Bernoulli **7**(6), 2001, 847–872

Asymptotic behaviour of the sample autocovariance and autocorrelation function of the AR(1) process with ARCH(1) errors

MILAN BORKOVEC

ITG Inc., 44 Farnsworth Street, Boston MA 02210, USA. E-mail: mborkove@itginc.com

We consider a stationary AR(1) process with ARCH(1) errors given by the stochastic difference equation

$$X_t = \alpha X_{t-1} + \sqrt{\beta + \lambda X_{t-1}^2}\varepsilon_t, \qquad t \in \mathbb{N},$$

where the (ε_t) are independent and identically distributed symmetric random variables. In contrast to ARCH and GARCH processes, AR(1) processes with ARCH(1) errors are not solutions of linear stochastic recurrence equations and there is no obvious way to embed them into such equations. However, we show that they still belong to the class of stationary sequences with regularly varying finite-dimensional distributions and therefore the theory of Davis and Mikosch can be applied. We present a complete analysis of the weak limit behaviour of the sample autocovariance and autocorrelation functions of (X_t), $(|X_t|)$ and (X_t^2). The results in this paper can be seen as a natural extension of results for ARCH(1) processes.

Keywords: ARCH model; autoregressive process; extremal index; geometric ergodicity; heavy tails; multivariate regular variation; point processes; sample autocovariance function; strong mixing

1. Introduction

Over the last two decades, there has been a great deal of interest in modelling real data using time series models which exhibit features such as long-range dependence, nonlinearity and heavy tails. Many data sets in econometrics, finance and telecommunications have these common characteristics. In particular, they appear to be reconcilable with the assumption of heavy-tailed marginal distributions. Examples are file lengths, the CPU time to complete a job, the length of on–off cycles in telecommunications, and the log-returns of stock indices, share prices and exchange rates in finance.

The feature of (non)linearity can be often detected by considering the sample autocorrelation functions (ACFs) of a time series, their absolute values and squares. By studying these quantities, many financial time series have been shown to be nonlinear. Indeed, the sample ACFs of log-returns are typically negligible at almost all lags (at least at lags higher than 3) whereas the sample ACFs of their corresponding absolute values and squares are significantly not zero for small lags and have a slow decay to zero when the lag size is increasing. Linear processes do not capture this behaviour in their ACFs and thus are inappropriate models for this type of financial data. In the case of a time series with infinite second moment of the marginal distribution, (non)linearity can sometimes be recognized only by examining the sample ACF of the time series. A time series with infinite second moment which can be represented as a moving average process has the property that the sample ACF at lag h converges in probability to a constant $\rho(h)$, although the mathematical correlation typically does not exist (Davis and Resnick 1985; 1986). On the other hand, for many nonlinear heavy-tailed sequences, the sample ACF at lag h converges in distribution to a non-degenerate random variable. Resnick and van den Berg (2000) propose a test for (non)linearity of a given infinite-variance time series based on subsample stability of the sample ACF.

The phenomenon of random limits of sample ACFs was first observed in the context of infinite-variance bilinear processes by Davis and Resnick (1996) and Resnick (1997). Davis and Mikosch (1998) studied the weak limit behaviour of a large variety of nonlinear processes with regularly varying marginal distributions which satisfy a weak mixing condition and some additional technical assumptions. They showed that the sample autocovariance function (ACVF) and ACF of such processes with infinite fourth but finite second moment have a rate of convergence to the true ACVF and ACF that becomes slower the closer the marginal distributions are to an infinite second moment. In cases of an infinite second moment, the limits of the sample ACVF and ACF are non-deterministic. Examples which belong to the framework of Davis and Mikosch (1998) are the autoregressive (AR) conditionally heteroscedastic processes of order 1 (ARCH(1)), the simple bilinear processes with light-tailed noise (Basrak *et al*. 1999) and the GARCH(1, 1) processes (Mikosch and Stărică 2000). Finally, Davis *et al*. (1999) embedded the three aforementioned processes into a larger class of processes which still satisfy the conditions for the theory of Davis and Mikosch (1998). These processes have in common that they can be described by using solutions to multivariate linear stochastic recurrence equations; equations of this form have been considered by Kesten (1973) and Goldie (1991) and include the important family of the squares of GARCH processes.

The general theory of Davis and Mikosch (1998) could, however, be applied to a different class of processes with a different structure than studied in Davis *et al*. (1999). The present paper addresses the AR processes of order 1 with ARCH(1) errors – or ARARCH(1, 1) processes for short – which do no fit the above-mentioned framework. The class of AR (or more generally AR moving average or ARMA) models with ARCH errors was first proposed by Weiss (1984), who found them to be successful in modelling 13 different US macroeconomic time series. AR models with ARCH errors are the simplest examples of models which can be written by a random recurrence equation of the form

$$X_t = \mu_t + \sigma_t \varepsilon_t, \qquad t \in \mathbb{N}, \tag{1.1}$$

where ε_t are independent and identically distributed (i.i.d.) innovations with mean zero, μ_t is the conditional expectation of X_t (which may or may not depend on t) and the volatility σ_t describes the change of (conditional) variance. Because of the non-constant conditional variance, models of the form (1.1) are often referred to as conditionally heteroscedastic models. Empirical work has confirmed that such models fit many types of financial data (log-returns, exchange rates, etc.). In this paper, we concentrate on the ARARCH(1, 1) process in order to have a Markov structure and thus make the model analytically tractable. It is defined by specifying μ_t and σ_t as follows:

$$\mu_t = \alpha X_{t-1} \quad \text{and} \quad \sigma_t^2 = \beta + \lambda X_{t-1}^2, \tag{1.2}$$

where $\alpha \in \mathbb{R}$ and $\beta, \lambda > 0$. Note that for $\alpha = 0$ we obtain precisely the ARCH(1) model introduced by Engle (1982).

There is a double motivation in studying the sample ACVF and ACF of an ARARCH(1, 1) process. First, an ARARCH(1, 1) process is a natural mixture of an AR(1) and an ARCH(1) process. Therefore, the results of this paper can be seen as a generalization of results for the aforementioned two processes. The weak limit behaviour of the ARCH(1) process was studied by Davis and Mikosch (1998). For $\lambda = 0$, the process defined by (1.1) and (1.2) is an AR(1) process. A summary of results on the asymptotic theory of the sample ACFs of AR processes can be found, for instance, in Brockwell and Davis (1991, Sections 7.2 and 13.3) or Embrechts *et al.* (1997, Section 7.3). Furthermore, ARARCH(1, 1) processes are not solutions of linear stochastic recurrence equations and there is also no obvious way to embed them into such equations. However, we show that the processes still belong to weakly dependent stationary sequences which have regularly varying finite-dimensional distributions. One conclusion of this paper is that ARARCH (1, 1) processes serve as one of the simplest examples of sequences which do not fulfill the setting of Davis *et al.* (1999) but to which the theory of Davis and Mikosch (1998) can still be applied.

The paper is organized as follows. In Section 2 we introduce the ARARCH(1, 1) process (X_t) and consider some of its basic theoretical properties. The weak convergence of some point processes associated with the sequences (X_t), $(|X_t|)$ and (X_t^2) is investigated in Section 3. Finally, in Section 4 we present results concerning the weak convergence of the sample ACVF and ACF of (X_t), $(|X_t|)$ and (X_t^2).

2. Preliminaries

We consider an ARARCH(1, 1) model defined by the stochastic difference equation

$$X_t = \alpha X_{t-1} + \sqrt{\beta + \lambda X_{t-1}^2}\,\varepsilon_t, \qquad t \in \mathbb{N}, \tag{2.1}$$

where the (ε_t) are i.i.d. random variables, $\alpha \in \mathbb{R}$, $\beta, \lambda > 0$ and the parameters α and λ satisfy in addition the inequality

$$\mathrm{E}(\log|\alpha + \sqrt{\lambda}\varepsilon|) < 0, \tag{2.2}$$

which is a necessary and sufficient condition for the existence and uniqueness of a stationary distribution. Here ε is a generic random variable with the same distribution as ε_t. Throughout this paper, we assume the same conditions for ε as in Borkovec and Klüppelberg (2001):

Assumption G. *ε is symmetric with continuous Lebesgue density $p(x)$ and distribution function $H(x)$. Furthermore, ε has full support $\mathbb{R}$, and the second moment of ε exists.*

Assumption T.

(*a*) $p(x) \geqslant p(x')$ *for every* $0 \leqslant x < x'$.

(*b*) *The lower and upper Matuszewska indices of* $\overline{H}$ *are equal, i.e.*

$$-\infty \leqslant \gamma := \lim_{v\to\infty} \frac{\log \limsup_{x\to\infty} \overline{H}(vx)/\overline{H}(x)}{\log v}$$

$$= \lim_{v\to\infty} \frac{\log \liminf_{x\to\infty} \overline{H}(vx)/\overline{H}(x)}{\log v} < -2.$$

(*c*) *If* $\gamma = -\infty$ *then, for all* $\delta > 0$, *there exist constants* $q \in (0, 1)$ *and* $x_0 > 0$ *such that, for all* $x > x_0$ *and* $t > x^q$,

$$p\left(\frac{x \pm \alpha t}{\sqrt{\lambda t^2}}\right) \geqslant (1-\delta) p\left(\frac{x \pm \alpha t}{\sqrt{\beta + \lambda t^2}}\right). \tag{2.3}$$

If $\gamma > -\infty$ *then, for all* $\delta > 0$, *there exist constants* $x_0 > 0$ *and* $T > 0$ *such that, for all* $x > x_0$ *and* $t > T$, (2.3) *holds*.

There exists a wide class of distributions which satisfy these assumptions. Examples are the normal distribution, the Laplace distribution and the Student t distribution. Assumption T is needed to determine the tail of the stationary distribution. For further details concerning these assumptions and for examples, we refer to Borkovec and Klüppelberg (2001). Note that the process (X_t) is evidently a homogeneous Markov chain with state space $\mathbb{R}$ equipped with the Borel σ-algebra. We will write $P_x(\cdot)$, $\mathrm{E}_x(\cdot)$ to denote probabilities and expectations computed under the assumption that $X_0 = x$. Moreover, $P^t(x, \cdot)$ denotes the t-step transition probability of (X_t), that is, $P^t(x, \cdot) = P(X_t \in \cdot | X_0 = x)$ for every $x \geqslant 0$, $t \in \mathbb{N}$. The next theorem gathers together some results on (X_t). The proofs can be found in Borkovec and Klüppelberg (2001). Subjects of interest are the existence (and uniqueness) of a stationary distribution, the tail behaviour of the stationary distribution and the geometric ergodicity. A real-valued Markov chain (X_t) is called *geometrically ergodic* if (X_t) has a unique stationary distribution π and there exists a $\rho \in (0, 1)$ such that, for every $x \in \mathbb{R}$,

$$\rho^{-t} \| P^t(x, \cdot) - \pi \| \to 0, \qquad \text{as } n \to \infty,$$

where $\|\cdot\|$ denotes total variation of signed measures on $\mathcal{B}((-\infty, \infty))$.

Theorem 2.1. *Consider the process* (X_t) *in* (2.1) *with* (ε_t) *satisfying Assumption G and with parameters* α *and* λ *satisfying* (2.2). *Then the following assertions hold:*

(*a*) (X_t) *is geometrically ergodic with unique stationary distribution* π. *In particular,* (X_t) *is strongly mixing with geometric rate of convergence* $\phi_X(h)$, $h \geqslant 0$, *that is, there exists a constant* $C > 0$ *such that, for every* $h \geqslant 0$,

$$\sup_{A\in\sigma(X_s, s\leqslant 0), B\in\sigma(X_s, s\geqslant h)} |P(A\cap B) - P(A)P(B)| := \phi_X(h) \leqslant C\rho^h. \tag{2.4}$$

The stationary distribution π *is continuous and symmetric*.

(*b*) *Let* $\overline{F}(x) = P(X > x)$, $x \geqslant 0$, *be the right tail of the stationary distribution function, and suppose that Assumption T also holds. Then*

$$\overline{F}(x) \sim cx^{-\kappa}, \qquad x \to \infty, \tag{2.5}$$

where

$$c = \frac{1}{2\kappa} \frac{\mathrm{E}(|\alpha|X| + \sqrt{\beta + \lambda X^2}\varepsilon|^{\kappa} - |(\alpha + \sqrt{\lambda}\varepsilon|X\|^{\kappa})}{\mathrm{E}(|\alpha + \sqrt{\lambda}\varepsilon|^{\kappa} \log|\alpha + \sqrt{\lambda}\varepsilon|)} \tag{2.6}$$

and κ *is given as the unique positive solution to*

$$\mathrm{E}(|\alpha + \sqrt{\lambda}\varepsilon|^{\kappa}) = 1. \tag{2.7}$$

Furthermore, the unique positive solution κ *is less than 2 if and only if* $\alpha^2 + \lambda \mathrm{E}(\varepsilon^2) > 1$.

Remark 2.2. (a) Note that $\mathrm{E}(|\alpha + \sqrt{\lambda}\varepsilon|^{\kappa})$ is a function of κ, α and λ. It can be shown that for fixed λ, the exponent κ is decreasing in $|\alpha|$. This means that the tails of the distribution of X become heavier as $|\alpha|$ increases. In particular, for $\alpha \neq 0$ the ARARCH(1, 1) process has heavier tails than the ARCH(1) process.

(b) The strong mixing property automatically implies that the sequence (X_t) satisfies the condition $\mathcal{A}(a_n)$. The condition $\mathcal{A}(a_n)$ is a mixing condition frequently used in connection with point process theory and was introduced by Davis and Hsing (1995). See (3.8) for the definition.

(c) The argument for the symmetry of the stationary distribution is very simple. Since the Markov chain has a unique stationary distribution the Markov chain with the same transition probabilities and initial condition $X_0 = 0$ converges in distribution to the stationary distribution. However, using the recursive definition of this Markov chain, one can write down a representation of the distribution explicitly. It is an infinite series of conditionally independent variables which has a symmetric distribution.

Throughout the remainder of this paper, we always assume that $X_0 \sim \pi$, that is, the ARARCH(1, 1) process is stationary. Note that, because of the symmetry and continuity of X_0 and ε, for every $t \in \mathbb{N}_0$, sign (X_t) and $|X_t|$ are independent and $P(\mathrm{sign}(X_t) = 1) = P(\mathrm{sign}(X_t) = -1) = \frac{1}{2}$.

It appears that the Markov chains which we define next will be relevant to the investigation of the limit behaviour of the sample ACVF and ACF of the stationary processes (X_t), $(|X_t|)$ and (X_t^2): let (Y_t) and $(\tilde{X}_t)$ be Markov chains given by the random recurrence equations

$$Y_t = \alpha Y_{t-1} + \sqrt{\lambda Y_{t-1}^2}\varepsilon_t, \qquad t \in \mathbb{N}, \tag{2.8}$$

and

$$\tilde{X}_t = |\alpha \tilde{X}_{t-1} + \sqrt{\beta + \lambda \tilde{X}_{t-1}^2}\tilde{\varepsilon}_t|, \qquad t \in \mathbb{N}, \tag{2.9}$$

respectively, where the constants $\alpha \in \mathbb{R}$, $\beta, \lambda > 0$ and the sequence (ε_t) are the same as in (2.1), $(\tilde{\varepsilon}_t) \stackrel{d}{=} (\varepsilon_t)$, $Y_0 = X_0$ almost surely and $\tilde{X}_0 = |X_0|$ almost surely.

A classical random walk argument shows that, due to (2.2), $Y_t \to 0$ a.s. as $t \to \infty$. Thus, under our assumptions, the Markov chain (Y_t) has (in contrast to (X_t)) only a degenerate stationary distribution. Nevertheless, we will see in the proof of Proposition 3.1 that (X_t) and (Y_t) are related in the sense that the sequences $(X_0, \ldots, X_m)$ and $(Y_0, \ldots, Y_m)$, $m \in \mathbb{N}$ arbitrary, are both *jointly regularly varying* with the same index $\kappa > 0$ and the same well-specified spectral measure P_Θ; that is, there exists a sequence (a_n) such that

$$\lim_{n\to\infty} nP\left(|(X_0, \ldots, X_m)| > ta_n, \frac{(X_0, \ldots, X_m)}{|(X_0, \ldots, X_m)|} \in B\right)$$

$$= \lim_{n\to\infty} nP\left(|(Y_0, \ldots, Y_m)| > ta_n, \frac{(Y_0, \ldots, Y_m)}{|(Y_0, \ldots, Y_m)|} \in B\right) = t^{-\kappa} P_\Theta(B),$$

for every $t > 0$ and $B \in \mathscr{B}(\mathscr{S}^m)$ with $P_\Theta(\partial B) = 0$, where $\mathscr{S}^m$ denotes the unit sphere in $\mathbb{R}^{m+1}$ with respect to the norm $|\cdot|$.

The next lemma shows that $(|X_t|)$ and $(\tilde{X}_t)$ have the same distribution.

Lemma 2.3. *Let $X_0 \sim \pi$. Then, under $P_{|X_0|,\mathrm{sign}(X_0)}$, $(|X_t|) \stackrel{d}{=} (\tilde{X}_t)$. More precisely, for every $t \in \mathbb{N}_0$ and $A_0, \ldots, A_t \in \mathcal{B}([0, \infty))$,*

$$P_{|X_0|,\mathrm{sign}(X_0)}(|X_0| \in A_0, \ldots, |X_t| \in A_t) = P_{|X_0|,\mathrm{sign}(X_0)}(|X_0| \in A_0, \tilde{X}_1 \in A_1, \ldots, \tilde{X}_t \in A_t)$$

$$= P_{\tilde{X}_0}(\tilde{X}_0 \in A_0, \tilde{X}_1 \in A_1, \ldots, \tilde{X}_t \in A_t).$$

Proof. Because of the symmetry and continuity of X_0 and ε, for every $t \in \mathbb{N}_0$, $|X_t|$ and $\mathrm{sign}(X_t)$ are independent and

$$P_{|X_0|,\mathrm{sign}(X_0)}(\mathrm{sign}(X_t) = 1) = P_{|X_0|,\mathrm{sign}(X_0)}(\mathrm{sign}(X_t) = -1) = \tfrac{1}{2}. \tag{2.10}$$

Suppose next that we even have

$$\mathrm{sign}(X_t) \text{ is independent of } |X_t|, |X_{t-1}|, |X_{t-2}|, \ldots, |X_0|, \qquad t \in \mathbb{N}_0. \tag{2.11}$$

By the symmetry of ε and (2.11), we then conclude that, for every $t \in \mathbb{N}$, $x_0, \ldots, x_{t-1} \geqslant 0$ and $A \in \mathcal{B}([0, \infty))$,

$$P_{|X_0|,\mathrm{sign}(X_0)}(|X_t| \in A \,|\, |X_{t-1}| = x_{t-1}, \ldots, |X_0| = x_0)$$

$$= P_{|X_0|,\mathrm{sign}(X_0)}(|X_t| \in A \,|\, |X_{t-1}| = x_{t-1}) = P\left(\alpha x_{t-1} + \sqrt{\beta + \lambda x_{t-1}^2}\varepsilon_t| \in A\right)$$

and, in particular,

$$P_{|X_0|,\mathrm{sign}(X_0)}(|X_t| \in A \,|\, |X_{t-1}| = x_{t-1}) = P\big(\tilde{X}_t \in A | \tilde{X}_{t-1} = x_{t-1}\big); \tag{2.12}$$

that is, supposing assumption (2.11) holds, the processes $(|X_t|)$ and $(\tilde{X}_t)$ are Markov chains with the same transition kernel and the same stationary distribution. Since $\mathrm{sign}(X_0)$ is symmetric and independent of $|X_0|$, the statement follows.

It remains to show (2.11). We use complete induction for the proof. The base case $t = 0$ is fulfilled due to the assumption that X_0 is symmetric. Now choose arbitrary $t \in \mathbb{N}$. Using the base case of the induction and the independence of $|X_{t-1}|$ and $\operatorname{sign}(X_{t-1})$, we obtain that, for every $A_i \in \mathcal{B}([0, \infty))$, $i = 0, 1, \ldots, t$,

$$\begin{aligned}
&P_{|X_0|,\operatorname{sign}(X_0)}(|X_0| \in A_0, \ldots, |X_{t-1}| \in A_{t-1}, |X_t| \in A_t, \operatorname{sign}(X_t) = 1) \\
&\quad = E_{|X_0|,\operatorname{sign}(X_0)}(1_{\{|X_0| \in A_0, \ldots, |X_{t-1}| \in A_{t-1}\}} \\
&\qquad \times P_{|X_0|,\operatorname{sign}(X_0)}(\alpha|X_{t-1}|\operatorname{sign}(X_{t-1}) + \sqrt{\beta + \lambda|X_{t-1}|^2}\varepsilon_t > 0, |X_t| \in A_t \mid \\
&\qquad\qquad |X_{t-1}|, \ldots, |X_1|, |X_0|)) \\
&\quad = E_{|X_0|,\operatorname{sign}(X_0)}(1_{\{|X_0| \in A_0, \ldots, |X_{t-1}| \in A_{t-1}\}} \\
&\qquad \times P_{|X_0|,\operatorname{sign}(X_0)}(\alpha|X_{t-1}|\operatorname{sign}(X_{t-1}) + \sqrt{\beta + \lambda|X_{t-1}|^2}\varepsilon_t < 0, |X_t| \in A_t \mid \\
&\qquad\qquad |X_{t-1}|, \ldots, |X_1|, |X_0|)) \\
&\quad = P_{|X_0|,\operatorname{sign}(X_0)}(|X_0| \in A_0, \ldots, |X_{t-1}| \in A_{t-1}, |X_t| \in A_t, \operatorname{sign}(X_t) = -1) \qquad (2.14)
\end{aligned}$$

and thus

$$\begin{aligned}
&P_{|X_0|,\operatorname{sign}(X_0)}(|X_0| \in A_0, \ldots, |X_{t-1}| \in A_{t-1}, |X_t| \in A_t, \operatorname{sign}(X_t) = 1) \\
&\quad = P_{|X_0|,\operatorname{sign}(X_0)}(|X_0| \in A_0, \ldots, |X_{t-1}| \in A_{t-1}, |X_t| \in A_t, \operatorname{sign}(X_t) = -1) \\
&\quad = \tfrac{1}{2} P_{|X_0|,\operatorname{sign}(X_0)}(|X_0| \in A_0, \ldots, |X_{t-1}| \in A_{t-1}, |X_t| \in A_t),
\end{aligned}$$

which, in combination with (2.10) completes the proof. □

3. Weak convergence of some point processes associated with the ARARCH(1, 1) process

In this section we formulate results on the weak convergence of point processes of the form

$$N_n = \sum_{t=1}^{n} \delta_{\mathbf{X}_t^{(m)}/a_n}, \qquad n = 1, 2, \ldots, \tag{3.1}$$

where $\mathbf{X}_t^{(m)}$ are random row vectors with arbitrary dimension $m + 1 \in \mathbb{N}$ whose components are closely related to the ARARCH(1, 1) process (X_t) defined in the previous section, and (a_n) is a normalizing sequence of positive numbers. The main result in this section is summarized in Theorem 3.7. The proof of this result is essentially an application of the theory in Davis and Mikosch (1998). Proposition 3.1 gathers together some properties of $(\mathbf{X}_t^{(m)})$ which we need for the proof of Theorem 3.7.

We follow the notation and the point process theory in Davis and Mikosch (1998) and Kallenberg (1983), respectively. The state space of the point processes considered is $\overline{\mathbb{R}}^{m+1}\backslash\{\mathbf{0}\}$. Write $\mathcal{M}$ for the collection of Radon counting measures on $\overline{\mathbb{R}}^{m+1}\backslash\{\mathbf{0}\}$ with null measure o. This means that $\mu \in \mathcal{M}$ if and only if μ is of the form $\mu = \sum_{i=1}^{\infty} n_i \delta_{\mathbf{x}_i}$, where $n_i \in \{1, 2, 3, \ldots\}$, the $\mathbf{x}_i \in \overline{\mathbb{R}}^{m+1} \backslash \{\mathbf{0}\}$ are distinct and $\#\{i | |\mathbf{x}_i| > y\} < \infty$ for all $y > 0$.

Recall that (X_t) is the stationary ARARCH(1, 1) process given by (2.1). (ε_t) satisfies Assumptions G and T, and the parameters α and λ are chosen such that (2.2) holds. We start by specifying the random row vectors $(\mathbf{X}_t^{(m)})$ and the normalizing constants (a_n) in (3.1) and by introducing some auxilary quantities in order to be in the framework of Davis and Mikosch (1998). For $m \in \mathbb{N}_0$, define

$$\mathbf{X}_t^{(m)} = (X_t, X_{t+1}, \ldots, X_{t+m}), \qquad t \in \mathbb{Z},$$

$$\mathbf{Z}_0^{(m)} = \left(r_0, (\alpha r_0 + \sqrt{\lambda}\varepsilon_1), \ldots, (\alpha r_0 + \sqrt{\lambda}\varepsilon_1) \prod_{s=1}^{m-1} (a + \sqrt{\lambda} r_s \varepsilon_{s+1}) \right)$$

and

$$\mathbf{Z}_t^{(m)} = (\alpha r_0 + \sqrt{\lambda}\varepsilon_1) \left(\prod_{s=1}^{t-1} (\alpha + \sqrt{\lambda} r_s \varepsilon_{s+1}), \ldots, \prod_{s=1}^{m+t-1} (\alpha + \sqrt{\lambda} r_s \varepsilon_{s+1}) \right), \qquad t \in \mathbb{N},$$

where $r_s = \operatorname{sign}(Y_s)$, (Y_s) is the process in (2.8) and $\prod_{i=1}^{0} = 1$. Note that the r_s are dependent, symmetric random variables.

Moreover, for $k \in \mathbb{N}_0$ arbitrary but fixed, define the stochastic vectors

$$\mathbf{X}_{-k}^{(m)}(2k+1) = (\mathbf{X}_{-k}^{(m)}, \mathbf{X}_{-k+1}^{(m)}, \ldots, \mathbf{X}_{k}^{(m)})$$

and

$$\mathbf{Z}_0^{(m)}(2k+1) = (\mathbf{Z}_0^{(m)}, \mathbf{Z}_1^{(m)}, \ldots, \mathbf{Z}_{2k}^{(m)}).$$

Analogously to Davis and Mikosch (1998), we take $|\cdot|$ to be the maximum norm in $\mathbb{R}^{m+1}$, that is,

$$|\mathbf{x}| = |(x_0, \ldots, x_m)| = \max_{i=0,\ldots,m} |x_i|.$$

We are now ready to define the sequence (a_n) in (3.1). Let $(a_n^{(k,m)})$ be a sequence of positive numbers such that

$$P(|X| > a_n^{(k,m)}) \sim (n\mathrm{E}(|\mathbf{Z}_0^{(m)}(2k+1)|^{\kappa}))^{-1}, \qquad \text{as } n \to \infty. \tag{3.2}$$

For $k = 0$, we write $a_n = a_n^{(0,m)}$. Note that because of (2.5) one can choose $a_n^{(k,m)}$ as

$$a_n^{(k,m)} = (2c\mathrm{E}(|\mathbf{Z}_0^{(m)}(2k+1)|^{\kappa}))^{1/\kappa} n^{1/\kappa}, \qquad n \geqslant 1. \tag{3.3}$$

From (3.3) one can readily see that a_n and $a_n^{(k,m)}$ differ asymptotically only by the constant $\delta := (\mathrm{E}(|\mathbf{Z}_0^{(m)}(2k+1)|^{\kappa})/\mathrm{E}(|\mathbf{Z}_0^{(m)}|^{\kappa}))^{1/\kappa}$, that is, $a_n^{(k,m)}/a_n \to \delta$, as $n \to \infty$.

With these observations and notation we can state the next two propositions.

Proposition 3.1. *Let (X_t) be the stationary ARARCH*(1, 1) process given by (2.1) *and assume that the conditions of Theorem* 2.1 *hold. Then:*

(*a*) $(\mathbf{X}_t^{(m)})$ *is strongly mixing with geometric rate of convergence* $\phi_{\mathbf{X}^{(m)}}(h) = \phi_X(h-m)$, $h \geqslant m$, *where* $\phi_X(\cdot)$ *is given in* (2.4).

(*b*) $\mathbf{X}_{-k}^{(m)}(2k+1)$ *is jointly regularly varying with index* $\kappa > 0$*; more precisely,*

$$nP\left(|\mathbf{X}_{-k}^{(m)}(2k+1)| > ta_n^{(k,m)}, \frac{\mathbf{X}_{-k}^{(m)}(2k+1)}{|\mathbf{X}_{-k}^{(m)}(2k+1)|} \in \cdot\right) \tag{3.4}$$

$$\xrightarrow{v} t^{-k} \frac{\mathrm{E}(|\mathbf{Z}_0^{(m)}(2k+1)|^\kappa 1_{\{\mathbf{Z}_0^{(m)}(2k+1)/|\mathbf{Z}_0^{(m)}(2k+1)| \in \cdot\}})}{\mathrm{E}(|\mathbf{Z}_0^{(m)}(2k+1)|^\kappa)}, \qquad t > 0,$$

as $n \to \infty$, *where the symbol* $\xrightarrow{v}$ *stands for vague convergence on the Borel* σ*-field of the unit sphere* $\mathcal{S} := \mathcal{S}^{(2k+1)(m+1)-1}$ *of* $\mathbb{R}^{(2k+1)(m+1)}$ *with respect to the maximum norm* $|\cdot|$.

Remark 3.2. (a) In the spirit of Davis and Mikosch (1998), the jointly regular varying property of $\mathbf{X}_{-k}^{(m)}(2k+1)$ can also be expressed in the more familiar way

$$nP\left(|\mathbf{X}_{-k}^{(m)}(2k+1)| > ta_n^{(k,m)}, \frac{\mathbf{X}_{-k}^{(m)}(2k+1)}{|\mathbf{X}_{-k}^{(m)}(2k+1)|} \in \cdot\right) \xrightarrow{v} t^{-\kappa} P_\Theta(\cdot), \qquad \text{as } n \to \infty, \tag{3.5}$$

where $P_\Theta = \tilde{P} \circ (\theta_{-k}^{(k)}, \ldots, \theta_k^{(k)})^{-1}$, $\theta_j^{(k)} = \mathbf{Z}_{k+j}^{(m)}/|\mathbf{Z}_0^{(m)}(2k+1)|$, $j = -k, \ldots, k$, and $\mathrm{d}\tilde{P} = |\mathbf{Z}_0^{(m)}(2k+1)|^\kappa/\mathrm{E}(|\mathbf{Z}_0^{(m)}(2k+1)|^\kappa)dP$. In the following we will use the latter notation.

(b) Due to statement (b) in Proposition 3.1, the positive sequence (a_n) in (3.2) with $k = 0$ can also be characterized by

$$\lim_{n\to\infty} nP(|\mathbf{X}_0^{(m)}| > a_n) = 1.$$

Proof. (a) This is an immediate consequence of the strong mixing property of (X_t) which is stated in Theorem 2.1(a) and the fact that strong mixing is characterized by the underlying σ-field.

(b) Define $\mathbf{Y}_t^{(m)} := (Y_t, Y_{t+1}, \ldots, Y_{t+m})$, $t \in \mathbb{N}_0$, and $\mathbf{Y}_0^{(m)}(2k+1) = (\mathbf{Y}_0^{(m)}, \mathbf{Y}_1^{(m)}, \ldots, \mathbf{Y}_{2k}^{(m)})$, where (Y_t) is the process given in (2.8). Using the definition of the process (Y_t) and of the stochastic vectors $\mathbf{Z}_t^{(m)}$, it can readily be seen that

$$\mathbf{Y}_0^{(m)}(2k+1) = |X_0|\mathbf{Z}_0^{(m)}(2k+1), \tag{3.6}$$

where $|\mathbf{X}_0|$ and $\mathbf{Z}_0^{(m)}(2k+1)$ are independent, non-negative random variables. Moreover, $|X_0|$ is regularly varying with index $\kappa > 0$ and $\mathrm{E}(|\mathbf{Z}_0^{(m)}(2k+1)|^\kappa) < \infty$. Thus, a result of Breiman (1965) yields that, for every $t > 0$,

$$
\begin{aligned}
& nP\left(|\mathbf{Y}_0^{(m)}(2k+1)| > ta_n^{(k,m)}, \frac{\mathbf{Y}_0^{(m)}(2k+1)}{|\mathbf{Y}_0^{(m)}(2k+1)|} \in \cdot \right) \\
& = nP\left(|X_0|\,|\mathbf{Z}_0^{(m)}(2k+1)| > ta_n^{(k,m)}, \frac{\mathbf{Z}_0^{(m)}(2k+1)}{|\mathbf{Z}_0^{(m)}(2k+1)|} \in \cdot \right) \\
& \sim nP(|X_0| > ta_n^{(k,m)}) \mathrm{E}(|\mathbf{Z}_0^{(m)}(2k+1)|^\kappa 1_{\{\mathbf{Z}_0^{(m)}(2k+1)/|\mathbf{Z}_0^{(m)}(2k+1)| \in \cdot\}}) \\
& \sim t^{-\kappa} \frac{\mathrm{E}(|\mathbf{Z}_0^{(m)}(2k+1)|^\kappa 1_{\{\mathbf{Z}_0^{(m)}(2k+1)/|\mathbf{Z}_0^{(m)}(2k+1)| \in \cdot\}})}{\mathrm{E}(|\mathbf{Z}_0^{(m)}(2k+1)|^\kappa)}, \qquad \text{as } n \to \infty,
\end{aligned}
$$

where we have used (2.5) and (3.2) in the fourth line. Therefore, $\mathbf{Y}_0^{(m)}(2k+1)$ is jointly regularly varying with index $\kappa > 0$ and the spectral measure is the same as in (3.5).

Now note that if there are two vectors $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A}$ is regularly varying and $P(|\mathbf{A} - \mathbf{B}| > x) = o(P(|\mathbf{A}| > x))$ as $x \to \infty$, then $\mathbf{B}$ is regularly varying with the same index and spectral measure as $\mathbf{A}$ (see, for example, Davis *et al*. 1999, Remark 5.11). Thus, setting $\mathbf{A} = \mathbf{Y}_0^{(m)}(2k+1)$, $\mathbf{B} = \mathbf{X}_0^{(m)}(2k+1)$ and $x = a_n$, it remains to show that

$$
\frac{P(|\mathbf{X}_0^{(m)}(2k+1) - \mathbf{Y}_0^{(m)}(2k+1)| > a_n)}{P(|\mathbf{Y}_0^{(m)}(2k+1)| > a_n)} \sim \delta^{-\kappa} nP(|\mathbf{X}_0^{(m)}(2k+1) - \mathbf{Y}_0^{(m)}(2k+1)| > a_n) \to 0, \qquad \text{as } n \to \infty,
$$

which follows from the joint regular variation of $\mathbf{Y}_0^{(m)}(2k+1)$ and the fact that $a_n^{(k,m)} \sim \delta a_n$ as $n \to \infty$. However, since the vectors $\mathbf{X}_0^{(m)}(2k+1)$ and $\mathbf{Y}_0^{(m)}(2k+1)$ are finite-dimensional, it suffices to show that, for every $s \in \mathbb{N}_0$,

$$
nP(|X_s - Y_s| > a_n) \to 0, \qquad \text{as } n \to \infty.
$$

Using the recurrence equations both for X_s and Y_s s times, one can easily see that, for every $s \in \mathbb{N}$,

$$
|X_s - Y_s| \leqslant \sqrt{\beta} \sum_{j=1}^{s} \prod_{i=j+1}^{s} (\alpha + \sqrt{\lambda}|\varepsilon_i|)|\varepsilon_j| =: E.
$$

Therefore, the expression $nP(|X_s - Y_s| > a_n)$ can be estimated by $nP(E > a_n)$ and the latter converges to zero because

$$
\mathrm{E}(E^\kappa) \leqslant 2^{\kappa-1} (\sqrt{\beta})^\kappa \sum_{j=1}^{s} \prod_{i=j+1}^{s} \mathrm{E}((\alpha + \sqrt{\lambda}|\varepsilon_i|)^\kappa) \mathrm{E}(|\varepsilon_j|^\kappa) < \infty.
$$

□

Proposition 3.3. *Let* (p_n) *be an increasing sequence such that* $p_n/n \to 0$ *as* $n \to \infty$*. Then, for every* $y > 0$*,*

$$\lim_{p\to\infty}\limsup_{n\to\infty} P\left(\bigvee_{p\leqslant|t|\leqslant p_n} |\mathbf{X}_t^{(m)}| > a_n y \,|\, |\mathbf{X}_0^{(m)}| > a_n y\right) = 0. \tag{3.7}$$

Remark 3.4. In the case of a strongly mixing process, the condition $p_n/n \to 0$, as $n \to \infty$, in combination with

$$\frac{n\phi_{\mathbf{X}^{(m)}}(\sqrt{p_n})}{p_n} \to 0, \qquad \text{as } n \to \infty,$$

is sufficient to guarantee that (p_n) is a $\mathcal{A}(a_n)$-separating sequence, that is,

$$\mathrm{E}\exp\left(-\sum_{t=1}^{n} f\left(\frac{\mathbf{X}_t^{(m)}}{a_n}\right)\right) - \left(\mathrm{E}\exp\left(-\sum_{t=1}^{p_n} f\left(\frac{\mathbf{X}_t^{(m)}}{a_n}\right)\right)\right)^{k_n} \to 0, \qquad \text{as } n \to \infty, \tag{3.8}$$

where $k_n = [n/p_n]$ and f is an arbitrary bounded non-negative step function on $\overline{\mathbb{R}}^m\backslash\{\mathbf{0}\}$ (see also the comments in Remark 3.2 in Davis *et al*. 1999).

In order to prove Proposition 3.3 we need the following lemma.

Lemma 3.5. *Let* $(\tilde{Z}_t) := (\log \tilde{X}_t^2)$ *and* $(S_t^{(a)})$, $a > 0$, *be the random walk given by*

$$S_t^{(a)} = S_{t-1}^{(a)} + \log V_t^{(a)},$$

where

$$V_t^{(a)} := \left(\alpha + \sqrt{\beta \mathrm{e}^{-a} + \lambda}\varepsilon_t\right)^2 - 2\alpha\sqrt{\beta}\mathrm{e}^{-a/2}\varepsilon_t 1_{\{\varepsilon_t<0\}} \geqslant 0, \qquad t \in \mathbb{N},$$

and $S_0^{(a)} \equiv 0$. *Choose* $a > 0$ *large such that*

$$\mathrm{E}\left(V^{(a)^{\kappa/4}}\right) < 1. \tag{3.9}$$

If $\tilde{Z}_0 > a$ *and* $N_a := \inf\{\nu \geqslant 1 | \tilde{Z}_\nu \leqslant a\}$, *then*

$$\tilde{Z}_t \leqslant \tilde{Z}_0 + S_t^{(a)}, \qquad \textit{for every } t \leqslant N_a \textit{ a.s}, \tag{3.10}$$

and the random walk $(S_t^{(a)})$ *has negative drift.*

Remark 3.6. (a) We can achieve (3.9) since (2.2) implies that $\mathrm{E}(|\alpha + \sqrt{\lambda}\varepsilon|^u) < 1$ for all $u \in (0, \kappa)$ and

$$\mathrm{E}\left(V^{(a)^{\kappa/4}}\right) \to \mathrm{E}\left(|\alpha + \sqrt{\lambda}\varepsilon|^{\kappa/2}\right), \qquad \text{as } a \to \infty,$$

by the dominated convergence theorem.

(b) Since $(\tilde{X}_t)$ is given by the stochastic recurrence equation (2.10), the process $(\tilde{Z}_t)$ can be recursively written as

$$\tilde{Z}_t = \tilde{Z}_{t-1} + \log\left((\alpha + \sqrt{\beta \mathrm{e}^{-\tilde{Z}_{t-1}} + \lambda}\varepsilon_t)^2\right), \qquad t \in \mathbb{N}, \tag{3.11}$$

where $\tilde{Z}_0 = \log X_0^2$ a.s.

Proof of Lemma 3.5. Let $x \geqslant a$ be arbitrary. If $\varepsilon \geqslant 0$ it can be readily seen that

$$(\alpha + \sqrt{\beta e^{-x} + \lambda\varepsilon})^2 \leqslant (\alpha + \sqrt{\beta e^{-a} + \lambda\varepsilon})^2. \tag{3.12}$$

Now consider $\varepsilon < 0$. Then,

$$(\alpha + \sqrt{\beta e^{-x} + \lambda\varepsilon})^2 - (\alpha + \sqrt{\beta} e^{-a} + \lambda\varepsilon)^2 \leqslant 2\alpha\sqrt{\beta} e^{-a/2}(-\varepsilon). \tag{3.13}$$

From (3.12) and (3.13), we obtain

$$(\alpha + \sqrt{\beta e^{-x} + \lambda\varepsilon})^2 \leqslant (\alpha + \sqrt{\beta e^{-a} + \lambda\varepsilon})^2 \left(1 - \frac{2\alpha\sqrt{\beta} e^{-a/2}\varepsilon}{(\alpha + \sqrt{\beta e^{-x} + \lambda\varepsilon})^2} 1_{\{\varepsilon<0\}}\right). \tag{3.14}$$

Taking logarithms on both sides and using the additive structure (3.11) proves (3.10). Finally, from (3.9) and Jensen's inequality we conclude that $(S_t^{(a)})$ has negative drift. □

Proof of Proposition 3.3. We start by rewriting the probability in statement (c).

$$\begin{aligned} &P\left(\bigvee_{p\leqslant|t|\leqslant p_n} |\mathbf{X}_t^{(m)}| > a_n y \,|\, |\mathbf{X}_0^{(m)}| > a_n y\right) \\ &= P\left(\max_{-p_n\leqslant t\leqslant -p+m} |X_t| > a_n y \,|\, \max_{0\leqslant j\leqslant m} |X_j| > a_n y\right) \\ &\quad + P\left(\max_{p\leqslant t\leqslant p_n+m} |X_t| > a_n y \,|\, \max_{0\leqslant j\leqslant m} |X_j| > a_n y\right) \\ &=: (J_1) + (J_2). \end{aligned}$$

In what follows we consider only (J_1); (J_2) can be treated in a similar way. First, note that

$$\begin{aligned} (J_1) &\leqslant \sum_{j=0}^{m} \frac{P(\max_{-p_n\leqslant t\leqslant -p+m} |X_t| > a_n y,\, |X_j| > a_n y)}{P(|X_j| > a_n y)} \frac{P(|X_j| > a_n y)}{P(\max_{0\leqslant j\leqslant m} |X_j| > a_n y)} \\ &\leqslant \sum_{j=0}^{m} P\left(\max_{-p_n-j\leqslant t\leqslant -p+m-j} |X_t| > a_n y \,|\, |X_0| > a_n y\right) \\ &\leqslant (m+1) P\left(\max_{-p_n-m\leqslant t\leqslant -p+m} |X_t| > a_n y \,|\, |X_0| > a_n y\right) \\ &\leqslant (m+1) \sum_{t=-p_n-m}^{-p+m} P(|X_t| > a_n y \,|\, |X_0| > a_n y). \end{aligned}$$

Moreover, by the stationarity of (X_t) and substitution, we obtain

$$\begin{aligned}(J_1) &\leqslant (m+1)\sum_{t=-p_n-m}^{-p+m} P(|X_{-t}| > a_n y\,|\,|X_0| > a_n y)\\ &= (m+1)\sum_{t=p-m}^{p_n+m} P(|X_t| > a_n y\,|\,|X_0| > a_n y).\end{aligned} \tag{3.15}$$

Using Lemma 2.3 and the notation in Lemma 3.5 with $a > 0$ such that (3.9) holds, we can rewrite the last expression in (3.15) as

$$\begin{aligned}(m+1)\Bigg(&\sum_{t=p-m}^{p_n+m} P(\tilde{Z}_t > \log(a_n y)^2,\, N_a < p-m|\tilde{Z}_0 > \log(a_n y)^2).\\ &+\sum_{t=p-m}^{p_n+m} P(\tilde{Z}_t > \log(a_n y)^2,\, p-m \leqslant N_a < p_n + m|\tilde{Z}_0 > \log(a_n y)^2)\\ &+\sum_{t=p-m}^{p_n+m} P(\tilde{Z}_t > \log(a_n y)^2,\, N_a > p_n + m|\tilde{Z}_0 > \log(a_n a)^2)\Bigg)\\ &=: (m+1)((K_1)+(K_2)+(K_3)).\end{aligned}$$

It remains to prove that all the K_i are negligible as $n \to \infty$. We follow the proof in Borkovec (2000, pp. 202–205). Note first that, because of the continuity of the transition probability of $\tilde{Z}_t$ and (2.6), there exist constants $C > 0$ and $N \in \mathbb{N}$ such that, for any $n > N$, $x \in [-n, a]$ and $s \in \mathbb{N}$,

$$nP(\tilde{Z}_s > \log(a_n y)^2|\tilde{Z}_0 = x) \leqslant C. \tag{3.16}$$

This enables us now to estimate K_1. Conditioning on $\tilde{Z}_{p-m-1}$ yields

$$\begin{aligned}K_1 = &\sum_{t=p-m}^{p_n+m} \mathrm{E}\Big(1_{\{N_a<p-m\}}1_{\{\tilde{Z}_{p-m-1}\geqslant -n\}}P(\tilde{Z}_t > \log(a_n y)^2|\tilde{Z}_{p-m-1})|\tilde{Z}_0 > \log(a_n y)^2\Big)\\ &+\sum_{t=p-m}^{p_n+m} \mathrm{E}\Big(1_{\{N_a<p-m\}}1_{\{\tilde{Z}_{p-m-1}<-n\}}P(\tilde{Z}_t > \log(a_n y)^2|\tilde{Z}_{p-m-1})|\tilde{Z}_0 > \log(a_n y)^2\Big)\\ =: &\, L_1 + L_2.\end{aligned}$$

By (3.16),

$$L_1 \leqslant \sum_{t=p-m}^{p_n+m} \frac{1}{n} \mathrm{E}\Big(1_{\{N_a<p-m\}} 1_{\{\tilde{Z}_{p-m-1} \geqslant -n\}} nP(\tilde{Z}_t > \log(a_n y)^2 | \tilde{Z}_{p-m-1}) \mid \tilde{Z}_0 > \log(a_n y)^2\Big)$$

$$\leqslant \sum_{t=p-m}^{p_n+m} \frac{C}{n} \mathrm{E}\Big(1_{\{N_a<p-m\}} 1_{\{\tilde{Z}_{p-m-1} \geqslant -n\}} \mid \tilde{Z}_0 > \log(a_n y)^2\Big) \tag{3.17}$$

$$\leqslant C \frac{p_n+m}{n} \to 0, \qquad \text{as } n \to \infty,$$

since $p_n = o(n)$. Furthermore, setting $B_\nu := \{\tilde{Z}_1 > a, \ldots, \tilde{Z}_{\nu-1} > a\}$ for any $\nu = 2, 3, 4, \ldots$ and $B_1 = \Omega$, we obtain

$$L_2 \leqslant \sum_{\nu=1}^{p-m-1} \sum_{t=p-m}^{p_n+m} \mathrm{E}\left(1_{\{Na=\nu\}} 1_{\{\tilde{Z}_\nu<-n\}} \mid \tilde{Z}_0 > \log(a_n y)^2\right)$$

$$\leqslant \sum_{\nu=1}^{p-m-1} (p_n+m) \mathrm{E}\left(1_{B_\nu} P(\tilde{Z}_\nu < -n | \tilde{Z}_{\nu-1}) \mid \tilde{Z}_0 > \log(a_n y)^2\right)$$

$$\leqslant \sum_{\nu=1}^{p-m-1} (p_n+m) \mathrm{E}\left(1_{B_\nu} P((\alpha + \sqrt{\beta \mathrm{e}^{\tilde{Z}_{\nu-1}} + \lambda}\varepsilon_\nu)^2 < -n - \tilde{Z}_{\nu-1} \mid \tilde{Z}_{\nu-1}) \mid \tilde{Z}_0 > \log(a_n y)^2\right)$$

$$\leqslant \sum_{\nu-1}^{p-m-1} (p_n+m) \mathrm{E}\left(1_{B_\nu} P\left(\frac{-\mathrm{e}^{-(n+a)/2} - \alpha}{\sqrt{\beta \mathrm{e}^{-\tilde{Z}_{\nu-1}} + \lambda}} < \varepsilon_\nu < \frac{\mathrm{e}^{-(n+a)/2} - \alpha}{\sqrt{\beta \mathrm{e}^{-\tilde{Z}_{\nu-1}} + \lambda}} \mid \tilde{Z}_{\nu-1}\right) \mid \tilde{Z}_0 > \log(a_n y)^2\right)$$

$$\leqslant \text{const. } p(p_n+m)\mathrm{e}^{-(n+a)/2} \to 0, \qquad \text{as } n \to \infty,$$

and therefore, with (3.17), $K_1 \to 0$ as $n \to \infty$.

Next, we find an upper bound for $\limsup_{n\to\infty} K_3$. Note first that, by the Markov inequality, for $z \geqslant 0$, $t \in \mathbb{N}$ arbitrary and $\eta := (V^{(a)^{\kappa/4}})$

$$P(S_t^{(a)} > -z) \leqslant P\Big(e^{(\kappa/4)S_t^{(a)}} > e^{-(\kappa/4)z}\Big) = P\left(\prod_{s=1}^t V_s^{(a)^{\kappa/4}} > e^{-(\kappa/4)z}\right) \leqslant e^{(\kappa/4)z}\eta^t. \tag{3.18}$$

Thus, from Lemma 3.5, (3.18) and the exponential tail behaviour of $\tilde{Z}_0$,

$$\begin{aligned}
\limsup_{n\to\infty} K_3 &\leqslant \limsup_{n\to\infty} \sum_{t=p-m}^{p_n+m} P(N_a \geqslant p_n+m,\ \tilde{Z}_0 + S_t^{(a)} > \log(a_n y)^2 \mid \tilde{Z}_0 > \log(a_n y)^2) \\
&\leqslant \limsup_{n\to\infty} \sum_{t=p-m}^{p_n+m} P(\tilde{Z}_0 + S_t^{(a)} > \log(a_n y)^2 \mid \tilde{Z}_0 > \log(a_n y)^2) \\
&\leqslant C \sum_{t=p-m}^{\infty} \int_0^\infty P(S_t^{(a)} > -z) \frac{\kappa}{2} e^{-(\kappa/2)z}\, \mathrm{d}z \qquad (3.19) \\
&\leqslant 2C \sum_{t=p-m}^{\infty} \eta^t = 2C \frac{\eta^{p-m}}{1-\eta}, \qquad \text{for some constant } C > 0.
\end{aligned}$$

Finally,

$$\begin{aligned}
K_2 \leqslant& \sum_{t=p-m}^{p_n+m} \sum_{v=p-m}^{t-1} P(N_a = v,\ \tilde{Z}_t > \log(a_n y)^2 \mid \tilde{Z}_0 > \log(a_n y)^2) \\
&+ \sum_{t=p-m}^{p_n+m} \sum_{v=t}^{p_n+m-1} P(N_a = v,\ \tilde{Z}_t > \log(a_n y)^2 \mid \tilde{Z}_0 > \log(a_n y)^2) \\
=:& M_1 + M_2.
\end{aligned}$$

Similarly to K_1 and K_3, respectively, we derive that

$$\limsup_{n\to\infty} M_1 = 0,$$

$$\limsup_{n\to\infty} M_2 = 2\tilde{C} \frac{\eta^{p-m}}{1-\eta}, \qquad \text{for some } \tilde{C} > 0.$$

Combining everything and letting $p \to \infty$, the statement follows. □

Proposition 3.1 proved some properties for $(\mathbf{X}_t^{(m)})$ and $(\mathbf{X}_t^{(m)}(2k+1))$ which turn out to be exactly the required assumptions in Davis and Mikosch (1998) for weak convergence of point processes of the form (3.1). If we define

$$\tilde{\mathcal{M}} = \{\mu \in \mathcal{M} \mid \mu(\{\mathbf{x} \mid |\mathbf{x}| > 1\}) = 0 \text{ and } \mu(\{\mathbf{x} | \mathbf{x} \in \mathcal{S}^m\}) > 0\}$$

and if we let $\mathcal{B}(\tilde{\mathcal{M}})$ be the Borel σ-field of $\tilde{\mathcal{M}}$, then the following theorem is an immediate consequence of Proposition 3.1.

Theorem 3.7. *Assume (X_t) is the stationary ARARCH$(1, 1)$ process satisfying the conditions of Theorem 2.1. Then*

$$N_n^{\mathbf{X}} = \sum_{t=1}^n \delta_{\mathbf{X}_t/a_n} \xrightarrow{v} N^{\mathbf{X}} = \sum_{i=1}^\infty \sum_{j=1}^\infty \delta_{P_i \mathbf{Q}_{ij}}, \qquad (3.20)$$

where $\mathbf{X}_t = \mathbf{X}_t^{(m)}$ *and* $\sum_{i=1}^{\infty}\delta_{P_i}$ *is a Poisson process on* $(0, \infty]$ *with intensity*

$$\nu(\mathrm{d}y) = \kappa^2 \int_1^{\infty} P\left(\sup_{k\geq 1}\prod_{s=1}^{k}(\alpha + \sqrt{\lambda}\varepsilon_s) \leq t^{-1}\right) t^{-\kappa-1}\,\mathrm{d}t y^{-\kappa-1}\,\mathrm{d}y.$$

The process $\sum_{i=1}^{\infty}\delta_{P_i}$ *is independent of the sequence of i.i.d. point processes* $\sum_{j=1}^{\infty}\delta_{\mathbf{Q}_{ij}}$, $i \geq 1$, *with joint distribution* Q *on* $(\tilde{\mathcal{M}}, \mathcal{B}(\tilde{\mathcal{M}}))$*, where* Q *is the weak limit of*

$$\frac{\tilde{\mathrm{E}}((|\theta_0^{(k)}|^{\kappa} - \bigvee_{j=1}^{k}|\theta_j^{(k)}|^{\kappa})_+ 1_{\{\cdot\}}(\sum_{|t|\leq k}\delta_{\theta_t^{(k)}}))}{\tilde{\mathrm{E}}(|\theta_0^{(k)}|^{\kappa} - \bigvee_{j=1}^{k}|\theta_j^{(k)}|^{\kappa})_+} \tag{3.21}$$

as $k \to \infty$*, and the limit exists.* $\tilde{\mathrm{E}}$ *is the expectation with respect to the probability measure* $\mathrm{d}\tilde{P}$ *defined in Remark* 3.2(*a*).

Remark 3.8. Analogous results can be found for the vectors

$$|\mathbf{X}_t| = |\mathbf{X}_t^{(m)}| = (|X_t|, \ldots, |X_{t+m}|) \quad \text{and} \quad \mathbf{X}_t^2 = \mathbf{X}_t^{(m)^2} = (X_t^2, \ldots, X_{t+m}^2), \qquad t \in \mathbb{Z},\ m \in \mathbb{N},$$

by using (3.20) and the continuous mapping theorem. Thus, under the same assumptions as in Theorem 3.7, we have

$$N_n^{|\mathbf{X}|} = \sum_{t=1}^{n}\delta_{|\mathbf{X}_t|/a_n} \xrightarrow{v} \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\delta_{P_i|\mathbf{Q}_{ij}|}$$

and

$$N_n^{\mathbf{X}^2} = \sum_{t=1}^{n}\delta_{\mathbf{X}_t^2/a_n} \xrightarrow{v} \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\delta_{P_i\mathbf{Q}_{ij}^2},$$

where the sequences (P_i), $(\mathbf{Q}_{ij})$ are the same as above and

$$|\mathbf{Q}_{ij}|^l = (|Q_{ij}^{(0)}|^l, |Q_{ij}^{(1)}|^l, \ldots, |Q_{ij}^{(m)}|^l), \qquad l = 1, 2.$$

Proof. The proof is a simple application of Theorem 2.8 in Davis and Mikosch (1998). Because of Proposition 3.1, all assumptions of the theorem are satisfied. Moreover, the extremal index $\gamma = \lim_{k\to\infty}\mathrm{E}(|\theta_0^{(k)}|^{\kappa} - \bigvee_{j=1}^{k}|\theta_j^{(k)}|^{\kappa})_+/\mathrm{E}|\theta_0^{(k)}|^{\kappa}$ of the ARARCH(1, 1) process is specified by the formula (see Borkovec 2000)

$$\gamma = \kappa\int_1^{\infty} P\left(\sup_{k\geq 1}\prod_{s=1}^{k}(\alpha + \sqrt{\lambda}\varepsilon_s) \leq t^{-1}\right) t^{\kappa-1}\,\mathrm{d}t.$$

□

4. Asymptotic behaviour of the sample ACVF and ACF

In what follows we derive the limit behaviour of the sample ACVF and ACF of the stationary ARARCH(1, 1) process considered in the previous sections. The point process results of Section 3 will be helpful.

Define the sample ACVF of (X_t) by

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

and the corresponding sample ACF by

$$\rho_{n,X}(h) = \frac{\gamma_{n,X}(h)}{\gamma_{n,X}(0)}, \qquad h = 0, 1, \ldots.$$

The sample ACVF and ACF for $(|X_t|)$ and (X_t^2) are given in the same way. Moreover, we write

$$\gamma_X(h) = \mathrm{E}(X_0 X_h), \qquad \gamma_{|X|^l}(h) = \mathrm{E}(|X_0|^l |X_h|^l)$$

and

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)},$$

$$\rho_{|X|^l}(h) = \frac{\gamma_{|X|^l}(h)}{\gamma_{|X|^l}(0)}, \qquad l = 1, 2,\ h = 0, 1, \ldots,$$

if these quantities exist. If this is the case a straightforward calculation yields

$$\gamma_X(h) = \alpha^h \gamma_X(0) = \frac{\alpha^h \beta}{1 - \alpha^2 - \lambda \mathrm{E}(\varepsilon^2)}$$

and

$$\gamma_{X^2}(h) = (\alpha^2 + \lambda \mathrm{E}(\varepsilon^2))^h \gamma_{X^2}(0) + \beta \mathrm{E}(\varepsilon^2)\gamma_X(0)\sum_{j=0}^{h-1}(\alpha^2 + \lambda \mathrm{E}(\varepsilon^2))^j, \qquad h \geqslant 0,$$

where

$$\gamma_{X^2}(0) = \frac{2\beta\gamma_X(0)(3\alpha^2 \mathrm{E}(\varepsilon^2) + \lambda \mathrm{E}(\varepsilon^4)) + \beta \mathrm{E}(\varepsilon^4)}{1 - \alpha^4 - 6\alpha^2\lambda \mathrm{E}(\varepsilon^2) - \lambda^2 \mathrm{E}(\varepsilon^4)}.$$

Mean-corrected versions for the sample ACVF and ACF can also be investigated. However, one can show (with the same approach as in the proof of Theorem 4.1) that the limits remain the same (see also Remark 3.6 of Davis and Mikosch 1998).

In order to state our results we have to introduce several mappings. Let $\delta > 0$, $\mathbf{x}_t = (x_t^{(0)}, \ldots, x_t^{(m)}) \in \overline{\mathbb{R}}^{m+1}\backslash\{\mathbf{0}\}$ and define the mappings

$$T_{h,k,\delta} : \mathcal{M} \to \overline{\mathbb{R}}$$

by

$$T_{-1,-1,\delta}(\sum_{t=1}^{\infty} n_t \delta_{\mathbf{x}_t}) = \sum_{t=1}^{\infty} n_t 1_{\{|x_t^{(0)}|>\delta\}},$$

$$T_{h,k,\delta}(\sum_{t=1}^{\infty} n_t \delta_{\mathbf{x}_t}) = \sum_{t=1}^{\infty} n_t x_t^{(h)} x_t^{(k)} 1_{\{|x_t^{(0)}|>\delta\}}, \qquad h,\ k \geqslant 0,$$

where $n_t \in \mathbb{N}_0$ for any $t \geqslant 1$. Since the set $\{\mathbf{x} \in \overline{\mathbb{R}}^{m+1} \backslash \{\mathbf{0}\} \mid |x^{(h)}| > \delta\}$ is bounded for any $h = 0, \ldots, m$ the mappings are a.s. continuous with respect to the limit point processes $N^{\mathbf{X}}$, $N^{|\mathbf{X}|}$ and $N^{\mathbf{X}^2}$. Consequently, by the continuous mapping theorem, we have in particular

$$T_{-1,-1,\delta}(N_n^{\mathbf{X}}) = \sum_{t=1}^{n} 1_{\{|X_t^{(0)}|>\delta\}} \xrightarrow{d} T_{-1,-1,\delta}(N^{\mathbf{X}}) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} 1_{\{|P_i Q_{ij}^{(0)}|>\delta\}} \tag{4.1}$$

and, for any $h,\ k \geqslant 0$,

$$T_{h,k,\delta}(N_n^{\mathbf{X}}) = \sum_{t=1}^{n} X_t^{(h)} X_t^{(k)} 1_{\{|X_t^{(0)}|>\delta\}} \xrightarrow{d} T_{h,k,\delta}(N^{\mathbf{X}}) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} P_i^2 Q_{ij}^{(h)} Q_{ij}^{(k)} 1_{\{|P_i Q_{ij}^{(0)}|>\delta\}}. \tag{4.2}$$

Note that, with obvious modifications, both (4.1) and (4.2) also hold for $N_n^{|\mathbf{X}|}$ and $N^{|\mathbf{X}|}$ or for $N_n^{\mathbf{X}^2}$ and $N^{\mathbf{X}^2}$. The following theorem collects the weak limit results of the sample ACVF and ACF of (X_t), $(|X_t|)$ and (X_t^2) depending on the tail index $\kappa > 0$. The weak limits turn out to be infinite-variance stable random vectors in parts 1 and 2 of Theorem 4.1. However, they are only functionals of point processes and have no explicit representation. Therefore, the results are only of a qualitative nature and explicit asymptotic confidence bounds for the sample ACVFs and ACFs cannot be constructed.

Theorem 4.1. *Assume* (X_t) *is the stationary* $ARARCH(1,1)$ *process satisfying the conditions of Theorem* 2.1 *with* $\mathrm{E}(\varepsilon^2) = 1$. *Let* $\kappa > 0$ *be the tail index in* (2.7) *and* (a_n) *be the sequence satisfying* (3.3) *for* $k = 0$. *Then the following statements hold:*

1. (*a*) *If* $\kappa \in (0, 2)$, *then*

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

$$(\rho_{n,X}(h))_{h=1,\ldots,m} \xrightarrow{d} \left(\frac{V_h^X}{V_0^X}\right)_{h=1,\ldots,m},$$

and

$$\left(n a_n^{-2} \gamma_{n,|X|}(h)\right)_{h=0,\ldots,m} \xrightarrow{d} \left(V_h^{|X|}\right)_{h=0,\ldots,m},$$

$$(\rho_{n,|X|}(h))_{h=1,\ldots,m} \xrightarrow{d} \left(\frac{V_h^{|X|}}{V_0^{|X|}}\right)_{h=1,\ldots,m},$$

where the vectors $\left(V_0^X, \ldots, V_m^X\right)$ *and* $\left(V_0^{|X|}, \ldots, V_m^{|X|}\right)$ *are jointly* $(\kappa/2)$*-stable in* $\mathbb{R}^{m+1}$ *with point process representation*

$$V_h^X = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} P_i^2 Q_{ij}^{(0)} Q_{ij}^{(h)}$$

and

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

respectively.

(*b*) *If* $\kappa \in (0, 4)$, *then*

$$(na_n^{-4}\gamma_{n,X^2}(h))_{h=0,\ldots,m} \xrightarrow{d} (V_h^{X^2})_{h=0,\ldots,m},$$

$$(\rho_{n,X^2}(h))_{h=1,\ldots,m} \xrightarrow{d} \left(\frac{V_h^{X^2}}{V_0^{X^2}}\right)_{h=1,\ldots,m},$$

where $(V_0^{X^2}, \ldots, V_m^{X^2})$ *is jointly* $(\kappa/4)$*-stable in* $\mathbb{R}^{m+1}$ *with point process representation*

$$V_h^{X^2} = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} P_i^4 (Q_{ij}^{(0)} Q_{ij}^{(h)})^2, \qquad h = 0, \ldots, m.$$

2. (*a*) *If* $\kappa \in (2, 4)$ *and* $\mathrm{E}(\varepsilon^4) < \infty$, *then*

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

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

and

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

$$(na_n^{-2}(\rho_{n,|X|}(h) - \rho_{|X|}(h)))_{h=1,\ldots,m} \xrightarrow{d} \gamma_{|X|}^{-1}(0)(V_h^{|X|} - \rho_{|X|}(h)V_0^{|X|})_{h=1,\ldots,m},$$

where the vectors $(V_0^X, \ldots, V_m^X)$ *and* $(V_0^{|X|}, \ldots, V_m^{|X|})$ *are jointly* $(\kappa/2)$*-stable in* $\mathbb{R}^{m+1}$ *with*

$$V_0^X = \tilde{V}_0^X(1 - (\alpha^2 + \lambda))^{-1},$$

$$V_m^X = \tilde{V}_m^X + \alpha V_{m-1}^X, \qquad m \geqslant 1,$$

and

$$V_0^{|X|} = V_0^X,$$

$$V_m^{|X|} = \tilde{V}_m^{|X|} + \mathrm{E}(|\alpha + \sqrt{\lambda}\varepsilon|)V_{m-1}^{|X|}, \qquad m \geqslant 1.$$

Furthermore, $(\tilde V_0^X, \ldots, \tilde V_m^X)$ *and* $(\tilde V_0^{|X|}, \ldots, \tilde V_m^{|X|})$ *are the distributional limits of*

$$(T_{1,1,\delta}(N^{\mathbf{X}}) - (\alpha^2+\lambda)T_{0,0,\delta}(N^{\mathbf{X}}), (T_{0,h,\delta}(N^{\mathbf{X}}) - \alpha T_{0,h-1,\delta}(N^{\mathbf{X}}))_{h=1,\ldots,m})$$

and

$$(T_{1,1,\delta}(N^{|\mathbf{X}|}) - (\alpha^2+\lambda)T_{0,0,\delta}(N^{|\mathbf{X}|}), (T_{0,h,\delta}(N^{|\mathbf{X}|}) - \mathrm{E}(|\alpha+\sqrt{\lambda}\varepsilon|)T_{0,h-1,\delta}(N^{|\mathbf{X}|}))_{h=1,\ldots,m}),$$

respectively, as $\delta \to \infty$.

(*b*) *If* $\kappa \in (4, 8)$ *and* $\mathrm{E}(\varepsilon^8) < \infty$, *then*

$$\left(na_n^{-4}(\gamma_{n,X^2}(h) - \gamma_{X^2}(h))\right)_{h=0,\ldots,m} \xrightarrow{d} \left(V_h^{X^2}\right)_{h=0,\ldots,m},$$

$$\left(na_n^{-4}(\rho_{n,X^2}(h) - \rho_{X^2}(h))\right)_{h=1,\ldots,m} \xrightarrow{d} \gamma_{X^2}^{-1}(0)\left(V_h^{X^2} - \rho_X(h)V_0^{X^2}\right)_{h=1,\ldots,m},$$

where $\left(V_0^{X^2}, \ldots, V_m^{X^2}\right)$ *is jointly* $(\kappa/4)$*-stable in* $\mathbb{R}^{m+1}$ *with*

$$V_0^{X^2} = \tilde V_0^{X^2}\left(1 - \left(\alpha^4 + 6\alpha^2\lambda + \lambda^2\mathrm{E}(\varepsilon^4)\right)\right)^{-1},$$

$$V_m^{X^2} = \tilde V_m^{X^2} + (\alpha^2+\lambda)V_{m-1}^{X^2}, \qquad m \geqslant 1,$$

and $\left(\tilde V_0^{X^2}, \ldots, \tilde V_m^{X^2}\right)$ *is the distributional limit of*

$$\Big(T_{1,1,\delta}\left(N^{\mathbf{X}^2}\right) - \left(\alpha^4 + 6\alpha^2\lambda + \lambda^2\mathrm{E}(\varepsilon^4)\right)T_{0,0,\delta}\left(N^{\mathbf{X}^2}\right),$$
$$\left(T_{0,h,\delta}\left(N^{\mathbf{X}^2}\right) - (\alpha^2+\lambda)T_{0,h-1,\delta}\left(N^{\mathbf{X}^2}\right)\right)_{h=1,\ldots,m}\Big),$$

as $\delta \to 0$.

3. (*a*) *If* $\kappa \in (4, \infty)$, *then*

$$\left(n^{1/2}(\gamma_{n,X}(h) - \gamma_X(h))\right)_{h=0,\ldots,m} \xrightarrow{d} (G_h^X)_{h=0,\ldots,m},$$

$$\left(n^{1/2}(\rho_{n,X}(h) - \rho_X(h))\right)_{h=1,\ldots,m} \xrightarrow{d} \gamma_X^{-1}(0)\left(G_h^X - \rho_X(h)G_0^X\right)_{h=1,\ldots,m}$$

and

$$\left(n^{1/2}(\gamma_{n,|X|}(h) - \gamma_{|X|}(h))\right)_{h=0,\ldots,m} \xrightarrow{d} (G_h^{|X|})_{h=0,\ldots,m},$$

$$\left(n^{1/2}(\rho_{n,X}(h) - \rho_X(h))\right)_{h=1,\ldots,m} \xrightarrow{d} \gamma_X^{-1}(0)\left(G_h^{|X|} - \rho_X(h)G_0^{|X|}\right)_{h=1,\ldots,m},$$

where the limits are multivariate Gaussian with mean zero.

(*b*) *If* $\kappa \in (8, \infty)$, *then*

$$\left(n^{1/2}(\gamma_{n,X^2}(h) - \gamma_{X^2}(h))\right)_{h=0,\ldots,m} \xrightarrow{d} (G_h^{X^2})_{h=0,\ldots,m}$$

$$\left(n^{1/2}(\rho_{n,X^2}(h) - \rho_{X^2}(h))\right)_{h=1,\ldots,m} \xrightarrow{d} \gamma_{X^2}^{-1}(0)\left(G_h^{X^2} - \rho_{X^2}(h)G_0^{X^2}\right)_{h=1,\ldots,m},$$

where the limits are multivariate Gaussian with mean zero.

Remark 4.2. (a) Theorem 4.1 is a generalization of results for the ARCH(1) process. See Davis and Mikosch (1998), who use a different approach which does not extend to the general case because of the autoregressive part of (X_t). See Figure 1.

(b) The assumption $\sigma^2 := \mathrm{E}(\varepsilon^2) = 1$ in the theorem is not a restriction. In cases where the second moment is different from unity, consider the process $(\hat{X}_t)$ defined by the stochastic recurrence equation

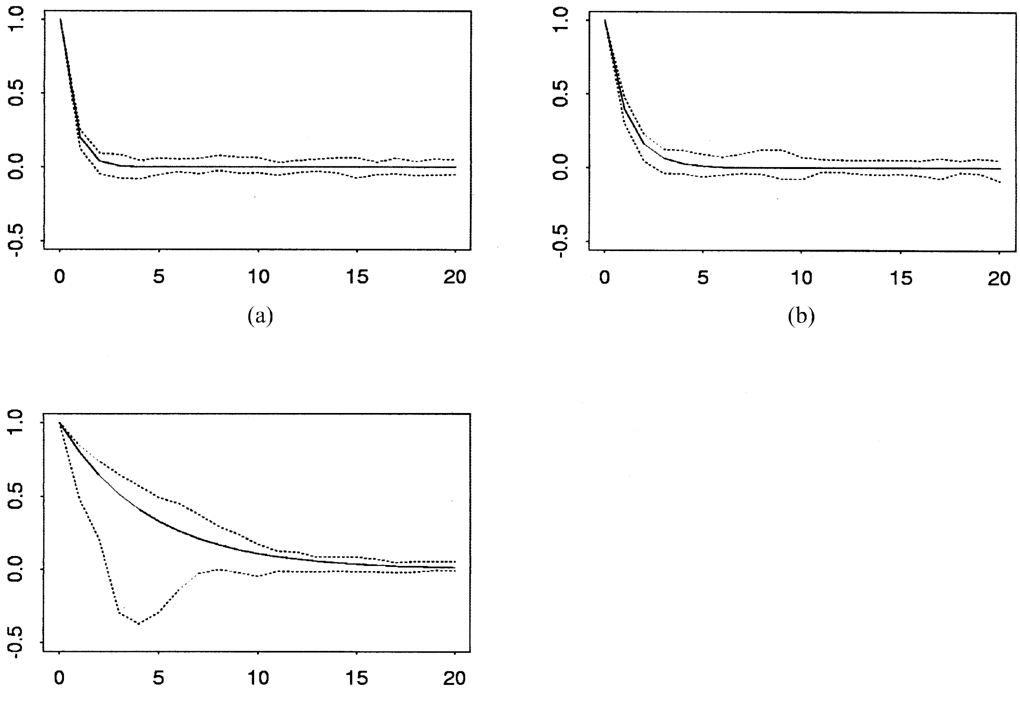

Figure 1. Limit behaviour of the sample ACFs of an ARARCH(1, 1) process with standard normal distributed innovations (ε_t) for three different parameter choices: (a) $\alpha = 0.2$, $\beta = 1$, $\lambda = 0.4$; (b) $\alpha = 0.4$, $\beta = 1$, $\lambda = 0.6$; (c) $\alpha = 0.8$, $\beta = 1$, $\lambda = 0.6$. In the first case $\kappa = 5.49$, in the second $\kappa = 2.87$ and in the last $\kappa = 1.35$. The solid lines denote the theoretical ACFs. The dotted lines indicate the 5% and 95% quantiles of the distributions of the sample ACFs at fixed lags. The underlying simulated sample paths have length 1000. The confidence bands were derived from 1000 independent simulations of the sample ACFs at these lags. The plots confirm the different limit behaviours of the sample ACFS as described in this paper.

$$\hat{X}_t = \alpha \hat{X}_{t-1} + \frac{\sqrt{\beta/\sigma^2 + \lambda \hat{X}_{t-1}^2}\,\varepsilon_t}{\sigma}, \qquad t \in \mathbb{N},$$

where the notation is the same as for the process (X_t) in (2.1). Note that $(\hat{X}_t) = (X_t/\sigma^2)$. Since the assumptions in the theorem do not depend on the parameter β, the results hold for $(\hat{X}_t)$ and hence they also hold for (X_t), replacing the limits $(V_h^X, V_h^{|X|}, V_h^{X^2})_{h=0,\ldots,m}$ with $\sigma^4(V_h^X, V_h^{|X|}, V_h^{X^2})_{h=0,\ldots,m}$ and $(G_h^X, G_h^{|X|}, G_h^{X^2})_{h=0,\ldots,m}$ with $\sigma^4(G_h^X, G_h^{|X|}, G_h^{X^2})_{h=0,\ldots,m}$, respectively.

(c) Note that the description of the distributional limits in part 2 of Theorem 4.1 is different than in Theorem 3.5 of Davis and Mikosch (1998). In the latter theorem the condition

$$\lim_{\delta \downarrow 0} \limsup_{n\to\infty} \operatorname{var}\left(a_n^{-2} \sum_{t=1}^{n-h} X_t X_{t+h} 1_{\{|X_t X_{t+h}| \leq a_n^2 \delta\}} \right) = 0$$

is required. However, this condition is in general very hard to check. Therefore, we choose another way and establish the convergence in distribution of the sample ACVF directly from the point process convergence in Theorem 3.7.

Proof. Statements 1(a) and 1(b) are immediate consequences of Theorem 3.5(1) of Davis and Mikosch (1998). Note that all conditions in this theorem are fulfilled because of Proposition 3.1 and Theorem 3.7. Statements 3(a) and 3(b) for the sample ACVFs follows from standard limit theorems for strongly mixing sequences (see Ibragimov and Linnik 1971, Chapter 18). The limit behaviour for the ACFs can be shown in the same way as, for example, in Davis and Mikosch (1998, p. 2062).

It remains to show 2(a) and 2(b). We restrict ourselves to the case $(|X_t|)$ and only establish joint convergence of $(\gamma_{n,|X|}(0), \gamma_{n,|X|}(1))$. All other cases can be treated similarly, or even more easily. Recall that, under $P_{|X_0|,\mathrm{sign}(X_0)}$, $(\tilde{X}_t) \stackrel{d}{=} (|X_t|)$, where the process $(\tilde{X}_t)$ is defined in (2.10). Thus, it is sufficient to study the sample ACVF of the process $(\tilde{X}_t)$.

We start by rewriting $\gamma_{n,\tilde{X}}(0)$ using the recurrence structure of $(\tilde{X}_t)$:

$$\begin{aligned} n a_n^{-2}(\gamma_{n,\tilde{X}}(0) - \gamma_{\tilde{X}}(0)) &= a_n^{-2} \sum_{t=1}^{n} \left(\tilde{X}_{t+1}^2 - \mathrm{E}(\tilde{X}^2) \right) \\ &= (\alpha^2 + \lambda) a_n^{-2} \sum_{t=1}^{n} \left(\tilde{X}_t^2 - \mathrm{E}(\tilde{X}^2) \right) \\ &\quad + a_n^{-2} \sum_{t=1}^{n} \left(2\alpha \tilde{X}_t \sqrt{\beta + \lambda \tilde{X}_t^2}\,\varepsilon_{t+1} + (\beta + \lambda \tilde{X}_t^2)(\varepsilon_{t+1}^2 - 1) \right). \end{aligned}$$

We conclude that, for any $\delta > 0$,

$$
\begin{aligned}
&(1-(\alpha^2+\lambda))na_n^{-2}(\gamma_{n,\tilde{X}}(0)-\gamma_{\tilde{X}}(0))\\
&= a_n^{-2}\sum_{t=1}^{n}(\beta+\lambda\tilde{X}_t^2)(\varepsilon_{t+1}^2-1)1_{\{\tilde{X}_t\leqslant a_n\delta\}} + 2\alpha a_n^{-2}\sum_{t=1}^{n}\tilde{X}_t\sqrt{\beta+\lambda\tilde{X}_t^2}\varepsilon_{t+1}1_{\{\tilde{X}_t\leqslant a_n\delta\}}\\
&\quad + a_n^{-2}\sum_{t=1}^{n}\left(\tilde{X}_t\sqrt{\beta+\lambda\tilde{X}_t^2}\varepsilon_{t+1}+(\beta+\lambda\tilde{X}_t^2)(\varepsilon_{t+1}^2-1)\right)1_{\{\tilde{X}_t>a_n\delta\}}+o_P(1)\\
&=: (I_1)+(I_2)+(I_3)+o_P(1).
\end{aligned}
$$

We show first that (I_1) and (I_2) converge in probability to zero. Note that the summands in (I_1) are uncorrelated. Therefore,

$$
\begin{aligned}
\operatorname{var}(I_1) &= a_n^{-4}\sum_{t=1}^{n}\operatorname{var}\left((\beta+\lambda\tilde{X}_t^2)1_{\{|\tilde{X}_t|\leqslant a_n\delta\}}(\varepsilon_{t+1}^2-1)\right)\\
&\leqslant a_n^{-4}\sum_{t=1}^{n}\mathrm{E}\left((\beta+\lambda\tilde{X}_t^2)^2 1_{\{|\tilde{X}_t|\leqslant a_n\delta\}}\right)\mathrm{E}\left((\varepsilon_{t+1}^2-1)^2\right)\\
&\sim \text{const.}\,\delta^{4-\kappa}, \qquad \text{as } n\to\infty,\\
&\to 0, \qquad \text{as } \delta\downarrow 0,
\end{aligned}
$$

where the asymptotic equivalence comes from Karamata's theorem on regular variation and the tail behaviour of the stationary distribution of $(\tilde{X}_t)$. Note that the condition $\mathrm{E}(\varepsilon^4)<\infty$ is crucial. Analogously, one can show that

$$
\lim_{\delta\downarrow 0}\lim_{n\to\infty}\operatorname{var}(I_2)=0.
$$

Next we consider (I_3). From (2.10) we obtain

$$
\begin{aligned}
(I_3) &= a_n^{-2}\sum_{t=1}^{n}\tilde{X}_{t+1}^2 1_{\{\tilde{X}_t>a_n\delta\}} - (\alpha^2+\lambda)a_n^{-2}\sum_{t=1}^{n}\tilde{X}_t^2 1_{\{\tilde{X}_t>a_n\delta\}} - \beta a_n^{-2}\sum_{t=1}^{n}1_{\{\tilde{X}_t>a_n\delta\}}\\
&\stackrel{d}{=} T_{1,1,\delta}(N_n^{|\mathbf{X}|}) - (\alpha^2+\lambda)T_{0,0,\delta}(N_n^{|\mathbf{X}|}) - \beta a_n^{-2}T_{-1,-1,\delta}(N_n^{|\mathbf{X}|})\\
&\stackrel{d}{\to} T_{1,1,\delta}(N^{|\mathbf{X}|}) - (\alpha^2+\lambda)T_{0,0,\delta}(N^{|\mathbf{X}|}),
\end{aligned} \tag{4.3}
$$

where the limit has expectation zero. Finally, following the same arguments as in Davis and Hsing (1995, pp. 897–898), the right-hand side in (4.3) converges in distribution to a $(\kappa/2)$-stable random variable, as $\delta\to 0$.

Now consider $\gamma_{n,\tilde{X}}(1)$. We proceed as above and write

$$
\begin{aligned}
na_n^{-2}\big(\gamma_{n,\tilde{X}}(1)-\gamma_{\tilde{X}}(1)\big) &= a_n^{-2}\sum_{t=1}^{n-1}\tilde{X}_t\tilde{X}_{t+1}-\mathrm{E}(\tilde{X}_0\tilde{X}_1)\\
&= a_n^{-2}\sum_{t=1}^{n-1}\big(f_{\varepsilon_{t+1}}(\tilde{X}_t)-\mathrm{E}(f_{\varepsilon_{t+1}}(\tilde{X}_t))\big)\\
&\quad + a_n^{-2}\sum_{t=1}^{n-1}\Big(\tilde{X}_t^2|\alpha+\sqrt{\lambda}\varepsilon_{t+1}|-\mathrm{E}(\tilde{X}^2)\mathrm{E}(|\alpha+\sqrt{\lambda}\varepsilon|)\Big)\\
&=: (J_1)+(J_2),
\end{aligned}
$$

where $f_z(y)=y(|\alpha y+\sqrt{\beta+\lambda y^2}z|-|\alpha y+\sqrt{\lambda}yz|)$ for any $y\geqslant 0$ and $z\in\mathbb{R}$. First, we show that (J_1) converges in probability to zero. Observe for that purpose that

$$
\begin{aligned}
\operatorname{var}\left(\left|a_n^{-2}\sum_{t=1}^{n-1}f_{\varepsilon_{t+1}}(\tilde{X}_t)-\mathrm{E}(f_{\varepsilon_{t+1}}(\tilde{X}_t))\right|\right) &\leqslant \operatorname{var}\left(a_n^{-2}\sum_{t=1}^{n}|f_{\varepsilon_{t+1}}(\tilde{X}_t)|\right) \qquad (4.4)\\
&= a_n^{-4}\sum_{t=1}^{n}\sum_{s=1}^{n}\operatorname{cov}(|f_{\varepsilon_{t+1}}(\tilde{X}_t)|,\,|f_{\varepsilon_{s+1}}(\tilde{X}_s)|).
\end{aligned}
$$

Now note that $|f_z(y)|\leqslant y\sqrt{\beta}|z|$ for any $y\geqslant 0$ and $z\in\mathbb{R}$. Therefore, and since $\kappa>2$, there exists a $\mu>0$ such that

$$
\mathrm{E}(|f_\varepsilon(\tilde{X})|^{2+\mu})\leqslant\sqrt{\beta}\mathrm{E}(|\varepsilon|^{2+\mu})\mathrm{E}(|\tilde{X}|^{2+\mu})<\infty. \qquad (4.5)
$$

Because of (4.5) and the geometric strong mixing property of $(\tilde{X}_t)$, all assumptions of Theorem 17.2.2 of Ibragimov and Linnik (1971) are satisfied and we can bound (4.4) by

$$
\text{const.}\ a_n^{-4}n\sum_{s=0}^{n-1}(\rho^{\mu/(2+\mu)})^s, \qquad (4.6)
$$

which converges to zero as $n\to\infty$ since $\kappa<4$. Next we rewrite (J_2) and obtain

$$
\begin{aligned}
(J_2) &= \mathrm{E}(|\alpha+\sqrt{\lambda}\varepsilon|)na_n^{-2}\big(\gamma_{n,\tilde{X}}(0)-\gamma_{\tilde{X}}(0)\big)\\
&\quad + a_n^{-2}\sum_{t+1}^{n-1}\tilde{X}_t^2 1_{\{\tilde{X}_t\leqslant a_n\delta\}}\Big(|\alpha+\sqrt{\lambda}\varepsilon_{t+1}|-\mathrm{E}(|\alpha+\sqrt{\lambda}\varepsilon|\Big)\\
&\quad + a_n^{-2}\sum_{t=1}^{n-1}\tilde{X}_t^2 1_{\{\tilde{X}_t>a_n\delta\}}\Big(|\alpha+\sqrt{\lambda}\varepsilon_{t+1}|-\mathrm{E}(|\alpha+\sqrt{\lambda}\varepsilon|)\Big)\\
&= (K_1)+(K_2)+(K_3).
\end{aligned}
$$

By (4.3), $K_1)\xrightarrow{d}T_{1,1,\delta}(N^{|\mathbf{X}|})-(\alpha^2+\lambda)T_{0,0,\delta}(N^{|\mathbf{X}|})$. Moreover, using the same arguments as before one can show that $\lim_{\delta\downarrow 0}\lim_{n\to\infty}\operatorname{var}(K_2)=0$. Hence $(K_2)=o_P(1)$. It remains to consider (K_3). We begin with the decomposition

$$(K_3) = a_n^{-2}\sum_{t=1}^{n-1}\tilde{X}_t 1_{\{\tilde{X}_t > a_n\delta\}}\left(|\alpha\tilde{X}_t + \sqrt{\lambda}\tilde{X}_t\varepsilon_{t+1}| - |\alpha\tilde{X}_t + \sqrt{\beta + \lambda\tilde{X}_t^2}\varepsilon_{t+1}|\right)$$

$$+ a_n^{-2}\sum_{t=1}^{n-1}\tilde{X}_{t+1}\tilde{X}_t 1_{\{\tilde{X}_t > a_n\delta\}} - a_n^{-2}\sum_{t=1}^{n-1}\tilde{X}_t^2 1_{\{\tilde{X}_t > a_n\delta\}}\mathrm{E}(|\alpha + \sqrt{\lambda}\varepsilon|).$$

Proceeding the same way as in (4.4)–(4.6) the first term converges in probability to zero. Thus,

$$(K_3) \stackrel{d}{=} o_P(1) + T_{0,1,\delta}(N_n^{|\mathbf{X}|}) - \mathrm{E}(|\alpha + \sqrt{\lambda}\varepsilon|)T_{0,0,\delta}(N_n^{|\mathbf{X}|})$$

$$\stackrel{d}{\to} T_{0,1,\delta}(N^{|\mathbf{X}|}) - \mathrm{E}(|\alpha + \sqrt{\lambda}\varepsilon|)T_{0,0,\delta}(N^{|\mathbf{X}|}),$$

where the limit has zero mean and converges again to a $(\kappa/2)$-stable random variable as $\delta \downarrow 0$. Since for the distributional convergence only the point process convergence and the continuous mapping theorem have been used, it is immediate that the same kind of argument yields the joint convergence of the sample autocovariances to a $(\kappa/2)$-stable limit as described in the statement. Finally, the asymptotic behaviour of the sample ACF can be shown in the same way as in Davis and Mikosch (1998, p. 2062). □

Acknowledgement

I gratefully acknowledge the hospitality of the school of Operations Research and Industrial Engineering of Cornell University where the bulk of this work took place. The research was supported by a Swiss National Science Foundation grant. Furthermore, I am very grateful to Thomas Mikosch and Claudia Klüppelberg for useful comments. Last, but not least, I am particularly grateful to one referee for helpful suggestions which improved the paper considerably.

References

Basrak, B., Davis, R.A. and Mikosch, T. (1999) The sample ACF of a simple bilinear process. *Stochastic Process. Appl.*, **83**, 1–14.

Borkovec, M. (2000) Extremal behavior of the autoregressive process with ARCH(1) errors. *Stochastic Process. Appl.*, **85**, 189–207.

Borkovec, M. and Klüppelberg, C. (2001) The tail of the stationary distribution of an autoregressive process with ARCH(1) errors. *Ann. Appl. Probab.* To appear.

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

Brockwell, P.J. and Davis, R.A. (1991) *Time Series: Theory and Methods*. New York: Springer-Verlag.

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 Mikosch, T. (1998) The sample autocorrelations of heavy-tailed processes with application to ARCH. *Ann. Statist.*, **26**, 2049–2080.

Davis, R.A. and Resnick, S.I. (1985) More limit theory for the sample correlation function of moving averages. *Stochastic Process. Appl.*, **20**, 257–279.

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

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

Davis, R.A., Mikosch, T. and Basrak, B. (1999) The sample ACF of multivariate stochastic recurrence equations with applications to GARCH. Technical Report.

Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance*. Berlin: Springer-Verlag.

Engle, R.F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* **50**, 987–1008.

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

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

Kallenberg, O. (1983) *Random Measures*, 3rd edition. Berlin: Akademie-Verlag.

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

Mikosch, T. and Stărică, C. (2000) Limit theory for the sample autocorrelations and extremes of a GARCH(1, 1) process, to appear in: *Ann. Statist.*, **28**, 1427–1451.

Resnick, S.I. (1997) Heavy tail modeling and teletraffic data. With discussion and a rejoinder by the author. *Ann. Statist.*, **25**, 1805–1869.

Resnick, S.I. and van den Berg, E. (2000) A test for nonlinearity of time series with infinite variance. *Extremes*, **3**, 145–172.

Weiss, A.A. (1984) ARMA models with ARCH errors. *J. Time Ser. Anal.*, **3**, 129–143.

Received February 2000 and revised July 2001