligned} Y_t^{(0)} &= Z_t, \\ Y_t^{(j)} &= \left(\prod_{i=1}^{j-1} Z_{t-i} \right) Z_{t-j}^2, \quad j \geq 1, \end{aligned}$$

so that

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

We use the convention that $\prod_{i=1}^0 = 1$. 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, \dots, Y_k are nonnegative random variables (but not necessarily independent or identically distributed). If Y_1 has distribution F satisfying (2.1) and if as $x \rightarrow \infty$,*

$$(2.6) \quad \frac{P[Y_i > x]}{1 - F(x)} \rightarrow c_i, \quad i = 1, \dots, k,$$

and

$$(2.7) \quad \frac{P[Y_i > x, Y_j > x]}{1 - F(x)} \rightarrow 0, \quad i \neq j,$$

then

$$\frac{P[\sum_{i=1}^k Y_i > x]}{1 - F(x)} \rightarrow c_1 + \dots + c_k.$$

PROOF. Let $k = 2$. Define b_n to satisfy

$$n(1 - F(b_n)) \rightarrow 1, \quad n \rightarrow \infty,$$

and on $(0, \infty]$ define the measure ν by $\nu(x, \infty] = x^{-\alpha}$. The definition of b_n yields vague convergence

$$nP[Y_i/b_n \in \cdot] \rightarrow_v c_i \nu$$

in the space of measures on $(0, \infty]$ and (2.6) implies

$$nP[b_n^{-1}(Y_1, Y_2) \in (dx, dy)] \rightarrow_v c_1 \nu(dx) \varepsilon_0(dy) + c_2 \varepsilon_0(dx) \nu(dy)$$

in the space of measures on $[0, \infty]^2 \setminus \{\mathbf{0}\}$. The proof is completed as in Resnick [(1987), page 227].

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 $x \rightarrow \infty$, for all $k > j \geq 1$,

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

and

$$(ii) \quad \frac{P[Y_t^{(j)} > x, Y_t^{(k)} > x]}{P[Y_t^{(k)} > x]} \rightarrow 0.$$

PROOF. A result of Breiman (1965) [see also Resnick (1986)] says that if ξ is a nonnegative random variable satisfying (2.1) and if η is another nonnegative random variable independent of ξ satisfying $E\eta^\gamma < \infty$ for some $\gamma > \alpha$, then

$$P[\eta\xi > x] \sim E\eta^\alpha P[\xi > x], \quad x \rightarrow \infty.$$

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

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

The result (i) now easily follows.

For (ii), observe that

$$\begin{aligned} &P[Y_t^{(j)} > x, Y_t^{(k)} > x] \\ &= P\left[\left(\prod_{i=1}^{j-1} Z_{t-i}\right)Z_{t-j}^2 > x, \left(\prod_{i=1}^{j-1} Z_{t-i}\right)\left(\prod_{i=j}^{k-1} Z_{t-i}\right)Z_{t-k}^2 > x\right] \\ &=: P[AZ_{t-j}^2 > x, ABZ_{t-k}^2 > x] \\ &\leq P[A \leq \varepsilon, AZ_{t-j}^2 > x] + P[A > \varepsilon, AZ_{t-j}^2 > x, ABZ_{t-k}^2 > x] \\ &\leq P[A1_{[A \leq \varepsilon]}Z_{t-j}^2 > x] + P\left[Z_{t-j}^2 > \frac{x}{\varepsilon}, BZ_{t-k}^2 > \frac{x}{\varepsilon}\right] \\ &= I + II. \end{aligned}$$

Now I is handled by the result of Breiman quoted above:

$$\limsup_{x \rightarrow \infty} \frac{P[A1_{[A \leq \varepsilon]}Z_{t-j}^2 > x]}{P[Z_1^2 > x]} = E(A1_{[A \leq \varepsilon]})^{\alpha/2} \rightarrow 0, \quad \varepsilon \rightarrow 0.$$

For II we have

$$\frac{II}{P[Z_1^2 > x]} \sim \frac{\varepsilon^{\alpha/2} P[Z_1^2 > x] E(B)^{\alpha/2} \varepsilon^{\alpha/2} P[Z_1^2 > x]}{P[Z_1^2 > x]} \rightarrow 0.$$

This completes the proof of (ii).

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

Combining Lemmas 2.1 and 2.2 yields the following corollary.

COROLLARY 2.3. *If $\{Z_t\}$ satisfies (2.1), then*

$$(2.8) \quad P\left[\sum_{j=1}^k c^j Y_t^{(j)} > x\right] \sim \sum_{j=1}^k c^{j\alpha/2} (EZ_1^{\alpha/2})^{j-1} P[Z_1^2 > x], \quad x \rightarrow \infty.$$

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

COROLLARY 2.4. *If $\{Z_t\}$ satisfies (2.1) and c satisfies (2.2), then*

$$(2.9) \quad \lim_{x \rightarrow \infty} \frac{P\left[\sum_{j=1}^{\infty} c^j Y_t^{(j)} > x\right]}{P[Z_1^2 > x]} = \sum_{j=1}^{\infty} c^{j\alpha/2} (EZ_1^{\alpha/2})^{j-1} = \frac{c^{\alpha/2}}{(1 - c^{\alpha/2} EZ_1^{\alpha/2})}.$$

PROOF. The proof follows closely the argument of Cline (1983) outlined in Resnick [(1987), page 228]. Clearly for any $k \geq 1$,

$$P\left[\sum_{j=1}^k c^j Y_t^{(j)} > x\right] \leq P\left[\sum_{j=1}^{\infty} c^j Y_t^{(j)} > x\right],$$

so that applying Corollary 2.3,

$$\liminf_{x \rightarrow \infty} \frac{P\left[\sum_{j=1}^{\infty} c^j Y_t^{(j)} > x\right]}{P[Z_1^2 > x]} \geq \frac{P\left[\sum_{j=1}^k c^j Y_t^{(j)} > x\right]}{P[Z_1^2 > x]} = \sum_{j=1}^k c^{j\alpha/2} (EZ_1^{\alpha/2})^{j-1}.$$

Letting $k \rightarrow \infty$ 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

$$\lim_{x \rightarrow \infty} \frac{\sum_{j=1}^{\infty} P\left[c^j Y_t^{(j)} > x\right]}{P[Z_1^2 > x]} = \frac{c^{\alpha/2}}{(1 - c_1^{\alpha/2} EZ_1^{\alpha/2})}.$$

To get the upper bound for (2.9) we proceed as in Resnick [(1987), page 229]. Assuming for convenience that $0 < \alpha < 1$ (with a similar Hölder argument when this assumption is not true) we must show

$$\limsup_{x \rightarrow \infty} \frac{\sum_{j=1}^{\infty} E c^j W_j Z_{t-j}^2 \mathbf{1}_{[W_j Z_{t-j}^2 \leq x]}}{x P[Z_1^2 > x]} \leq (\text{const}) \sum_{j=1}^{\infty} c^j E W_j^{\delta}$$

for some $\delta < \alpha$, where

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

However,

$$\frac{\sum_{j=1}^{\infty} E c^j W_j Z_j^2 \mathbf{1}_{\{W_j Z_{t-j}^2 \leq x\}}}{x P[Z_1^2 > x]} = \sum_{j=1}^{\infty} c^j \int_0^{\infty} w \frac{E(Z_{t-j}^2 \mathbf{1}_{\{Z_{t-j}^2 \leq x/w\}})}{E(Z_1^2 \mathbf{1}_{\{Z_1^2 \leq x\}})} F_{W_j}(dw) \frac{E(Z_1^2 \mathbf{1}_{\{Z_1^2 \leq x\}})}{x P[Z_1^2 > x]}$$

and applying Potter’s inequalities [Bingham, Goldie and Teugels (1987); Resnick (1987); Geluk and de Haan (1987)], this is bounded by

$$\begin{aligned} &= \sum_{j=1}^{\infty} c^j \left[\int_0^1 w (w^{-1})^{1-\alpha+\alpha-\delta} F_{W_j}(dw) (\text{const}) + \int_1^{\infty} u F_{W_j}(dw) \right] \\ &\leq \sum_{j=1}^{\infty} c^j 2 (\text{const}) E W_j^{\delta} < \infty \end{aligned}$$

and the desired result follows. The rest of the proof mimics the material in Resnick [(1987), pages 229–230]. □

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

$$(3.1) \quad X_t = c X_{t-1} Z_{t-1} + Z_t,$$

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

$$(3.2) \quad P[|Z_1| > x] = x^{-\alpha} L(x), \quad \alpha > 0, L(x) \text{ is regularly varying,}$$

and

$$(3.3) \quad \frac{P[Z_1 > x]}{P[|Z_1| > x]} \rightarrow p \quad \text{and} \quad \frac{P[Z_1 < -x]}{P[|Z_1| > x]} \rightarrow q,$$

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

$$(3.4) \quad |c|^{\alpha/2} E|Z_1|^{\alpha/2} < 1.$$

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

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

where

$$(3.5) \quad Y_t^{(j)} = \begin{cases} Z_t, & \text{if } j = 0, \\ \left(\prod_{i=1}^{j-1} Z_{t-i} \right) Z_{t-j}^2, & \text{if } j \geq 1. \end{cases}$$

The object of interest in this section is the sequence of point processes based on the points $\{b_n^{-2} X_t, t = 1, \dots, n\}$, where b_n is the $1 - n^{-1}$ quantile of $|Z_1|$,

that is,

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

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

$$m = \sum_{i=1}^{\infty} \varepsilon_{x_i},$$

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

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

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 $m_n \in M_p(\mathbb{E})$ converges 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 $m_n(f) \rightarrow m_0(f)$, where $m_n(f) = \int_{\mathbb{E}} f dm_n$. The nonnegative continuous functions with compact support will be denoted by $C_K^+(\mathbb{E})$.

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

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 (1996) to the case when Z_i may have negative values.

PROPOSITION 3.1. *Suppose the marginal distribution F of the iid sequence $\{Z_i\}$ satisfies (3.2)–(3.3) and m is a fixed positive integer. Suppose further that $\sum_{s=1}^{\infty} \varepsilon_{j_s}$ is $\text{PRM}(\mu)$, where $\mu(dx) = \alpha(px^{-\alpha-1}1_{[x>0]}) + q(-x)^{-\alpha-1}1_{[x<0]}$ dx and $\{U_{kl}, U'_{kl}, k \geq 1, l \geq 1\}$ are iid with distribution F . If $\mathbf{e}_i \in [-\infty, \infty]^m$ denotes the basis element with i th component equal to 1 and the rest zero and $\mathbb{E}_m = [-\infty, \infty]^m \setminus \{\mathbf{0}\}$, then*

$$\begin{aligned} \sum_{t=1}^n \varepsilon_{(b_n^{-1}(Z_{t-i}, i=1, \dots, m), Z_{t-j}, j=1, \dots, m)} &\Rightarrow \sum_{s=1}^{\infty} \varepsilon_{(j_s \mathbf{e}_1, \text{sgn}(j_s)^\infty, U'_{s1}, \dots, U'_{s, m-1})} \\ &+ \sum_{s=1}^{\infty} \varepsilon_{(j_s \mathbf{e}_2, U_{s1}, \text{sgn}(j_s)^\infty, U'_{s1}, \dots, U'_{s, m-2})} \\ &+ \dots + \sum_{s=1}^{\infty} \varepsilon_{(j_s \mathbf{e}_m, U_{s, m-1}, \dots, U_{s1}, \text{sgn}(j_s)^\infty)} \end{aligned}$$

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

Now for $k = 1, \dots, m$, consider the point processes defined on space $\mathbb{E}_1 := [-\infty, \infty] \setminus 0$ given by

$$I_n^k = \sum_{t=1}^m \varepsilon_{b_n^{-2} Y_t^{(k)}},$$

where $Y_t^{(k)}$ is as defined in (3.5). We first establish the joint convergence of $(I_n^{(1)}, \dots, I_n^{(m)})$ on $M_p^m(\mathbb{E}_1)$.

PROPOSITION 3.2. *Under the assumptions of Proposition 3.1, we have*

$$(I_n^{(1)}, \dots, I_n^{(m)}) \Rightarrow (I^{(1)}, \dots, I^{(m)})$$

on $M_p^m(\mathbb{E}_1)$, where $I^k = \sum_{s=1}^{\infty} \varepsilon_{J_s^2(U_{s,1}, U_{s,2}, \dots, U_{s,k-1})}$.

PROOF. For $k \in \{1, \dots, m\}$, the restriction

$$g(x_1, \dots, x_m, u_1, \dots, u_m) = (x_k, u_1, \dots, u_{k-1})$$

is a continuous mapping from $\mathbb{E}_m \times [-\infty, \infty]^m$ into $\mathbb{E}_1 \times [-\infty, \infty]^{k-1}$ with the property that $g^{-1}(K)$ is compact for every compact $K \subset \mathbb{E}_1 \times [-\infty, \infty]^{k-1}$. This mapping, therefore, induces a continuous mapping [see Resnick (1987)] from $M_p(\mathbb{E}_m \times [-\infty, \infty]^m)$ into $M_p(\mathbb{E}_1 \times [-\infty, \infty]^{k-1})$ and hence

$$(3.7) \quad \tilde{I}_n^k := \sum_{t=1}^n \varepsilon_{(b_n^{-1} Z_{t-k}, Z_{t-1}, \dots, Z_{t-k+1})} \Rightarrow \tilde{I}^k := \sum_{s=1}^{\infty} \varepsilon_{(J_s, U_{s,k-1}, \dots, U_{s,1})},$$

where the convergence is joint in $k = 1, \dots, m$. If M and $-M$ are continuity points of F , then this convergence also holds for these point processes when restricted to the set $\mathbb{E}_1 \times [-M, M]^{k-1}$. That is,

$$(3.8) \quad \begin{aligned} I_n^{k,M}(\cdot) &:= \tilde{I}_n^k(\cdot \cap (\mathbb{E}_1 \times [-M, M]^{k-1})) \\ &\Rightarrow I^{k,M}(\cdot) := \tilde{I}^k(\cdot \cap (\mathbb{E}_1 \times [-M, M]^{k-1})) \end{aligned}$$

jointly for $k = 1, \dots, m$ in $M_p^m(\mathbb{E}_1 \times [-\infty, \infty]^{k-1})$.

Now consider the mapping $f_k: \mathbb{E}_1 \times [-M, M]^{k-1} \mapsto \mathbb{E}_1$ defined by

$$f_k(x, u_1, \dots, u_{k-1}) = \begin{cases} x^2 u_1 \cdots u_k, & \text{if } \forall_{i=1}^{k-1} |u_i| < \infty, \\ 17, & \text{otherwise.} \end{cases}$$

Observe that if K is the compact set in \mathbb{E}_1 given by $\{x: |x| \geq b\}$, then $f_k^{-1}(K) \cap (\mathbb{E}_1 \times [-M, M]^{k-1}) \subset \{x: |x| > (b/M^{k-1})^{1/2}\} \times [-M, M]^{k-1}$ which is compact in $\mathbb{E}_1 \times [-\infty, \infty]^{k-1}$. It follows that $f_k^{-1}(K)$ restricted to $\mathbb{E}_1 \times [-M, M]^{k-1}$ is compact for any compact subset K of \mathbb{E}_1 , and since f_k is continuous on the support of $\tilde{I}^{k,M}$ a.s., we have by Corollary 1.2 of Resnick (1986) and (3.8) that

$$(\tilde{I}_n^{1,M} \circ f_1^{-1}, \dots, \tilde{I}_n^{m,M} \circ f_m^{-1}) \Rightarrow (\tilde{I}^{1,M} \circ f_1^{-1}, \dots, \tilde{I}^{m,M} \circ f_m^{-1})$$

in $M_p^m(\mathbb{E}_1)$. Since the point processes $I^k = \tilde{I}^k \circ f_k^{-1}$ are well defined Poisson processes, we have as $M \rightarrow \infty$,

$$(\tilde{I}^{1,M} \circ f_1^{-1}, \dots, \tilde{I}^{m,M} \circ f_m^{-1}) \Rightarrow (I^1, \dots, I^m)$$

pointwise in the vague metric a.s. Noting that $I_n^k = \tilde{I}_n^k \circ f_k^{-1}$, the conclusion of the proposition will follow [see Theorem 4.2 in Billingsley (1968)] once we show that for each k and any $\eta > 0$,

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[\rho \left(\tilde{I}_n^{k, M} \circ f_k^{-1}, \tilde{I}_n^k \circ f_k^{-1} \right) > \eta \right] = 0.$$

By the form of the metric ρ , it is enough to show for any $h \in C_K^+(\mathbb{E}_1)$ that

$$(3.9) \quad \lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[\left| \tilde{I}_n^{k, M} \circ f_k^{-1}(h) - I_n^k \circ f_k^{-1}(h) \right| > \eta \right] = 0.$$

If the support of h is contained in the set $G_\delta = \{x: |x| > \delta\}$, then the probability in (3.9) is bounded by $P[\tilde{I}_n^k(G_{\delta'} \times K_M^c) \geq 1]$, 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 $n \rightarrow \infty$ to

$$P[\tilde{I}^k(G_{\delta'} \times K_M^c) \geq 1] \rightarrow 0 \quad \text{as } M \rightarrow \infty,$$

which completes the proof. \square

PROPOSITION 3.3. *On the space $M_p(\mathbb{E}_m)$ with vague metric ρ_m ,*

$$\rho_m \left(\sum_{t=1}^n \varepsilon_{b_n^{-2}(Y_t^{(1)}, \dots, Y_t^{(m)})}, \sum_{l=1}^m \sum_{t=1}^n \varepsilon_{b_n^{-2} Y_t^{(l)} \mathbf{e}_l} \right) \rightarrow_P 0.$$

PROOF. Let $B = (b_1, c_1] \times \dots \times (b_m, c_m]$ 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 Davis and Resnick (1985a), page 181]. If B is bounded away from each of the coordinate axes, then

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

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 (1985a) 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 $\sum_{s=1}^\infty \varepsilon_{j_s}$ is PRM(μ) with μ given in Proposition 3.1 and $\{U_{s,k}, s \geq 1, k \geq 1\}$ are iid with distribution F , then:*

(i) *In $M_p(\mathbb{E}_1)$,*

$$\sum_{t=1}^n \varepsilon_{b_n^{-2} X_t} \Rightarrow \sum_{k=1}^\infty \sum_{s=1}^\infty \varepsilon_{j_s^2 c^k W_{s,k}},$$

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 $M_p(\mathbb{E}_{h+1})$,

$$\sum_{t=1}^n \mathcal{E} b_n^{-2}(X_t, X_{t-1}, \dots, X_{t-h}) \Rightarrow \sum_{k=1}^{\infty} \sum_{s=1}^{\infty} \mathcal{E} j_s^2(c^k W_{s,k}, c^{k-1} W_{s,k-1}, \dots, c^{k-h} W_{s,k-h}).$$

PROOF. (i) Propositions 3.2 and 3.3 imply that

$$(3.10) \quad \sum_{t=1}^n \mathcal{E} b_n^{-2}(Y_t^{(1)}, \dots, Y_t^{(m)}) \Rightarrow \sum_{k=1}^m \sum_{s=1}^{\infty} \mathcal{E} j_s^2(U_{s,1}, \dots, U_{s,k-1}) \mathbf{e}_k = \sum_{k=1}^m \sum_{s=1}^{\infty} \mathcal{E} j_s^2 W_{s,k} \mathbf{e}_k$$

on $M_p(\mathbb{E}_m)$. Now the map

$$(y_1, \dots, y_m) \mapsto \sum_{k=1}^m c^k y_k$$

induces a continuous map from $M_p(\mathbb{E}_m) \mapsto M_p(\mathbb{E}_1)$ and so by the continuous mapping theorem applied to the convergence in (3.10) we obtain

$$\sum_{t=1}^n \mathcal{E} b_n^{-2} \sum_{k=1}^m c^k Y_t^{(k)} \Rightarrow \sum_{k=1}^m \sum_{s=1}^{\infty} \mathcal{E} j_s^2 c^k W_{s,k}$$

in $M_p(\mathbb{E}_1)$. As $m \rightarrow \infty$,

$$\sum_{k=1}^m \sum_{s=1}^{\infty} \mathcal{E} j_s^2 c^k W_{s,k} \rightarrow \sum_{k=1}^{\infty} \sum_{s=1}^{\infty} \mathcal{E} j_s^2 c^k W_{s,k}$$

pointwise in the vague metric and so by Theorem 4.1 in Billingsley (1968), it suffices to show that for any $\eta > 0$ and $f \in C_K^+(\mathbb{E}_1)$,

$$(3.11) \quad \lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[\left| \sum_{t=1}^n f \left(b_n^{-2} \sum_{k=1}^m c^k Y_t^{(k)} \right) - \sum_{t=1}^n f(b_n^{-2} X_t) \right| > \eta \right] = 0.$$

To prove (3.11) first note that

$$\begin{aligned} & P \left[b_n^{-2} \bigvee_{t=1}^n \left| \sum_{k=1}^m c^k Y_t^{(k)} - X_t \right| > \eta \right] \\ & \leq P \left[b_n^{-2} \bigvee_{t=1}^n \left(|Y_t^{(0)}| + \sum_{k=m+1}^{\infty} |c^k| |Y_t^{(k)}| \right) > \eta \right] \\ & \leq nP \left[b_n^{-2} |Y_1^{(0)}| > \eta/2 \right] + nP \left[b_n^{-2} \sum_{k=m+1}^{\infty} |c^k| |Y_t^{(k)}| > \eta/2 \right], \end{aligned}$$

which, by Corollary 2.4,

$$\begin{aligned} &\rightarrow 0 + (\eta/2)^{-\alpha/2} \sum_{j=m+1}^{\infty} |c|^{j\alpha/2} (\mathbf{E}|Z_1|^{\alpha/2})^{j-1} \\ &\rightarrow 0 \quad \text{as } m \rightarrow \infty. \end{aligned}$$

The remainder of the proof of (3.11) is now identical to the argument given for (2.11) in Davis and Resnick (1985a) 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

$$\begin{aligned} &(Y_t^{(1)}, \dots, Y_t^{(m)}, Y_{t-1}^{(1)}, \dots, Y_{t-1}^{(m-1)}) \\ &= (Y_t^{(1)}, Z_{t-1}(Y_{t-1}^{(1)}, \dots, Y_{t-1}^{(m-1)}), Y_{t-1}^{(1)}, \dots, Y_{t-1}^{(m-1)}). \end{aligned}$$

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

$$\sum_{t=1}^n \mathcal{E}_{b_n^{-2}(Y_t^{(1)}, \dots, Y_t^{(m)}, Y_{t-1}^{(1)}, \dots, Y_{t-1}^{(m-1)})} \Rightarrow \sum_{k=1}^m \sum_{s=1}^{\infty} \mathcal{E}_{j_s^2(W_{s,k} \mathbf{e}_k + W_{s,k-1} \mathbf{e}_{m+k-1})}$$

in $M_p(\mathbb{E}_{2m-1})$, where the \mathbf{e}_i are the unit basis elements in \mathbb{R}^{2m-1} . Then, using the continuous mapping of $M_p(\mathbb{E}_{2m-1}) \mapsto M_p(\mathbb{E}_2)$ induced by the function

$$(x_1, \dots, x_m, u_1, \dots, u_{m-1}) \mapsto \left(\sum_{k=1}^m c^k x_k, \sum_{k=1}^{m-1} c^k u_k \right),$$

we obtain

$$\sum_{t=1}^n \mathcal{E}_{b_n^{-2}(\sum_{k=1}^m c^k Y_t^{(k)}, \sum_{k=1}^{m-1} c^k Y_{t-1}^{(k)})} \Rightarrow \sum_{k=1}^m \sum_{s=1}^{\infty} \mathcal{E}_{j_s^2(c^k W_{s,k}, c^{k-1} W_{s,k-1})}.$$

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 nonpositive 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 Theorem 3.4(i) 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

$M_n = \max\{X_1, \dots, X_n\}$ and note that $\{b_n^{-2}M_n \leq x\} = \{N_n(x, \infty) = 0\}$, where N_n is the point process $N_n = \sum_{t=1}^n \varepsilon b_n^{-2} X_t$. It follows that

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

where $N = \sum_{k=1}^{\infty} \sum_{s=1}^{\infty} \varepsilon_j^2 c^k W_{s,k}$. Now the event $\{N(x, \infty) = 0\}$ is equivalent to the event that none of the points $\{j_s^2 c^k W_{s,k}, s \geq 1, k \geq 1\}$ exceeds x . The latter can be expressed as the set $\cap_{s=1}^{\infty} \{j_s^2 V_s \leq x\}$, where $V_s = \bigvee_{k=1}^{\infty} (c^k W_{s,k})$ and since $\{j_s^2 V_s, s \geq 1\}$ are the points of a PRM on $(0, \infty]$ with mean measure $\nu(x, \infty) = EV_1^{\alpha/2} x^{-\alpha/2}$ [see Resnick (1986), (4.4)], 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 α of regular variation is less than 4, 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_1 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

$$b_n^{-2}S_n \Rightarrow S := \sum_{k=1}^{\infty} \sum_{s=1}^{\infty} j_s^2 c^k W_{s,k} = \sum_{s=1}^{\infty} j_s^2 A_s,$$

where $A_s = \sum_{k=1}^{\infty} c^k W_{s,k}$. [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 $\text{Var}(Z_t) < \infty$ and the symmetry of Z_t allows for a simpler expression of the limit random variable in terms of the points $\{j_s c^k W_{s,k}\}$ of the limit point process. Applying Theorem 3.2(ii) to the truncated sequence $\{X_t^{(m)} = \sum_{k=0}^m c^k Y_t^{(k)}\}$, we obtain

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

where $\mu_n = E(Z_1^2 1_{[Z_1^2 \leq b_n^2]})$ and

$$S_k = \begin{cases} \lim_{\varepsilon \rightarrow 0} \sum_{s=1}^{\infty} j_s^2 1_{[j_s^2 > \varepsilon]} - \alpha(\alpha - 2)^{-1}(\varepsilon^{1-\alpha/2} - 1), & \text{if } k = 1, \\ \sum_{s=1}^{\infty} j_s^2 c^k W_{s,k}, & \text{if } k > 1. \end{cases}$$

[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 $\eta > 0$,

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

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

$$\begin{aligned} & b_n^{-2} \left(S_n - \sum_{t=1}^n X_t^{(m)} \right) \\ (4.2) \quad &= b_n^{-2} \sum_{t=1}^n \sum_{k=m+1}^{\infty} c^k Y_t^{(k)} \\ &= b_n^{-2} \sum_{t=1}^n \sum_{k=m+1}^{\infty} c^k Y_t^{(k)} \mathbf{1}_{[Z_{t-k}^2 \leq b_n^2 \delta]} + b_n^{-2} \sum_{t=1}^n \sum_{k=m+1}^{\infty} c^k Y_t^{(k)} \mathbf{1}_{[Z_{t-k}^2 > b_n^2 \delta]}. \end{aligned}$$

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

$$\begin{aligned} (4.3) \quad & n b_n^{-2} \sum_{k=m+1}^{\infty} |c|^k (E|Z_1|)^{k-1} E(Z_1^2) \mathbf{1}_{[Z_1^2 \leq b_n^2 \delta]} \\ &= (\text{const}) \sum_{k=m+1}^{\infty} r^k n b_n^{-2} E(Z_1^2 \mathbf{1}_{[Z_1^2 \leq b_n^2 \delta]}), \end{aligned}$$

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

$$\begin{aligned} (4.4) \quad & (\text{const}) r^m n b_n^{-2} (b_n^2 \delta) P[Z_1^2 > b_n^2 \delta] \rightarrow (\text{const}) r^m \delta^{1-\alpha/2} \quad (\text{as } n \rightarrow \infty) \\ & \rightarrow 0 \quad (\text{as } m \rightarrow \infty). \end{aligned}$$

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

$$\begin{aligned} & b_n^{-4} \sum_{t=1}^n \sum_{k=m+1}^{\infty} c^{2k} \text{Var}(Y_t^{(k)} \mathbf{1}_{[Z_t^2 \leq b_n^2 \delta]}) \\ & \leq n b_n^{-4} \sum_{k=m+1}^{\infty} c^{2k} (EZ_1^2)^{k-1} E(Z_1^4 \mathbf{1}_{[Z_1^2 \leq b_n^2 \delta]}). \end{aligned}$$

Using Karamata's theorem again, the right-hand side is

$$\begin{aligned} & \leq (\text{const}) \gamma^m \delta^2 n P[Z_1^2 > b_n^2 \delta] \\ & \rightarrow (\text{const}) \gamma^m \delta^{2-\alpha/2} \quad (\text{as } n \rightarrow \infty) \\ & \rightarrow 0 \quad (\text{as } m \rightarrow \infty), \end{aligned}$$

where $\gamma = E(cZ_1)^2 < 1$. 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 $\{\hat{\rho}_H(l), l = 1, \dots, h\}$ for integers

$h = 1, 2, \dots$. Recall that $\hat{\rho}_H(l)$ was defined in Section 1 to be

$$\hat{\rho}_H(l) = \frac{\sum_{t=1}^{n-l} X_t X_{t+l}}{\sum_{t=1}^n X_t^2}.$$

In Davis and Resnick (1985a, 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 nonlinear process that sample correlations converge in distribution to nondegenerate 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 $h = 1, 2, \dots$ that*

$$(\hat{\rho}_H(l), l = 1, \dots, h) \Rightarrow (L_i, i = 1, \dots, h)$$

in \mathbb{R}^h , where in the notation of Theorem 3.4

$$L_i = \frac{\sum_{s=1}^{\infty} \sum_{k=1}^{\infty} j_s^4 c^{2k-i} W_{s,k} W_{s,k-i}}{\sum_{s=1}^{\infty} \sum_{k=1}^{\infty} j_s^4 c^{2k} W_{s,k}^2}, \quad i = 1, \dots, h.$$

PROOF. Theorem 3.4(ii) implies

$$(4.5) \quad \left(\sum_{t=1}^n \mathcal{E}_{b_n^{-2}}(X_t, X_{t-l}), l = 1, \dots, h \right) \Rightarrow \left(\sum_{s=1}^{\infty} \sum_{k=1}^{\infty} \mathcal{E}_{j_s^2}(c^k W_{s,k}, c^{k-l} W_{s,k-l}), l = 1, \dots, h \right)$$

in $M_p^h(\mathbb{E}_2)$. 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):

$$(4.6) \quad \sum_{t=1}^n \mathcal{E}_{b_n^{-2}}(X_t, X_{t-1}) \Rightarrow \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} \mathcal{E}_{j_s^2}(c^k W_{s,k}, c^{k-1} W_{s,k-1}).$$

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

$$\sum_{t=1}^n \mathcal{E}_{b_n^{-2}}(X_t, X_{t-1})(\cdot \cap \mathbb{E}_{\delta}) \Rightarrow \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} \mathcal{E}_{j_s^2}(c^k W_{s,k}, c^{k-1} W_{s,k-1})(\cdot \cap \mathbb{E}_{\delta}).$$

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

$$\begin{aligned} & \sum_{t=2}^n \mathbf{1}_{[|X_t| \vee |X_{t-1}| > b_n^2 \delta]} \mathcal{E} b_n^{-4}(X_t X_{t-1}) \\ & \Rightarrow \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} \mathbf{1}_{[|j_s^2 c^k W_{s,k}| \vee j_s^2 c^{k-1} W_{s,k-1}| > \delta]} \mathcal{E} j_s^4 c^{2k-1} W_{s,k} W_{s,k-1}. \end{aligned}$$

Each point process on the previous line has only a finite number of points and applying the summation functional we get

$$\begin{aligned} \gamma_{n,\delta}(1) & := \sum_{t=2}^n \mathbf{1}_{[|X_t| \vee |X_{t-1}| > b_n^2 \delta]} b_n^{-4}(X_t X_{t-1}) \\ & \Rightarrow \gamma_{\infty,\delta}(1) := \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} \mathbf{1}_{[|j_s^2 c^k W_{s,k}| \vee j_s^2 c^{k-1} W_{s,k-1}| > \delta]} j_s^4 c^{2k-1} W_{s,k} W_{s,k-1}. \end{aligned}$$

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

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

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

$$(4.9) \quad \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} j_s^4 c^{2k-1} W_{s,k} W_{s,k-1} = \sum_{s=1}^{\infty} j_s^4 B_s,$$

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

$$(4.10) \quad \begin{aligned} \mathbf{E}|B_s|^{\alpha/4} & \leq \sum_{k=1}^{\infty} |c|^{(2k-1)\alpha/4} \mathbf{E}|W_{s,k-1}|^{\alpha/2} \mathbf{E}|Z_1|^{\alpha/4} \\ & \leq (\text{const}) \sum_{k=1}^{\infty} (\mathbf{E}|cZ_1|^{\alpha/2})^{k-1} < \infty. \end{aligned}$$

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

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

$$P\left[b_n^{-4} \sum_{t=1}^n |X_t X_{t-1}| \mathbf{1}_{[|X_t| \vee |X_{t-1}| \leq b_n^2 \delta]} > \eta\right] \leq \frac{n}{b_n^4} \mathbf{E}|X_2 X_1| \mathbf{1}_{[|X_2| \vee |X_1| \leq b_n^2 \delta]} / \eta,$$

which by Cauchy's inequality is dominated by

$$\frac{n}{b_n^4} E|X_1|^2 \mathbf{1}_{\{|X_1| \leq b_n^2 \delta\}} / \eta$$

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

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \frac{n}{b_n^4} E|X_1|^2 \mathbf{1}_{\{|X_1| \leq b_n^2 \delta\}} / \eta = 0,$$

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 $\gamma_{n,0}(0) = \sum_{t=2}^n X_t X_{t-1} / b_n^2$. 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. \square

REFERENCES

- BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.
- BINGHAM, N., GOLDIE, C. and TEUGELS, J. (1987). Regular variation. *Encyclopedia of Mathematics and Its Applications*. Cambridge Univ. Press.
- BREIMAN, L. (1965). On some limit theorems similar to the arc-sin law. *Theory Probab. Appl.* **10** 323–331.
- BROCKWELL, P. and DAVIS, R. (1991). *Time Series: Theory and Methods*, 2nd ed. Springer, New York.
- CLINE, D. (1983). Estimation and linear prediction for regression, autoregression and ARMA with infinite variance data. Thesis, Dept. Statistics, Colorado State Univ.
- DAVIS, R. A. and HSING, T. (1995). Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Probab.* **23** 879–917.
- DAVIS, R. A. and RESNICK, S. (1985a). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.* **13** 179–195.
- DAVIS, R. A. and RESNICK, S. (1985b). More limit theory for the sample correlation function of moving averages. *Stochastic Process. Appl.* **20** 257–279.
- DAVIS, R. A. and RESNICK, S. (1986). Limit theory for the sample covariance and correlation functions of moving averages. *Ann. Statist.* **14** 533–558.
- DUFFY, D., MCINTOSH, A., ROSENSTEIN, M. and WILLINGER, W. (1993). Analyzing telecommunications traffic data from working common channel signaling subnetworks. In *Proceedings of the 25th Interface Conference, San Diego, CA* 156–165. Interface Foundation of North America.
- DUFFY, D., MCINTOSH, A., ROSENSTEIN, M. and WILLINGER, W. (1994). Statistical analysis of CCSN/SS7 traffic data from working CCS subnetworks. *IEEE Journal on Selected Areas in Communications* **12** 544–551.
- FEIGIN, P. and RESNICK, S. (1994). Limit distributions for linear programming time series estimators. *Stochastic Process. Appl.* **51** 135–166.
- FEIGIN, P. and RESNICK, S. (1996). Pitfalls of fitting autoregressive models for heavy-tailed time series. Unpublished manuscript. Available at <http://www.orie.cornell.edu/trlist/trlist.html> as TR1163.ps.Z.
- FEIGIN, P., KRATZ, M. and RESNICK, S. (1996). Parameter estimation for moving averages with positive innovations. *Ann. Appl. Probab.* **6** 1157–1190.

- FEIGIN, P., RESNICK, S. and STĂRICĂ, C. (1995). Testing for independence in heavy tailed and positive innovation time series. *Stochastic Models* **11** 587–612.
- GELUK, J. and DE HAAN, L. (1987). *Regular Variation, Extensions and Tauberian Theorems*. *CWI Tract* **40**. Center for Mathematics and Computer Science, Amsterdam.
- KALLENBERG, O. (1983). *Random Measures*, 3rd ed. Akademie-Verlag, Berlin.
- LIU, J. (1989). On the existence of a general multiple bilinear time series. *J. Time Series Anal.* **10** 341–355.
- MEIER-HELLSTERN, K., WIRTH, P., YAN, Y. and HOEFLIN, D. (1991). Traffic models for ISDN data users: office automation application. In *Teletraffic and Datatrafic in a Period of Change. Proceedings of the 13th ITC* (A. Jensen and V. B. Iversen, eds.) 167–192. North-Holland, Amsterdam.
- RESNICK, S. (1986). Point processes, regular variation and weak convergence. *Adv. in Appl. Probab.* **18** 66–138.
- RESNICK, S. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer, New York.
- RESNICK, S. (1995). Heavy tail modelling and teletraffic data. Unpublished manuscript. (Available at <http://www.orie.cornell.edu/trlist/trlist.html> as TR1134.ps.Z.)
- WILLINGER, W., TAQQU, M., SHERMAN, R. and WILSON, D. (1995). Self-similarity through high-variability: statistical analysis of ethernet LAN traffic at the source level. Preprint.

DEPARTMENT OF STATISTICS
COLORADO STATE UNIVERSITY
FT. COLLINS, COLORADO 80523
E-MAIL: rdavis@stat.colostate.edu

CORNELL UNIVERSITY
SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
ETC BUILDING
ITHACA, NEW YORK 14853
E-MAIL: sid@orie.cornell.edu