Extremal behavior of the autoregressive process with ARCH(1) errors

Milan Borkovec*

Abstract

We investigate the extremal behavior of a special class of autoregressive processes with ARCH(1) errors given by the stochastic difference equation

$$X_n = \alpha X_{n-1} + \sqrt{\beta + \lambda X_{n-1}^2}\,\varepsilon_n\,, \quad n \in \mathbb{N},$$

where $(\varepsilon_n)_{n\in\mathbb{N}}$ are i.i.d. random variables. The extremes of such processes occur typically in clusters. We give an explicit formula for the extremal index and the probabilities for the length of a cluster.

AMS 1991 Subject Classifications: primary: 60G70, 60J05

secondary: 60F05, 60G55

Keywords: ARCH model, autoregressive process, compound Poisson process, coupling, extremal behavior, extremal index, Fréchet distribution, heavy tail, heteroskedastic homogeneous Markov process, recurrent Harris chain, separating sequence, strong mixing

*Center of Mathematical Sciences, Munich University of Technology, D-80290 Munich, Germany,
email: borkovec@mathematik.tu-muenchen.de, http://www-m4.mathematik.tu-muenchen.de/m4

1 Introduction

Random recurrence equations have been used in numerous fields of applied probability. We refer for instance to Kesten (1973), Vervaat (1979) and Embrechts and Goldie (1994). Stochastic models in finance are an important field of application for random recurrence equations. Over the last years a variety of these models have been suggested as appropriate models for financial time series (see e.g. Priestley (1988), Tong (1990), Taylor (1995)). Due to the random recurrence structure, many of these models possess the property that their conditional variance depends on the past information (conditional heteroskedasticity). Empirical work has confirmed that such models fit quite many types of financial data. The most known examples of volatility models in finance with random recurrence structure are autoregressive conditionally heteroskedastic processes (ARCH). These models were introduced by Engle (1982). They serve as special exchange rate or asset price models and are very popular in econometrics. In a series of papers, the ARCH models have been analyzed and generalized, see for instance the survey article by Bollerslev, Chou and Kroner (1992) and the statistical review paper by Shephard (1996).

The class of autoregressive (AR) models with ARCH errors proposed by Weiss (1984) are a natural extension of ARCH processes. These models are also called SETAR-ARCH models (self-exciting autoregressive). They are defined by the random recurrence equation

$$X_n = f(X_{n-1}, ..., X_{n-k}) + \sigma_n \varepsilon_n, \quad n \geq k\,, \tag{1.1}$$

where f is a linear function in its arguments, the innovations $(\varepsilon_n)_{n\in\mathbb{N}}$ are i.i.d. symmetric random variables with mean zero and σ_n is given by

$$\sigma_n^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j X_{n-j}^2\,, \quad \alpha_0 > 0,\, \alpha_1, ..., \alpha_p \geq,\, \alpha_p > 0, \tag{1.2}$$

for some $p \geq 1$. These models combine the advantages of AR models which target more on the conditional mean of X_n given the past and ARCH models which concentrate on the conditional variance of X_n (given the past). Autoregressive models with ARCH errors capture the structure of financial data quite well, i.e. the tendency of volatility clustering and the fact that unconditional price and return distributions tend to have fatter tails than the normal distribution. Statistical and/or probabilistic properties of such models have been investigated by Weiss (1984), Diebolt and Guégan (1990), Maercker (1997) and Borkovec and Klüppelberg (1998).

In the present paper we study the extremal behavior of AR processes with ARCH errors. We focus on the AR(1) process with ARCH(1) errors, i.e. $f(X_{n-1}, ..., X_{n-k}) = \alpha X_{n-1}$ for some

$\alpha \in \mathbb{R}$ and σ_n is given in (1.2) with $p = 1$. This Markovian model is analytically tractable and serves as a prototype for the larger class of models (1.1). Furthermore, in the special case $\alpha = 0$ we get just the ARCH(1) model of Engle (1982) and hence our results for the extremes will be an extension of the results in de Haan, Resnick, Rootzén and de Vries (1989).

Extremal behavior of a Markov process $(X_n)_{n\in\mathbb{N}}$ is for instance manifested in the asymptotic behavior of the maxima

$$M_n = \max_{1\le k\le n} X_k\,, \quad n \ge 1\,.$$

The limit behavior of M_n is a well-studied problem in extreme value theory. Two review paper on this and related problems are Rootzén (1988) and Perfekt (1994). For a general overview of extremes of Markov processes, see also Leadbetter, Lindgren and Rootzén (1983) and the references therein. Loosly speaking, under quite general mixing conditions, one can show that for n and x large

$$P(M_n \le x) \approx F^{n\theta}(x)\,, \tag{1.3}$$

where F is the stationary distribution function of $(X_n)_{n\in\mathbb{N}}$ and $\theta \in (0,1)$ is a constant called *extremal index.* A natural interpretation of θ is that of the reciprocal of mean cluster size (see e.g. Embrechts, Klüppelberg and Mikosch (1997, Chapter 6) and the references therein). The practical implication of (1.3) is that dependence in data does often not invalidate the application of classical extreme value theory. There are many methods for determing the extremal index. However, most are very technical and often useless in practice. An alternative is then to estimate θ from the data.

For the AR(1) process with ARCH(1) errors we derive an explicit formula for the extremal index. We furthermore investigate the point process of exceedances of a high threshold u of $(X_n)_{n\in\mathbb{N}}$ which characterizes the extremal behavior of the process in detail. This point process converges in distribution to a compound Poisson process with a well-specified intensity and a well-specified distribution of the size of the jumps.

The paper is organized as follows: in section 2 we present the model and introduce the required assumptions on the innovations $(\varepsilon_n)_{n\in\mathbb{N}}$. The conditions are the same as in Borkovec and Klüppelberg (1998), namely the so-called general conditions and the technical conditions $(D.1)-(D.3)$. The general conditions guarantee the existence of a stationary version of $(X_n)_{n\in\mathbb{N}}$ whereas $(D.1)-(D.3)$ allow us to describe the tail behavior of the stationary distribution. We present furthermore some results on the AR(1) process with ARCH(1) errors $(X_n)_{n\in\mathbb{N}}$ and on the related process $(Z_n)_{n\in\mathbb{N}} = (\ln(X_n^2))_{n\in\mathbb{N}}$. It turns out that the process $(Z_n)_{n\in\mathbb{N}}$ is crucial for the study of the extremal behavior of $(X_n)_{n\in\mathbb{N}}$. We show in Lemma 2.3 that $(Z_n)_{n\in\mathbb{N}}$ behaves above a high threshold asymptotically as a random walk with negative drift which can be completely specified. Theorem 2.1 collects some known results on the AR(1) process with ARCH(1) errors $(X_n)_{n\in\mathbb{N}}$ which were proved in Borkovec and Klüppelberg (1998). In particular, the stationary distribution of $(X_n)_{n\in\mathbb{N}}$ has a Pareto-like tail. Section 3 contains the main results (Theorem 3.1) concerning the extremal behavior of $(X_n)_{n\in\mathbb{N}}$. We interprete these results and present some simulations. We conclude the paper in section 4 with the proof of Theorem 3.1.

2 Preliminaries

We consider an autoregressive model of order 1 with autoregressive conditional heteroskedastic errors of order 1 (AR(1) model with ARCH(1) errors) which is defined by the stochastic difference equation

$$X_n = \alpha X_{n-1} + \sqrt{\beta + \lambda X_{n-1}^2}\varepsilon_n\,, \quad n \in \mathbb{N}, \tag{2.1}$$

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

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

This condition is required to guarantee the existence and uniqueness of a stationary distribution. Let ε be a generic random variable with the same distribution as ε_n. Throughout this paper, we assume the same conditions for ε as in Borkovec and Klüppelberg (1998). These are the so-called *general conditions*:

$$\begin{aligned} \varepsilon \text{ is symmetric with continuous Lebesgue density } p(x)\,, \\ \varepsilon \text{ has full support } \mathbb{R}\,, \\ \text{the second moment of } \varepsilon \text{ exists}\,. \end{aligned} \tag{2.3}$$

and the *technical conditions* $(D.1) - (D.3)$:

(D.1) $p(x) \geq p(x')$ for any $0 \leq x < x'$.

(D.2) For any $c \geq 0$ there exists a constant $q = q(c) \in (0,1)$ and functions $f_+(c,\cdot)$, $f_-(c,\cdot)$ with $f_+(c,x), f_-(c,x) \to 1$ as $x \to \infty$ such that for any $x > 0$ and $t > x^q$

$$p(\frac{x+c+\alpha t}{\sqrt{\beta+\lambda t^2}}) \geq p(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}})\, f_+(c,x)\,,$$
$$p(\frac{x+c-\alpha t}{\sqrt{\beta+\lambda t^2}}) \geq p(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}})\, f_-(c,x)\,.$$

(D.3) There exists a constant $\eta > 0$ such that

$$p(x) = o(x^{-(N+1+\eta+3q)/(1-q)})\,, \quad \text{as } x \to \infty\,,$$

where $N := \inf\{u \geq 0\,;\, E(|\sqrt{\lambda}\varepsilon|^u) > 2\}$ and q is the constant in $(D.2)$.

There exists a wide class of distributions which satisfy these assumptions. Examples are the normal distribution, the Laplace distribution or the Students distribution. Conditions $(D.1)-(D.3)$ are necessary for determing the tail of the stationary distribution. For further details concerning the conditions and examples we refer to Borkovec and Klüppelberg (1998). Note that the process $(X_n)_{n\in\mathbb{N}}$ is evidently a homogeneous Markov chain with state space $\mathbb{R}$ equipped with the Borel σ-algebra. The transition kernel density is given by

$$P(X_1 \in dy \,|\, X_0 = x) = \frac{1}{\sqrt{\beta+\lambda x^2}}\, p(\frac{y-\alpha x}{\sqrt{\beta+\lambda x^2}})dy\,, \tag{2.4}$$

The next theorem collects some results on $(X_n)_{n\in\mathbb{N}}$ from Borkovec and Klüppelberg (1998).

Theorem 2.1 *Consider the process $(X_n)_{n\in\mathbb{N}}$ in (2.1) with $(\varepsilon_n)_{n\in\mathbb{N}}$ satisfying the general conditions (2.3) and with parameters α and λ satisfying (2.2). Then the following assertions hold:*

(a) Let ν be the normalized Lebesgue-measure $\nu(\cdot) := \lambda(\cdot \cap [-M,M])/\lambda([-M,M])$. Then $(X_n)_{n\in\mathbb{N}}$ is an aperiodic positive ν-recurrent Harris chain with regeneration set $[-M,M]$ for M large enough. In particular, there exists a constant $C \in (0,1)$ such that for any Borel-measurable set B and $x \in [-M,M]$

$$P(X_1 \in B \,|\, X_0 = x) \geq C\,\nu(B)\,. \tag{2.5}$$

(b) $(X_n)_{n\in\mathbb{N}}$ is geometric ergodic. In particular, $(X_n)_{n\in\mathbb{N}}$ has a unique stationary distribution and satisfies the strong mixing condition with geometric rate of convergence. The stationary df is continuous and symmetric.

(c) Let $\overline{F}(x) = P(X > x)$, $x \geq 0$, be the right tail of the stationary df and the conditions $(D.1) - (D.3)$ are in addition fulfilled. Then

$$\overline{F}(x) \sim c x^{-\kappa}, \quad x \to \infty, \tag{2.6}$$

where

$$c = \frac{1}{2\kappa} \frac{E\left(\left|\alpha|X| + \sqrt{\beta + \lambda X^2}\varepsilon\right|^{\kappa} - \left|(\alpha + \sqrt{\lambda}\varepsilon)|X|\right|^{\kappa}\right)}{E\left(|\alpha + \sqrt{\lambda}\varepsilon|^{\kappa} \ln|\alpha + \sqrt{\lambda}\varepsilon|\right)} \tag{2.7}$$

and κ is given as the unique positive solution to

$$E(|\alpha + \sqrt{\lambda}\varepsilon|^{\kappa}) = 1. \tag{2.8}$$

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

Remark 2.2 (a) Note that $E(|\alpha + \sqrt{\lambda}\varepsilon|^{\kappa})$ is a function of κ, α and λ. It can be shown that for $\varepsilon \sim N(0,1)$ and fixed λ, the exponent κ is decreasing in $|\alpha|$. This means that the distribution of X gets heavier tails when $|\alpha|$ increases. In particular, the AR(1) process with ARCH(1) errors has for $\alpha \neq 0$ heavier tails than the ARCH(1) process (see also Table 3 in Borkovec and Klüppelberg (1998)).

(b) Theorem 2.1 is crucial for investigating the extremal behavior of $(X_n)_{n\in\mathbb{N}}$. The strong mixing property includes automatically that the sequence $(X_n)_{n\in\mathbb{N}}$ satisfies the conditions $D(u_n)$ and $\Delta(u_n)$. The condition $D(u_n)$ is a frequently used mixing condition due to Leadbetter et al. (1983) whereas the slightly stronger condition $\Delta(u_n)$ was introduced by Hsing (1984). Loosely speaking, $D(u_n)$ and $\Delta(u_n)$ give the "degree of independence" of extremes situated far apart from each other. This property together with (2.6) implies that the maximum of the process $(X_n)_{n\in\mathbb{N}}$ belongs to the domain of attraction of a Fréchet distribution. We will specify the normalizing constants of the maxima and the limit distribution in section 3. $\square$

In order to study the extremal behavior of $(X_n)_{n\in\mathbb{N}}$ and $(X_n^2)_{n\in\mathbb{N}}$ we define the auxiliary process $(Z_n)_{n\in\mathbb{N}} := (\ln(X_n^2))_{n\in\mathbb{N}}$ which is again a regenerative, strongly mixing process. Since $(X_n)_{n\in\mathbb{N}}$ follows (2.1) the process $(Z_n)_{n\in\mathbb{N}}$ satisfies the stochastic difference equation

$$Z_n = Z_{n-1} + \ln\Big((\alpha + \sqrt{\beta\, e^{-Z_{n-1}} + \lambda}\, \varepsilon_n)^2\Big), \quad n \in \mathbb{N}, \tag{2.9}$$

where $(\varepsilon_n)_{n\in\mathbb{N}}$ are i.i.d. random variables that satisfy the general conditions and $(D.1) - (D.3)$, the constants are the same as in our old process $(X_n)_{n\in\mathbb{N}}$ and Z_0 equals $\ln(X_0^2)$ a.s.. Note that the process $(Z_n)_{n\in\mathbb{N}}$ is independent of the sign of the parameter α since ε_n is symmetric. Hence we may w.l.o.g. in the following assume that $\alpha \geq 0$. We will see that $(Z_n)_{n\in\mathbb{N}}$ can be bounded by two random walks $(S_n^{l,a})_{n\in\mathbb{N}}$ and $(S_n^{u,a})_{n\in\mathbb{N}}$ from below and above, respectively. This result is essential for the study of the extremal behavior of $(X_n)_{n\in\mathbb{N}}$. Via results for $(Z_n)_{n\in\mathbb{N}}$, we prove for instance that the regenerative process $(X_n)_{n\in\mathbb{N}}$ has finite mean recurrence times which allow us to consider only the extremal behavior of the stationary process $(X_n)_{n\in\mathbb{N}}$. The process $(Z_n)_{n\in\mathbb{N}}$ will be also important in the proof of Lemma 4.1. For the construction of the two random walks $(S_n^{l,a})_{n\in\mathbb{N}}$ and $(S_n^{u,a})_{n\in\mathbb{N}}$ we need some more definitions. With the same notation as before, let

$$A_a := \{\omega \,|\, \frac{-\alpha}{\sqrt{\beta\, e^{-a}+\lambda}-\sqrt{\beta}\, e^{-a/2}} \leq \varepsilon(\omega) \leq \frac{-\alpha}{\sqrt{\beta\, e^{-a}+\lambda}+\sqrt{\beta}\, e^{-a/2}}\}\,, \tag{2.10}$$

$$p(a,\alpha,\beta,\lambda,\varepsilon) := \ln\Big((\alpha+\sqrt{\beta\, e^{-a}+\lambda}\,\varepsilon)^2\Big)\,,$$

$$q(a,\alpha,\beta,\lambda,\varepsilon) := \ln\Big(1-\frac{2\alpha\sqrt{\beta}e^{-a/2}\varepsilon}{(\alpha+\sqrt{\beta\, e^{-a}+\lambda}\,\varepsilon)^2}1_{\{\varepsilon<0\}}\Big)\,, \tag{2.11}$$

$$r(a,\alpha,\beta,\lambda,\varepsilon) := \ln\Big(1-\frac{\beta\varepsilon^2 e^{-a}}{(\alpha+\sqrt{\beta\, e^{-a}+\lambda}\,\varepsilon)^2}1_{\{\varepsilon<0\}}\Big)\,.$$

Note that $q(a,\alpha,\beta,\lambda,\varepsilon), r(a,\alpha,\beta,\lambda,\varepsilon) \to 0$ a.s. for $a\to\infty$. Now define

$$S_n^{l,a} := \sum_{j=1}^n U_j^a \quad \text{and} \quad S_n^{u,a} := \sum_{j=1}^n V_j^a, \quad n\in\mathbb{N}, \tag{2.12}$$

where

$$\begin{aligned} U_j^a \;\; := \;\; -\infty \cdot 1_{A_a} &+ \Big(p(a,\alpha,\beta,\lambda,\varepsilon_j)+r(a,\alpha,\beta,\lambda,\varepsilon_j)\Big)\cdot 1_{A_a^c\cap\{\varepsilon_j<0\}} \\ &+ \ln(\alpha+\sqrt{\lambda}\varepsilon_j)^2 \cdot 1_{\{\varepsilon_j\geq 0\}} \end{aligned} \tag{2.13}$$

and

$$V_j^a := p(a,\alpha,\beta,\lambda,\varepsilon_j)+q(a,\alpha,\beta,\lambda,\varepsilon_j) \tag{2.14}$$

for some $a\geq 0$. The following lemma shows that the random walks defined in (2.12)-(2.14) are really upper and lower bounds for $(Z_n)_{n\in\mathbb{N}}$ above a high level.

Lemma 2.3 *Let a be large enough, $N_a := \inf\{j\geq 1 \,|\, Z_j\leq a\}$ and $Z_0>a$. Then*

$$Z_0+S_k^{l,a} \leq Z_k \leq Z_0+S_k^{u,a} \quad \textit{for any } k\leq N_a \textit{ a.s.} \tag{2.15}$$

Proof. We prove only the lower bound. The proof of the upper bound is similar but easier. Let $x \geq a$ be arbitrary. If $\varepsilon \geq 0$ it is obvious that

$$(\alpha + \sqrt{\beta\, e^{-x} + \lambda}\,\varepsilon)^2 \geq (\alpha + \sqrt{\lambda}\,\varepsilon)^2\,. \tag{2.16}$$

Consider now $\varepsilon < 0$, then

$$\begin{aligned}
&(\alpha + \sqrt{\beta\, e^{-x} + \lambda}\,\varepsilon)^2 - (\alpha + \sqrt{\beta\, e^{-a} + \lambda}\,\varepsilon)^2 \\
&= \; 2\alpha(-\varepsilon)\Big(\sqrt{\beta\, e^{-a} + \lambda} - \sqrt{\beta\, e^{-x} + \lambda}\Big) - \beta(e^{-a} - e^{-x})\,\varepsilon^2 \\
&\geq \; -\beta\, e^{-a}\,\varepsilon^2\,.
\end{aligned} \tag{2.17}$$

Note that we have a non-trivial lower bound of $(\alpha + \sqrt{\beta\, e^{-x} + \lambda}\,\varepsilon)^2$ if and only if

$$(\alpha + \sqrt{\beta\, e^{-a} + \lambda}\,\varepsilon)^2 - \beta\, e^{-a}\,\varepsilon^2 > 0\,. \tag{2.18}$$

It is straightforward that (2.18) is equivalent to

$$\varepsilon > \frac{-\alpha}{\sqrt{\beta\, e^{-a} + \lambda} + \sqrt{\beta}\, e^{-a/2}} \quad \text{or} \quad \varepsilon < \frac{-\alpha}{\sqrt{\beta\, e^{-a} + \lambda} - \sqrt{\beta}\, e^{-a/2}}\,. \tag{2.19}$$

From (2.16), (2.17) and (2.19), we obtain

$$\Big(\alpha + \sqrt{\beta\, e^{-x} + \lambda}\,\varepsilon(\omega)\Big)^2 \geq \begin{cases} (\alpha + \sqrt{\lambda}\,\varepsilon(\omega))^2\,, & \omega \in \{\varepsilon \geq 0\} \\ (\alpha + \sqrt{\beta\, e^{-a} + \lambda}\,\varepsilon(\omega))^2 - \beta\, e^{-a}\,\varepsilon(\omega)^2\,, & \omega \in A_a^c \cap \{\varepsilon < 0\} \\ 0\,, & \omega \in A_a \end{cases} \tag{2.20}$$

Now take logarithms and use the additive structure (2.9) of $(Z_n)_{n\in\mathbb{N}}$. □

Remark 2.4 (a) If a is large enough then $S_n^{u,a}$ and $S_n^{l,a}$ are random walks with negative drift.
Proof. Note that

$$\begin{aligned}
E(V_1^a) \;&=\; E\left(p(a,\alpha,\beta,\lambda,\varepsilon_1) + q(a,\alpha,\beta,\lambda,\varepsilon_1)\right) \\
&=\; E\left(\ln\left((\alpha + \sqrt{\beta\, e^{-a} + \lambda}\,\varepsilon_1)^2 + 2\alpha\sqrt{\beta}\, e^{-a/2}\,(-\varepsilon_1)\, 1_{\{\varepsilon_1<0\}}\right)\right) \\
&\to\; E(\ln(\alpha + \sqrt{\lambda}\varepsilon_1)^2) < 0\,, \quad \text{as } a \to \infty\,,
\end{aligned}$$

where we used the dominated convergence theorem and (2.2) in the last step. Hence for a large enough the statement follows. □

(b) Let $(S_n)_{n\in\mathbb{N}} := \left(\sum_{j=1}^n \ln\Big((\alpha + \sqrt{\lambda}\varepsilon_j)^2\Big)\right)_{n\in\mathbb{N}}$. For $a \uparrow \infty$ we have

$$S_k^{l,a} \xrightarrow{P} S_k \quad \text{and} \quad S_k^{u,a} \xrightarrow{a.s.} S_k\,, \tag{2.21}$$

for any $k \in \mathbb{N}$, i.e. both random walks converge at least in probability to the same random walk. Furthermore,

$$\sup_{k\geq 1} S_k^{l,a} \xrightarrow{d} \sup_{k\geq 1} S_k \quad \text{and} \quad \sup_{k\geq 1} S_k^{u,a} \xrightarrow{a.s.} \sup_{k\geq 1} S_k\,. \tag{2.22}$$

Proof. The a.s. convergence of $(S_n^{u,a})_{n\in\mathbb{N}}$ and $\sup_{k\geq 1} S_k^{u,a}$ is straightforward since p, q and r converge a.s.. Consider therefore the lower random walk $(S_n^{l,a})_{n\in\mathbb{N}}$. Note that for $a \uparrow \infty$

$$P(A_a) \to 0$$

and hence

$$1_{A_a^c\cap\{\varepsilon<0\}} \xrightarrow{P} 1_{\{\varepsilon<0\}} \quad \text{and} \quad 1_{A_a\cap\{\varepsilon<0\}} \xrightarrow{P} 0\,. \tag{2.23}$$

Furthermore,

$$p(a,\alpha,\beta,\lambda,\varepsilon_1) + r(a,\alpha,\beta,\lambda,\varepsilon_1) \xrightarrow{a.s.} \ln\Big((\alpha+\sqrt{\lambda}\varepsilon_1)^2\Big)\,, \tag{2.24}$$

and therefore (2.21) holds. Finally we note that

$$\begin{aligned} E\max(0,U_1^a) &= E\max\Big(0,\Big(p(a,\alpha,\beta,\lambda,\varepsilon_1)+r(a,\alpha,\beta,\lambda,\varepsilon_1)\Big)\,1_{A_a^c\cap\{\varepsilon_1<0\}}\Big) \\ &\quad + E\max\Big(0,\ln(\alpha+\sqrt{\lambda}\varepsilon_1)^2\,1_{\{\varepsilon_1\geq 0\}}\Big) \\ &\to E\max(0,\ln(\alpha+\sqrt{\lambda}\varepsilon_1)^2), \quad \text{as } a\to\infty\,, \end{aligned} \tag{2.25}$$

where we used (2.23), (2.24) and the dominated convergence theorem. By Borovkov (1976), Theorem 22, p.53, (2.21) and (2.25) we derive that

$$\sup_{k\geq 1} S_k^{l,a} \xrightarrow{d} \sup_{k\geq 1} S_k\,.$$

□

Lemma 2.3 characterizes the behavior of the process $(Z_n)_{n\in\mathbb{N}}$ above a high treshold a and hence also the behavior of $(X_n^2)_{n\in\mathbb{N}}$. This is the key to what follows: the process $(S_n)_{n\in\mathbb{N}}$ will determine completely the extremal behavior of (X_n^2). Recall from Theorem 2.1 that $(X_n)_{n\in\mathbb{N}}$ is Harris recurrent with regeneration set $[-e^{a/2}, e^{a/2}]$ for a large enough. Thus there exists in particular a renewal point process T_0, T_1, T_2, ... which describes the regenerative structure of $(X_n)_{n\in\mathbb{N}}$.

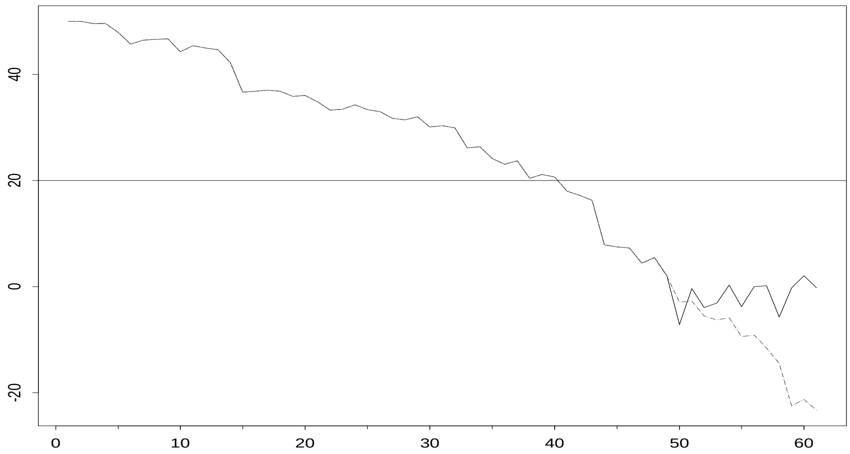

Figure 1: Simulated sample path of $(Z_n)_{n\in\mathbb{N}}$ with parameters $\alpha = 0.6, \beta = 1, \lambda = 0.4$ and starting point $Z_0 = 50$ (solid line) and the corresponding random walks $(S_n^{l,a})_{n\in\mathbb{N}}$ and $(S_n^{u,a})_{n\in\mathbb{N}}$ with $a = 20$ (dotted lines), respectively. Note that the random walks are hardly distinguishable from each other and $(Z_n)_{n\in\mathbb{N}}$ for $n \leq 47$. Hence they are extremely good bounds above the level $a = 20$. If the process falls far below the level 20 they are still very close, but are no longer bounds for $(Z_n)_{n\in\mathbb{N}}$. The picture also confirms our statement that the random walks have negative drift and converge to the same limit.

Corollary 2.5 *The renewal point process $(T_n)_{n\in\mathbb{N}_0}$ which describes the regenerative structure of $(X_n)_{n\in\mathbb{N}}$ is aperiodic and has finite mean recurrence times $C_0 = T_0$ and $C_1 = T_1 - T_0$.*

Proof. The renewal process can be constructed in the following way (see e.g. Asmussen (1989), Section VI.3 for some background on regenerative processes):
Define

$$\tau_1 := \inf\{k \geq 1 \,|\, X_k \in [-e^{a/2}, e^{a/2}]\} = \inf\{k \geq 1 \,|\, Z_k \leq a\} = N_a$$

and $\tau_{i+1} := \inf\{k > \tau_i \,|\, Z_k \leq a\}$ for $i = 1, 2, 3, \ldots$. Since $(Z_n)_{n\in\mathbb{N}}$ is above the level a dominated by the random walk with negative drift $(S_n^{u,a})_{n\in\mathbb{N}}$ and

$$\sup_{x\in(-\infty,a]} E(\max(0, Z_1) \,|\, Z_0 = x) < \infty\,, \tag{2.26}$$

it follows that $\tau_1, \tau_2, \tau_3, \ldots$ are well defined and have finite expectations. Now let $M_1 := \inf\{i \geq 1 \,|\, I_{\tau_i} = 1\}$ and $M_{j+1} := \inf\{i > M_j \,|\, I_{\tau_i} = 1\}$ for $j = 1, 2, 3, \ldots$ with $P(I_1 = 1) = 1 - P(I_1 = 0) = C$ and independent of $(X_n)_{n\in\mathbb{N}}$ where C is the constant in (2.7). Note that

$$P(M_j - M_{j-1} = i) = C(1 - C)^{i-1} \quad \text{for } i,\, j = 1, 2, \ldots \text{ and } M_0 = 0\,. \tag{2.27}$$

From Asmussen (1989), p.151 and (2.5), the renewal process $(T_n)_{n\geq 0}$ is now given by

$$T_n := \tau_{M_{n+1}} + 1\,, \quad n \geq 0\,,$$

and hence, by (2.27)

$$E(C_0) = E(T_0) \leq E(\tau_{M_1+1}) \leq const\, E(M_1 + 1) < \infty\,.$$

Similar calculation shows that $E(C_1) < \infty$ as well. Since the transition density of $(Z_n)_{n\in\mathbb{N}}$ is positive and continuous it follows finally that C_1 is aperiodic. □

As a consequence of Corollary 2.5 we may suppose in the following that the process $(X_n)_{n\in\mathbb{N}}$ is stationary. One can show by a coupling argument that for any probability measure μ and any sequence $(u_n)_{n\in\mathbb{N}}$

$$\left|P^{\mu}\Big(\max_{1\leq k\leq n} X_k \leq u_n\Big) - P^{\pi}\Big(\max_{1\leq k\leq n} X_k \leq u_n\Big)\right| \to 0, \quad \text{as } n \to \infty\,,$$

where P^{μ} denotes the probability law for $(X_n)_{n\in\mathbb{N}}$ when X_0 starts with distribution μ and π is the stationary distribution. For the coupling argument one needs explicitly that the process $(X_n)_{n\in\mathbb{N}}$ is regenerative and that the embedded renewal process is aperiodic and has finite mean recurrence time. We refer to Lindvall (1992, Chapter II and III) for further details.

3 Extremal behavior of the AR(1) process with ARCH(1) errors

In this section we present the main results concerning the extremal behavior of the AR(1) process with ARCH(1) errors and the accompanying squared process. Let $(\widehat{X}_n)_{n\in\mathbb{N}}$ be the associated independent process of $(X_n)_{n\in\mathbb{N}}$, i.e. $\widehat{X}_1, \widehat{X}_2, ...$ are i.i.d. random variables with the stationary distribution function of $(X_n)_{n\in\mathbb{N}}$. From (2.6) and classical extreme value theory we obtain

$$\lim_{n\to\infty} P(n^{-1/\kappa} \max_{1\leq k\leq n} \widehat{X}_k \leq x) = \exp(-c\, x^{-\kappa})\,, \quad x \geq 0\,, \tag{3.1}$$

hence the maximum of the associated independent process $(\widehat{X}_n)_{n\in\mathbb{N}}$ belongs to the domain of attraction of a Fréchet distribution. In the dependent case we prove a similar result. The limit distribution is still a Fréchet distribution but a constant θ occurs in the exponent. θ is called the *extremal index* of the process $(X_n)_{n\in\mathbb{N}}$ and is a measure of local dependence amongst the exceedances over a high threshold by the process $(X_n)_{n\in\mathbb{N}}$. It has a natural interpretation as the reciprocal of the mean cluster size. In order to describe the extremes in more detail, we also consider the point process $(N_n)_{n\in\mathbb{N}}$ of exceedances of an appropriately chosen high threshold u_n given by

$$N_n(\cdot) := \#\{k/n \in \cdot \,|\, X_k > u_n \,,\, k \in \{1, ..., n\}\,\} \tag{3.2}$$

and show that this point process converges to a compound Poisson process N. We derive the intensity and the distribution of the jumps which we denote by $(\pi_k)_{k\in\mathbb{N}}$. Note that in the extreme value theory for strong mixing processes the jumps equal the lengths of clusters of exceedances. For further background we refer to Leadbetter et al. (1983), Rootzén (1988) or Embrechts et al. (1997, Section 8.1). For the ARCH(1) process it was convenient to investigate first the squared process. This is not the case for our model since we have a completely different structure due to the autoregressive part of $(X_n)_{n\in\mathbb{N}}$. Nevertheless, only for the squared process $(X_n^2)_{n\in\mathbb{N}}$ a comparison with results in the ARCH(1) case (see de Haan et al. (1989)) is possible. The following theorem collects our results.

Theorem 3.1 *(a) Suppose $(X_n)_{n\in\mathbb{N}}$ is given by equation (2.1) with $(\varepsilon_n)_{n\in\mathbb{N}}$ satisfying the general conditions (2.3) and $(D.1)-(D.3)$ with parameters α and λ satisfying (2.2) and $X_0 \sim \mu$. Then*

$$\lim_{n\to\infty} P^\mu(n^{-1/\kappa} \max_{1\le j\le n} X_j \le x) = \exp(-c\theta x^{-\kappa})\,, \quad x \ge 0\,, \tag{3.3}$$

where P^μ denotes the law for $(X_n)_{n\in\mathbb{N}}$ when X_0 starts with the distribution μ, κ solves the equation $E(|\alpha+\lambda\varepsilon|^\kappa)=1$, c is defined by (2.7) and

$$\theta = \kappa \int_1^\infty P(\sup_{k\ge1} \prod_{i=1}^k (\alpha+\sqrt{\lambda}\varepsilon_i) \le y^{-1}) y^{-\kappa-1} dy\,.$$

For $x \in \mathbb{R}$, let N_n be the point process of exceedances of the threshold $u_n = n^{1/\kappa}x$ by $X_1, ..., X_n$ given by (3.2). Then

$$N_n \quad \stackrel{d}{\to} \quad N, \quad n\to\infty\,,$$

where N is a compound Poisson process with intensity $c\theta x^{-\kappa}$ and cluster probabilities

$$\pi_k = \frac{\theta_k - \theta_{k+1}}{\theta}\,, \quad k \in \mathbb{N}, \tag{3.4}$$

where

$$\theta_k = \kappa \int_1^\infty P(\#\{j\ge1 \,|\, \prod_{i=1}^j (\alpha+\sqrt{\lambda}\varepsilon_i) > y^{-1}\} = k-1) y^{-\kappa-1} dy\,, \quad k\in\mathbb{N}\,.$$

In particular, $\theta_1 = \theta$.

(b) Let $(X_n)_{n\in\mathbb{N}}$ *be the AR(1)-process with ARCH(1)-errors in (a) and* $(X_n^2)_{n\in\mathbb{N}}$ *the squared process. Then*

$$\lim_{n\to\infty} P^\mu(n^{-2/\kappa} \max_{1\le j\le n} X_j^2 \le x) = \exp(-2c\theta^{(2)} x^{-\kappa/2}), \quad x \ge 0, \tag{3.5}$$

where κ, c *are the same constants as in* (a) *and*

$$\theta^{(2)} = \frac{\kappa}{2} \int_1^\infty P(\sup_{k\ge 1} \prod_{i=1}^k (\alpha + \sqrt{\lambda}\varepsilon_i)^2 \le y^{-1}) y^{-\frac{\kappa}{2}-1} dy\,.$$

For $x \in \mathbb{R}$, *let* $N_n^{(2)}$ *be the point process of exceedances of the threshold* $u_n = n^{2/\kappa}x$ *by* $X_1^2, ..., X_n^2$. *Then*

$$N_n^{(2)} \quad \stackrel{d}{\to} \quad N^{(2)}, \quad n \to \infty,$$

where $N^{(2)}$ *is a compound Poisson process with intensity* $2c\theta^{(2)} x^{-\kappa/2}$ *and cluster probabilities*

$$\pi_k^{(2)} = \frac{\theta_k^{(2)} - \theta_{k+1}^{(2)}}{\theta^{(2)}}, \quad k \in \mathbb{N}, \tag{3.6}$$

where

$$\theta_k^{(2)} = \frac{\kappa}{2} \int_1^\infty P(\#\{j \ge 1 \,|\, \prod_{i=1}^j (\alpha + \sqrt{\lambda}\varepsilon_i)^2 > y^{-1}\} = k-1) y^{-\frac{\kappa}{2}-1} dy\,, \quad k \in \mathbb{N}.$$

In particular, $\theta_1^{(2)} = \theta^{(2)}$.

Remark 3.2 (a) Theorem 3.1 is a generalization of the result of de Haan et al. (1989) in the ARCH(1) case (i.e. $\alpha = 0$). They use a different approach which does not extend to the general case because of the autoregressive part of $(X_n)_{n\in\mathbb{N}}$.

(b) Note that for the squared process one can describe the extremal index and the cluster probabilities by the random walk $(S_n)_{n\in\mathbb{N}}$, namely

$$\theta_k^{(2)} = \frac{\kappa}{2} \int_0^\infty P(\#\{j \ge 1 \,|\, S_j > -x\} = k-1)\, e^{-\frac{\kappa}{2}x}\, dx\,, \quad k \in \mathbb{N}.$$

The description of the extremal behavior of $(X_n^2)_{n\in\mathbb{N}}$ by the random walk $(S_n)_{n\in\mathbb{N}}$ is to be expected since by Lemma 2.3 and Remark 2.4 the process $(Z_n)_{n\in\mathbb{N}} = (\ln(X_n^2))_{n\in\mathbb{N}}$ behaves above a high threshold asymptotically like $(S_n)_{n\in\mathbb{N}}$. Unfortunately, this link fails for $(X_n)_{n\in\mathbb{N}}$. Another possibility for proving statement (b) is to follow the work of Hooghiemstra and Meester

(1995) using the regenerative structure of $(Z_n)_{n\in\mathbb{N}}$, Lemma 2.3, Corollary 2.5 and Remark 2.4(b).

(c) Analogous to de Haan et al. (1989) we may construct "estimators" for the extremal indices $\theta^{(2)}$ and $\theta_k^{(2)}$ of $(X_n^2)_{n\in\mathbb{N}}$, respectively, by

$$\widehat{\theta}^{(2)} = \frac{1}{N}\sum_{i=1}^{N} 1_{\{\sup_{1\le j\le m} S_j^{(i)} \le -E_\kappa^{(i)}\}}$$

and

$$\widehat{\theta}_k^{(2)} = \frac{1}{N}\sum_{i=1}^{N} 1_{\{\sum_{j=1}^{m} 1_{\{S_j^{(i)} > -E_\kappa^{(i)}\}} = k-1\}}, \quad \text{for } k \in \mathbb{N},$$

where N denotes the number of simulated sample paths of $(S_n)_{n\in\mathbb{N}}$, $E_\kappa^{(i)}$ are i.i.d. exponential random variables with intensity κ and m is chosen large enough. These estimators can be studied as in the case $\alpha = 0$ and $\varepsilon \sim N(0,1)$ in de Haan et al. (1989). In particular,

$$\frac{\theta^{(2)} - \widehat{\theta}^{(2)}}{(\theta^{(2)}(1-\theta^{(2)})/N)^{1/2}}$$

is approximately $N(0,1)$ distributed. Because of Remark 3.2(b) this approach is not possible for $(X_n)_{n\in\mathbb{N}}$. We choose as "estimators" for θ and θ_k for $(X_n)_{n\in\mathbb{N}}$

$$\widehat{\theta} = \frac{1}{N}\sum_{i=1}^{N} 1_{\{\sup_{1\le j\le m} \prod_{l=1}^{j}(\alpha+\sqrt{\lambda}\,\varepsilon_l) \le 1/P_\kappa^{(i)}\}} \tag{3.7}$$

and

$$\widehat{\theta}_k = \frac{1}{N}\sum_{i=1}^{N} 1_{\{\sum_{j=1}^{m} 1_{\{\prod_{l=1}^{j}(\alpha+\sqrt{\lambda}\,\varepsilon_l) > 1/P_\kappa^{(i)}\}} = k-1\}}, \quad \text{for } k \in \mathbb{N}, \tag{3.8}$$

where N denotes the number of simulated paths of $(\prod_{l=1}^{n}(\alpha + \sqrt{\lambda}\,\varepsilon_l))_{n\in\mathbb{N}}$, $P_\kappa^{(i)}$ are i.i.d. Pareto-distributed random variables with intensity κ, i.e. with distribution function $G(x) = 1 - x^{-\kappa}$, $x \ge 0$, and m is large enough. These are suggestive estimators since $\prod_{l=1}^{n}(\alpha + \sqrt{\lambda}\,\varepsilon_l) \to 0$ a.s. as $n \to \infty$ because of assumption (2.2).

(d) Note that the extremal index θ of $(X_n)_{n\in\mathbb{N}}$ is not symmetric in the parameter α (see Table 1). This observation is intuitively obvious since for $\alpha > 0$ the clustering is stronger by the autoregressive part than for $\alpha < 0$.

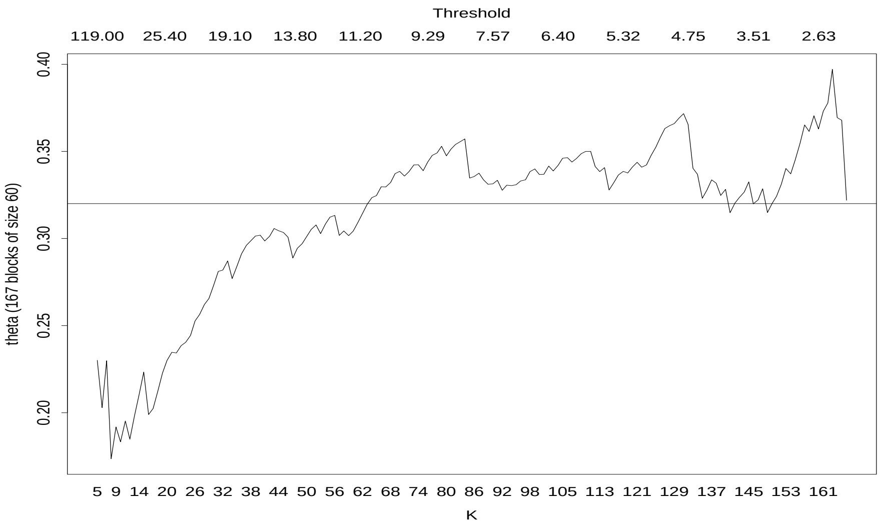

Figure 2: Estimated extremal index of a simulated sample path of $(X_n)_{0\le n\le 10000}$ with parameters $\alpha = 0.8, \beta = 1, \lambda = 0.6$ and $\varepsilon \sim N(0,1)$ using the blocks method for the data (see Embrechts et al. (1997), Section 8.1). The length of a block is chosen as 60. The solid line is the numerically computed extremal index using (3.7), see also Table 1.

α \ λ	0.2	0.4	0.6	0.8	1.0	1.2	1.5	2.0	2.5	3.0	3.5
-1.2	-	0.001	0.001	0.003	0.004	0.001	0.000	-	-	-	-
-1	0.15	0.19	0.19	0.16	0.13	0.09	0.05	0.01	-	-	-
-0.8	0.56	0.47	0.41	0.34	0.26	0.21	0.13	0.05	0.01	-	-
-0.6	0.86	0.71	0.61	0.50	0.41	0.33	0.22	0.10	0.03	0.00	-
-0.4	0.96	0.85	0.71	0.60	0.50	0.40	0.30	0.14	0.06	0.01	-
-0.2	0.98	0.89	0.77	0.65	0.56	0.47	0.33	0.18	0.07	0.02	0.00
0	0.98	0.89	0.78	0.65	0.55	0.45	0.33	0.18	0.08	0.02	0.00
0.2	0.94	0.82	0.72	0.61	0.52	0.43	0.32	0.18	0.07	0.02	0.00
0.4	0.85	0.72	0.63	0.53	0.45	0.37	0.28	0.13	0.06	0.01	-
0.6	0.68	0.55	0.48	0.41	0.35	0.29	0.21	0.10	0.03	0.00	-
0.8	0.39	0.34	0.32	0.27	0.22	0.19	0.12	0.05	0.01	-	-
1.0	0.09	0.14	0.13	0.13	0.11	0.08	0.04	0.01	-	-	-
1.2	-	0.000	0.001	0.003	0.004	0.001	0.000	-	-	-	-

Table 1: "Estimated" extremal index θ of $(X_n)_{n\in\mathbb{N}}$ in the case $\varepsilon \sim N(0,1)$. We chose $N = m = 2000$. Note that the extremal index decreases as $|\alpha|$ increases and that we have no symmetry in α.

$\vert\alpha\vert$ \ λ	0.2	0.4	0.6	0.8	1.0	1.2	1.5	2.0	2.5	3.0	3.5
0	0.95	0.80	0.65	0.52	0.41	0.31	0.22	0.11	0.04	0.01	0.00
0.2	0.94	0.77	0.62	0.49	0.38	0.31	0.22	0.10	0.04	0.01	0.00
0.4	0.84	0.67	0.55	0.43	0.35	0.26	0.19	0.08	0.03	0.01	-
0.6	0.67	0.52	0.41	0.34	0.25	0.18	0.14	0.06	0.02	0.00	-
0.8	0.38	0.31	0.26	0.20	0.16	0.13	0.08	0.03	0.00	-	-
1.0	0.09	0.12	0.11	0.10	0.07	0.05	0.03	0.01	-	-	-
1.2	-	0.000	0.001	0.001	0.000	0.000	0.000	-	-	-	-

Table 2: "Estimated" extremal index $\theta^{(2)}$ of $(X_n^2)_{n\in\mathbb{N}}$ dependent on $|\alpha|$ and λ in the case $\varepsilon \sim N(0,1)$. We chose $N = m = 2000$. Note that the extremal index decreases as $|\alpha|$ increases.

$\vert\alpha\vert$	λ	θ	π_1	π_2	π_3	π_4	π_5	π_6
0	0.2	0.954	0.959	0.037	0.004	0.000	0.000	0.000
0	0.6	0.651	0.682	0.186	0.092	0.018	0.010	0.008
0	1	0.406	0.455	0.233	0.135	0.054	0.044	0.023
0.4	0.2	0.844	0.853	0.122	0.018	0.004	0.002	0.001
0.4	0.6	0.553	0.610	0.201	0.095	0.054	0.015	0.008
0.4	1	0.342	0.431	0.216	0.107	0.066	0.045	0.023
0.8	0.2	0.378	0.445	0.184	0.159	0.071	0.057	0.011
0.8	0.6	0.255	0.328	0.202	0.145	0.088	0.012	0.045
0.8	1	0.152	0.237	0.178	0.099	0.092	0.053	0.010

Table 3: "Estimated" extremal index $\theta^{(2)}$ and cluster probabilities $(\pi_k)_{1\le k\le 6}$ of $(X_n^2)_{n\in\mathbb{N}}$ dependent on α and λ in the case $\varepsilon \sim N(0,1)$. We chose $N = m = 2000$.

α	λ	θ	π_1	π_2	π_3	π_4	π_5	π_6
0	0.2	0.974	0.973	0.027	0.000	0.000	0.000	0.000
0	0.6	0.781	0.799	0.147	0.036	0.012	0.005	0.001
0	1	0.549	0.607	0.188	0.107	0.036	0.034	0.017
-0.4	0.2	0.962	0.962	0.037	0.001	0.000	0.000	0.000
0.4	0.2	0.853	0.867	0.103	0.026	0.002	0.002	0.000
-0.4	0.6	0.715	0.747	0.168	0.048	0.026	0.006	0.002
0.4	0.6	0.624	0.676	0.182	0.066	0.040	0.019	0.012
-0.4	1	0.497	0.540	0.210	0.115	0.075	0.040	0.004
0.4	1	0.445	0.533	0.185	0.080	0.109	0.032	0.017
-0.8	0.2	0.572	0.626	0.185	0.111	0.026	0.033	0.001
0.8	0.2	0.386	0.470	0.172	0.148	0.062	0.068	0.006
-0.8	0.6	0.414	0.520	0.159	0.134	0.072	0.043	0.016
0.8	0.6	0.314	0.443	0.156	0.110	0.087	0.073	0.041
-0.8	1	0.273	0.429	0.137	0.126	0.106	0.016	0.012
0.8	1	0.224	0.346	0.132	0.114	0.129	0.045	0.004

Table 4: "Estimated" extremal index θ and cluster probabilities $(\pi_k)_{1\le k\le 6}$ of $(X_n)_{n\in\mathbb{N}}$ dependent on α and λ in the case $\varepsilon \sim N(0,1)$. We chose $N = m = 2000$. Note that the extremal index for $\alpha > 0$ is much larger than for $\alpha < 0$.

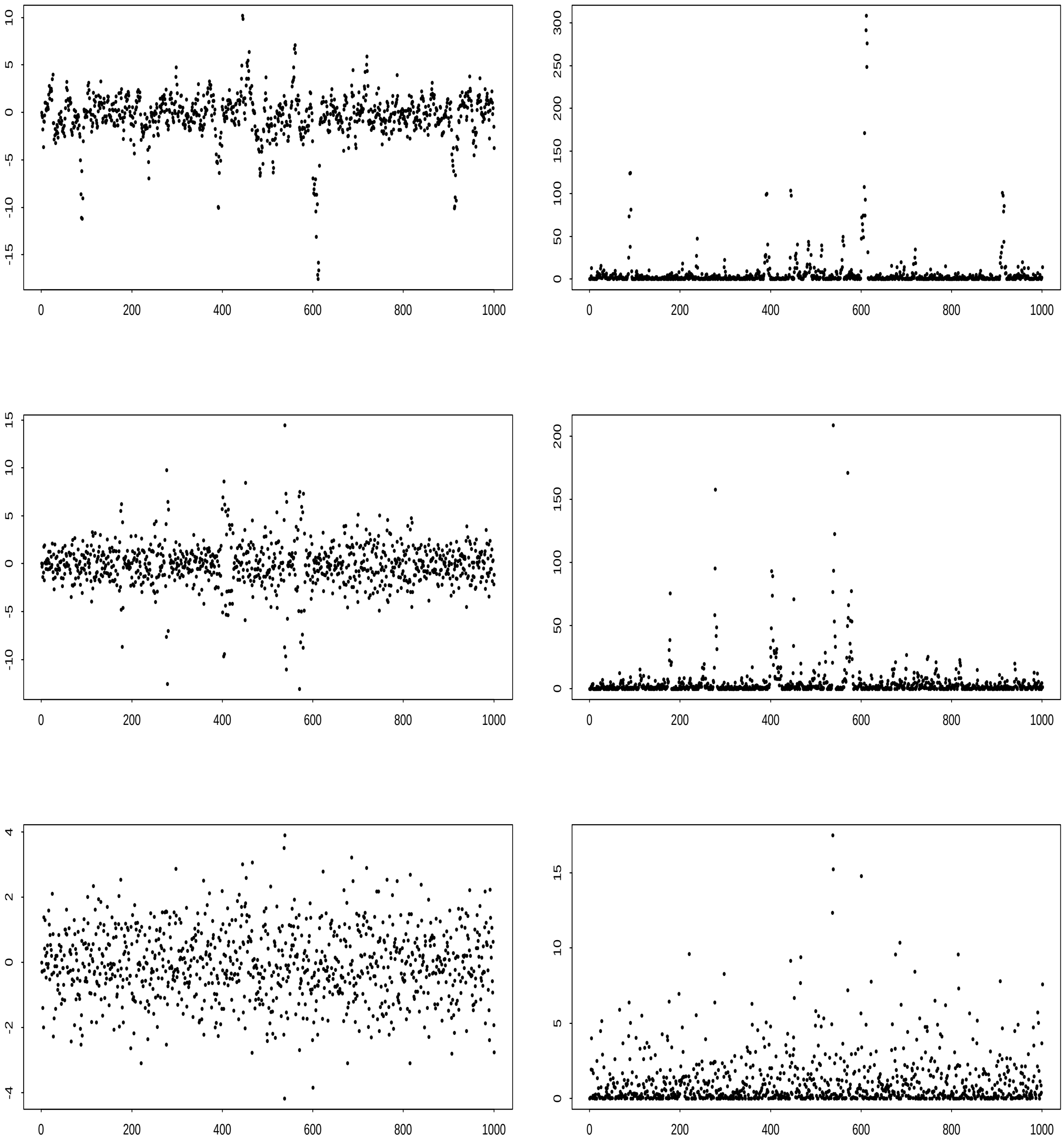

Figure 3: Simulated sample path of $(X_n)_{n\in\mathbb{N}}$ with parameters $\alpha = 0.8, \beta = 1, \lambda = 0.2$ (top, left), of $(X_n^2)_{n\in\mathbb{N}}$ with the same parameters (top, right), of $(X_n)_{n\in\mathbb{N}}$ with parameters $\alpha = -0.8, \beta = 1, \lambda = 0.2$ (middle, left), of $(X_n^2)_{n\in\mathbb{N}}$ with the same parameters (middle, right), of $(X_n)_{n\in\mathbb{N}}$ with parameters $\alpha = 0, \beta = 1, \lambda = 0.2$ (bottom, left) and of $(X_n^2)_{n\in\mathbb{N}}$ with the same parameters (bottom,right) in the case $\varepsilon \sim N(0,1)$. All simulations are based on the same simulated noise sequence $(\varepsilon_n)_{n\in\mathbb{N}}$.

4 The proof of Theorem 3.1

The proof of Theorem 3.1 will be an application of results in Perfekt (1994) (see also the Appendix). In order to apply these results we need to check the assumptions in Theorem A1.1 and A1.2. The next lemma provides a technical property for the squared AR(1) process with ARCH(1) errors $(X_n^2)_{n\in\mathbb{N}}$. It is the most restrictive assumption in Perfekt (1994).

Lemma 4.1 *Let $(p_n)_{n\in\mathbb{N}}$ be an increasing sequence such that*

$$\frac{p_n}{n} \to 0 \quad \textit{and} \quad \frac{n\gamma(\sqrt{p_n})}{p_n} \to 0 \quad \textit{as } n \to \infty\,, \tag{4.1}$$

where γ is the mixing function of $(X_n)_{n\in\mathbb{N}}$. Then for $u_n = n^{2/\kappa}x$

$$\lim_{p\to\infty} \limsup_{n\to\infty} P(\max_{p\le j\le p_n} X_j^2 > u_n \mid X_0^2 > u_n) = 0\,. \tag{4.2}$$

Remark 4.2 (a) The strong mixing condition is a property of the underlying $\sigma-$field of a process. Hence γ is also the mixing function of $(X_n^2)_{n\in\mathbb{N}}$ and $(Z_n)_{n\in\mathbb{N}}$ and we may work in all these cases with the same sequence $(p_n)_{n\in\mathbb{N}}$.

(b) In the case of a strong mixing process, conditions (4.1) are sufficient to guarantee that $(p_n)_{n\in\mathbb{N}}$ is a $\Delta(u_n)$-separating sequence. The notion of a $\Delta(u_n)-$separating sequence was first introduced by O'Brian (1989) and describes somehow the interval length needed to accomplish asymptotic independence of extremal events over a high level u_n in separate intervals. For a definition see also Perfekt (1994). Note that $(p_n)_{n\in\mathbb{N}}$ is in the case of a strong mixing process independent of (u_n).

Proof. Note that

$$\begin{aligned} P(\max_{p\le j\le p_n} X_j^2 > u_n \mid X_0^2 > u_n) &= P(N_a < p,\, \max_{p\le j\le p_n} X_j^2 > u_n \mid X_0^2 > u_n) \\ &\quad + P(p \le N_a < p_n,\, \max_{p\le j\le p_n} X_j^2 > u_n \mid X_0^2 > u_n) \\ &\quad + P(N_a \ge p_n,\, \max_{p\le j\le p_n} X_j^2 > u_n \mid X_0^2 > u_n) \\ &=: I_1 + I_2 + I_3\,, \end{aligned} \tag{4.3}$$

where $N_a = \inf\{j \ge 1 \mid Z_j \le a\} = \inf\{j \ge 1 \mid X_j^2 \le e^a\}$ as in Lemma 2.3. In order to get upper bounds of I_1, I_2 and I_3 we show first that there exist constants $C > 0$ and $N \in \mathbb{N}$ such that for any $n > N$, $x \in [e^{-n}, e^a]$ and $k \in \mathbb{N}$

$$n\, P(X_k^2 > u_n \mid X_0^2 = x) \,\le\, C\,. \tag{4.4}$$

Assume that (4.4) does not hold. Choose $C, N > 0$ arbitrary and $\eta > 0$ small. Because of the continuity of the transition probability (i.e. equicontinuity on compact sets), there exist $n > N$, $x \in [e^{-n}, e^a], k \in \mathbb{N}$ and $\delta = \delta(\eta) > 0$ such that for any $y \in (x - \delta, x + \delta) \cap [e^{-n}, e^a]$

$$n\, P(X_k^2 > u_n \,|\, X_0^2 = y) \,>\, C - \eta\,. \tag{4.5}$$

Let F_{X^2} denote the stationary df of $(X_n^2)_{n \in \mathbb{N}}$. By Theorem 2.1 we have that

$$\lim_{n \to \infty} n\, \overline{F}_{X^2}(u_n) = 2\, c\, x^{-\kappa/2}\,, \tag{4.6}$$

where c is given by the formula in (2.7) and κ is the solution of (2.8). Furthermore, by (4.5) we have

$$\begin{aligned}
n\, \overline{F}_{X^2}(u_n) &= \int_{(-\infty,\infty)} n\, P(X_k^2 > u_n \,|\, X_0^2 = y) dF_{X^2}(y) \\
&\geq \int_{(x-\delta,x+\delta)\cap[e^{-n},e^a]} n\, P(X_k^2 > u_n \,|\, X_0^2 = y) dF_{X^2}(y) \\
&> (C - \eta)\, P(X_0^2 \in (x - \delta, x + \delta) \cap [e^{-n}, e^a]) \\
&\geq (C - \eta)\, D\,,
\end{aligned}$$

where $D := \inf_{z \in [0, e^a]} (F_{X^2}(z + \delta) - F_{X^2}(z)) > 0$ because F_{X^2} is continuous. Since $C > 0$ is arbitrary this is a contradiction to (4.6).

Now we estimate (4.3).

$$\begin{aligned}
I_1 &\leq \sum_{l=1}^{p-1} P\Big(N_a = l,\, \max_{p \leq j \leq p_n} X_j^2 > u_n \,\Big|\, X_0^2 > u_n\Big) \\
&\leq \sum_{l=1}^{p-1} \sum_{j=l+1}^{p_n} P\Big(N_a = l,\, X_j^2 > u_n \,\Big|\, X_0^2 > u_n\Big) \\
&= \sum_{l=1}^{p-1} \sum_{j=l+1}^{p_n} E\Big(1_{\{N_a = l\}}\, P(X_j^2 > u_n \,|\, X_l^2) \,\Big|\, X_0^2 > u_n\Big) \\
&= \sum_{l=1}^{p-1} \sum_{j=l+1}^{p_n} E\Big(1_{\{N_a = l\}}\, 1_{\{X_l^2 \geq e^{-n}\}}\, P(X_j^2 > u_n \,|\, X_l^2) \,\Big|\, X_0^2 > u_n\Big) \\
&\quad + \sum_{l=1}^{p-1} \sum_{j=l+1}^{p_n} E\Big(1_{\{N_a = l\}}\, 1_{\{X_l^2 < e^{-n}\}}\, P(X_j^2 > u_n \,|\, X_l^2) \,\Big|\, X_0^2 > u_n\Big) \\
&=: J_1 \,+\, J_2\,.
\end{aligned} \tag{4.7}$$

Furthermore, by (4.4),

$$J_1 \leq \sum_{l=1}^{p-1} \sum_{j=l+1}^{p_n} \frac{1}{n} E\Big(1_{\{N_a = l\}}\, 1_{\{X_l^2 \geq e^{-n}\}}\, n\, P(X_j^2 > u_n \,|\, X_l^2) \,\Big|\, X_0^2 > u_n\Big)$$

$$
\begin{aligned}
&\le \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} \frac{C}{n} E\Big(1_{\{N_a=l\}}\,1_{\{X_l^2\ge e^{-n}\}}\,\Big|\,X_0^2>u_n\Big) \\
&\le \sum_{j=1}^{p_n} \frac{C}{n} P(N_a<p\,|\,X_0^2>u_n) \qquad (4.8)\\
&\le C\,\frac{p_n}{n} \\
&\to 0\,,\quad \text{as } n\to\infty\,,
\end{aligned}
$$

since $p_n=o(n)$. Similarly, with $B_l:=\{X_1^2>e^a,...,X_{l-1}^2>e^a\}$ for any $l=2,3,4,...$ and $B_1=\Omega$, we obtain

$$
\begin{aligned}
J_2 &\le \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\Big(1_{\{N_a=l\}}\,1_{\{X_l^2<e^{-n}\}}\,\Big|\,X_0^2>u_n\Big)\\
&= \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\Big(1_{B_l}\,P(X_l^2<e^{-n}\,|X_{l-1}^2)\,\Big|\,X_0^2>u_n\Big)\\
&= \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\left(1_{B_l}\,P\Big((\alpha X_{l-1}+\sqrt{\beta+\lambda X_{l-1}^2}\,\varepsilon_l)^2<e^{-n}\,\Big|X_{l-1}^2\Big)\,\Big|\,X_0^2>u_n\right)\\
&= \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\left(1_{B_l\cap\{X_{l-1}>0\}}\,P\Big(\frac{-e^{-n/2}/X_{l-1}-\alpha}{\sqrt{\beta/X_{l-1}^2+\lambda}}<\varepsilon_l<\frac{e^{-n/2}/X_{l-1}-\alpha}{\sqrt{\beta/X_{l-1}^2+\lambda}}\Big)\,\Big|\,X_0^2>u_n\right)\\
&\quad+\sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\left(1_{B_l\cap\{X_{l-1}<0\}}\,P\Big(\frac{e^{-n/2}/X_{l-1}+\alpha}{\sqrt{\beta/X_{l-1}^2+\lambda}}<\varepsilon_l<\frac{-e^{-n/2}/X_{l-1}+\alpha}{\sqrt{\beta/X_{l-1}^2+\lambda}}\Big)\,\Big|\,X_0^2>u_n\right)\\
&= \sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\left(1_{B_l\cap\{X_{l-1}>0\}}\,P\Big(\frac{-e^{-n/2-a/2}-\alpha}{\sqrt{\lambda}}<\varepsilon_l<\frac{e^{-n/2-a/2}-\alpha}{\sqrt{\lambda}}\Big)\,\Big|\,X_0^2>u_n\right)\\
&\quad+\sum_{l=1}^{p-1}\sum_{j=l+1}^{p_n} E\left(1_{B_l\cap\{X_{l-1}<0\}}\,P\Big(\frac{-e^{-n/2-a/2}+\alpha}{\sqrt{\lambda}}<\varepsilon_l<\frac{e^{-n/2-a/2}+\alpha}{\sqrt{\lambda}}\Big)\,\Big|\,X_0^2>u_n\right)\\
&\le 2\,const\,p\,p_n\,e^{-n/2-a/2}\\
&\to 0\,,\quad \text{as } n\to\infty\,,
\end{aligned}
$$

and therefore with (4.8) $I_1\to 0$ as $n\to\infty$.

Now we estimate $\limsup_{n\to\infty} I_3$. Note first that by the Markov inequality

$$
\begin{aligned}
P\Big(\max_{p\le j\le p_n} S_j^{u,a}>-z\Big) &\le \sum_{j=p}^{p_n} P\Big(e^{\frac{\kappa}{4}S_j^{u,a}}>e^{-\frac{\kappa}{4}z}\Big)\\
&= \sum_{j=p}^{p_n} P\Big(\prod_{m=1}^{j}\Big((\alpha+\sqrt{\beta e^{-a}+\lambda}\,\varepsilon_m)^2-2\alpha\sqrt{\beta}e^{-a/2}\varepsilon_m 1_{\{\varepsilon_m<0\}}\Big)^{\kappa/4}>e^{-\frac{\kappa}{4}z}\Big)
\end{aligned}
$$

$$
\begin{aligned}
&\leq \; e^{\frac{\kappa}{4}z}\sum_{j=p}^{p_n} E\Bigg(\Big((\alpha+\sqrt{\beta\,e^{-a}+\lambda}\,\varepsilon_1)^2-2\alpha\sqrt{\beta}e^{-a/2}\varepsilon_1 1_{\{\varepsilon_1<0\}}\Big)^{\kappa/4}\Bigg)^j \\
&\leq \; e^{\frac{\kappa}{4}z}\sum_{j=p}^{p_n}\eta^j\,, \qquad (4.9)
\end{aligned}
$$

where $\eta<1$ such that $E\Big(\Big((\alpha+\sqrt{\beta\,e^{-a}+\lambda}\,\varepsilon_1)^2-2\alpha\sqrt{\beta}e^{-a/2}\varepsilon_1 1_{\{\varepsilon_1<0\}}\Big)^{\kappa/4}\Big)\leq\eta$ for a large enough. This is possible because of (2.2) which implies that $E(|\alpha+\sqrt{\lambda}\,\varepsilon_1|^u)<1$ for all $u\in(0,\kappa)$ and the fact that

$$
E\Big(\Big((\alpha+\sqrt{\beta\exp(-a)+\lambda}\,\varepsilon_1)^2-2\alpha\sqrt{\beta}e^{-a/2}\varepsilon_1 1_{\{\varepsilon_1<0\}}\Big)^{\kappa/4}\Big)\to E\Big(|\alpha+\sqrt{\lambda}\,\varepsilon_1|^{\kappa/2}\Big)\,,\quad a\to\infty
$$

by the dominated convergence theorem. Thus from Theorem 2.1, Lemma 2.3, (4.9) and a large enough,

$$
\begin{aligned}
\limsup_{n\to\infty} I_3 &\leq \limsup_{n\to\infty} P(N_a\geq p_n,\ \max_{p\leq j\leq p_n} Z_0+S_j^{u,a}>\ln u_n\,|\,Z_0>\ln u_n) \\
&\leq \limsup_{n\to\infty} P(\max_{p\leq j\leq p_n} Z_0+S_j^{u,a}>\ln u_n\,|\,Z_0>\ln u_n) \\
&= \limsup_{n\to\infty}\int_0^\infty P(\max_{p\leq j\leq p_n} S_j^{u,a}>-z)\frac{\kappa}{2}e^{-\frac{\kappa}{2}z}dz \qquad (4.10)\\
&\leq 2\sum_{j=p}^{\infty}\eta^j \;=\; 2\frac{\eta^{p-1}}{1-\eta}\,.
\end{aligned}
$$

Finally, note that

$$
\begin{aligned}
I_2 &\leq P(p\leq N_a<p_n,\ \max_{N_a<j\leq p_n} X_j^2>u_n\,|\,X_0^2>u_n)+P(p\leq N_a<p_n,\ \max_{p\leq j\leq N_a} X_j^2>u_n\,|\,X_0^2>u_n) \\
&=: K_1+K_2\,.
\end{aligned}
$$

Similarly as for I_1 and I_3, respectively, we derive that

$$
\limsup_{n\to\infty} K_1=0 \quad\text{and}\quad \limsup_{n\to\infty} K_2 \;=\; 2\frac{\eta^{p-1}}{1-\eta}\,.
$$

Now plugging all together and letting $p\to\infty$ the statement follows. □

Corollary 4.3 *Let $(p_n)_{n\in\mathbb{N}}$ be the same sequence as in Lemma 4.1. Then $(p_n)_{n\in\mathbb{N}}$ is also a $\Delta(u_n)-$separating sequence for $(X_n)_{n\in\mathbb{N}}$, where $u_n=n^{1/\kappa}x$ and $x\in\mathbb{R}$ arbitrary and*

$$
\lim_{p\to\infty}\limsup_{n\to\infty} P(\max_{p\leq j\leq p_n} X_j>u_n\,|\,X_0>u_n)=0\,. \qquad (4.11)
$$

Proof. Because of Remark 4.2(a) and (b), it is straightforward that $(p_n)_{n\in\mathbb{N}}$ is a $\Delta(u_n)-$separating sequence for $(X_n)_{n\in\mathbb{N}}$. Note furthermore that

$$\begin{aligned} P(\max_{p\le j\le p_n} X_j^2 > u_n^2 \,|\, X_0^2 > u_n^2) &= \frac{P(\max_{p\le j\le p_n} X_j^2 > u_n^2,\, X_0^2 > u_n^2)}{P(X_0^2 > u_n^2)} \\ \ge \quad \frac{P(\max_{p\le j\le p_n} X_j > u_n,\, X_0 > u_n)}{P(X_0 > u_n) + P(X_0 < -u_n)} &= \frac{1}{2} P(\max_{p\le j\le p_n} X_j > u_n \,|\, X_0 > u_n) \end{aligned}$$

and hence the statement follows using Lemma 4.1. □

Now we are finally able to prove Theorem 3.1.

Proof of Theorem 3.1. The proof is an application of Theorem A1.2. We prove only statement (a), statement (b) follows along the same lines using Theorem A1.1. As stated already we may assume w.l.o.g. that $(X_n)_{n\in\mathbb{N}}$ is stationary. Let $x \in \mathbb{R}$ be arbitrary. Note that

$$\lim_{u\to\infty} \frac{P(X_0 > u + \frac{1}{\kappa} u\, x)}{P(X_0 > u)} = \begin{cases} \infty & , \quad 1 + \frac{1}{\kappa} x \le 0 \\ (1 + \frac{1}{\kappa} x)^{-\kappa} & , \quad 1 + \frac{1}{\kappa} x > 0 \end{cases}$$

and

$$\lim_{u\to\infty} P(\frac{X_1}{u} \le x \,|\, X_0 = u) = P(\alpha + \sqrt{\lambda}\,\varepsilon \le x)\,.$$

By Corollary 4.3 and the strong mixing property of $(X_n)_{n\in\mathbb{N}}$ all assumptions of Theorem A1.2 are fulfilled and we have that the extremal index θ is given by

$$\begin{aligned} \theta &= \int_1^\infty P(\#\{j \ge 1 \,|\, (\prod_{i=1}^{j} (\alpha + \sqrt{\lambda}\,\varepsilon_i)) Y_0 > 1\} = 0 \,|\, Y_0 = y)\, \kappa\, y^{-\kappa-1} dy \\ &= \int_1^\infty P(\max_{j\ge 1} (\prod_{i=1}^{j} (\alpha + \sqrt{\lambda}\,\varepsilon_i) \le y^{-1})\, \kappa\, y^{-\kappa-1} dy\,. \end{aligned}$$

The cluster probabilities can be determined in the same way and hence the statement follows.
□

A1 Appendix

The theorem below gives the extremal properties of a fairly large class of stationary Markov chains. The original version can be found in Perfekt (1994, Theorem 3.2, p. 538). We present a simplified version of Perfekt's result which can be directly applied to our situation.

Theorem A1.1 *Suppose $(X_n)_{n\in\mathbb{N}}$ is a stationary Markov chain which satisfies for some $\gamma \in (-\infty, \infty)$ the following properties*

(i)

$$\lim_{u\uparrow x_F} \frac{1 - F(u + g(u)x)}{1 - F(u)} = (1 - \gamma\, x)_+^{1/\gamma},\ x \in (-\infty, \infty),$$

where F is the stationary df, $x_F := \sup\{x\,;\, F(x) < 1\}$, $y_+ := \max\{0, y\}$ and

$$x_F = \infty \quad \text{and} \quad g(u) = -\gamma\, u \qquad \text{if } \gamma < 0$$

$$x_F < \infty \quad \text{and} \quad g(u) = \gamma(x_F - u) \quad \text{if } \gamma > 0$$

If $\gamma = 0$, then the auxiliary function g is unique up to asymptotic equivalence and strictly positive on (x_0, x_F) for some $x_0 < x_F$.

(ii)

$$\lim_{u\to x_F} P\Big((1 - \gamma\, \frac{(X_1 - u)}{g(u)})_+^{-1/\gamma} \leq x \,|\, X_0 = u\Big) = H(x)$$

for some df H on $[0, \infty)$.

Let furthermore $(A_n)_{n\in\mathbb{N}}$ be an i.i.d. sequence with marginal df H and let Y_0 be a random variable independent of $(A_n)_{n\in\mathbb{N}}$. Define the tail chain $(Y_n)_{n\in\mathbb{N}}$ by $Y_n = A_n\, Y_{n-1}$ for $n \geq 1$ and denote by P^μ the law of $(Y_n)_{n\in\mathbb{N}}$ when Y_0 has distribution μ. Assume $\mu(dx) = x^{-2}dx$, $x > 1$ and let $(u_n(\tau))$ be a sequence which satisfies

$$\lim_{n\to\infty} n(1 - F(u_n(\tau))) \to \tau\,.$$

(a) Assume $D(u_n(\tau))$ holds for each $\tau > 0$. If for some τ_0 there is a $D(u_n(\tau_0))$-separating sequence $(p_n)_{n\in\mathbb{N}}$ such that

$$\lim_{p\to\infty} \limsup_{n\to\infty} P(\max_{p\leq j\leq p_n} X_j > u_n \,|\, X_0 > u_n) = 0 \tag{A.1}$$

holds with $u_n = u_n(\tau)$ then $(X_n)_{n\in\mathbb{N}}$ has extremal index θ given by

$$\theta = P^\mu(\#\{n \geq 1 \,|\, Y_n > 1\} = 0)\,.$$

(b) Suppose $(X_n)_{n\in\mathbb{N}}$ has extremal index $\theta > 0$ and, for some $\tau_1 > 0$ satisfies $\Delta(u_n(\sigma\tau_1))$ for each $\sigma > 0$. Suppose further there is a $\Delta(u_n(\tau_1))$-separating sequence $(p_n)_{n\in\mathbb{N}}$ such that (A.1) holds with $u_n = u_n(\tau_1)$. Then, for each $\sigma > 0$, $N_n^{\sigma\tau_1} := \#\{k \in \{1,...,n\} \,|\, k/n \in \cdot,\, X_k > u_n(\sigma\tau_1)\}$ converges in distribution to a compound Poisson process N with intensity $\theta\sigma\tau_1$ and jump probabilities π_i given by

$$\pi_i = \frac{1}{\theta}\Big(P^\mu(\#\{n \geq 1 \,|\, Y_n > 1\} \;=\; i-1) - P^\mu(\#\{n \geq 1 \,|\, Y_n > 1\} \;=\; i)\Big)\,, \quad i \in \mathbb{N}\,.$$

The next theorem is an extension of Theorem A1.1. In some cases it is easier to apply then the last one.

Theorem A1.2 *(Extension of Theorem 3.2 of Perfekt (1994), p. 543) Suppose $(X_n)_{n\in\mathbb{N}}$ is a stationary Markov chain which satisfies*

$$\lim_{u\uparrow x_F} \frac{1-F(u+g(u)x)}{1-F(u)} = (1-\gamma\, x)_+^{1/\gamma}\,, \quad x \in (-\infty,\infty),$$

where F is the stationary df, $x_F := \sup\{x\,;\, F(x) < 1\} = \infty$, $y_+ := \max\{0, y\}$ and

$$g(u) = -\gamma\, u \text{ for some } \gamma < 0\,.$$

Suppose furthermore that $\inf\{x\,;\, F(x) > 0\} = -\infty$ and that

$$\lim_{u\to\infty} P(\frac{X_1}{u} \leq x \,|\, X_0 = u) = H(x)\,,$$

for some df H on $(-\infty,\infty)$. Let $(A_n)_{n\in\mathbb{N}}$ be an i.i.d. sequence with distribution H and define the tail chain through $Y_n = A_n Y_{n-1}$, $n \geq 1$, Y_0 being independent of $(A_n)_{n\in\mathbb{N}}$. Then, if $(X_n)_{n\in\mathbb{N}}$ satisfies the conditions in (a) and (b) of Theorem A1.1, the result of the theorem holds with the initial distribution μ given by $\mu(dx) := |\gamma|^{-1}x^{1/\gamma-1}dx$, for $x > 1$.

Acknowledgement

I wish to express my gratitude to Claudia Klüppelberg for her constant support and advice. I am also grateful to Alex McNeil for an Splus program which estimates the extremal index by the blocks method.

References

[1] Asmussen, S. (1987) *Applied Probability and Queues.* Wiley, Chichester.

[2] Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992) ARCH modelling in finance: a review of the theory and empirical evidence. *Journ. Econometrics* **52**, 5-59.

[3] Borkovec, M. and Klüppelberg, C. (1998) The tail of the stationary distribution of an autoregressive process with ARCH(1) errors. *Preprint.*

[4] Borovkov, A. A. (1976) *Stochastic Processes in Queueing Theory.* Springer, New York.

[5] Diebolt, J. and Guégan, D. (1990) Probabilistic properties of the general nonlinear Markovian process of order one and applications to time series modelling. *Rapp. Tech.* **125**, L.S.T.A., Paris 6.

[6] Embrechts, P. and Goldie, C.M. (1994) Perpetuities and random equations. In: Mandl, P., Huskova, M. (Eds.) Asymptotic Statistics. *Proceedings of the 5th Prague Symposium.* September 4–9, 75–86, Physica-Verlag, Heidelberg.

[7] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Heidelberg.

[8] Engle, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation. *Econometrica* **50**, 987–1007.

[9] Haan, L. de, Resnick, S.I., Rootzén, H. and Vries, C.G. de (1989) Extremal behavior of solutions to a stochastic difference equation with applications to ARCH processes. *Stoch. Proc. Appl.* **32**, 213–224.

[10] Hooghiemstra, G. and Meester L.E. (1995) Computing the extremal index of special Markov chains and queues. *Stoch. Proc. Appl.* **65**, 171–185.

[11] Hsing, T. (1984) Point processes associated with extreme value theory. *Dissertation, Technical Report 83, Center for Stochastic Processes, Univ. North Carolina.*

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

[13] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983) *Extremes and Related Properties of Random Sequences and Processes.* Springer, Berlin.

[14] Lindvall, T. (1992) *Lectures on the Coupling Method.* Wiley, New York.

[15] Maercker, G. (1997) *Statistical Inference in Conditional Heteroskedastic Autoregressive Models.* Shaker, Aachen. (Dissertation, University of Braunschweig)

[16] O'Brien, G. L. (1987) Extreme values for stationary and Markov sequences. *Ann. Prob.* **15**, 281–291.

[17] Perfekt, R. (1994) Extremal behavior of stationary Markov chains with applications. *Ann. Appl. Prob.* **4**, 529–548.

[18] Priestley, M.B. (1988) *Non-linear and Non-stationary Time Series.* Academic Press, New York.

[19] Rootzén, H. (1988) Maxima and exceedances of stationary Markov chains. *Adv. Appl. Prob.* **20**, 371–390.

[20] Shephard, N. (1996) Statistical aspects of ARCH and stochastic volatility. In: Cox, D.R., Hinkley, D.V. and Barndorff-Nielsen, O.E. (Eds.), Likelihood, Time Series with Econometric and other Applications. Chapman and Hall, London.

[21] Tong, H. (1990) *Non-linear Time Series – A Dynamical System Approach* Oxford University Press, Oxford.

[22] Vervaat, W. (1979) On a stochastic difference equation and a representation of nonnegative infinitely divisible random variables. *Adv. Appl. Prob.* **11**, 750–783.

[23] Weiss, A.A. (1984) ARMA models with ARCH errors. *J.T.S.A.* **3**, 129–143.