Stochastic Processes and their Applications 83 (1999) 1–14

stochastic processes and their applications

www.elsevier.com/locate/spa

The sample ACF of a simple bilinear process

Bojan Basrak[a], Richard A. Davis[b,*], Thomas Mikosch[a]

[a] *Department of Mathematics, P.O. Box 800, University of Groningen, NL-9700 AV Groningen, The Netherlands*

[b] *Department of Statistics, Colorado State University, Fort Collins, CO 80523, USA*

Received 13 March 1998; received in revised form 10 February 1999; accepted 12 February 1999

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, Acta Math. 131, 207–248) 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 non-normal limits. © 1999 Elsevier Science B.V. All rights reserved.

MSC: primary 62M10; secondary 60F05; 60G10; 60G55

Keywords: 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 (1995) and further in Davis and Mikosch (1998) in order to analyse 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:

$$X_t = aX_{t-1} + bX_{t-1}Z_{t-1} + Z_t, \tag{1.1}$$

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.

* Corresponding author.

E-mail address: rdavis@stat.colostate.edu (R.A. Davis)

0304-4149/99/$ - see front matter © 1999 Elsevier Science B.V. All rights reserved.
PII: S0304-4149(99)00013-7

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 modelled 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, 1985,1986). For non-linear 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 (1996) under the assumptions that $a = 0$ and the Z_t's are random variables from a distribution with regularly varying tail with index α. 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 non-linear time series. On the other hand, the sample ACF can be a useful tool for detecting non-linearities 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 ACFs are noticeably different, then this suggests that a non-linear model might be appropriate. See Davis and Resnick (1996) 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 (1995) and its application to the analysis of the sample ACF in Davis and Mikosch (1998). 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 $\mathscr{A}(a_n)$. Below we give both of them.

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

$$nP(|\boldsymbol{X}|>tx_n,\ \boldsymbol{X}/|\boldsymbol{X}|\in\cdot)\stackrel{v}{\rightarrow}t^{-\alpha}P(\boldsymbol{\theta}\in\cdot),\quad t>0. \tag{2.1}$$

The symbol $\stackrel{v}{\rightarrow}$ denotes vague convergence on the Borel σ-field of $\mathbb{S}^{m-1}$. 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. Let (a_n) be a sequence of positive numbers such that

$$nP(|\boldsymbol{X}|>a_n)\rightarrow 1,\quad n\rightarrow\infty. \tag{2.2}$$

We introduce the *mixing condition* $\mathscr{A}(a_n)$: for a stationary sequence of random vectors $\boldsymbol{X}_t$ with values in $\mathbb{R}^m$ we say that the condition $\mathscr{A}(a_n)$ holds if there exists a sequence of positive integers (r_n) such that $r_n\rightarrow\infty$, $k_n=[n/r_n]\rightarrow\infty$ as $n\rightarrow\infty$ and

$$E\exp\left\{-\sum_{t=1}^{n}f(\boldsymbol{X}_t/a_n)\right\}-\left(E\exp\left\{-\sum_{t=1}^{r_n}f(\boldsymbol{X}_t/a_n)\right\}\right)^{k_n}\rightarrow 0, \tag{2.3}$$

for every bounded, non-negative step function f on $\bar{\mathbb{R}}^m\backslash\{\boldsymbol{0}\}$ with bounded support. Condition $\mathscr{A}(a_n)$ is indeed very weak and is implied by various known mixing condition, in particular, by the strong mixing condition (cf. Leadbetter and Rootzén, 1988). We will use this fact later.

Point process techniques have played a major role in the analysis of stationary processes $(\boldsymbol{X}_t)$ satisfying (2.1) and (2.3). Let

$$N_n=\sum_{t=1}^{n}\varepsilon_{\boldsymbol{X}_t/a_n},\quad n=1,2,\ldots,$$

be the point process constructed from the sequence $(\boldsymbol{X}_t)$, where (a_n) is given by (2.2) and ε_x represents unit point measure at the point x. We write o for the null measure on $\bar{\mathbb{R}}^m\backslash\{\boldsymbol{0}\}$.

The following result, which corresponds to Theorem 2.8 of Davis and Mikosch (1998), 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 $v(v(\mathrm{d}y)=\gamma\alpha y^{-\alpha-1}\boldsymbol{I}_{\mathbb{R}_+}(y)\,\mathrm{d}y)$. For each point P_i of the Poisson process, there is a point process of clusters, $\sum_{j=1}^{\infty}\varepsilon_{\boldsymbol{Q}_{ij}}$ defined on $\bar{\mathbb{R}}^m\backslash\{\boldsymbol{0}\}$ such that $\max_j|\boldsymbol{Q}_{ij}|=1$ a.s. If Q denotes the distribution of the point process of clusters, then it is assumed that the sequence of point processes, $\sum_{j=1}^{\infty}\varepsilon_{\boldsymbol{Q}_{ij}}$, $i\geqslant 1$, is iid with distribution Q. The limit cluster point process then takes the form $N=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\varepsilon_{P_i\boldsymbol{Q}_{ij}}$. The measures v and Q are given in Theorem 2.2 and Corollary 2.5 of Davis and Mikosch (1998).

Theorem 2.1. *Let* $(\boldsymbol{X}_t)$ *be a stationary sequence of random vectors. Assume that the* $(2k+1)m$*-dimensional vector* $(\boldsymbol{X}_{-k},\ldots,\boldsymbol{X}_k)$ *is jointly regularly varying with index*

$\alpha > 0$ for each k. Let $(\boldsymbol{\theta}_{-k},\ldots,\boldsymbol{\theta}_k)$ be the random vector with values in the unit sphere $\mathbb{S}^{(2k+1)m-1}$ that appears in the definition of regular variation. Assume that condition $\mathscr{A}(a_n)$ holds for $(\boldsymbol{X}_t)$ and that

$$\lim_{k\to\infty}\limsup_{n\to\infty} P\left(\bigvee_{k\leqslant |t|\leqslant r_n} |\boldsymbol{X}_t| > a_n y \,\middle|\, |\boldsymbol{X}_0| > a_n y\right) = 0 \quad \text{for every } y>0\,. \tag{2.4}$$

Then the limit

$$\gamma = \lim_{k\to\infty} E\left(|\boldsymbol{\theta}_0^{(k)}|^\alpha - \bigvee_{j=1}^{k} |\boldsymbol{\theta}_j^{(k)}|^\alpha\right)_+ \Big/ E|\boldsymbol{\theta}_0^{(k)}|^\alpha \tag{2.5}$$

exists. If $\gamma > 0$, then $N_n \xrightarrow{\mathrm{d}} N \neq o$ where the limit point process has the representation,

$$N = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \varepsilon_{P_i \boldsymbol{Q}_{ij}},$$

and Q is the weak limit of

$$E\left(|\boldsymbol{\theta}_0^{(k)}|^\alpha - \bigvee_{j=1}^{k} |\boldsymbol{\theta}_j^{(k)}|^\alpha\right)_+ I.\left(\sum_{|t|\leqslant k} \varepsilon_{\boldsymbol{\theta}_t^{(k)}}\right) \Big/ E\left(|\boldsymbol{\theta}_0^{(k)}|^\alpha - \bigvee_{j=1}^{k} |\boldsymbol{\theta}_j^{(k)}|^\alpha\right)_+,$$

as $k\to\infty$, which exists.

For a stationary sequence (X_t) of random variables we define the *sample autocovariance function* (ACVF) by

$$\gamma_{n,X}(h) = n^{-1}\sum_{t=1}^{n-h} X_t X_{t+h}, \quad h \geqslant 0,$$

and the corresponding *sample ACF* by

$$\rho_{n,X}(h) = \gamma_{n,X}(h)/\gamma_{n,X}(0), \quad h\geqslant 1.$$

If $EX_0^2 < \infty$, the ACVF $\gamma_X(h) = EX_0X_h$ 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 (1998).

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 $\boldsymbol{X}_t(m) = (X_t,\ldots,X_{t+m})$, $t\in\mathbb{Z}$, satisfies the conditions of Theorem 2.1, so that*

$$N_n = \sum_{t=1}^{n} \varepsilon_{\boldsymbol{X}_t/a_n} \xrightarrow{\mathrm{d}} N = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \varepsilon_{P_i\boldsymbol{Q}_{ij}}.$$

(i) *If $\alpha\in(0,2)$, then*

$$(na_n^{-2}\gamma_{n,X}(h))_{h=0,\ldots,m} \xrightarrow{\mathrm{d}} (V_h)_{h=0,\ldots,m},$$

$$(\rho_{n,X}(h))_{h=1,\ldots,m} \xrightarrow{\mathrm{d}} (V_h/V_0)_{h=1},\ldots,m,$$

where

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

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* $h=0,\ldots,m$,

$$\lim_{\varepsilon\to 0}\limsup_{n\to\infty}\operatorname{var}\left(a_n^{-2}\sum_{t=1}^{n-h}X_tX_{t+h}I_{\{|X_tX_{t+h}|\leqslant a_n^2\varepsilon\}}\right)=0, \tag{2.6}$$

then

$$(n\,a_n^{-2}(\gamma_{n,X}(h)-\gamma_X(h)))_{h=0,\ldots,m}\xrightarrow{\mathrm{d}}(V_h)_{h=0,\ldots,m}, \tag{2.7}$$

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

$$\left(\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}P_i^2Q_{ij}^{(0)}Q_{ij}^{(h)}I_{(\varepsilon,\infty]}(P_i^2|Q_{ij}^{(0)}Q_{ij}^{(h)}|)-\int_{B_{\varepsilon,h}}x^{(0)}x^{(h)}\tau(\mathrm{d}\boldsymbol{x})\right)_{h=0,\ldots,m},$$

$$B_{\varepsilon,h}=\{\boldsymbol{x}\in\mathbb{R}^{m+1}:\ \varepsilon<|x_0x_h|\},$$

as $\varepsilon\to 0$, *and* τ *is the measure on* $\bar{\mathbb{R}}^m\setminus\{\mathbf{0}\}$ *such that* $nP(\boldsymbol{X}_t(m)/a_n\in\cdot)\xrightarrow{\mathrm{v}}\tau(\cdot)$.

Moreover,

$$(na_n^{-2}(\rho_{n,X}(h)-\rho_X(h)))_{h=1,\ldots,m}\xrightarrow{\mathrm{d}}\gamma_X^{-1}(0)(V_h-\rho_X(h)V_0)_{h=1,\ldots,m}. \tag{2.8}$$

Remark 2.3. The conclusions of the theorem remain valid if $\gamma_{n,X}(h)$ and $\rho_{n,X}(h)$ are replaced by their respective mean-corrected versions defined by $\tilde{\gamma}_{n,X}(h)=n^{-1}\sum_{t=1}^{n-h}(X_t-\bar{X}_n)(X_{t+h}-\bar{X}_n)$ and $\tilde{\rho}_{n,X}(h)=\tilde{\gamma}_{n,X}(h)//\tilde{\gamma}_{n,X}(0)$, where $\bar{X}_n$ is the sample mean. (If $\alpha\in(2,4)$, then $\gamma_X(h)$ must also be replaced by its mean-corrected version.) To see this equivalence, we have $\gamma_{n,X}(h)-\tilde{\gamma}_{n,X}(h)=\bar{X}_n^2+\bar{X}_n\mathrm{O_p}(1)$. A routine analysis shows that this difference is $\mathrm{o_p}(n^{-1}a_n^2)$ if $\alpha\in(0,2)$ and is $(EX_1)^2+\mathrm{o_p}(n^{-1}a_n^2)$ if $\alpha\in(2,4)$, from which the result follows by a Slutsky argument.

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

$$X_t=Y_{t-1}+Z_t,\quad t\in\mathbb{Z}, \tag{3.1}$$

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

$$Y_t=A_tY_{t-1}+B_t,\quad t\in\mathbb{Z}, \tag{3.2}$$

where the (A_t,B_t)s are iid pairs of random variables, $A_t=a+bZ_t$ and $B_t=A_tZ_t$. 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 works of Kesten (1973), Vervaat (1979) and Goldie (1991). The facts needed are best summarized in the following theorem of Kesten (1973).

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 λ.*

Then there exists a random variable Y, independent of A and B, such that $Y \stackrel{d}{=} AY + B$. 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 $Y_t \stackrel{d}{\to} Y$, independent of the starting value Y_0. In particular, if $Y_0 \stackrel{d}{=} Y$ 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 α provided $\alpha > 0$ is the solution of the equation

$$E|a + bZ_1|^\alpha = 1. \tag{3.3}$$

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 α.

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 $h \geqslant 0$ set $\mathbf{Y}_t = (Y_t, \ldots, Y_{t+h})$ and let the sequence of normalizing constants (a_n) be chosen such that $nP(|\mathbf{Y}| > a_n) \to 1$. Then the conditions of Theorem 2.1 are satisfied and hence*

$$N_{n,\mathbf{Y}} = \sum_{t=1}^{n} \varepsilon_{\mathbf{Y}_t / a_n} \stackrel{d}{\to} N.$$

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

$$N_{n,\mathbf{X}} = \sum_{t=1}^{n} \varepsilon_{\mathbf{X}_t / a_n} \stackrel{d}{\to} N,$$

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

Remark 3.3. Perfekt (1994) deals with the convergence of the point processes of exceedances of (a_n) by the points of a Markov chain (Y_t) (also the multilevel case). Turkman and Turkman (1997) study the convergence of such processes for solutions to random difference equations; see also Example 4.2 in Perfekt (1994). This kind of point process is tailored for the purpose of extreme value theory. Here we consider the more general point processes $N_{n,\mathbf{Y}}$ since their convergence implies the convergence of vectors of products of components of $\boldsymbol{Y}_t$ from which the weak convergence of the sample ACF can be established. As a matter of fact, the assumptions and parts of the proof for the point process convergence in the mentioned papers are similar to ours.

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

(i) If $\boldsymbol{Y}_t^{(h)}(m) = (\boldsymbol{Y}_t, \ldots, \boldsymbol{Y}_{t+m})$, then the random vector $\boldsymbol{Y}_t^{(h)}(m)$ is regularly varying with index α, i.e., there exists a sequence (y_n) and a random vector $\boldsymbol{\theta}$ such that

$$nP\left(|\boldsymbol{Y}_t^{(h)}(m)| > t y_n, \frac{\boldsymbol{Y}_t^{(h)}(m)}{|\boldsymbol{Y}_t^{(h)}(m)|} \in \cdot \right) \xrightarrow{\text{v}} t^{-\alpha} P(\boldsymbol{\theta} \in \cdot).$$

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

$$\boldsymbol{Y}_t(m) = \boldsymbol{Y}_t^{(1)}(m) = ((Y_t, Y_{t+1}), \ldots, (Y_{t+m}, Y_{t+m+1})),$$

we see that $\boldsymbol{Y}_t(m)$ satisfies the relation

$$\begin{aligned} \boldsymbol{Y}_t(m) &= Y_t((1, A_{t+1}), (A_{t+1}, A_{t+1}A_{t+2}), \ldots, (A_{t+1} \cdots A_{t+m}, A_{t+1} \cdots A_{t+m+1})) + \boldsymbol{R}_t \\ &= Y_t \boldsymbol{A}_t + \boldsymbol{R}_t, \end{aligned}$$

where

$$\boldsymbol{A}_t = ((1, A_1), (A_1, A_1 A_2), \ldots, (A_1 \cdots A_m, A_1 \cdots A_{m+h})),$$

and the remainder term $\boldsymbol{R}_t$ does not contribute to the asymptotic behaviour of the tail of $\boldsymbol{Y}_t(m)$. To show that the product $Y_t \boldsymbol{A}_t$ is regularly varying we use a result of Breiman (1965); see also Davis and Mikosch (1998): assume ξ is a non-negative random variable with a regularly varying tail of index $\tilde{\alpha} > 0$ and η is another non-negative random variable independent of ξ with $E\eta^{\kappa} < \infty$ for some $\kappa > \tilde{\alpha}$, then

$$P(\eta\xi > x) \sim E\eta^{\tilde{\alpha}} P(\xi > x), \quad x \to \infty.$$

Applying this result with $\xi = |Y_t|$, $\eta = |\boldsymbol{A}_t| I_{\{\boldsymbol{A}_t/|\boldsymbol{A}_t| \in B\}}$, where B is any Borel subset of $\mathbb{S}^{m-1}$, and choosing y_n as the $1 - n^{-1}$ quantile of the distribution of $|\boldsymbol{Y}_1(m)|$, we obtain

$$\begin{aligned} nP(|\boldsymbol{Y}_1(m)| > y_n t, \boldsymbol{Y}_1(m)/|\boldsymbol{Y}_1(m)| \in B) &= nP(\xi\eta > y_n t) \\ &\sim nE\eta^{\alpha} P(\xi > y_n t) \to t^{-\alpha} E\eta^{\alpha}/E|\boldsymbol{A}_1|^{\alpha}. \end{aligned}$$

We have used the property $nP(|Y_1| > y_n t) \to t^{-\alpha}/E|\boldsymbol{A}_1|^{\alpha}$, which follows from the relation

$$nP(|\boldsymbol{Y}_1(m)| > y_n t) \sim nP(|Y_1| > y_n t) E|\boldsymbol{A}_1|^{\alpha} \to t^{-\alpha}.$$

(ii) We show that $\boldsymbol{Y}_t^{(h)}(m)$ satisfies $\mathscr{A}(a_n)$ 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 (1993), see also the proofs of Lemmas 5.2 and 5.3 in Davis and Mikosch (1998).

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

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

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 $E(V(Y_1)\,|\,Y_0=x_0)=E|A_1x_0+B_1|^s+1$ is a bounded function of x_0 on every compact set, for instance on $[-M,M]$. For $x\in[-M,M]^c$ we have

$$E(V(Y_1)\,|\,Y_0=x_0)=E|A_1x_0+B_1|^s+1\leqslant|x_0|^sE|A_1|^s+E|B_1|^s+1\leqslant|x_0|^sc+K,$$

where K is a fixed constant. If we choose M sufficiently large and $\delta>0$ sufficiently small, for instance such that $K\leqslant(1-\delta-c)M^s$, sufficient conditions for V-uniform ergodicity of (Y_t) are satisfied, see Theorem 16.0.1 in Meyn and Tweedie (1993). 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 (1998), $(\boldsymbol{Y}_t(m))$ is also strongly mixing. A routine argument (see for example Leadbetter and Rootzén, 1988) shows that strong mixing implies $\mathscr{A}(a_n)$.

(iii) Condition (2.4) for $(\boldsymbol{Y}_t)$ follows along the lines of proof of (4.10) in Perfekt (1994) by observing that it suffices to show (2.4) for the one-dimensional 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 (1988), it suffices to show that

$$N_{n,\boldsymbol{Y}}(f)-N_{n,\boldsymbol{X}}(f)=\sum_{t=1}^n f(a_n^{-1}\boldsymbol{Y}_t)-\sum_{t=1}^n f(a_n^{-1}\boldsymbol{X}_t)=o_P(1),$$

for every bounded, continuous, non-negative function f on $\bar{\mathbb{R}}^m\backslash\{\boldsymbol{0}\}$ with compact support. We show that $E|N_{n,\boldsymbol{Y}}(f)-N_{n,\boldsymbol{X}}(f)|\to 0$. Since f has compact support there exists $\beta>0$ such that $f(\boldsymbol{x})=0$ for any $\boldsymbol{x}$ with $|\boldsymbol{x}|>\beta$. Now, obviously,

$$\begin{aligned} &E|N_{n,\boldsymbol{Y}}(f)-N_{n,\boldsymbol{X}}(f)| \\ &\quad\leqslant\sum_{t=1}^n E|f(a_n^{-1}\boldsymbol{Y}_{t-1})-f(a_n^{-1}\boldsymbol{X}_t)|+o(1) \\ &\quad=nE[|f(a_n^{-1}\boldsymbol{Y}_0)-f(a_n^{-1}\boldsymbol{X}_1)|I_{[\beta,\infty]}(a_n^{-1}|\boldsymbol{Y}_0|\vee a_n^{-1}|\boldsymbol{X}_1|)]+o(1). \end{aligned} \tag{3.4}$$

Since the tail of the distribution of Z_1 decreases exponentially, it follows easily from (2.2) and (3.1) that $P(a_n^{-1}|\boldsymbol{Y}_0|\vee a_n^{-1}|\boldsymbol{X}_1|\geqslant\beta)=O(1/n)$ and for any $\delta>0$, $nP(a_n^{-1}|\boldsymbol{Z}|>\delta)\to 0$, where $\boldsymbol{Z}=(Z_1,\ldots,Z_{1+h})$. Also, since f is bounded and uniformly continuous, there exists M such that $f(\boldsymbol{x})<M$ for all $\boldsymbol{x}$, and for every $\varepsilon>0$ there is a $\delta>0$ such that $|\boldsymbol{x}-\boldsymbol{y}|<\delta$ implies $|f(\boldsymbol{x})-f(\boldsymbol{y})|<\varepsilon$. Intersecting the expectation in (3.4) with the sets $\{a_n^{-1}|\boldsymbol{Z}|\leqslant\delta\}$ and $\{a_n^{-1}|\boldsymbol{Z}|>\delta\}$, the lim sup of the right-hand side of (3.4) is bounded by

$$\limsup_{n\to\infty}(\varepsilon nP(a_n^{-1}|\boldsymbol{Y}_0|\vee a_n^{-1}|\boldsymbol{X}_1|\geqslant\beta)+2MnP(a_n^{-1}|\boldsymbol{Z}|>\delta))\leqslant\varepsilon\,(\text{const}),$$

which can be made arbitrarily small for suitably chosen ε. □

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 $E|a+bZ_1|^\alpha=1$. 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 α. 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 (1971).

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

$$(na_n^{-2}\gamma_{n,X}(h))_{h=0,\dots,m}\xrightarrow{\text{d}}(V_h)_{h=0,\dots,m},$$

$$(\rho_{n,X}(h))_{h=1,\dots,m}\xrightarrow{\text{d}}(V_h/V_0)_{h=1,\dots,m},$$

where $(V_h)_{h=0,\dots,m}$ is the α-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 (Figs. 1 and 2).

(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 (1998). First we show that

$$na_n^{-2}[\gamma_{n,X}(h)-\gamma_{n,Y}(h)-E(X_0X_h)+E(Y_0Y_h)]\xrightarrow{P}0. \tag{4.1}$$

The difference in (4.1) can be written as

$$\begin{aligned}&na_n^{-2}[\gamma_{n,X}(h)-\gamma_{n,Y}(h)-E(Z_0Z_h)-E(Z_1Y_h)]\\&\quad=a_n^{-2}\sum_{t=1}^{n}(Z_tZ_{t+h}-E(Z_0Z_h))+a_n^{-2}\sum_{t=1}^{n}Z_{t+h}Y_{t-1}\\&\qquad+a_n^{-2}\sum_{t=1}^{n}(Z_tY_{t-1+h}-E(Z_1Y_h))+o_P(1).\end{aligned}\tag{4.2}$$

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

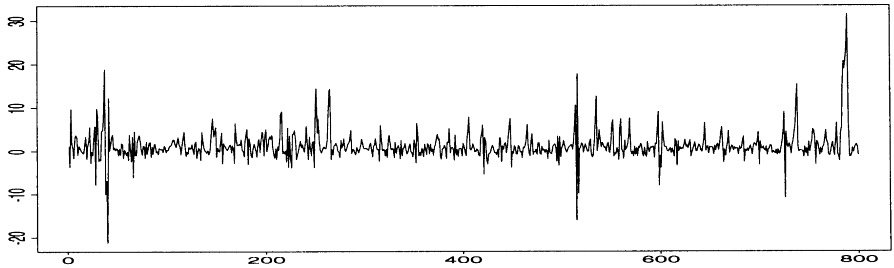

Fig. 1. A realization of the simple bilinear process with $a = 0.2$ and $b = 1$, $\alpha = 1.68$.

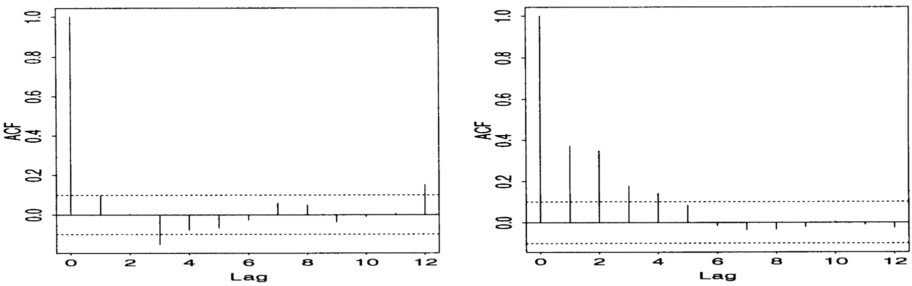

Fig. 2. Sample ACF based on the first and the second half of the time series from Fig. 1.

treatment. Using the recursions for (Y_t) and the identity, $E(Z_1 Y_1) = bEY_1 + a$, we have the decomposition

$$\begin{aligned} a_n^{-2} \sum_{t=1}^{n} (Z_t Y_t - E(Z_1 Y_1)) &= a_n^{-2} \left(a \sum_{t=1}^{n} (Z_t^2 - 1) + b \sum_{t=1}^{n} Z_t^3 + \sum_{t=1}^{n} (aZ_t + b(Z_t^2 - 1)) Y_{t-1} \right. \\ &\quad \left. + b \sum_{t=1}^{n} (Y_{t-1} - EY_{t-1}) \right) \\ &= a_n^{-2} [K_n^{(1)} + K_n^{(2)} + K_n^{(3)} + K_n^{(4)}]. \end{aligned}$$

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

$$\mathrm{var}(a_n^{-2} K_n^{(3)}) = n a_n^{-4} \,\mathrm{var}(aZ_1 + b(Z_1^2 - 1)) \to 0,$$

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

$$\begin{aligned} & na_n^{-2}(\gamma_{n,Y}(0) - EY_1^2) \\ &\quad = a_n^{-2}\sum_{t=1}^{n}[(A_tY_{t-1}+B_t)^2 - E(A_tY_{t-1}+B_t)^2] \\ &\quad = a_n^{-2}\sum_{t=1}^{n}(A_t^2Y_{t-1}^2 - E[A_t^2Y_{t-1}^2]) + a_n^{-2}\sum_{t=1}^{n}2(A_tB_t - E[A_1B_1])Y_{t-1} \\ &\qquad + a_n^{-2}\sum_{t=1}^{n}2E(A_1B_1)(Y_{t-1}-EY_1) + a_n^{-2}\sum_{t=1}^{n}(B_t^2 - EB_t^2). \end{aligned} \tag{4.3}$$

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 $\varepsilon > 0$, write

$$\begin{aligned} J &= a_n^{-2}\sum_{t=1}^{n}(A_t^2 - EA_t^2)Y_{t-1}^2 I_{\{|Y_{t-1}|\leqslant a_n\varepsilon\}} + a_n^{-2}\sum_{t=1}^{n}(A_t^2 - EA_t^2)Y_{t-1}^2 I_{\{|Y_{t-1}|> a_n\varepsilon\}} \\ &\quad + a_n^{-2}EA_t^2\sum_{t=1}^{n}(Y_{t-1}^2 - EY_{t-1}^2) = J_1 + J_2 + J_3. \end{aligned}$$

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

$$\begin{aligned} \mathrm{var}(J_1) &= na_n^{-4}\,\mathrm{var}(Y_0^2(A_1^2 - EA_t^1)I_{\{|Y_0|\leqslant a_n\varepsilon\}}) \\ &\leqslant \mathrm{const}\, na_n^{-4}EY_0^4 I_{\{|Y_0|\leqslant a_n\varepsilon\}} \\ &\sim \mathrm{const}\, na_n^{-4}(a_n\varepsilon)^4 P(|Y_0| > a_n\varepsilon) \\ &\to \mathrm{const}\,\varepsilon^{4-\alpha} \quad \text{as } n\to\infty \\ &\to 0 \quad \text{as } \varepsilon \to 0. \end{aligned} \tag{4.4}$$

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

$$T_{0,\varepsilon}\left(\sum_{i=1}^{\infty} n_i\varepsilon_{\boldsymbol{x}_i}\right) = \sum_{i=1}^{\infty} n_i(x_i^{(0)})^2 I_{\{|x_i^{(0)}|>\varepsilon\}},$$

$$T_{1,\varepsilon}\left(\sum_{i=1}^{\infty} n_i\varepsilon_{\boldsymbol{x}_i}\right) = \sum_{i=1}^{\infty} n_i(x_i^{(1)})^2 I_{\{|x_i^{(0)}|>\varepsilon\}},$$

$$T_{h,\varepsilon}\left(\sum_{i=1}^{\infty} n_i\varepsilon_{\boldsymbol{x}_i}\right) = \sum_{i=1}^{\infty} n_i x_i^{(0)} x_i^{(h-1)} I_{\{|x_i^{(0)}|>\varepsilon\}}, \quad h\geqslant 2,$$

where we denote $\boldsymbol{x}_t = (x_t^{(0)},\ldots,x_t^{(m)}) \in \bar{\mathbb{R}}^{m+1}\setminus\{\mathbf{0}\}$. Using the fact that the set $\{\boldsymbol{x}\in\bar{\mathbb{R}}^2\setminus\{\mathbf{0}\}: |x^{(0)}| > \varepsilon\}$ is bounded and the point process $N_n = \sum_{t=1}^{n}\varepsilon_{a_n^{-1}(Y_t,\ldots,Y_{t+m})}$ is converging according to Theorem 2.2 we have

$$J_2 = T_{1,\varepsilon}N_n - EA_1^2 T_{0,\varepsilon}N_n + o_P(1)$$
$$\xrightarrow{d} T_{1,\varepsilon}N - (a^2+b^2)T_{0,\varepsilon}N =: S_0(\varepsilon,\infty). \tag{4.5}$$

By observing that $ES_0(\varepsilon,\infty)=0$ and using (4.4) and the arguments of Davis and Hsing (1995) in the proof of Theorem 3.1 we conclude that $S_0(\varepsilon,\infty)\xrightarrow{d} V_0^*$ as $\varepsilon\to 0$, where V_0^* is an α-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 $\gamma_{n,Y}(1)$

$$na_n^{-2}(\gamma_{n,Y}(1) - EY_0Y_1)$$
$$= a_n^{-2}\sum_{t=1}^{n}(A_{t+1}Y_t^2 - E(A_{t+1}Y_t) + o_P(1)$$
$$= a_n^{-2}\sum_{t=1}^{n} Y_t^2(A_{t+1}-EA_{t+1}) + a_n^{-2}EA_{t+1}\sum_{t=1}^{n}(Y_t^2 - EY_t^2) + o_P(1).$$

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

$$a_n^{-2}\sum_{t=1}^{n}(A_{t+1}-EA_{t+1})Y_t^2 I_{\{|Y_t|\leqslant a_n\varepsilon\}} + a_n^{-2}\sum_{t=1}^{n}(A_{t+1}-EA_{t+1})Y_t^2 I_{\{|Y_t|>a_n\varepsilon\}}$$
$$=: I_1 + I_2,$$

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

$$I_2 = T_{2,\varepsilon}N_n - EA_1T_{1,\varepsilon}N_n + o_P(1)$$
$$\xrightarrow{d} T_{2,\varepsilon}N - EA_1T_{1,\varepsilon}N =: S_1(\varepsilon,\infty). \tag{4.6}$$

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

$$na_n^{-2}(\gamma_{n,Y}(1) - EY_0Y_1)\xrightarrow{d} V_1^* + (EA_1)V_0 =: V_1.$$

Thus we conclude

$$na_n^{-2}(\gamma_{n,Y}(h) - EY_0Y_h)_{h=0,\dots,m}\xrightarrow{d}(V_h)_{h=0,\dots,m}$$

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

$$na_n^{-2}\left(\rho_{n,Y}(h) - \frac{EY_0Y_h}{EY_0^2}\right)_{h=1,\dots,m}\xrightarrow{d}\frac{1}{EY_0^2}(V_h)_{h=1,\dots,m}.$$

This and (4.1) imply

$$na_n^{-2}(\gamma_{n,X}(h) - EX_0X_h)_{h=0,\dots,m}\xrightarrow{d}(V_h)_{h=0,\dots,m}$$

and

$$na_n^{-2}\left(\rho_{n,X}(h)-\frac{EX_0X_h}{EX_0^2}\right)_{h=1,\dots,m}\stackrel{\mathrm{d}}{\rightarrow}\frac{1}{EX_0^2}(V_h)_{h=1,\dots,m}.$$

(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, 1971) to show that the following limit holds

$$(n^{1/2}(\gamma_{n,X}(h)-EX_0X_h))_{h=0,\dots,m}\stackrel{\mathrm{d}}{\rightarrow}(G_h)_{h=0,\dots,m},$$

where $(G_h)_{h=0,\dots,m}$ is a random vector with mean zero from a multivariate normal distribution. Therefore we further have

$$n^{1/2}\left(\rho_{n,X}(h)-\frac{EX_0X_h}{EX_0^2}\right)_{h=1,\dots,m}\stackrel{\mathrm{d}}{\rightarrow}\frac{1}{EX_0^2}(G_h)_{h=1,\dots,m}.$$

Our analysis shows that the simple bilinear process exhibits similar phenomena as already noticed for other non-linear processes, like the ARCH(1) process with light-tailed input. This is due to the fact that the analysis of both processes can be reduced to the solution of a random difference equation to which Kesten's (1973) theory applies. 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 α and having sufficiently many finite moments.

We finish with two short remarks.

Remark 4.1. The proof of Theorem 3.2 can be adjusted to cover the bilinear process from Davis and Resnick (1996). They take a noise sequence (Z_t) of iid random variables having regularly varying tail probabilities with index α. They assume in addition $a=0$ 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 4.2. The constant γ defined in Theorem 2.1 is the extremal index of the sequence $(|Y_t|)$, see Davis and Mikosch (1998), and therefore the extremal index of the sequence $(|X_t|)$ so that one obtains a result similar to Turkman and Turkman (1998), except that they calculate the extremal index of (X_t). Namely, the extremal index γ is given by the following expression:

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

(see Remark 4.3 of Davis and Mikosch, 1998).

Acknowledgements

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

Breiman, L., 1965. On some limit theorems similar to the arc-sin law. Theory Probab. Appl. 10, 323–331.

Daley, D.J., Vere-Jones, D., 1988. An Introduction to the Theory of Point Processes. Springer, New York.

Davis, R.A., 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., Resnick, S.I., 1985. More limit theory for the sample correlation function of moving averages. Stoch. Proc. Appl. 20, 257–279.

Davis, R.A., Resnick, S.I., 1986. Limit theory for the sample covariance and correlation functions of moving averages. Ann. Statist. 14, 533–558.

Davis, R.A., Resnick, S.I., 1996. Limit theory for bilinear processes with heavy tailed noise. Ann. Appl. Probab. 6, 1191–1210.

Davis, R.A., Mikosch, T., 1998. The sample ACF of heavy-tailed stationary processes with applications to ARCH. Ann. Statist. 26, 2049–2080.

Goldie, C.M., 1991. Implicit renewal theory and tails of solutions of random equations. Ann. Appl. Probab. 1, 126–166.

Ibragimov, I.A., Linnik, Yu.V., 1971. Independent and Stationary Sequences of Random Variables. Wolters-Noordhoff, Groningen.

Kesten, H., 1973. Random difference equations and renewal theory for products of random matrices. Acta Math. 131, 207–248.

Leadbetter, M.R., Rootzén, H., 1988. Extremal theory for stochastic processes. Ann. Probab. 16, 431–478.

Meyn, S.P., Tweedie, R.L., 1993. Markov Chains and Stochastic Stability. Springer, London.

Perfekt, R., 1994. Extremal behaviour of stationary Markov chains with applications. Ann. Appl. Probab. 4, 529–548.

Turkman, K.F., Turkman, M.A.A., 1997. Extremes of bilinear time series models. J. Time Ser. Anal. 18, 305–319.

Vervaat, W., 1979. On a stochastic difference equation and a representation of a non-negative infinitely divisible random variables. Adv. Appl. Prob. 11, 750–783.