The Sample ACF of a Simple Bilinear Process

BOJAN BASRAK,
University of Groningen

RICHARD. A. DAVIS,
Colorado State University

and

THOMAS MIKOSCH,
University of Groningen

March 9, 1998

ABSTRACT

We consider a simple bilinear process $X_t = aX_{t-1} + bX_{t-1}Z_{t-1} + Z_t$, where $(Z_t)$ is a sequence of iid $N(0,1)$ random variables. It follows from a result by Kesten (1973) that $X_t$ has a distribution with regularly varying tails of index $\alpha > 0$ provided the equation $E|a+bZ_1|^{u} = 1$ has the solution $u = \alpha$. We study the limit behaviour of the sample autocorrelations and autocovariances of this heavy-tailed non-linear process. Of particular interest is the case when $\alpha < 4$. If $\alpha \in (0,2)$ we prove that the sample autocorrelations converge to non-degenerate limits. If $\alpha \in (2,4)$ we prove joint weak convergence of the sample autocorrelations and autocovariances to an infinite variance $\alpha/2$-stable limit vector.

AMS 1991 Subject Classification: Primary: 62M10 Secondary: 60F05 60G10 60G55

Key Words and Phrases. Sample autocorrelation, sample autocovariance, heavy tails, infinite variance, stable distribution, convergence of point processes, mixing condition, stochastic recurrence equation, bilinear process.
1 Introduction

Our intention is to use the machinery developed in Davis and Hsing [3] and further in Davis and Mikosch [8] in order to analyze a simple bilinear process and the limiting behaviour of its sample autocorrelation function (abbreviated as ACF). A stationary sequence \((X_t)_{t \in \mathbb{Z}}\) of random variables is called a simple bilinear process if it satisfies the following recursive relation,

\[
X_t = aX_{t-1} + bX_{t-1}Z_{t-1} + Z_t,
\]

where \((Z_t)\) is an iid noise sequence and \(a, b\) are real constants. For the purpose of this presentation let us assume \(Z_t \sim N(0,1)\), although our arguments can be applied to a wider class of noise distributions.

It has been generally acknowledged in the econometrics and applied financial literature that many financial time series such as log-returns of share prices, stock indices, and exchange rates, exhibit stochastic volatility and heavy-tailedness. These features cannot be adequately modeled via a linear time series model. Nonlinear models, such as the bilinear process (1.1) and the ARCH models, have been proposed to capture these and other characteristics. In order for a linear time series model to possess heavy-tailed marginal distributions, it is necessary for the input noise sequence to be heavy-tailed. Interestingly, in this situation, the sample ACF has a number of desirable properties, even if the underlying sequence has infinite variance (see Davis and Resnick [5], [7]). For nonlinear models, heavy-tailed marginals can be obtained when the system is injected with light-tailed marginals such as with normal noise. Unlike the linear process case, however, the sample ACF may no longer be of any value for estimating the parameters of the model.

Model (1.1) was studied in Davis and Resnick [7] under the assumptions that \(a = 0\) and the \(Z_t s\) are random variables from a distribution with regularly varying tail with index \(\alpha\). Not surprisingly, the marginal distribution also has heavy-tails and is in fact regularly varying with index \(\alpha/2\). In the case \(\alpha \in (0, 4)\), which corresponds to an infinite variance process, they showed that the sample autocorrelations \(\rho_{n,X}(h)\), without any normalization, converge jointly in distribution to some non-degenerate random vector. We show that the same phenomenon holds for the case of light-tailed inputs. This similarity in the asymptotic behaviour of the sample ACF for the two situations is quite striking. In the case when the marginal distribution has a finite variance but infinite fourth moment, we show that the sample ACF has an asymptotic non-normal stable distribution.

The fact that the sample ACF, without any centering or rescaling, may have random limits suggests that it should be used with caution for modelling heavy-tailed nonlinear time series. On
the other hand, the sample ACF can be a useful tool for detecting nonlinearities in the process. For example if the data set is split into two contiguous pieces, then the sample ACF computed for both segments should look nearly the same if the data can be modelled as a linear process. If the plots of the two ACF’s are noticeably different, then this suggests that a nonlinear model might be appropriate. See Davis and Resnick [7] for further remarks on this point.

In Section 2, we review point process results required for establishing the limit theory for the sample ACF of the simple bilinear model. In Section 3, we apply these results to the model (1.1). In particular, we show that the finite–dimensional distributions are regularly varying and prove convergence for the sequence of point processes constructed from the bilinear process. In Section 4, we give the limit theory for the sample ACF.

2 Background results

Our results are based on the theory given in Davis and Hsing [3] and its application to the analysis of the sample ACF in Davis and Mikosch [8]. In our arguments we use some of the ideas of the latter paper in which the sample ACF of an ARCH(1) process was treated. Two basic conditions were imposed on the time series \((X_t)\): regular variation of the finite–dimensional distributions of the sequence \((X_t)\) and a mild mixing condition \(\mathcal{A}(a_n)\). Below we give both of them.

The distribution of the random vector \(X = (X_1, \ldots, X_m)\) is **jointly regularly varying** with index \(\alpha > 0\) if there exists a sequence of constants \(x_n\) and a random vector \(\theta \in \mathbb{S}^{m-1}\), where \(\mathbb{S}^{m-1}\) denotes the unit sphere in \(\mathbb{R}^m\) with respect to the norm \(\| \cdot \|\), such that

\[
\mathbb{P}(\|X\| > tx_n, X/\|X\| \in \cdot) \overset{\nu}{\to} t^{-\alpha} \mathbb{P}(\theta \in \cdot), \quad t > 0.
\]

The symbol \(\overset{\nu}{\to}\) denotes vague convergence on the Borel \(\sigma\)-field of \(\mathbb{S}^{m-1}\). It can be shown that (2.1) is equivalent to

\[
\frac{\mathbb{P}(\|X\| > tx_n, X/\|X\| \in \cdot)}{\mathbb{P}(\|X\| > x)} \overset{\nu}{\to} t^{-\alpha} \mathbb{P}(\theta \in \cdot), \quad t > 0.
\]

In the above definition one can take an arbitrary norm \(\| \cdot \|\) on \(\mathbb{R}^m\). However, for our purposes it is natural to choose the max–norm given by

\[
\|x\| = |(x_1, \ldots, x_m)| = \max_{i=1,\ldots,m} |x_i|.
\]

Let \((a_n)\) be a sequence of positive numbers such that

\[
\mathbb{P}(\|X\| > a_n) \to 1, \quad n \to \infty.
\]
We introduce the **mixing condition** \( \mathcal{A}(a_n) \): for a stationary sequence of random vectors \( X_t \) with values in \( \mathbb{R}^m \) we say that the condition \( \mathcal{A}(a_n) \) holds if there exists a sequence of positive integers \( (r_n) \) such that \( r_n \to \infty \), \( k_n = [n/r_n] \to \infty \) as \( n \to \infty \) and

\[
E \exp \left\{ - \sum_{t=1}^{n} f(X_t/a_n) \right\} - \left( E \exp \left\{ - \sum_{t=1}^{r_n} f(X_t/a_n) \right\} \right)^{k_n} \to 0,
\]

for every bounded, non-negative step function \( f \) on \( \mathbb{R}^m \setminus \{0\} \) with bounded support.

Condition \( \mathcal{A}(a_n) \) is indeed very weak and is implied by various known mixing condition, in particular, by the strong mixing condition (cf. Leadbetter and Rootzén [14]). We will use this fact later.

Point process techniques have played a major role in the analysis of stationary processes \( (X_t) \) satisfying (2.1) and (2.3). For background on point processes we refer to Kallenberg [12] and Resnick [16]. Let

\[
N_n = \sum_{t=1}^{n} \varepsilon_{X_t/a_n}, \quad n = 1, 2, \ldots,
\]

be the point process constructed from the sequence \( (X_t) \), where \( (a_n) \) is given by (2.2) and \( \varepsilon_x \) represents unit point measure at the point \( x \). We write \( o \) for the null measure on \( \mathbb{R}^m \setminus \{0\} \).

The following result, which corresponds to Theorem 2.8 of Davis and Mikosch [8] for the case \( m > 1 \) and to Theorem 2.7 of Davis and Hsing [3] for the case \( m = 1 \), characterizes the limiting behaviour of the point process \( N_n \) for mixing sequences that have regularly varying finite-dimensional distributions. First, the clusters are anchored by a Poisson point process, denoted by \( \sum_{i=1}^{\infty} \varepsilon_{P_i} \), on \( \mathbb{R}_+ \) with intensity measure \( \nu(dy) = \gamma \alpha y^{-\alpha - 1} I_{\mathbb{R}_+}(y)dy \). For each point \( P_i \) of the Poisson process, there is a point process of clusters, \( \sum_{j=1}^{\infty} \varepsilon_{Q_{ij}} \) defined on \( \mathbb{R}^m \setminus \{0\} \) such that \( \max_j |Q_{ij}| = 1 \) a.s. If \( Q \) denotes the distribution of the point process of clusters, then it is assumed that the sequence of point processes, \( \sum_{j=1}^{\infty} \varepsilon_{Q_{ij}} \), \( i \geq 1 \), is iid with distribution \( Q \). The limit cluster point process then takes the form \( N = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_iQ_{ij}} \). The measures \( \nu \) and \( Q \) are given in Theorem 2.2 and Corollary 2.5 of Davis and Mikosch [8].

**Theorem 2.1** Let \( (X_t) \) be a stationary sequence of random vectors. Assume that the \( (2k+1)m \)-dimensional vector \( (X_{-k}, \ldots, X_k) \) is jointly regularly varying with index \( \alpha > 0 \) for each \( k \). Let \( (\theta_{-k}, \ldots, \theta_k) \) be the random vector with values in the unit sphere \( S^{(2k+1)m-1} \) that appears in the definition of regular variation. Assume that condition \( \mathcal{A}(a_n) \) holds for \( (X_t) \) and that

\[
\lim_{k \to \infty} \limsup_{n \to \infty} P \left( \bigvee_{k \leq |t| \leq r_n} |X_t| > a_n y \bigg| |X_0| > a_n y \right) = 0 \quad \text{for every} \ y > 0.
\]
Then the limit
\begin{equation}
\gamma = \lim_{k \to \infty} \frac{E \left( |\theta_0^{(k)}|^\alpha - \sqrt{k} |\theta_j^{(k)}|^\alpha \right)_+}{E|\theta_0^{(k)}|^\alpha}
\end{equation}
exists.

In the case $\gamma = 0$, $N_n \xrightarrow{d} 0$.

If $\gamma > 0$, then $N_n \xrightarrow{d} N \neq 0$ where the limit point process has the representation,
\[ N = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i Q_{ij}}, \]
and $Q$ is the weak limit of
\[ E \left( |\theta_0^{(k)}|^\alpha - \sqrt{k} |\theta_j^{(k)}|^\alpha \right)_+ / \left( \sum_{i \leq k} \varepsilon_{\theta_i^{(k)}} \right) \]
as $k \to \infty$, which exists.

For a stationary sequence $(X_t)$ of random variables we define the sample autocovariance function (ACVF)
\[ \gamma_{n,X}(h) = n^{-1} \sum_{i=1}^{n-h} X_t X_{t+h}, \quad h \geq 0, \]
and the corresponding sample ACF
\[ \rho_{n,X}(h) = \gamma_{n,X}(h) / \gamma_{n,X}(0), \quad h \geq 1. \]

If $EX_0^2 < \infty$, the ACVF $\gamma_X(h) = EX_0 X_h$ and ACF $\rho_X(h) = \gamma_X(h) / \gamma_X(0)$ of the sequence $(X_t)$ at lag $h$ are well defined. The following result describes the asymptotic behaviour of the sample ACVF and the sample ACF under suitable conditions. It is Theorem 3.5 in Davis and Mikosch [8].

**Theorem 2.2** Let $(X_t)$ be a strictly stationary sequence of random variables. Assume for some fixed $m$ that the sequence of the random vectors $X_t(m) = (X_t, \ldots, X_{t+m})$, $t \in \mathbb{Z}$, satisfies the conditions of Theorem 2.1, so that
\[ N_n = \sum_{t=1}^{n} \varepsilon_{X_t/a_n} \xrightarrow{d} N = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i Q_{ij}}. \]

(i) If $\alpha \in (0, 2)$, then
\[ (n a_n^{-2} \gamma_{n,X}(h))_{h=0, \ldots, m} \xrightarrow{d} (V_h)_{h=0, \ldots, m}, \]
\[ (\rho_{n,X}(h))_{h=1, \ldots, m} \xrightarrow{d} (V_h/V_0)_{h=1, \ldots, m}, \]
where

\[ V_h = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i^2 Q_{i,j}^{(0)} Q_{i,j}^{(h)}, \quad h = 0, \ldots, m. \]

The vector \((V_0, \ldots, V_m)\) is jointly \(\alpha/2\)-stable in \(\mathbb{R}^{m+1}\).

(ii) If \(\alpha \in (2, 4)\) and for \(h = 0, \ldots, m,\)

\[
(2.6) \quad \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{\text{var} \left( a_n^{-2} \sum_{t=1}^{n-h} X_t X_{t+h} I_{\{|X_t X_{t+h}| \leq a_n^2 \epsilon\}} \right)}{} = 0,
\]

then

\[
(2.7) \quad (n a_n^{-2} (\gamma_n X(h) - \gamma X(h)))_{h=0, \ldots, m} \xrightarrow{d} (V_h)_{h=0, \ldots, m},
\]

where \((V_0, \ldots, V_m)\) is the distributional limit of

\[
\left( \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i^2 Q_{i,j}^{(0)} Q_{i,j}^{(h)} I_{\{\epsilon, \infty\}}(P_i^2 Q_{i,j}^{(0)} Q_{i,j}^{(h)}) \right) - \int_{B_{\epsilon,h}} x^{(0)} x^{(h)} \tau(dx),
\]

\[
B_{\epsilon,h} = \{ x \in \mathbb{R}^{m+1} : \epsilon < |x_0 x_h| \},
\]

as \(\epsilon \to 0\), and \(\tau\) is the measure on \(\mathbb{R}^m \setminus \{0\}\) such that \(nP(X_\epsilon(m)/a_n \in \cdot) \xrightarrow{\nu} \tau(\cdot)\) Moreover,

\[
(2.8) \quad (n a_n^{-2} (\rho_n X(h) - \rho X(h)))_{h=1, \ldots, m} \xrightarrow{d} \gamma^{-1}_X(0) (V_h - \rho X(h) V_0)_{h=1, \ldots, m}.
\]

### 3 The bilinear process and difference equations

Before we apply the results of the preceding section to the bilinear model we first consider some of its properties. The bilinear process \((1.1)\) can be written as

\[
(3.1) \quad X_t = Y_{t-1} + Z_t, \quad t \in \mathbb{Z},
\]

where \(Y_t = (a + bZ_t) X_t\). The process \((Y_t)\) satisfies the random difference equation

\[
(3.2) \quad Y_t = A_t Y_{t-1} + B_t, \quad t \in \mathbb{Z},
\]

where the \((A_t, B_t)\)s are iid pairs of random variables, \(A_t = a + bZ_t\) and \(B_t = A_t Z_t\). It is not difficult to see that the stationary solution to \((1.1)\) exists if we can find the stationary solution to \((3.2)\).

Equations of type \((3.2)\) have been extensively studied for years, see for instance Kesten [13], Vervaat [19], Goldie [10]. The facts needed are best summarized in the following theorem of Kesten [13].
Theorem 3.1 Let $A, B$ be random variables on a common probability space. Assume the following conditions hold:

i) there exists a number $\alpha > 0$ such that $E|A|^\alpha = 1$, $E|A|^\alpha \ln^+ |A| < \infty$ and $0 < E|B|^\alpha < \infty$.

ii) The conditional law of $\ln |A|$ given $\{A \neq 0\}$ is non-arithmetic, i.e. it is not concentrated on $\{n\lambda : n \in \mathbb{Z}\}$ for any $\lambda$.

Then there exists a random variable $Y$, independent of $A$ and $B$, such that $Y \overset{d}{=} AY + B$. Moreover, there exist non-negative constants $C_+, C_-$ such that

$$P(Y > t) \sim C_+ t^{-\alpha}, \quad P(Y < -t) \sim C_- t^{-\alpha} \quad \text{as } t \to \infty,$$

where $C_+ + C_- > 0$ if and only if for each $c \in \mathbb{R}$

$$P(B = (1 - A)c) < 1.$$

Assume in addition that $(Y_t)$ is a sequence of random variables satisfying the recursive relation (3.2), where the iid pairs $(A_t, B_t)$ have the same distribution as $(A, B)$. Then $Y_t \overset{d}{=} Y$, independently of the starting value $Y_0$. In particular, if $Y_0 \overset{d}{=} Y$ then $(Y_t)$ is a stationary sequence.

A consequence of this theorem is that the $Y_t$s in (3.2) have regularly varying tails with index $\alpha$ provided $\alpha > 0$ is the solution of the equation

$$(3.3) \quad E|a + bZ_1|^\alpha = 1.$$ 

Since the tail of a normal random variable decays faster than exponentially we conclude that the tails of the random variables $X_t$ in (3.1) are regularly varying with the same index $\alpha > 0$. From now on we always assume that (3.3) holds for some positive $\alpha$.

Theorem 3.2 Assume that $(Y_t)$ is a stationary solution of (3.2) and that $E|A_1|^\alpha = 1$ holds for some $\alpha > 0$. For $h \geq 0$ set $Y_t = (Y_t, \ldots, Y_{t+h})$ and let the sequence of normalizing constants $(a_n)$ be chosen such that $nP(|Y| > a_n) \to 1$. Then the conditions of Theorem 2.1 are satisfied and hence

$$N_{n,Y} = \sum_{t=1}^{n} \varepsilon_{Y_t/a_n} \overset{d}{=} N.$$ 

In addition, if $(X_t)$ is the bilinear process given in (1.1), then

$$N_{n,X} = \sum_{t=1}^{n} \varepsilon_{X_t/a_n} \overset{d}{=} N,$$

where $X_t = (X_t, \ldots, X_{t+h})$. 

6
Proof. We divide the proof into four parts. In part (i) regular variation of the finite-dimensional distributions of \((Y_t)\) is proved, in part (ii) strong mixing of the process \((Y_t)\) is established, in part (iii) condition (2.4) is verified for \((Y_t)\), and in part (iv) the relation \(N_{n,X} - N_{n,Y} = o_p(1)\) is shown.

(i) If \(Y_t^{(h)}(m) = (Y_t, \ldots, Y_{t+m})\), then the random vector \(Y_t^{(h)}(m)\) is regularly varying with index \(\alpha\), i.e., there exists a sequence \((y_n)\) and a random vector \(\theta\) such that

\[
n P \left( \frac{|Y_t^{(h)}(m)|}{|Y_t^{(h)}(m)|} > t y_n, \frac{Y_t^{(h)}(m)}{|Y_t^{(h)}(m)|} \in \cdot \right) \sim t^{-\alpha} P(\theta \in \cdot).
\]

For simplicity of the presentation we only provide the argument for the case \(h = 1\). Writing

\[
Y_t(m) = Y_t^{(1)}(m) = ((Y_t, Y_{t+1}), \ldots, (Y_{t+m}, Y_{t+m+1})),
\]

we see that \(Y_t(m)\) satisfies the relation

\[
Y_t(m) = Y_t((1, A_{t+1}), (A_{t+1}, A_{t+1}A_{t+2}), \ldots, (A_{t+1} \cdots A_{t+m}, A_{t+1} \cdots A_{t+m+1})) + R_t
\]

\[
= Y_t A_t + R_t,
\]

where

\[
A_t = ((1, A_1), (A_1, A_1A_2), \ldots, (A_1 \cdots A_m, A_1 \cdots A_{m+h}))
\]

and the remainder term \(R_t\) does not contribute to asymptotic behaviour of the tail of \(Y_t(m)\). To show that \(Y_t A_t\) is regularly varying we use a result of Breiman [1]; see also Davis and Mikosch [8]: assume \(\xi\) is a non-negative random variable with a regularly varying tail of index \(\bar{\alpha} > 0\) and \(\eta\) is another non-negative random variable independent of \(\xi\) and with \(E\eta^{\alpha} < \infty\) for some \(\kappa > \bar{\alpha}\), then

\[
P(\eta \xi > x) \sim E\eta^{\bar{\alpha}} P(\xi > x), \quad x \to \infty.
\]

Applying this result with \(\xi = |Y_t|\), \(\eta = |A_t|I_{(A_t, |A_t| \in B})\), where \(B\) is any Borel subset of \(S^{m-1}\), and choosing \(y_n\) as the \(1 - n^{-1}\) quantile of the distribution of \(|Y_1(m)|\), we obtain

\[
n P(|Y_1(m)| > y_n t, Y_1(m) / |Y_1(m)| \in B) = n P(\xi \eta > y_n t) \sim n E \eta^{\alpha} P(\xi > y_n t) \to t^{-\alpha} E \eta^{\alpha} / |A_1|^\alpha.
\]

We have used the property \(n P(|Y_1| > y_n t) \to t^{-\alpha} / E |A_1|^\alpha\), which follows from the relation

\[
n P(|Y_1(m)| > y_n t) \sim n P(|Y_1| > y_n t) E |A_1|^\alpha \to t^{-\alpha}.
\]

(ii) We show that \(Y_t^{(h)}(m)\) satisfies \(A(a_n)\) by proving that the sequence \((Y_t)\) is strongly mixing. We prove that \((Y_t)\) is a \(V\)-uniformly ergodic sequence. For the definition of \(V\)-uniform ergodicity
and other details we refer to Meyn and Tweedie [15], see also the proofs of Lemmas 5.2 and 5.3 in Davis and Mikosch [8].

Define $h(r) = E|A_1|^r$. We calculate the first and second derivatives of the function $h$:

$$h'(r) = E(|A_1|^r \ln |A_1|), \quad h''(r) = E(|A_1|^r \ln^2 |A_1|).$$

Therefore $h'' > 0$ so that the function $h$ is twice differentiable and strictly convex on $(0, \infty)$. Since $h(0) = h(\alpha) = 1$ there must be some $s \in (0, \min\{1, \alpha\})$ and a $c \in (0, 1)$ such that $c = h(s) < 1$. For such an $s$ define $V(x) = |x|^s + 1$. Obviously $E(V(Y_1)|Y_0 = x_0) = E|A_1 x_0 + B_1|^s + 1$ is a bounded function of $x_0$ on every compact set, for instance on $[-M, M]$. For $x \in [-M, M]^c$ we have

$$E(V(Y_1)|Y_0 = x_0) = E|A_1 x_0 + B_1|^s + 1 \leq |x_0|^s E|A_1|^s + E|B_1|^s + 1 \leq |x_0|^s c + K,$$

where $K$ is a fixed constant. If we choose $M$ sufficiently large and $\delta > 0$ sufficiently small, for instance such that $K \leq (1 - \delta - c) M^\delta$, sufficient conditions for $V$-uniform ergodicity of $(Y_t)$ are satisfied, see Theorem 16.0.1 in Meyn and Tweedie [15]. As in Section 16.1.2 of the same book, we conclude that the process $(Y_t)$ is strongly mixing at a geometric rate and, as in Lemma 5.3 of Davis and Mikosch [8], $(Y_t(m))$ is also strongly mixing. A routine argument (see for example Leadbetter and Rootzén [14]) shows that strong mixing implies $A(a_n)$.

(iii) Now we prove that condition (2.4) holds. First observe that

$$Y_t = Y_0 \prod_{j=1}^t A_j + \sum_{k=1}^t B_k \prod_{j=k+1}^t A_j =: Y_0 I_t^{(1)} + I_t^{(2)}.$$

Then

$$P(|Y_t| > a_n y \mid |Y_0| > a_n y) \leq P(|Y_0 I_t^{(1)}| > a_n y/2 \mid |Y_0| > a_n y) + P(|I_t^{(2)}| > a_n y/2 \mid |Y_0| > a_n y).$$

By Markov’s inequality and Karamata’s theorem we can check that the $\limsup_{n \to \infty}$ of the first term on the right–hand side is bounded by

$$\limsup_{n \to \infty} E|I_t^{(1)}|^{\delta} (2/y)^{\delta} a_n^{-\delta} E(|Y_0|^{\delta} I(|Y_0| > a_n y))/P(|Y_0| > a_n y) \leq C u^\delta,$$

where we may take any $\delta \in (0, \alpha)$ such that $u = E|A_1|^\delta < 1$ and $C$ is a constant. Since $I_t^{(2)}$ is independent of $Y_0$ it is not difficult to see that

$$P \left( |I_t^{(2)}| > a_n y/2 \mid |Y_0| > a_n y \right) = O(n^{-1}).$$
Therefore
\[
\lim_{k \to \infty} \limsup_{n \to \infty} P \left( \max_{k \leq |t| \leq r_n} |Y_t| > a_n y \mid |Y_0| > a_n y \right) \\
\leq \lim_{k \to \infty} \limsup_{n \to \infty} 2(m+1) \sum_{t=k}^{r_n+m} P \left( |Y_t| > a_n y \mid |Y_0| > a_n y \right) P(|Y_0| > a_n y) / P(|Y_0| > a_n y) \\
\leq \lim_{k \to \infty} (const) \sum_{t=k}^{\infty} u^t \\
= 0.
\]

Hence (2.4) holds for the sequence \((Y_t)\).

The assumptions of Theorem 2.1 are now immediate from (i)-(iii) and the fact that \(\gamma > 0\) in (2.5) (which follows from the first part of the proof).

(iv) According to Proposition 9.1.VII of Daley and Vere-Jones [2], it suffices to show that
\[
N_{n,Y}(f) - N_{n,X}(f) = \sum_{t=1}^{n} f(a_n^{-1} Y_t) - \sum_{t=1}^{n} f(a_n^{-1} X_t) = o_P(1),
\]
for every bounded, continuous, non-negative function \(f\) on \(\mathbb{R}^m \setminus \{0\}\) with compact support. We show that \(E|N_{n,Y}(f) - N_{n,X}(f)| \to 0\). Since \(f\) has compact support there exists \(\beta > 0\) such that \(f(x) = 0\) for any \(x\) with \(|x| < \beta\). Now, obviously,
\[
E|N_{n,Y}(f) - N_{n,X}(f)| \\
\leq \sum_{t=1}^{n} E|f(a_n^{-1} Y_t) - f(a_n^{-1} X_t)| + o(1) \\
= nE \left[ |f(a_n^{-1} Y_0) - f(a_n^{-1} X_1)| I_{[\gamma, \infty]}(a_n^{-1} |Y_0|, a_n^{-1} |X_1|) \right] + o(1). \tag{3.4}
\]

Since the tail of the distribution of \(Z_1\) decreases exponentially, it follows easily from (2.2) and (3.1) that \(P(a_n^{-1} |Y_0| \vee a_n^{-1} |X_1| \geq \beta) = O(1/n)\) and for any \(\delta > 0\), \(nP(a_n^{-1} |Z| > \delta) \to 0\), where \(Z = (Z_1, \ldots, Z_{1+h})\). Also, since \(f\) is bounded and uniformly continuous, there exists \(M\) such that \(f(x) < M\) for all \(x\), and for every \(\epsilon > 0\) there is a \(\delta > 0\) such that \(|x-y| < \delta\) implies \(|f(x) - f(y)| < \epsilon\). Intersecting the expectation in (3.4) with the sets \(\{a_n^{-1} |Z| \leq \delta\}\) and \(\{a_n^{-1} |Z| > \delta\}\), the lim sup of the right-hand side of (3.4) is bounded by
\[
\limsup_{n \to \infty} (nE P(a_n^{-1} |Y_0| \vee a_n^{-1} |X_1| \geq \beta) + 2MnP(a_n^{-1} |Z| > \delta)) \leq \epsilon (const),
\]
which can be made arbitrarily small for suitably chosen \(\epsilon\). \(\square\)
4 The sample ACF of the simple bilinear process

Let \((X_t), (Y_t)\) and \((Z_t)\) be three sequences as described in the previous section. Assume there exists an \(\alpha > 0\) such that \(E|a + bZ_t|^{\alpha} = 1\). Our main results on the limit behaviour for the sample ACVF and ACF of \((X_t)\) are essentially direct applications of Theorems 3.2 and 2.2. In studying the limit behaviour of these functions, we distinguish three different cases with respect to the index \(\alpha\). The cases \(\alpha \in (0, 2)\) and \(\alpha \in (2, 4)\) can be treated according to Theorems 2.2 and 3.2, while for the case \(\alpha > 4\), we use the standard central limit theory for strongly mixing sequences, see e.g., Ibragimov and Linnik [11].

![Figure 4.1 A realization of the simple bilinear process with \(a = 0.2\) and \(b = 1\), \(\alpha = 1.68\).](image1)

![Figure 4.2 Sample ACF based on the first and the second half of the time series from Figure 4.1.](image2)

I The case \(\alpha \in (0, 2)\). A direct application of Theorem 2.2 immediately yields

\[
\left( n\sigma_n^{-2} \gamma_{n h}(h) \right)_{h=0, \ldots, m} \xrightarrow{d} (V_h)_{h=0, \ldots, m},
\]

\[
(\rho_{n h}(h))_{h=1, \ldots, m} \xrightarrow{d} (V_h/V_0)_{h=1, \ldots, m},
\]

where \((V_h)_{h=0, \ldots, m}\) is the \(\alpha\)-stable random vector defined in Theorem 2.2. Hence, the sample autocorrelations of a stationary bilinear process satisfying \(E|a + bZ_t|^{\alpha} = 1\) for some \(\alpha \in (0, 2)\) have
non-degenerate limit distribution without any normalization.

II The case $\alpha \in (2, 4)$. Assumption (2.6) in part (ii) of Theorem 2.2 is not easily verified. Therefore we take a different approach, as in Davis and Mikosch [8]. First we show that

\begin{equation}
na_n^{-2}[\gamma_{n,X}(h) - \gamma_{n,Y}(h) - E(X_0X_h) + E(Y_0Y_h)] \overset{p}{\to} 0.
\end{equation}

The difference in (4.1) can be written as

\begin{equation}
na_n^{-2}[\gamma_{n,X}(h) - \gamma_{n,Y}(h) - E(Z_0Z_h) - E(Z_1Y_h)]
= a_n^{-2} \sum_{t=1}^{n} (Z_t Z_{t+h} - E(Z_0Z_h)) + a_n^{-2} \sum_{t=1}^{n} Z_{t+h} Y_{t-1} + a_n^{-2} \sum_{t=1}^{n} (Z_t Y_{t-1+h} - E(Z_1Y_h)) + o_P(1).
\end{equation}

Note the first two terms on the right-hand side are sums of uncorrelated random variables and hence have variances of order $na_n^{-4}$. Since $a_n \sim Cn^{1/\alpha}$ for some constant $C > 0$, these variances converge to 0 so that the first two terms in (4.2) are $o_P(1)$. It remains to show that the third sum in (4.2) is also $o_P(1)$. For simplicity of presentation we restrict attention to the case $h = 1$, the other cases requiring a similar treatment. Using the recursions for $(Y_t)$ and the identity, $E(Z_1Y_1) = bEY_1 + a$, we have the decomposition

\begin{align*}
a_n^{-2} \sum_{t=1}^{n} (Z_t Y_t - E(Z_1Y_1)) &= a_n^{-2} \left( a \sum_{t=1}^{n} (Z_t^2 - 1) + b \sum_{t=1}^{n} Z_t^3 \
&\quad + \sum_{t=1}^{n} (aZ_t + b(Z_t^2 - 1)) Y_{t-1} + b \sum_{t=1}^{n} (Y_{t-1} - EY_{t-1}) \right) \\
&= a_n^{-2} \left[ K_n^{(1)} + K_n^{(2)} + K_n^{(3)} + K_n^{(4)} \right].
\end{align*}

Applying the CLT for iid and strongly mixing sequences, it follows that $a_n^{-2}K_n^{(i)} = o_P(1)$ for $i = 1, 2, 4$. Since

\[ \text{var}(a_n^{-2}K_n^{(3)}) = na_n^{-4}\text{var}(aZ_1 + b(Z_1^2 - 1)) \to 0, \]

we conclude that the left-hand side of (4.2) is also $o_P(1)$ which establishes (4.1) as claimed.

In view of (4.1), the limiting behaviour of the sample ACVF of the $(X_t)$ will be inherited by that of the sample ACVF based on the auxiliary process $(Y_t)$. For simplicity we restrict consideration to the ACVF at lags 0 and 1, the general case being a routine adaptation of the present argument. Using the recursive relation (3.2), we have

\begin{equation}
na_n^{-2}(\gamma_{n,Y}(0) - EY_1^2) = a_n^{-2} \sum_{t=1}^{n} [(A_t Y_{t-1} + B_t)^2 - E(A_t Y_{t-1} + B_t)^2]
\end{equation}
(4.3) \[ a_n^{-2} \sum_{t=1}^{n} (A_t^2 Y_{t-1}^2 - E[A_t^2 Y_{t-1}^2]) + a_n^{-2} \sum_{t=1}^{n} 2(A_t B_t - E[A_t B_t]) Y_{t-1} \]
\[ + a_n^{-2} \sum_{t=1}^{n} 2E(A_t B_t)(Y_{t-1} - EY_1) + a_n^{-2} \sum_{t=1}^{n} (B_t^2 - E B_t^2). \]

The second term in (4.3) is a sum of uncorrelated random variables and hence has variance converging to 0. By the CLT, the last two sums are also of order \( o_P(1) \). Denote the remaining term in (4.3) by \( J \). For an arbitrary \( \epsilon > 0 \), write

\[
J = a_n^{-2} \sum_{t=1}^{n} (A_t^2 - E A_t^2) Y_{t-1}^2 I_{\{|Y_{t-1}| \leq a_n \epsilon \}} + a_n^{-2} \sum_{t=1}^{n} (A_t^2 - E A_t^2) Y_{t-1}^2 I_{\{|Y_{t-1}| > a_n \epsilon \}} \]
\[ + a_n^{-2} E A_t^2 \sum_{t=1}^{n} (Y_{t-1}^2 - EY_{t-1}^2) \]
\[ = J_1 + J_2 + J_3. \]

We observe that \( J_3 \), up to a negligible error, is equal to the expression (4.3) multiplied by \( E A_t^2 = a^2 + b^2 \). Since the summands of \( J_1 \) are uncorrelated we have by Karamata’s theorem,

\[
\var(J_1) = n a_n^{-4} \var \left( Y_0^2 (A_1^2 - E A_1^2) I_{\{|y_0| \leq a_n \epsilon \}} \right) \]
\[ \leq \text{const} \ n a_n^{-4} E Y_0^2 I_{\{|y_0| \leq a_n \epsilon \}} \]
\[ \sim \text{const} \ n a_n^{-4} (a_n \epsilon)^4 P(|Y_0| > a_n \epsilon) \]
\[ \to \text{const} \ \epsilon^{4-a} \ \text{as} \ n \to \infty \]
(4.4)
\[ \to 0 \ \text{as} \ \epsilon \to 0. \]

We introduce a sequence of mappings from the measurable space of point processes to \( \mathbb{R} \)

\[
T_{0,\epsilon} \left( \sum_{i=1}^{\infty} n_i \xi_{x_i} \right) = \sum_{i=1}^{\infty} n_i (x_i^{(0)})^2 I_{\{|x_i^{(0)}| > \epsilon \}}, \]
\[
T_{1,\epsilon} \left( \sum_{i=1}^{\infty} n_i \xi_{x_i} \right) = \sum_{i=1}^{\infty} n_i (x_i^{(1)})^2 I_{\{|x_i^{(0)}| > \epsilon \}}, \]
\[
T_{h,\epsilon} \left( \sum_{i=1}^{\infty} n_i \xi_{x_i} \right) = \sum_{i=1}^{\infty} n_i x_i^{(0)} x_i^{(h-1)} I_{\{|x_i^{(0)}| > \epsilon \}}, \ h \geq 2, \]

where we denote \( x_t = (x_t^{(0)}, \ldots, x_t^{(m)}) \in \mathbb{R}^{m+1} \setminus \{0\} \). Using the fact that the set \( \{ x \in \mathbb{R}^2 \setminus \{0\} : |x^{(0)}| > \epsilon \} \) is bounded and the point process \( N_n = \sum_{i=1}^{n} \xi_{a_n^{-2}(Y_{t_i}, Y_{t_i+m})} \) is converging according to Theorem 2.2 we have

\[
J_2 = T_{1,\epsilon} N_n - E A_t^2 T_{0,\epsilon} N_n + o_P(1) \]

\[ (4.5) \quad \frac{\epsilon}{T_{\epsilon}} N - (a^2 + b^2) T_{\epsilon} N =: S_0(\epsilon, \infty). \]

By observing that \( ES_0(\epsilon, \infty) = 0 \) and using (4.4) and the arguments of Davis and Hsing [3] in the proof of Theorem 3.1 we conclude that \( S_0(\epsilon, \infty) \xrightarrow{d} V_0^* \) as \( \epsilon \to 0 \), where \( V_0^* \) is an \( \alpha \)-stable random variable. Now summarizing the facts above we see that

\[ na_n^{-2}(\gamma_n Y(0) - EY_1^2) \xrightarrow{d} (1 - EA_1^2) V_0^* = V_0. \]

In a similar way we obtain for \( \gamma_n Y(1) \)

\[
na_n^{-2}(\gamma_n Y(1) - EY_0 Y_1) \\
= \frac{a_n^{-2} \sum_{t=1}^{n} (A_{t+1} Y_t^2 - EA_{t+1} Y_t) + o_P(1)}{\sum_{t=1}^{n} Y_t^2 (A_{t+1} - EA_{t+1}) + a_n^{-2} EA_{t+1} \sum_{t=1}^{n} (Y_t^2 - EY_t^2) + o_P(1)}.
\]

The second term converges in distribution to \((EA_1) V_0\) as above while the first one is equal to

\[
a_n^{-2} \sum_{t=1}^{n} (A_{t+1} - EA_{t+1}) Y_t^2 I_{\{|Y_t| \leq a_n\epsilon\}} + a_n^{-2} \sum_{t=1}^{n} (A_{t+1} - EA_{t+1}) Y_t^2 I_{\{|Y_t| > a_n\epsilon\}} =: I_1 + I_2,
\]

where \( I_1 \) converges to 0 by Karamata's theorem, as already shown. As for \( I_2 \), we obviously have

\[ I_2 = T_{2, \epsilon} N - EA_1 T_{1, \epsilon} N + o_P(1) \]

\[ (4.6) \quad \xrightarrow{d} T_{2, \epsilon} N - EA_1 T_{1, \epsilon} N =: S_1(\epsilon, \infty). \]

Denoting the limit in distribution of \( S_1(\epsilon, \infty) \) as \( \epsilon \to 0 \) by \( V_1^* \), we observe

\[ na_n^{-2}(\gamma_n Y(1) - EY_0 Y_1) \xrightarrow{d} V_1^* + (EA_1) V_0 =: V_1. \]

Thus we conclude

\[ na_n^{-2}(\gamma_n Y(h) - EY_0 Y_h) \xrightarrow{d} (V_h)_{h=0, \ldots, m} \]

and since \( \sum_{t=1}^{n} Y_t^2 / n \xrightarrow{P} EY_0^2 \) by the ergodic theorem we obtain

\[ na_n^{-2} \left( \rho_n Y(h) - \frac{EY_0 Y_h}{EY_0^2} \right) \xrightarrow{d} \frac{1}{EY_0^2} (V_h)_{h=1, \ldots, m}. \]

This and (4.1) imply

\[ na_n^{-2}(\gamma_n X(h) - EY_0 X_h) \xrightarrow{d} (V_h)_{h=0, \ldots, m}. \]
and
\[ na_n^{-2} \left( \rho_{n,X}(h) - \frac{EX_0X_h}{EX_0^2} \right)_{h=1,\ldots,m} \overset{d}{\to} \frac{1}{EX_0^2} (V_h)_{h=1,\ldots,m}. \]

**III** Suppose now \( \alpha > 4 \). The sample ACVF and sample ACF of \( X_t \) have normal limit distributions in this case. One can use the standard limit theory for strongly mixing sequences (e.g., from Ibragimov and Linnik [11]) to show that the following limit holds
\[ \left( n^{1/2} (\gamma_{n,X}(h) - EX_0X_h) \right)_{h=0,\ldots,m} \overset{d}{\to} (G_h)_{h=0,\ldots,m}, \]
where \( (G_h)_{h=0,\ldots,m} \) is a random vector with mean zero from a multivariate normal distribution. Therefore we further have
\[ n^{1/2} \left( \rho_{n,X}(h) - \frac{EX_0X_h}{EX_0^2} \right)_{h=1,\ldots,m} \overset{d}{\to} \frac{1}{EX_0^2} (G_h)_{h=1,\ldots,m}. \]

Our analysis shows that the simple bilinear process exhibits similar phenomena as already noticed for other non-linear processes, like the ARCH(1) process. Namely, we witness that light-tailed input, our Gaussian noise \( (Z_t) \), can cause heavy-tailed output, the process \( X_t \), which does not necessarily have a finite second moment or even a finite expectation. It is also clear that the choice of Gaussian noise \( (Z_t) \) was almost arbitrary; one can substitute it by any iid sequence satisfying \( E|a + bZ_1|^{\alpha} = 1 \) for some positive \( \alpha \) and having sufficiently many finite moments. We also see that the autocorrelations have non-degenerate limits for \( \alpha < 2 \) which indicates that the sample ACF plot becomes unreliable in the analysis of possibly non-linear time series. We finish with two short remarks.

**Remark 4.3** The proof of Theorem 3.2 can be adjusted to cover the bilinear process from Davis and Resnick [7]. They take a noise sequence \( (Z_t) \) of iid random variables having regularly varying tail probabilities with index \( \alpha \). They assume in addition \( \alpha = 0 \) and \( |b|^{\alpha/2} E|Z_1|^{\alpha/2} < 1 \) in which case the process has the infinite series representation,
\[ X_t = Z_t + \sum_{j=1}^{\infty} b^j \left( \prod_{i=1}^{j-1} Z_{t-i} \right) Z_{t-j}^2. \]

The tail behaviour of the distribution of \( X_t \) is then obtained by first deriving the tail characteristics of the truncated infinite series. The distribution of the truncated series, and hence the infinite series,
is regularly varying with index $\alpha/2$. This argument does not work in the case of light-tailed noise, as the truncated series would also have light tails. Davis and Resnick also use a truncation argument to establish convergence of the associated sequence of point processes. The techniques of this paper, specifically Theorems 2.2 and 3.2, are directly applicable to their model.

**Remark 4.4** The constant $\gamma$ defined in Theorem 2.1 is the extremal index of the sequence $(|Y_t|)$, see Davis and Mikosch [8], and therefore the extremal index of the sequence $(|X_t|)$ so that one obtains a result similar to Turkman and Turkman [18], except that they calculate the extremal index of $(X_t)$. Namely, the extremal index $\gamma$ is given by the following expression

$$\gamma = \int_1^\infty P \left( \bigvee_{j=1}^\infty \prod_{i=1}^j |A_i| \leq y^{-1} \right) \alpha y^{-\alpha-1} dy$$

(see Remark 4.3 of Davis and Mikosch [8]).

**Acknowledgement.** This project commenced during a visit by Bojan Basrak and Thomas Mikosch to the Department of Statistics at Colorado State University in Fort Collins. They would like to express their gratitude to their colleagues for their excellent hospitality. Financial support for Bojan Basrak by the Dutch Science Foundation NWO and for Richard Davis by NSF DMS grant No. 9504596 are gratefully acknowledged.

**References**


