# Extremes of Autoregressive Threshold Processes

Claudia Brachner* Vicky Fasen† Alexander Lindner‡

**Abstract**

In this paper we study the tail and the extremal behavior of stationary solutions of autoregressive threshold (TAR) models. It is shown that a regularly varying noise sequence leads in general only to an O-regularly varying tail of the stationary solution. Under further conditions on the partition, it is however shown that $\mathrm{TAR}(S,1)$ models of order 1 with $S$ regimes have regularly varying tails, provided the noise sequence is regularly varying. In these cases, the finite dimensional distribution of the stationary solution is even multivariate regularly varying and its extremal behavior is studied via point process convergence. In particular, a TAR model with regularly varying noise can exhibit extremal clusters. This is in contrast to TAR models with noise in the maximum domain of attraction of the Gumbel distribution and which is either subexponential or in $\mathcal{L}(\gamma)$ with $\gamma > 0$. In that case it turns out that the tail of the stationary solution behaves like a constant times that of the noise sequence, regardless of the order and the specific partition of the TAR model, and that the process cannot exhibit clusters on high levels.

*AMS 2000 Subject Classifications:* primary: 60G70
secondary: 60G10, 60G55

*Keywords:* ergodic, exponential tail, extreme value theory, O-regular variation, point process, regular variation, SETAR process, subexponential distribution, tail behavior, TAR process

*Allianz Investment Management SE, Königinstraße 28, 80802 München, Germany

†Centre for Mathematical Sciences, Technische Universität München, Boltzmannstraße 3, 85747 Garching, Germany, email: fasen@ma.tum.de. Financial support by the Deutsche Forschungsgemeinschaft through a research grant is gratefully acknowledged.

‡Institut für Mathematische Stochastik, Technische Universität Braunschweig, Pockelsstraße 14, 38106 Braunschweig, Germany, email: a.lindner@tu-bs.de

## 1 Introduction

A *(self exciting) threshold autoregressive* (TAR or SETAR) *model of order* $q$ *with* $S$ *regimes* is a piecewise AR($q$) process with different regimes, where the current regime depends on the size of the past observations. More precisely, we will consider the following model: let $(Z_k)_{k\in\mathbb{N}_0}$ be an independent and identically distributed (i. i. d.) noise sequence, let $q, p, S, d_1, \ldots, d_p \in \mathbb{N}$ with $d_1 < \ldots < d_p$, let $\{J_i : i = 1, \ldots, S\}$ be a partition of $\mathbb{R}^p$ into pairwise disjoint Borel sets and $\alpha_i$, $i = 1, \ldots, S$, as well as $\beta_{ij}$, $i = 1, \ldots, S$, $j = 1, \ldots, q$, be real coefficients. Then we call a process $(X_k)_{k\in\mathbb{N}_0}$ satisfying

$$X_k = \sum_{i=1}^{S} \{\alpha_i + \sum_{j=1}^{q} \beta_{ij} X_{k-j}\} \mathbf{1}_{\{(X_{k-d_1},\ldots,X_{k-d_p})\in J_i\}} + Z_k, \quad k \geq \max\{q, d_p\}, \tag{1.1}$$

and for which the starting vector $(X_0, \ldots, X_{\max\{q,d_p\}-1})$ is independent of $(Z_{k+h})_{h\in\mathbb{N}_0}$, a TAR($S, q$) process. The current regime at time $k$ is determined by the vector $(X_{k-d_1}, \ldots, X_{k-d_p})$ of the past observations, and within each regime $(X_k)$ follows an AR($q_i$) process with $q_i := \max\{j \in \{1, \ldots, q\} : \beta_{ij} \neq 0\}$. Sometimes one also allows the variance of the noise to be regime dependent, by replacing $Z_k$ in (1.1) by $\sigma_i Z_k$, where $\sigma_i$ depends on the past in the same way as $\alpha_i$ does, but in this paper we shall not consider such a specification. Autoregressive threshold models were introduced by Tong [28] in 1977 and were systematically presented by Tong and Lim [30], who used them as a model for the lynx data. Since then they have found various applications in many areas, such as financial economics, physics, population dynamics, or neural sciences, to name just a few; see the presentation in Fan and Yao [13] for further information and references. In particular, when used as a model for financial data, it is important to have information about the tail- and the extremal behavior of these models, since stylized facts of financial data are heavy tails and clusters on high levels. The present paper will investigate the tail and the extremal behavior of TAR models for various classes of driving noise sequences.

A somewhat related paper by Diop and Guegan [10] considers the tail behavior of threshold autoregressive stochastic volatility models. Observe, however, that the threshold model under consideration in [10] is governed by a different regime switching mechanism, where the regime is not determined by the size of the previous observation as for the TAR process, but by the sign of the volatility model.

The paper is organized as follows: in Section 2 we collect some known assumptions under which the model has a strictly stationary and geometrically ergodic solution. We then prove that under the given conditions the tail of the stationary solution can be estimated by the tail of a certain corresponding AR($q$) sequence, a lemma which will turn out to be a crucial ingredient for the determination of the tail behavior.

Next, in Section 3 we derive the tail and the extremal behavior of a TAR process with a regularly varying noise. It is shown that the tail of the TAR model is O-regularly varying but not necessarily regularly varying. For this reason we restrict our attention to the classical TAR($S$,1) model of order 1 with $S$ regimes, where the partition is a partition into $S$ intervals and the regime is determined by whether $X_{k-1}$ is in these intervals. In that case, we show that the stationary solution has a regularly varying tail, and even that the finite dimensional distributions are multivariate regularly varying. Furthermore, we derive the extremal behavior by point process convergence. A result is that the classical TAR($S$,1) process models extremal clusters.

Finally, Section 4 is again on general TAR($S, q$) models with noise which has at most an exponentially decreasing tail without being regularly varying. It is then shown that the tail of the TAR model is in the same class and its extremal behavior is determined. Here, the sequence of point processes converges to a Poisson random measure, which reflects, in contrast to the regularly varying case, the absence of extremal clusters. Some of the results presented in this paper can also be found in the diploma thesis [3] of the first named author.

Throughout the paper we shall denote $\mathbb{N} = \{1, 2, \ldots\}$, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, $\mathbb{R}_+ = (0, \infty)$, $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$ and use "$\stackrel{w}{\Longrightarrow}$" to denote weak convergence. For the integer part of a real number $x$ we write $\lfloor x \rfloor = \sup\{n \in \mathbb{Z} : n \leq x\}$. The Dirac measure at a point $x$ will be denoted by $\varepsilon_x$. For two strictly positive functions $g$ and $h$ and a constant $c \in [0, \infty)$ we write $g(t) \sim ch(t)$ as $t \to \infty$ if the quotient $g(t)/h(t)$ tends to $c$ as $t \to \infty$. For $\mathbf{x} \in \mathbb{R}^d$ we denote by $\mathbf{x}^T$ the transposed of $\mathbf{x}$ and by $\|\mathbf{x}\|$ the maximum norm of $\mathbf{x}$. As usual, the positive part and the negative part of $x \in \mathbb{R}$ is denoted by $x^+ = \max\{x, 0\}$ and $x^- = \max\{0, -x\}$, respectively. The tail of a distribution function $F$ will be written as $\overline{F} = 1 - F$.

## 2 Model assumptions and basic properties

In this paper we restrict our attention to noise sequences $(Z_k)$ which are subexponential (which includes regularly varying noise) or are in the class $\mathcal{L}(\gamma)$ with $\gamma > 0$, which includes tails of the form $\mathbb{P}(Z_1 > x) \sim K x^b \mathrm{e}^{-\gamma x}$ as $x \to \infty$ with $b \in \mathbb{R}$ and $K > 0$. When determining the tail behavior, we shall further assume the following classical *tail balance condition TB*, which is standard also for extreme value theory of linear ARMA processes, as presented, e.g., in Embrechts et al. [12], Section A.3.3.

**Condition TB.**
*There are constants $p^+, p^- \in [0,1]$ such that $p^+ + p^- = 1$ and the tail of $Z_1$ satisfies the tail balance condition*

$$\mathbb{P}(Z_1 > x) \sim p^+ \mathbb{P}(|Z_1| > x) \quad \textit{and} \quad \mathbb{P}(Z_1 < -x) \sim p^- \mathbb{P}(|Z_1| > x) \quad \textit{as } x \to \infty.$$

For the existence of strictly stationary solutions $(X_k)_{k\in\mathbb{N}_0}$ we use the following sufficient condition DC on the *distribution* of $Z_1$ and the size of the *coefficients*, which is sufficiently general for our purposes. That it is not a necessary condition for stationarity can be seen from the presentation in Chen and Tsay [6].

**Condition DC.**
*Let $(Z_k)_{k\in\mathbb{N}_0}$ be an i. i. d. sequence, whose marginal distribution has a Lebesgue density $h$ satisfying $\inf_{x\in K} h(x) > 0$ for every compact set $K \subset \mathbb{R}$. Furthermore, assume that $\mathbb{E}|Z_1|^{\min\{1,\eta\}} < \infty$ for some $\eta > 0$. Denoting $\alpha := \max_{k=1,\ldots,S} |\alpha_k|$ and $\beta_j := \max_{i=1,\ldots,S} |\beta_{ij}|$ assume further that $\beta := \sum_{j=1}^{q} \beta_j < 1$.*

Define the function $f : \mathbb{R}^l \to \mathbb{R}$ by

$$f(x_{k-1}, x_{k-2}, \ldots, x_{k-l}) := \sum_{i=1}^{S} \{\alpha_i + \sum_{j=1}^{q_i} \beta_{ij} x_{k-j}\} \, \mathbf{1}_{\{(x_{k-d_1},\ldots,x_{k-d_p})\in J_i\}},$$

where $l := \max\{q, d_p\}$, and denote $\overrightarrow{X_k} := (X_k, X_{k-1}, \ldots, X_{k-l+1})^T$. Then the threshold model (1.1) has the representation

$$X_k = f(\overrightarrow{X}_{k-1}) + Z_k \quad \text{ for } k \in \{l, l+1, \ldots\},$$

with $f(\overrightarrow{X}_{k-1})$ being independent of $Z_k$. Furthermore, we write $\overrightarrow{Z}_k := (Z_k, 0, \ldots, 0)^T$ and $\overrightarrow{f}(\overrightarrow{x}) := (f(\overrightarrow{x}), x_1, \ldots, x_{l-1})^T$ for $\overrightarrow{x} := (x_1, \ldots, x_l)^T \in \mathbb{R}^l$. Then $(\overrightarrow{X}_k)_{k\geq l-1}$ is a Markov chain, where

$$\overrightarrow{X}_k = \overrightarrow{f}(\overrightarrow{X}_{k-1}) + \overrightarrow{Z}_k \quad \text{ for } k \geq l.$$

The following lemma assures the existence of a geometrically ergodic strictly stationary solution. Geometric ergodicity was already proved by Chan and Tong [5] (cf. Tong [29], Example A1.2, p. 464), and An and Huang [1], Theorem 3.2 and Example 3.6, under the slightly more restrictive condition that $\mathbb{E}|Z_1| < \infty$. However, the proof in [1] can easily be generalized to the case where only finiteness of $\mathbb{E}|Z_1|^{\min\{1,\eta\}}$ for some $\eta > 0$ is assumed: simply replace the test function $(x_1, \ldots, x_p) \mapsto \max\{|x_1|, \ldots, |x_p|\}$ appearing in the proof of Theorem 3.2 in An and Huang [1] by $(x_1, \ldots, x_p) \mapsto (\max\{|x_1|, \ldots, |x_p|\})^{\min\{1,\eta\}}$. Also note that the proof in [1] carries over to the more general partitions considered in (1.1) without change. That geometric ergodicity of the stationary solution then implies strong mixing with geometrically decreasing mixing rate in the sense that

$$\sup_{\substack{A\in\sigma(X_j:j\leq m)\\ B\in\sigma(X_j:j\geq k+m)}} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)| \leq K\gamma^k\,, \quad k, m \in \mathbb{N}, \tag{2.1}$$

for some $0 < \gamma < 1$ and $K > 0$, then follows from Meyn and Tweedie [21], Theorem 16.1.5.

**Lemma 2.1** *Let $(X_k)_{k\in\mathbb{N}_0}$ be a TAR process as given in (1.1), and suppose that Condition DC holds. Then $(X_k)_{k\in\mathbb{N}_0}$ is geometrically ergodic and admits a unique strictly stationary solution, which is strongly mixing with geometrically decreasing mixing rate in the sense of (2.1) .*

The next lemma is crucial for the analysis of extremes of TAR models. It states that under our assumptions, there is a causal AR($q$) process whose stationary solution has a tail which is not smaller than that of the stationary solution of the TAR process.

**Lemma 2.2** *Suppose that condition DC holds and let $(X_k)_{k\in\mathbb{N}_0}$ be the stationary solution of the TAR model (1.1). Let $(\widetilde{Z}_k)_{k\in\mathbb{Z}}$ be an i. i. d. sequence such that $\widetilde{Z}_k = |Z_k| + \alpha$ for $k \in \mathbb{N}_0$. Denote by $(\widetilde{X}_k)_{k\in\mathbb{Z}}$ the unique strictly stationary solution of the causal AR($q$) process*

$$\widetilde{X}_k := \sum_{j=1}^{q} \beta_j \widetilde{X}_{k-j} + \widetilde{Z}_k, \quad k \in \mathbb{Z}. \tag{2.2}$$

*Then $(\widetilde{X}_k)_{k\in\mathbb{Z}}$ has the almost surely convergent MA representation*

$$\widetilde{X}_k = \sum_{j=0}^{\infty} \psi_j \widetilde{Z}_{k-j}, \tag{2.3}$$

*where $\psi_0 = 1$, $0 \le \psi_j < 1$ for $j \in \mathbb{N}$ and $(\psi_j)_{j\in\mathbb{N}_0}$ is bounded by a geometrically decreasing sequence, i. e. $\psi_j \le K\,\gamma^j$ for some $0 < \gamma < 1$ and $K > 0$. Furthermore, for any $m \in \mathbb{N}$, $k_1, \dots, k_m \in \mathbb{N}_0$ and $x_1, \dots, x_m \in \mathbb{R}$ it holds*

$$\mathbb{P}(|X_{k_1}| > x_1, \dots, |X_{k_m}| > x_m) \le \mathbb{P}(\widetilde{X}_{k_1} > x_1, \dots, \widetilde{X}_{k_m} > x_m). \tag{2.4}$$

**Proof.** Since the polynomial $\Phi(z) := 1 - \sum_{j=1}^{q} \beta_j z^j$ has no zeroes for $|z| \le 1$ as a consequence of $\sum_{j=1}^{q} \beta_j < 1$, it follows that the process (2.2) is causal. Expanding $\Phi(z)^{-1} = \sum_{j=0}^{\infty} \psi_j z^j$ in a power series around the origin gives $\psi_0 = 1$, $\psi_1 = \beta_1$ and the recursions $\psi_m = \sum_{j=\max\{0,m-q\}}^{m-1} \beta_{m-j}\psi_j$ for $m \ge 2$. A simple induction argument then shows that $0 \le \psi_j < 1$ for $j \ge 1$ and that $(\psi_m)$ can be dominated by an exponentially decreasing sequence. In particular, $\sum_{j=0}^{\infty} \psi_j^{\min\{1,\eta\}} < \infty$. Since $\mathbb{E}|\widetilde{Z}_1|^{\min\{1,\eta\}} < \infty$ by assumption, this gives almost sure convergence of (2.3).

Define the sequence $(X_k^*)_{k\in\mathbb{N}_0}$ by $X_0^* := |X_0|, \dots, X_{q-1}^* := |X_{q-1}|$, and

$$X_k^* := \sum_{j=1}^{q} \beta_j\, X_{k-j}^* + \widetilde{Z}_k, \quad k \ge q.$$

Then it follows by induction that

$$|X_k| \le X_k^* \quad \forall\, k \in \mathbb{N}_0, \tag{2.5}$$

since

$$|X_k| \le \max_{i=1,\dots,S} \alpha_i + \sum_{j=1}^{q} \max_{i=1,\dots,S} |\beta_{ij}|\,|X_{k-j}| + |Z_k| \le \sum_{j=1}^{q} \beta_j X^*_{k-j} + \widetilde{Z}_k = X^*_k\,.$$

Observe that $(X^*_{t_1+n}, \dots, X^*_{t_m+n})$ converges in distribution as $n \to \infty$ to the stationary solution $(\widetilde{X}_{t_1}, \dots, \widetilde{X}_{t_m})$ of the causal AR($q$) process (2.2), and that the topological boundary of the set $[x_1, \infty) \times \dots \times [x_m, \infty)$ has $\mathbb{P}_{(\widetilde{X}_{t_1},\dots,\widetilde{X}_{t_m})}$ measure zero as a consequence of the absolute continuity of the distribution of $Z_1$. Thus, we conclude from (2.5) that

$$\begin{aligned}
\mathbb{P}(X_{t_1} > x_1, \dots, X_{t_m} > x_m) &= \lim_{n\to\infty} \mathbb{P}(X_{t_1+n} > x_1, \dots, X_{t_m+n} > x_m) \\
\le \limsup_{n\to\infty} \mathbb{P}(X^*_{t_1+n} > x_1, \dots, X^*_{t_m+n} > x_m) &= \mathbb{P}(\widetilde{X}_{t_1} > x_1, \dots, \widetilde{X}_{t_m} > x_m),
\end{aligned}$$

showing (2.4) □

# 3 Regularly varying noise

Recall that a measurable function $f : (0, \infty) \to (0, \infty)$ is said to be *regularly varying* (at $\infty$) *with index* $-\kappa \in \mathbb{R}$, written $f \in \mathcal{R}_{-\kappa}$, if

$$\lim_{x\to\infty} \frac{f(xu)}{f(x)} = u^{-\kappa} \quad \forall\, u > 0. \tag{3.1}$$

Functions in $\mathcal{R}_0$ are also called *slowly varying* functions, and for $\kappa \ge 0$ it holds $f \in \mathcal{R}_{-\kappa}$ if and only if $f(x) = x^{-\kappa} L(x)$ for all $x > 0$ with a slowly varying function $L$. For a random variable $Z$ with distribution function $F_Z$ we also write $Z \in \mathcal{R}_{-\kappa}$ to indicate that $Z$ has a regularly varying tail, i. e. that $\overline{F}_Z \in \mathcal{R}_{-\kappa}$. Examples of distributions having regularly varying tails include Pareto distributions and $\alpha$-stable distributions with $\alpha \in (0, 2)$.

## 3.1 O-regular variation of TAR models

Unlike for linear models such as ARMA processes, stationary solutions of general TAR models with regularly varying noise give only O-regularly varying tails. Recall that a measurable function $f : (0, \infty) \to (0, \infty)$ is called *O-regularly varying* (at $\infty$), if

$$0 < \liminf_{x\to\infty} \frac{f(xu)}{f(x)} \le \limsup_{x\to\infty} \frac{f(xu)}{f(x)} < \infty \quad \forall\, u > 0.$$

Clearly, every regularly varying function is O-regularly varying. For TAR processes with regularly varying noise, we now have:

**Lemma 3.1 (O-regular variation)** *Suppose that conditions TB and DC hold, and let* $(X_k)_{k\in\mathbb{N}_0}$ *be a stationary version of the TAR process as given in (1.1). Suppose further that* $|Z_1| \in \mathcal{R}_{-\kappa}$ *for some* $\kappa > 0$*. Then*

$$p^+2^{-\kappa} \le \liminf_{x\to\infty} \frac{\mathbb{P}(X_0 > x)}{\mathbb{P}(|Z_1| > x)} \le \limsup_{x\to\infty} \frac{\mathbb{P}(X_0 > x)}{\mathbb{P}(|Z_1| > x)} \le \sum_{j=0}^{\infty} \psi_j^{\kappa}, \tag{3.2}$$

*where* $(\psi_j)_{j\in\mathbb{N}_0}$ *is given as in Lemma 2.2. In particular,* $\overline{F}_X$ *is O-regularly varying if* $p^+ > 0$*.*

**Proof.** From Lemma 2.2 and Resnick [24], Lemma 4.24, we obtain

$$\limsup_{x\to\infty} \frac{\mathbb{P}(X_0 > x)}{\mathbb{P}(|Z_1| > x)} \le \limsup_{x\to\infty} \frac{\mathbb{P}(\widetilde{X}_0 > x)}{\mathbb{P}(|Z_1| > x)} = \sum_{j=0}^{\infty} \psi_j^{\kappa}.$$

On the other hand, since $X_k - Z_k$ is independent of $Z_k$, it also holds

$$\liminf_{x\to\infty} \frac{\mathbb{P}(X_0 > x)}{\mathbb{P}(|Z_1| > x)} \ge \liminf_{x\to\infty} \frac{\mathbb{P}(Z_1 > 2x)}{\mathbb{P}(|Z_1| > x)} \mathbb{P}(X_1 - Z_1 > -x) = p^+2^{-\kappa}.$$

This gives (3.2), implying O-regular variation of the tail of $X_0$ if $p^+ > 0$. □

The next proposition shows that without specific assumptions on the partition, regular variation of the stationary distribution cannot be expected, even for a TAR(2,1) model.

**Proposition 3.2** *Let* $(Z_k)_{k\in\mathbb{N}_0}$ *be an i. i. d. sequence such that* $\mathbb{P}(Z_1 > x) \sim x^{-\kappa}$ *as* $x \to \infty$ *for some* $\kappa > 0$ *and that conditions DC and TB hold with* $p^+ > 0$*. For the partition* $J_1 := \bigcup_{m\in\mathbb{N}_0} \left(4^m, 4^{m+1/2}\right]$, $J_2 := \mathbb{R} \setminus J_1$*, consider the TAR(2,1) model*

$$X_k = \begin{cases} \beta_1 X_{k-1} + Z_k, & \text{for } X_{k-1} \in J_1, \\ Z_k, & \text{for } X_{k-1} \in J_2, \end{cases} \qquad k \in \mathbb{N},$$

*where* $0 < \beta_1 < 1$*. Then there are constants* $0 < c_1 < c_2 < \infty$ *such that the stationary solution* $(X_k)_{k\in\mathbb{N}_0}$ *of the TAR(2,1) model with distribution function* $F_X$ *satisfies*

$$c_1 x^{-\kappa} \le \overline{F}_X(x) \le c_2 x^{-\kappa}, \quad x \ge 1, \tag{3.3}$$

*but* $\overline{F}_X$ *is not regularly varying.*

**Proof.** Equation (3.3) follows immediately from Lemma 3.1. Let $L$ be the function satisfying

$$\overline{F}_X(x) = \mathbb{P}(X_0 > x) = L(x)\, x^{-\kappa}, \quad x > 0.$$

By (3.3) it follows that $c_1 \le L(x) \le c_2$ for $x \ge 1$. For $k \in \mathbb{N}_0$, define

$$W_k := \begin{cases} \beta_1 X_k, & \text{for } X_k \in J_1, \\ 0, & \text{for } X_k \in J_2. \end{cases}$$

We will show that the assumption that $\overline{F}_X$ is regularly varying gives an O-regularly varying tail of $W_{k-1}$ which is not regularly varying, and then obtain a contradiction of the tail behavior of $X_k = W_{k-1} + Z_k$, where $W_{k-1}$ and $Z_k$ are independent.

So we assume that $\overline{F}_X$ is regularly varying. Then it follows that $L$ must be slowly varying. We shall show first that there are constants $0 < d_1 \le d_2 < \infty$ such that

$$d_1 \, x^{-\kappa} \le \mathbb{P}(W_k > x) \le d_2 x^{-\kappa}, \quad x \ge 1. \tag{3.4}$$

Here, the right-hand inequality follows easily from $\mathbb{P}(W_k > x) \le \mathbb{P}(\beta_1 X_k > x)$ and Equation (3.3). For the left-hand inequality, observe that the regular variation of $\overline{F}_X$ implies

$$\begin{aligned} \lim_{x\to\infty} \frac{\mathbb{P}(X_k \in (x, 2x])}{\mathbb{P}(X_k \in (x, 4x])} &= \lim_{x\to\infty} \frac{\mathbb{P}(X_k > x) - \mathbb{P}(X_k > 2x)}{\mathbb{P}(X_k > x) - \mathbb{P}(X_k > 4x)} \\ &= \lim_{x\to\infty} \frac{[\mathbb{P}(X_k > x) - \mathbb{P}(X_k > 2x)]/\mathbb{P}(X_k > x)}{[\mathbb{P}(X_k > x) - \mathbb{P}(X_k > 4x)]/\mathbb{P}(X_k > x)} \\ &= \frac{1 - 2^{-\kappa}}{1 - 4^{-\kappa}}. \end{aligned}$$

For $x > 0$ denote by $m(x)$ the unique non-negative integer such that $4^{m(x)-1} < x/\beta_1 \le 4^{m(x)}$. Then, given $\varepsilon > 0$, it follows that for large enough $x \ge x(\varepsilon)$,

$$\begin{aligned} \mathbb{P}(W_k > x) &= \sum_{m=0}^{\infty} \mathbb{P}(X_k > x/\beta_1, X_k \in \left(4^m, 4^{m+1/2}\right]) \\ &\ge \sum_{m=m(x)}^{\infty} \mathbb{P}(X_k \in \left(4^m, 4^{m+1/2}\right]) \\ &\ge (1-\varepsilon) \frac{1 - 2^{-\kappa}}{1 - 4^{-\kappa}} \sum_{m=m(x)}^{\infty} \mathbb{P}(X_k \in \left(4^m, 4^{m+1}\right]) \\ &= (1-\varepsilon) \frac{1 - 2^{-\kappa}}{1 - 4^{-\kappa}} \mathbb{P}(X_k > 4^{m(x)}). \end{aligned}$$

The left-hand inequality in (3.4) then follows from the corresponding one in (3.3) and $4^{m(x)} < 4x/\beta_1$. Thus, it follows from (3.4) that we can write

$$\mathbb{P}(W_k > x) = r(x) \, x^{-\kappa}, \quad x \in \mathbb{R}, \tag{3.5}$$

where $d_1 \le r(x) \le d_2$ for $x \ge 1$. Now, let $(x_m)_{m\in\mathbb{N}}$ be a sequence of numbers such that $x_m/\beta_1 \in [4^{m+4/6}, 4^{m+5/6}]$. Then $\lambda x_m/\beta_1 \in (4^{m+1/2}, 4^{m+1})$ for every $\lambda \in (4^{-1/6}, 4^{1/6})$, so that

$$\mathbb{P}(W_k > \lambda x_m) = \mathbb{P}(X_k > \lambda x_m/\beta_1, X_k \in J_1) = \mathbb{P}(X_k > x_m/\beta_1, X_k \in J_1) = \mathbb{P}(W_k > x_m),$$

giving

$$r(\lambda x_m) = \lambda^\kappa\, r(x_m), \quad \forall\, \lambda \in (4^{-1/6}, 4^{1/6}), \quad m \in \mathbb{N}. \tag{3.6}$$

This implies in particular that $r$ is not slowly varying, so that the distribution function of $W_k$ cannot have a regularly varying tail by (3.5).

Choose $\delta > 0$ such that $4^{-1/6} < 1-\delta < 1+\delta < 4^{1/6}$, and let $x' := (1+\delta)x$, $x'' := (1-\delta)x$. Write $\mathbb{P}(Z_1 > x) = q(x)x^{-\kappa}$, so that $\lim_{x\to\infty} q(x) = 1$. Then, with exactly the same proof as in Feller [17], pp. 278, it follows that for given $\varepsilon > 0$ and large enough $x \ge x(\varepsilon)$, the tail of $X_k = W_{k-1} + Z_k$ satisfies

$$(1-\varepsilon)\,(r(x') + q(x'))\,(x')^{-\kappa} \le \mathbb{P}(X_k > x) \le (1+\varepsilon)\,(r(x'') + q(x''))\,(x'')^{-\kappa}. \tag{3.7}$$

Choosing $x_m$ as before, the left-hand inequality of (3.7) together with (3.6) show that for large enough $m$,

$$\begin{aligned} &\frac{\mathbb{P}(X_k > x_m)}{(r(x_m) + q(x_m))x_m^{-\kappa}} \\ &\quad \ge \quad (1-\varepsilon)(1+\delta)^{-\kappa}\, \frac{r(x_m)(1+\delta)^\kappa + q(x'_m)}{r(x_m) + q(x_m)} \\ &\quad = \quad (1-\varepsilon)(1+\delta)^{-\kappa} \left(1 + \frac{r(x_m)((1+\delta)^\kappa - 1) + q(x'_m) - q(x_m)}{r(x_m) + q(x_m)}\right). \end{aligned}$$

Using $d_1 \le r(x) \le d_2$ for $x \ge 1$, $\lim_{x\to\infty} q(x) = 1$ and $\lim_{\delta\to 0}(1+\delta)^\kappa - 1 = 0$, we conclude that

$$\liminf_{m\to\infty} \frac{\mathbb{P}(X_k > x_m)}{(r(x_m) + q(x_m))x_m^{-\kappa}} \ge 1.$$

A similar argument holds for the limes superior, so that

$$L(x_m) \sim (r(x_m) + q(x_m)), \quad m \to \infty. \tag{3.8}$$

Taking for $x_m$ the sequences $u_m := 4^{m+9/12}\beta_1$ and $v_m := 4^{m+10/12}\beta_1 = 4^{1/12}u_m$, it follows from (3.6) and (3.8) that

$$\frac{L(v_m)}{L(u_m)} \sim \frac{r(v_m) + q(v_m)}{r(u_m) + q(u_m)} \sim \frac{4^{\kappa/12} r(u_m) + q(u_m)}{r(u_m) + q(u_m)} = 1 + \frac{(4^{\kappa/12} - 1)r(u_m)}{r(u_m) + q(u_m)}, \quad m \to \infty,$$

and the latter does not converge to 1 as $m \to \infty$, since $d_1 \le r(u_m) \le d_2$ and $\lim_{x\to\infty} q(x) = 1$. Hence, $L$ cannot be slowly varying, contradicting the regular variation of $\overline{F}_X$. □

## 3.2 Regular variation of TAR(S,1) models with specific partitions

In this and the next subsection we restrict our attention to stationary TAR($S$,1) models with representation

$$X_k = \sum_{i=1}^{S} \{\alpha_i + \beta_i X_{k-1}\} \mathbf{1}_{\{X_{k-1} \in J_i\}} + Z_k \quad \text{for } k \in \mathbb{N}, \tag{3.9}$$

where $J_1 = (-\infty, r_1]$, $J_2 = (r_2, \infty)$ for some $r_1, r_2 \in \mathbb{R}$, $r_1 \le r_2$, and $\{J_i : i = 3, \dots, S\}$ is a measurable partition of $(r_1, r_2]$. For this model we are able to compute the tail behavior explicitly as we will show in the next lemma.

**Lemma 3.3 (Regular variation)** *Suppose that conditions TB and DC hold, and let $(X_k)_{k \in \mathbb{N}_0}$ be a stationary version of the TAR(S,1) process as given in (3.9). Suppose further that $|Z_1| \in \mathcal{R}_{-\kappa}$ for some $\kappa > 0$. Then $|X_0| \in \mathcal{R}_{-\kappa}$. More precisely, denoting*

$$\beta_i^+ := \begin{cases} \beta_i, & \beta_i > 0, \\ 0, & \beta_i \le 0, \end{cases} \quad \text{and} \quad \beta_i^- := \begin{cases} |\beta_i|, & \beta_i < 0, \\ 0, & \beta_i \ge 0, \end{cases}$$

*for $i = 1, 2$, it holds*

$$\lim_{x \to \infty} \frac{\mathbb{P}(X_0 > x)}{\mathbb{P}(|Z_1| > x)} = \frac{p^+ + p^-(\beta_1^-)^\kappa}{1 - (\beta_2^+)^\kappa - (\beta_1^-)^\kappa (\beta_2^-)^\kappa} =: \widetilde{p}^+ \tag{3.10}$$

*and*

$$\lim_{x \to \infty} \frac{\mathbb{P}(X_0 < -x)}{\mathbb{P}(|Z_1| > x)} = \frac{p^- + p^+(\beta_2^-)^\kappa}{1 - (\beta_1^+)^\kappa - (\beta_2^-)^\kappa (\beta_1^-)^\kappa} =: \widetilde{p}^-. \tag{3.11}$$

*In particular, $X_0 \in \mathcal{R}_{-\kappa}$ if $\widetilde{p}^+ > 0$. Furthermore, $\widetilde{p}^+ \ge p^+$ and $\widetilde{p}^- \ge p^-$.*

**Proof.** Let $x > 0$ be fixed and $(a_n)_{n \in \mathbb{N}}$ be a sequence of positive real numbers tending to $\infty$ as $n \to \infty$. Then for any $\delta > 0$, we can write

$$\begin{aligned} \mathbb{P}(X_1 > x a_n) &= \mathbb{P}(X_1 > x a_n, |Z_1| > \delta a_n, |X_0| > \delta a_n) \\ &\quad + \mathbb{P}(X_1 > x a_n, |Z_1| > \delta a_n, |X_0| \le \delta a_n) \\ &\quad + \mathbb{P}(X_1 > x a_n, |Z_1| \le \delta a_n, |X_0| \le \delta a_n) \\ &\quad + \mathbb{P}(X_1 > x a_n, |Z_1| \le \delta a_n, X_0 > \delta a_n) \\ &\quad + \mathbb{P}(X_1 > x a_n, |Z_1| \le \delta a_n, X_0 < -\delta a_n) \\ &= (I) + (II) + (III) + (IV) + (V), \quad \text{say.} \end{aligned} \tag{3.12}$$

We shall study the tail behavior of the five summands of (3.12). Using the independence of $Z_1$ and $X_0$, we obtain for the first term

$$\lim_{n \to \infty} \frac{(I)}{\mathbb{P}(|Z_1| > x a_n)} \le \lim_{n \to \infty} \frac{\mathbb{P}(|Z_1| > \delta a_n) \mathbb{P}(|X_0| > \delta a_n)}{\mathbb{P}(|Z_1| > x a_n)} = 0. \tag{3.13}$$

For the second term of (3.12), observe that

$$\begin{aligned}\limsup_{n\to\infty}\frac{\mathbb{P}(X_1>xa_n,|Z_1|>\delta a_n,|X_0|\le\delta a_n)}{\mathbb{P}(|Z_1|>xa_n)} &\le \limsup_{n\to\infty}\frac{\mathbb{P}(Z_1>(x-\beta\delta)a_n-\alpha)}{\mathbb{P}(|Z_1|>xa_n)}\\ &= p^+\frac{(x-\beta\delta)^{-\kappa}}{x^{-\kappa}}.\end{aligned}$$

On the other hand,

$$\begin{aligned}&\liminf_{n\to\infty}\frac{\mathbb{P}(X_1>xa_n,|Z_1|>\delta a_n,|X_0|\le\delta a_n)}{\mathbb{P}(|Z_1|>xa_n)}\\ &\ge \liminf_{n\to\infty}\frac{\mathbb{P}(Z_1>(x+\beta\delta)a_n+\alpha)-\mathbb{P}(Z_1>(x+\beta\delta)a_n+\alpha,|X_0|>\delta a_n)}{\mathbb{P}(|Z_1|>xa_n)}\\ &= p^+\frac{(x+\beta\delta)^{-\kappa}}{x^{-\kappa}},\end{aligned}$$

for $0<\delta<x$, and we conclude

$$\lim_{\delta\downarrow 0}\liminf_{n\to\infty}\frac{(II)}{\mathbb{P}(|Z_1|>xa_n)}=\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{(II)}{\mathbb{P}(|Z_1|>xa_n)}=p^+. \tag{3.14}$$

The third term of (3.12) is zero provided $\delta<x/2$. For the investigation of (IV) and (V) we define

$$\begin{aligned}\overline{A} &= \limsup_{u\to\infty}\frac{\mathbb{P}(X_0>u)}{\mathbb{P}(|Z_1|>u)}, & \underline{A} &= \liminf_{u\to\infty}\frac{\mathbb{P}(X_0>u)}{\mathbb{P}(|Z_1|>u)},\\ \overline{B} &= \limsup_{u\to\infty}\frac{\mathbb{P}(X_0<-u)}{\mathbb{P}(|Z_1|>u)}, & \underline{B} &= \liminf_{u\to\infty}\frac{\mathbb{P}(X_0<-u)}{\mathbb{P}(|Z_1|>u)}.\end{aligned}$$

All these terms are finite by Lemma 3.1. Then we obtain

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{(IV)}{\mathbb{P}(|Z_1|>xa_n)}\le\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{\mathbb{P}(\alpha_2+\beta_2^+X_0>(x-\delta)a_n)}{\mathbb{P}(|Z_1|>xa_n)}=\overline{A}(\beta_2^+)^\kappa, \tag{3.15}$$

and similarly,

$$\begin{aligned}\lim_{\delta\downarrow 0}\liminf_{n\to\infty}\frac{(IV)}{\mathbb{P}(|Z_1|>xa_n)} &\ge \lim_{\delta\downarrow 0}\liminf_{n\to\infty}\frac{\mathbb{P}(\alpha_2+\beta_2^+X_0>(x+\delta)a_n)}{\mathbb{P}(|Z_1|>xa_n)}\mathbb{P}(|Z_1|\le\delta a_n)\\ &= \underline{A}(\beta_2^+)^\kappa.\end{aligned} \tag{3.16}$$

The bounds of (V) are

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{(V)}{\mathbb{P}(|Z_1|>xa_n)}\le\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{\mathbb{P}(\alpha_1-\beta_1^-X_0>(x-\delta)a_n)}{\mathbb{P}(|Z_1|>xa_n)}=\overline{B}(\beta_1^-)^\kappa, \tag{3.17}$$

and

$$\begin{aligned}\lim_{\delta\downarrow 0}\liminf_{n\to\infty}\frac{(V)}{\mathbb{P}(|Z_1|>xa_n)} &\ge \lim_{\delta\downarrow 0}\liminf_{n\to\infty}\frac{\mathbb{P}(\alpha_1-\beta_1^-X_0>(x+\delta)a_n)}{\mathbb{P}(|Z_1|>xa_n)}\mathbb{P}(|Z_1|\le\delta a_n)\\ &= \underline{B}(\beta_1^-)^\kappa.\end{aligned} \tag{3.18}$$

Then (3.12)–(3.18) give

$$p^+ + \underline{A}(\beta_2^+)^\kappa + \underline{B}(\beta_1^-)^\kappa \le \underline{A} \le \overline{A} \le p^+ + \overline{A}(\beta_2^+)^\kappa + \overline{B}(\beta_1^-)^\kappa. \tag{3.19}$$

Since

$$-X_k = \sum_{i=1}^{S} (-\alpha_i) + \beta_i(-X_{k-1}) \mathbf{1}_{\{-X_{k-1} \in -J_i\}} - Z_k,$$

we obtain by symmetry

$$p^- + \underline{B}(\beta_1^+)^\kappa + \underline{A}(\beta_2^-)^\kappa \le \underline{B} \le \overline{B} \le p^- + \overline{B}(\beta_1^+)^\kappa + \overline{A}(\beta_2^-)^\kappa. \tag{3.20}$$

If $\beta_1 \ge 0$, then (3.19) gives

$$\underline{A} = \overline{A} = \frac{p^+}{1 - (\beta_2^+)^\kappa}.$$

In the case $\beta_1 < 0$ we obtain by (3.20),

$$p^- + \underline{A}(\beta_2^-)^\kappa \le \underline{B} \quad \text{and} \quad \overline{B} \le p^- + \overline{A}(\beta_2^-)^\kappa. \tag{3.21}$$

Inserting (3.21) in (3.19) yields

$$\frac{p^+ + p^-(\beta_1^-)^\kappa}{1 - (\beta_2^+)^\kappa - (\beta_2^-)^\kappa(\beta_1^-)^\kappa} \le \underline{A} \le \overline{A} \le \frac{p^+ + p^-(\beta_1^-)^\kappa}{1 - (\beta_2^+)^\kappa - (\beta_2^-)^\kappa(\beta_1^-)^\kappa},$$

which gives the result $\underline{A} = \overline{A} = \widetilde{p}^+$. Inserting this in (3.20) gives also $\underline{B} = \overline{B} = \widetilde{p}^-$. That $\widetilde{p}^+ \ge p^+$ and $\widetilde{p}^- \ge p^-$ is clear. □

Denote by $\|\cdot\|$ the maximum norm and by $\mathbb{S}^m = \{\mathbf{x} \in \mathbb{R}^{m+1} : \|\mathbf{x}\| = 1\}$ the unit sphere with respect to the maximum norm in $\mathbb{R}^{m+1}$. Recall that a random vector $\mathbf{Y} \in \mathbb{R}^{m+1}$ is *multivariate regularly varying with index* $-\kappa < 0$ (sometimes also termed *with index* $\kappa > 0$) if there exists a random vector $\Theta$ with values in $\mathbb{S}^m$ such that for every $x > 0$, the measures

$$\frac{\mathbb{P}(\|\mathbf{Y}\| > ux,\, \mathbf{Y}/\|\mathbf{Y}\| \in \cdot)}{\mathbb{P}(\|\mathbf{Y}\| > u)}$$

on $\mathcal{B}(\mathbb{S}^m)$ converge weakly to a measure $x^{-\kappa}\mathbb{P}(\Theta \in \cdot)$ as $u \to \infty$. The distribution of $\Theta$ is called the *spectral measure of* $\mathbf{Y}$ (with respect to the maximum norm). Multivariate regular variation can be defined with respect to any other norm on $\mathbb{R}^{m+1}$, but since all these definitions are equivalent (only the form of the spectral measure differs), we have chosen to work with the maximum norm which is particularly convenient for our calculations. It is further known that a random vector $\mathbf{Y}$ is regularly varying with index $-\kappa$ if and only if there exists a non-zero Radon measure $\sigma$ on $\overline{\mathbb{R}}^{m+1} \setminus \{\mathbf{0}\}$ with $\sigma(\overline{\mathbb{R}}^{m+1} \setminus \mathbb{R}^{m+1}) = 0$ and a sequence $(a_n)_{n\in\mathbb{N}}$ of positive numbers increasing to $\infty$ such that

$$n\mathbb{P}(a_n^{-1}\mathbf{Y} \in \cdot) \overset{v}{\Longrightarrow} \sigma(\cdot) \quad \text{as } n \to \infty, \tag{3.22}$$

where $\overset{v}{\Longrightarrow}$ denotes vague convergence on $\mathcal{B}(\overline{\mathbb{R}}^{m+1}\setminus\{\mathbf{0}\})$. For further information regarding multivariate regular variation, we refer to Resnick [23] or Basrak et al. [2].

Using Lemma 3.3 we will prove that the finite dimensional distributions $\mathbf{X}^{(m)} = (X_0, X_1, \ldots, X_m)^T \in \mathbb{R}^{m+1}$ of the stationary TAR($S$,1) process (3.9) are multivariate regularly varying for every $m \in \mathbb{N}_0$. For this, we need the definition of the following matrices: let

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} \beta_2^+ & \beta_1^- \\ \beta_2^- & \beta_1^+ \end{pmatrix} \in \mathbb{R}^{2\times 2},$$

and for $m \in \mathbb{N}_0$ define

$$\mathbf{C}^{(m)} = \begin{pmatrix} I & 0 & 0 & \cdots & 0 \\ B & I & 0 & \ddots & \vdots \\ B^2 & B & I & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ B^m & B^{m-1} & \cdots & B & I \end{pmatrix}, \qquad \mathbf{S}^{(m)} = \begin{pmatrix} 1 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & -1 & & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & -1 \end{pmatrix},$$

where $\mathbf{C}^{(m)} \in \mathbb{R}^{2(m+1)\times 2(m+1)}$ and $\mathbf{S}^{(m)} \in \mathbb{R}^{(m+1)\times 2(m+1)}$, and $B^m$ denotes the $m$'th power of $B$ (with $B^0 = I$). Finally, define

$$\widetilde{\mathbf{C}}^{(m)} := \mathbf{S}^{(m)}\mathbf{C}^{(m)} =: (\mathbf{c}_0^+, \mathbf{c}_0^-, \ldots, \mathbf{c}_m^+, \mathbf{c}_m^-) \in \mathbb{R}^{(m+1)\times 2(m+1)}. \tag{3.23}$$

The vectors $\mathbf{c}_k^+$ and $\mathbf{c}_k^-$ ($k = 0, \ldots, m$) will be used to describe the spectral measure of $(X_0, \ldots, X_m)^T$. Insight into their structure can be obtained by writing

$$B^m = \begin{pmatrix} b_{11}^{(m)} & b_{12}^{(m)} \\ b_{21}^{(m)} & b_{22}^{(m)} \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

for $m \in \mathbb{N}_0$. Then $b_{ij}^{(m)} \geq 0$ for $m \in \mathbb{N}_0$, $i, j = 1, 2$, and $B^m$ has at most one non-zero element in every column. Putting

$$b_1^{(m)} := b_{11}^{(m)} - b_{21}^{(m)} \quad \text{and} \quad b_2^{(m)} := b_{12}^{(m)} - b_{22}^{(m)} \quad \text{for } m \in \mathbb{N}_0, \tag{3.24}$$

we see that the $j$'th component $(\mathbf{c}_k^{\pm})_j$ of $\mathbf{c}_k^{\pm}$ satisfies

$$\begin{aligned} &(\mathbf{c}_k^+)_j = 0 \quad \text{for } j \leq k, && (\mathbf{c}_k^+)_j = b_1^{(j-k-1)} \quad \text{for } k+1 \leq j \leq m+1, \\ &(\mathbf{c}_k^-)_j = 0 \quad \text{for } j \leq k, && (\mathbf{c}_k^-)_j = b_2^{(j-k-1)} \quad \text{for } k+1 \leq j \leq m+1. \end{aligned}$$

It is easy to check that $(|b_1^{(m)}|)_{m\in\mathbb{N}_0}$ and $(|b_2^{(m)}|)_{m\in\mathbb{N}_0}$ are decreasing sequences, and that $|b_1^{(m)}| \le \beta^m$ and $|b_2^{(m)}| \le \beta^m$. In particular, it follows that $\mathbf{c}_k^\pm \in \mathbb{S}^m$ for $0 \le k \le m$. With these preparations, we can now show that the stationary version of the TAR$(S,1)$-model (3.9) is multivariate regularly varying.

**Theorem 3.4 (Multivariate regular variation)** *Suppose that conditions TB and DC hold, and let $(X_k)_{k\in\mathbb{N}_0}$ be a stationary version of the TAR(S,1) process as given in (3.9). Suppose further that $|Z_1| \in \mathcal{R}_{-\kappa}$ for some $\kappa > 0$. Then $\mathbf{X}^{(m)} = (X_0,\dots,X_m)^T$ is multivariate regularly varying with index $-\kappa$, and its spectral measure with respect to the maximum norm is given by*

$$\mathbb{P}(\Theta^{(m)} \in \cdot) = \frac{1}{\widetilde{p}^+ + \widetilde{p}^- + m}\left[\widetilde{p}^+\,\mathbf{1}_{\{\mathbf{c}_0^+\in\cdot\}} + \widetilde{p}^-\,\mathbf{1}_{\{\mathbf{c}_0^-\in\cdot\}} + \sum_{j=1}^m \left(p^+\,\mathbf{1}_{\{\mathbf{c}_j^+\in\cdot\}} + p^-\,\mathbf{1}_{\{\mathbf{c}_j^-\in\cdot\}}\right)\right] \quad (3.25)$$

*where $\widetilde{p}^+$ and $\widetilde{p}^-$ are defined as in Lemma 3.3 and $\mathbf{c}_j^\pm$ $(j = 0,\dots,m)$ as above.*

**Proof.** We define $\mathbf{Z}^{(m)} = \left(X_0^+, X_0^-, Z_1^+, Z_1^-, \dots, Z_m^+, Z_m^-\right)^T \in \mathbb{R}^{2(m+1)}$ for $m \in \mathbb{N}_0$, which by Lemma 3.3 is multivariate regularly varying of index $-\kappa$ with spectral measure

$$\mathbb{P}(\widetilde{\Theta}^{(m)} \in \cdot) = \frac{1}{\widetilde{p}^+ + \widetilde{p}^- + m}\left[\widetilde{p}^+\,\mathbf{1}_{\{\mathbf{e}_1\in\cdot\}} + \widetilde{p}^-\,\mathbf{1}_{\{\mathbf{e}_2\in\cdot\}} + \sum_{j=1}^m \left(p^+\,\mathbf{1}_{\{\mathbf{e}_{2j+1}\in\cdot\}} + p^-\,\mathbf{1}_{\{\mathbf{e}_{2j+2}\in\cdot\}}\right)\right],$$

where $\mathbf{e}_j \in \mathbb{R}^{2(m+1)}$, $j = 1,\dots,2(m+1)$, is one in the $j$-th component and else zero.

Furthermore, we define $(Y_k)_{k\in\mathbb{N}_0}$ by $Y_0 := X_0$ and

$$Y_k = \beta_2 Y_{k-1}\,\mathbf{1}_{\{Y_{k-1}>0\}} + \beta_1 Y_{k-1}\,\mathbf{1}_{\{Y_{k-1}<0\}} + Z_k = \beta_2 Y_{k-1}^+ - \beta_1 Y_{k-1}^- + Z_k\,, \quad k \in \mathbb{N}.$$

Let

$$\begin{aligned} \widetilde{\mathbf{W}} &:= \widetilde{\mathbf{C}}^{(m)}\mathbf{Z}^{(m)} =: (\widetilde{W}_0,\dots,\widetilde{W}_m), & \widetilde{\mathbf{Y}} &:= (Y_0,\dots,Y_m), \\ \mathbf{W} &:= \mathbf{C}^{(m)}\mathbf{Z}^{(m)} =: (W_0^+, W_0^-,\dots,W_m^+, W_m^-), & \mathbf{Y} &:= (Y_0^+, Y_0^-,\dots,Y_m^+, Y_m^-). \end{aligned}$$

Since $\widetilde{\mathbf{W}}$ is obtained from $\mathbf{Z}^{(m)}$ by a linear transformation, it is easy to see that $\widetilde{\mathbf{W}}$ is again multivariate regularly varying with spectral measure $\Theta^{(m)}$ as given in (3.25), see e.g. Proposition A.1 in Basrak et al. [2] and Fasen [14], Lemma 2.1. We shall show that $\widetilde{\mathbf{Y}}$ is regularly varying with the same spectral measure as $\widetilde{\mathbf{W}}$. Further, observe that by the definition of the matrix $\mathbf{C}^{(m)}$, it holds

$$\begin{aligned} &W_0^+ = X_0^+, \quad W_0^- = X_0^-, \quad \text{and} \\ &W_{k+1}^+ = \beta_2^+ W_k^+ + \beta_1^- W_k^- + Z_{k+1}^+, \quad W_{k+1}^- = \beta_2^- W_k^+ + \beta_1^+ W_k^- + Z_{k+1}^- \end{aligned}$$

for $k = 0,\dots,m-1$.

Let $(a_n)_{n\in\mathbb{N}}$ be a sequence of positive numbers increasing to $\infty$ such that $\lim_{n\to\infty} n\mathbb{P}(|Z_1| > a_n) = 1$, and let $0 < \delta < 1$. Then

$$\begin{aligned}\mathbb{P}(\|\widetilde{\mathbf{W}} - \widetilde{\mathbf{Y}}\| > 2\delta a_n) &\leq \mathbb{P}(\|\mathbf{W} - \mathbf{Y}\| > \delta a_n) \\ &\leq \sum_{k=1}^{m} [\mathbb{P}(|W_k^+ - Y_k^+| > \delta a_n) + \mathbb{P}(|W_k^- - Y_k^-| > \delta a_n)].\end{aligned} \tag{3.26}$$

Let $\Delta_k^\pm := W_k^\pm - Y_k^\pm$ for $k \in \mathbb{N}_0$. First, we will show that

$$\mathbb{P}(|\Delta_1^+| > \delta a_n) = o(\mathbb{P}(|Z_1| > a_n)) \quad \text{as } n \to \infty, \tag{3.27}$$

for any $\delta > 0$, and hence, by symmetry the same arguments lead to

$$\mathbb{P}(|\Delta_1^-| > \delta a_n) = o(\mathbb{P}(|Z_1| > a_n)) \quad \text{as } n \to \infty.$$

Then we use induction to prove that for any $\delta > 0$,

$$\mathbb{P}(|\Delta_k^\pm| > \delta a_n) = o(\mathbb{P}(|Z_1| > a_n)) \quad \text{as } n \to \infty,\ k \in \mathbb{N}. \tag{3.28}$$

Since the result is trivial if $\beta = 0$, i. e. $\beta_1 = \beta_2 = 0$, we shall assume that $\beta \neq 0$ from now on. In order to study (3.27), let $0 < \beta\delta_2 < \delta_1 < \delta$. Then

$$\begin{aligned}\mathbb{P}(|\Delta_1^+| > \delta a_n) &\leq \mathbb{P}(|\Delta_1^+| > \delta a_n, Z_1 > \delta_1 a_n, |Y_0| \leq \delta_2 a_n) \\ &\quad + \mathbb{P}(|\Delta_1^+| > \delta a_n, Z_1 < -\delta_1 a_n, |Y_0| \leq \delta_2 a_n) \\ &\quad + \mathbb{P}(|\Delta_1^+| > \delta a_n, |Z_1| \leq \delta_1 a_n, Y_0 > \delta_2 a_n) \\ &\quad + \mathbb{P}(|\Delta_1^+| > \delta a_n, |Z_1| \leq \delta_1 a_n, Y_0 < -\delta_2 a_n) \\ &\quad + \mathbb{P}(|\Delta_1^+| > \delta a_n, |Z_1| \leq \delta_1 a_n, |Y_0| \leq \delta_2 a_n) \\ &\quad + \mathbb{P}(|Z_1| > \delta_1 a_n, |Y_0| > \delta_2 a_n) \\ &=: (I) + (II) + (III) + (IV) + (V) + (VI), \text{ say.}\end{aligned} \tag{3.29}$$

The summand $(I)$ can be estimated by

$$(I) \leq \mathbb{P}(\beta|Y_0| > \delta a_n, Z_1 > \delta_1 a_n, |Y_0| \leq \delta_2 a_n) = 0, \tag{3.30}$$

since $\beta\delta_2 < \delta$. Similarly, we obtain

$$(II) \leq \mathbb{P}(\beta|Y_0| > \delta a_n, |Y_0| \leq \delta_2 a_n) = 0. \tag{3.31}$$

That also $(III) = 0$ can be seen from

$$\begin{aligned}(III) &\leq \mathbb{P}(|\Delta_1^+| > \delta a_n, |Z_1| \leq \delta_1 a_n, |\beta_2|Y_0 > \delta_1 a_n) \\ &\quad + \mathbb{P}(|\Delta_1^+| > \delta a_n, |Z_1| \leq \delta_1 a_n, \delta_2|\beta_2| a_n < |\beta_2|Y_0 \leq \delta_1 a_n) \\ &= 0 + 0 = 0\end{aligned} \tag{3.32}$$

if $\beta_2 \neq 0$, and from $W_1^+ = Y_1^+$ if $\beta_2 = 0$ and $Y_0 > 0$. By symmetry, also

$$(IV) = 0. \tag{3.33}$$

Provided $\delta_1, \delta_2$ are small, we further have

$$(V) = 0, \tag{3.34}$$

and finally we estimate

$$(VI) = \mathbb{P}(|Z_1| > \delta_1 a_n)\mathbb{P}(|Y_0| > \delta_2 a_n) = o(\mathbb{P}(|Z_1| > a_n)) \quad \text{as } n \to \infty. \tag{3.35}$$

Hence, (3.29)–(3.35) give

$$\mathbb{P}(|\Delta_1^+| > \delta a_n) = o(\mathbb{P}(|Z_1| > a_n)) \quad \text{as } n \to \infty. \tag{3.36}$$

Next, we assume (3.28) holds for some $k \in \mathbb{N}$ and every $\delta > 0$. Define $\overline{W}_{k+1}^+ := \beta_2^+ Y_k^+ + \beta_1^- Y_k^- + Z_{k+1}^+$ for $k \in \mathbb{N}_0$, and let $\delta_3 \in (\beta, 1)$. Then

$$\begin{aligned}
\mathbb{P}(|\Delta_{k+1}^+| > \delta a_n) \quad \leq \quad & \mathbb{P}(|\Delta_k^+| \leq \delta/2 a_n, |\Delta_k^-| \leq \delta/2 a_n, |\overline{W}_{k+1}^+ - Y_{k+1}^+| > (1-\delta_3)\delta a_n) \\
& + \mathbb{P}(|\Delta_k^+| \leq \delta/2 a_n, |\Delta_k^-| \leq \delta/2 a_n, |W_{k+1}^+ - \overline{W}_{k+1}^+| > \delta_3 \delta a_n) \\
& + \mathbb{P}(|\Delta_k^+| > \delta/2 a_n) + \mathbb{P}(|\Delta_k^-| > \delta/2 a_n) \\
=: \quad & (VII) + (VIII) + (XI) + (X), \quad \text{say.}
\end{aligned} \tag{3.37}$$

With exactly the same reasoning that led to (3.36) we obtain

$$(VII) \leq \mathbb{P}(|\overline{W}_{k+1}^+ - Y_{k+1}^+| > (1-\delta_3)\delta a_n) = o(\mathbb{P}(|Z_1| > a_n)) \quad \text{as } n \to \infty. \tag{3.38}$$

On the other hand,

$$(VIII) = \mathbb{P}(|\Delta_k^+| \leq \delta/2 a_n, |\Delta_k^-| \leq \delta/2 a_n, |\beta_2^+ \Delta_k^+ + \beta_1^- \Delta_k^-| > \delta_3 \delta a_n) = 0. \tag{3.39}$$

Furthermore,

$$(XI) = o(\mathbb{P}(|Z_1| > a_n)) \quad \text{and} \quad (X) = o(\mathbb{P}(|Z_1| > a_n)) \quad \text{as } n \to \infty, \tag{3.40}$$

by induction hypothesis. Equations (3.37)–(3.40) then give (3.28). Using (3.26) and (3.28) we hence obtain

$$\mathbb{P}(\|\widetilde{\mathbf{W}} - \widetilde{\mathbf{Y}}\| > 2\delta a_n) \leq \mathbb{P}(\|\mathbf{W} - \mathbf{Y}\| > \delta a_n) = o(\mathbb{P}(|Z_1| > a_n)) \quad \text{as } n \to \infty.$$

Since $\delta$ can be arbitrary small, $\widetilde{\mathbf{Y}}$ is multivariate regularly varying with the same spectral measure as $\widetilde{\mathbf{W}}$. Finally, $\|\mathbf{X}^{(m)} - \widetilde{\mathbf{Y}}\| \leq \sum_{k=1}^m \Big( (2\beta)^k \max\{|r_1|, |r_2|\} + \alpha \sum_{j=0}^{k-1} (2\beta)^j \Big)$, which follows again by induction, shows that $\mathbf{X}^{(m)}$ is multivariate regularly varying with the same spectral measure as $\widetilde{\mathbf{Y}}$ and hence, as $\widetilde{\mathbf{W}}$. □

## 3.3 Extremal behavior of the TAR(S,1) model with specific partitions

For a locally compact Hausdorff space $E$ we denote by $M_P(E)$ the space of all point measures on $E$. A point process is then a random element with values in $M_P(E)$, and a Poisson point process (= Poisson random measure) with mean measure $\vartheta$ will be denoted by PRM($\vartheta$). In extreme value theory, the space $E$ is often $\overline{\mathbb{R}} \setminus \{0\}$, $\mathbb{R}$, $[0,\infty) \times (\overline{\mathbb{R}} \setminus \{0\})$ or $[0,\infty) \times \mathbb{R}$, and the extremal behavior of a stationary sequence $(\xi_k)_{k\in\mathbb{N}}$ is described by the weak limit of the point process $\sum_{k=1}^{\infty} \varepsilon_{a_n^{-1}(\xi_k - b_n)}$ in $M_P(\overline{\mathbb{R}} \setminus \{0\})$ or $M_P(\mathbb{R})$, respectively, as $n \to \infty$, or still more informative by the weak limit of the point process $\sum_{k=1}^{\infty} \varepsilon_{(k/n, a_n^{-1}(\xi_k - b_n))}$ in $M_P([0,\infty)\times(\overline{\mathbb{R}}\setminus\{0\}))$ or $M_P([0,\infty)\times\mathbb{R})$, respectively, as $n \to \infty$. Here, $a_n > 0$ and $b_n \in \mathbb{R}$ are the norming constants of an associated i. i. d. sequence $(\widetilde{\xi}_k)_{k\in\mathbb{N}}$ with distribution $\widetilde{\xi}_1 \overset{d}{=} \xi_1$, such that

$$\exp\left(-\lim_{n\to\infty} n\mathbb{P}(\xi_1 > a_n x + b_n)\right) = \lim_{n\to\infty} \mathbb{P}\left(a_n^{-1}\left(\bigvee_{k=1}^{n} \widetilde{\xi}_k - b_n\right) \leq x\right) = G(x) \qquad (3.41)$$

for $x$ in the support of $G$, where $G$ is an extreme value distribution, i. e. either a Fréchet distribution, a Gumbel distribution or a Weibull distribution. For further information about extreme value theory and point processes, we refer to Resnick [23, 25]. Observe in particular that Equation (3.41) holds for $G$ being the Fréchet distribution $\Phi_\kappa(x) = \exp(-x^{-\kappa})\mathbf{1}_{(0,\infty)}(x)$ with $\kappa > 0$ if and only if $\xi_1$ is in $\mathcal{R}_{-\kappa}$, in which case the norming constants can be chosen to be $b_n = 0$ and $a_n$ subject to $\lim_{n\to\infty} n\mathbb{P}(\xi_1 > a_n) = 1$.

Now we describe the extremal behavior of the stationary TAR($S,1$) model (3.9) via point processes. Observe that the constants $a_n$ defined in (3.42) below are the norming constants of an i. i. d. sequence with the same distribution as $|X_0|$, rather than that of $X_0$.

**Theorem 3.5 (Point process behavior)** *Suppose that conditions TB and DC hold, and let $(X_k)_{k\in\mathbb{N}_0}$ be a stationary version of the TAR($S,1$) process as given in (3.9). Suppose further that $|Z_1| \in \mathcal{R}_{-\kappa}$ for some $\kappa > 0$. Let $0 < a_n \uparrow \infty$ be a sequence of constants such that*

$$\lim_{n\to\infty} n\mathbb{P}(|X_0| > a_n) = 1. \qquad (3.42)$$

*Then as $n \to \infty$,*

$$\sum_{k=1}^{\infty} \varepsilon_{(k/n, a_n^{-1} X_k)} \overset{w}{\Longrightarrow} \sum_{k=1}^{\infty}\sum_{j=0}^{\infty} \varepsilon_{(s_k, b_1^{(j)} P_k^+ + b_2^{(j)} P_k^-)} \qquad \text{in } M_P([0,\infty) \times (\overline{\mathbb{R}}\setminus\{0\})),$$

*where $b_1^{(j)}$ and $b_2^{(j)}$ are given by (3.24), $\sum_{k=1}^{\infty} \varepsilon_{(s_k, P_k)}$ is PRM($\vartheta$) in $M_P([0,\infty) \times (\overline{\mathbb{R}} \setminus \{0\}))$ with*

$$\vartheta(dt \times dx) = dt \times \kappa\theta(p^+ x^{-\kappa-1}\, \mathbf{1}_{(0,\infty)}(x) + p^-(-x)^{-\kappa-1}\, \mathbf{1}_{(-\infty,0)}(x))\, dx$$

*and*

$$\theta = \left( p^+ \sum_{j=0}^{\infty} |b_1^{(j)}|^\kappa + p^- \sum_{j=0}^{\infty} |b_2^{(j)}|^\kappa \right)^{-1} = (\widetilde{p}^+ + \widetilde{p}^-)^{-1} \in (0,1]\,.$$

*Here,* $\widetilde{p}^+$ *and* $\widetilde{p}^-$ *are given by* (3.10) *and* (3.11)*, respectively.*

**Proof.** We apply the results of Davis and Hsing [7], Theorem 2.7. By Lemma 2.1 $(X_k)_{k\in\mathbb{N}_0}$ is geometrically strongly mixing. Hence, the mixing condition $\mathcal{A}(a_n)$ of Davis and Hsing [7], p. 882, is satisfied meaning that there exists a sequence of positive integers $(v_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} v_n = \infty$, $\lim_{n\to\infty} v_n/n = 0$, and

$$\lim_{n\to\infty} \mathbb{E}\exp\left(-\sum_{j=1}^{n} f(X_j/a_n)\right) - \left(\mathbb{E}\exp\left(-\sum_{j=1}^{v_n} f(X_j/a_n)\right)\right)^{\lfloor n/v_n \rfloor} = 0$$

holds for all step functions $f$ on $\overline{\mathbb{R}} \setminus \{0\}$ with bounded support which is bounded away from 0. Furthermore, for $x > 0$,

$$\begin{aligned} &\mathbb{P}\left( \bigvee_{m\le k\le v_n} |X_k| > xa_n, |X_0| > xa_n \right) \\ &\quad\le \sum_{m\le k\le v_n} \mathbb{P}(|X_k| > xa_n, |X_0| > xa_n) \\ &\quad\le \sum_{m\le k\le v_n} \mathbb{P}\left( \beta^k |X_0| + \alpha \sum_{j=0}^{k-1} \beta^j + \sum_{j=1}^{k} \beta^{k-j} |Z_j| > xa_n, |X_0| > xa_n \right) \\ &\quad\le \sum_{m\le k\le v_n} \mathbb{P}\left( \beta^k |X_0| + \alpha/(1-\beta) + \sum_{j=0}^{\infty} \beta^j |Z_j| > xa_n, |X_0| > xa_n \right). \end{aligned} \tag{3.43}$$

Let $\delta \in (0, x)$. Then (3.43) and the independence of $X_0$ and $(Z_k)_{k\in\mathbb{N}_0}$ result in

$$\begin{aligned} &\mathbb{P}\left( \bigvee_{m\le k\le v_n} |X_k| > xa_n, |X_0| > xa_n \right) \\ &\quad\le \sum_{m\le k\le v_n} \mathbb{P}\left( \beta^k |X_0| + \alpha/(1-\beta) > \delta a_n \right) \\ &\qquad + \sum_{m\le k\le v_n} \mathbb{P}\left( \sum_{j=0}^{\infty} \beta^j |Z_j| > (x-\delta)a_n \right) \mathbb{P}(|X_0| > xa_n) \\ &\quad =: J_1(n) + J_2(n), \quad \text{say}. \end{aligned}$$

Since $|X_0| \in \mathcal{R}_{-\kappa}$ by Lemma 3.3, and since the limit in (3.1) is uniform for $u \in [u_0, \infty)$ for every $u_0 > 0$ (see e.g. Resnick [23], Proposition 0.5), it follows using dominated convergence that

$$\limsup_{m\to\infty} \lim_{n\to\infty} \frac{J_1(n)}{\mathbb{P}(|X_0| > xa_n)} \le \limsup_{m\to\infty} \sum_{k\ge m} \left( \frac{\beta^{-k}\delta}{x} \right)^{-\kappa} \le \limsup_{m\to\infty} \left( \frac{\delta}{x} \right)^{-\kappa} \sum_{k\ge m} (\beta^\kappa)^k = 0.$$

$J_2(n)$ can be estimated by

$$\limsup_{m\to\infty}\lim_{n\to\infty}\frac{J_2(n)}{\mathbb{P}(|X_0|>xa_n)}\le \limsup_{m\to\infty}\lim_{n\to\infty} v_n\mathbb{P}\left(\sum_{j=0}^{\infty}\beta^j|Z_j|>(x-\delta)a_n\right)=0,$$

since $\lim_{n\to\infty} v_n/n=0$ and since

$$\mathbb{P}(|X_0|>x)\sim(\widetilde{p}^++\widetilde{p}^-)\mathbb{P}(|Z_1|>x)\sim(\widetilde{p}^++\widetilde{p}^-)(1-\beta^\kappa)\mathbb{P}\left(\sum_{j=0}^{\infty}\beta^j|Z_j|>x\right),\quad x\to\infty.$$

Hence, we conclude

$$\lim_{m\to\infty}\lim_{n\to\infty}\mathbb{P}\left(\bigvee_{m\le k\le v_n}|X_k|>xa_n\,\middle|\,|X_0|>xa_n\right)=0,\quad x>0,$$

which is Equation (2.8) of Davis and Hsing [7]. Since $\mathbf{X}^{(2m+1)}$ is multivariate regularly varying of index $-\kappa$ and spectral measure described by the vector $\Theta^{(2m)}=(\Theta_j^{(2m)})_{j=0,\dots,2m}$ as given in Theorem 3.4 we obtain

$$\mathbb{E}|\Theta_m^{(2m)}|^\kappa=\frac{1}{\widetilde{p}^++\widetilde{p}^-+2m}\left(\widetilde{p}^+|b_1^{(m)}|^\kappa+\widetilde{p}^-|b_2^{(m)}|^\kappa+p^+\sum_{j=0}^{m-1}|b_1^{(j)}|^\kappa+p^-\sum_{j=0}^{m-1}|b_2^{(j)}|^\kappa\right).$$

Using further that $(|b_1^{(m)}|)_{m\ge0}$ and $(|b_2^{(m)}|)_{m\ge0}$ are non-increasing sequences, we get

$$\begin{aligned}
&\mathbb{E}\left(\bigvee_{k=m}^{2m}|\Theta_k^{(2m)}|^\kappa-\bigvee_{k=m+1}^{2m}|\Theta_k^{(2m)}|^\kappa\right)\\
&=\frac{1}{\widetilde{p}^++\widetilde{p}^-+2m}\Big[\widetilde{p}^+(|b_1^{(m)}|^\kappa-|b_1^{(m+1)}|^\kappa)+\widetilde{p}^-(|b_2^{(m)}|^\kappa-|b_2^{(m+1)}|^\kappa)\\
&\qquad+p^+\sum_{j=0}^{m-1}(|b_1^{(j)}|^\kappa-|b_1^{(j+1)}|^\kappa)+p^-\sum_{j=0}^{m-1}(|b_2^{(j)}|^\kappa-|b_2^{(j+1)}|^\kappa)\Big]\\
&=\frac{1-p^+|b_1^{(m)}|^\kappa-p^-|b_2^{(m)}|^\kappa+\widetilde{p}^+(|b_1^{(m)}|^\kappa-|b_1^{(m+1)}|^\kappa)+\widetilde{p}^-(|b_2^{(m)}|^\kappa-|b_2^{(m+1)}|^\kappa)}{\widetilde{p}^++\widetilde{p}^-+2m}.
\end{aligned}$$

Then

$$\lim_{m\to\infty}\frac{\mathbb{E}\left(\bigvee_{k=m}^{2m}|\Theta_k^{(2m)}|^\kappa-\bigvee_{k=m+1}^{2m}|\Theta_k^{(2m)}|^\kappa\right)}{\mathbb{E}|\Theta_m^{2m}|^\kappa}=\frac{1}{p^+\sum_{j=0}^\infty|b_1^{(j)}|^\kappa+p^-\sum_{j=0}^\infty|b_2^{(j)}|^\kappa}=\theta,$$

where $\theta\in(1-\beta^\kappa,1]$ since $|b_1^{(j)}|,|b_2^{(j)}|\le\beta^j$ and $|b_1^{(0)}|=|b_2^{(0)}|=1$.

Similarly, defining the probability measures $R_m$ on $M_P(\overline{\mathbb{R}}\setminus\{0\})$ by

$$R_m(\cdot)=\frac{\mathbb{E}\left(\left[\bigvee_{k=m}^{2m}|\Theta_k^{(2m)}|^\kappa-\bigvee_{k=m+1}^{2m}|\Theta_k^{(2m)}|^\kappa\right]\mathbf{1}\{\sum_{j=0}^{2m}\varepsilon_{\Theta_j^{(2m)}}\in\cdot\}\right)}{\mathbb{E}\left(\bigvee_{k=m}^{2m}|\Theta_k^{(2m)}|^\kappa-\bigvee_{k=m+1}^{2m}|\Theta_k^{(2m)}|^\kappa\right)}$$

it follows that $R_m$ converges weakly as $m \to \infty$ to the distribution of the point process $\sum_{j=0}^\infty \varepsilon_{Q_j} \in M_P(\overline{\mathbb{R}} \setminus \{0\})$, where

$$\sum_{j=0}^\infty \varepsilon_{Q_j} = \chi_1 \sum_{j=0}^\infty \varepsilon_{b_1^{(j)}} + (1-\chi_1) \sum_{j=0}^\infty \varepsilon_{b_2^{(j)}}$$

with $\mathbb{P}(\chi_1 = 1) = p^+$ and $\mathbb{P}(\chi_1 = 0) = p^-$. Hence, the assumptions of Davis and Hsing [7], Theorem 2.7, are all satisfied. Let $\sum_{k=1}^\infty \varepsilon_{(\widetilde{s}_k, \widetilde{P}_k)}$ be PRM($\widetilde{\vartheta}$) on $[0,\infty) \times (\overline{\mathbb{R}} \setminus \{0\})$ with

$$\widetilde{\vartheta}(dt \times dx) = dt \times \kappa\theta x^{-\kappa-1}\, \mathbf{1}_{(0,\infty)}(x)\, dx$$

and $(\sum_{j=0}^\infty \varepsilon_{Q_{kj}})_{k\in\mathbb{N}}$ be an i. i. d. sequence with $\sum_{j=0}^\infty \varepsilon_{Q_{kj}} \stackrel{d}{=} \sum_{j=0}^\infty \varepsilon_{Q_j}$, independent of $\sum_{k=1}^\infty \varepsilon_{(\widetilde{s}_k, \widetilde{P}_k)}$. Then Theorem 2.7, Corollary 2.4 and Remark 2.3 of Davis and Hsing [7] imply that

$$\sum_{k=1}^\infty \varepsilon_{(k/n, a_n^{-1}X_k)} \stackrel{w}{\Longrightarrow} \sum_{k=1}^\infty \sum_{j=0}^\infty \varepsilon_{(\widetilde{s}_k, Q_{kj}\widetilde{P}_k)} \quad \text{in} \quad M_P([0,\infty) \times (\overline{\mathbb{R}} \setminus \{0\})), \quad n \to \infty$$

(see also Lemma 4.1.2 in Hsing [18] for the connection between the convergence of the process $\sum_{k=1}^\infty \varepsilon_{(k/n, a_n^{-1}X_k)}$ and that of $\sum_{k=1}^\infty \varepsilon_{a_n^{-1}X_k}$ as $n \to \infty$). Since

$$\sum_{k=1}^\infty \sum_{j=0}^\infty \varepsilon_{(\widetilde{s}_k, Q_{kj}\widetilde{P}_k)} \stackrel{d}{=} \sum_{k=1}^\infty \sum_{j=0}^\infty \varepsilon_{(s_k, b_1^{(j)} P_k^+ + b_2^{(j)} P_k^-)},$$

the result follows apart from the representation $\theta = (\widetilde{p}^+ + \widetilde{p}^-)^{-1}$. To see the latter, suppose, e.g., that $\beta_1 \le 0$ and $\beta_2 > 0$. Then

$$\widetilde{p}^+ + \widetilde{p}^- = \frac{p^+ + p^-(-\beta_1)^\kappa}{1-\beta_2^\kappa} + p^- = p^+ \frac{1}{1-\beta_2^\kappa} + p^- \frac{1-\beta_2^\kappa + (-\beta_1)^\kappa}{1-\beta_2^\kappa}.$$

On the other hand,

$$b_1^{(j)} = \beta_2^j, \quad j \in \mathbb{N}_0, \qquad b_2^{(0)} = -1, \qquad b_2^{(j)} = (-\beta_1)\beta_2^{j-1}, \quad j \in \mathbb{N},$$

so that

$$p^+ \sum_{j=0}^\infty |b_1^{(j)}|^\kappa + p^- \sum_{j=0}^\infty |b_2^{(j)}|^\kappa = p^+ \frac{1}{1-\beta_2^\kappa} + p^- \frac{1-\beta_2^\kappa + (-\beta_1)^\kappa}{1-\beta_2^\kappa} = \widetilde{p}^+ + \widetilde{p}^-.$$

The other cases follow similarly. □

Having the point process convergence in Theorem 3.5, it is standard to derive many results about the asymptotic behavior of the stationary sequence $(X_k)_{k\in\mathbb{N}_0}$, such as convergence of the maxima to extremal processes, the asymptotic distribution of the order statistics or of exceedances over high thresholds, or the determination of the extremal index. We shall concentrate here on the latter. Recall that a stationary sequence $(\xi_k)_{k\in\mathbb{N}}$ has *extremal index* $\rho \in (0,1]$, if there exist norming constants $a_n > 0$ and $b_n \in \mathbb{R}$, and a non-degenerate distribution function $G$ such that (3.41) holds for an associated i. i. d. sequence $(\widetilde{\xi}_k)_{k\in\mathbb{N}}$ with $\widetilde{\xi}_1 \stackrel{d}{=} \xi_1$, and

$$\lim_{n\to\infty} \mathbb{P}\left( a_n^{-1} \left( \bigvee_{k=1}^{n} \xi_k - b_n \right) \le x \right) = G(x)^\rho \quad \text{for all } x \in \mathbb{R}.$$

Under weak mixing conditions (which are satisfied in our case), it is known that the reciprocal $\rho^{-1}$ of the extremal index can be interpreted as the mean cluster size of exceedances over high thresholds. In particular, an extremal index of size 1 says that high exceedances of a stationary sequence behave asymptotically like that of an i. i. d. sequence with the same marginals, while an extremal index which is less than 1 shows that clusters occur. We refer to Leadbetter et al. [20] and Embrechts et al. [12] for further information regarding the extremal index.

The following result gives the asymptotic behavior of the maxima of the stationary TAR$(S,1)$ model and its extremal index.

**Corollary 3.6** *Let the assumptions of Theorem 3.5 hold with* $(a_n)_{n\in\mathbb{N}}$ *as defined in* (3.42), *and denote* $M_n = \max_{k=1,\dots,n} X_k$ *for* $n \in \mathbb{N}$. *Then for* $x > 0$

$$\lim_{n\to\infty} \mathbb{P}(a_n^{-1} M_n \le x) = \exp(-\theta(p^+ + p^-(\beta_1^-)^\kappa)x^{-\kappa}) = \exp\left( -\frac{p^+ + p^-(\beta_1^-)^\kappa}{\widetilde{p}^+ + \widetilde{p}^-} x^{-\kappa} \right),$$

*with* $\widetilde{p}^+$ *and* $\widetilde{p}^-$ *as defined in* (3.10) *and* (3.11). *In particular, if* $\widetilde{p}^+ > 0$, *then* $(X_k)_{k\in\mathbb{N}_0}$ *has extremal index* $\rho = 1 - (\beta_2^+)^\kappa - (\beta_1^-)^\kappa(\beta_2^-)^\kappa$.

**Proof.** Applying the continuous mapping theorem for the functional

$$T_1 : M_P([0,\infty) \times (\overline{\mathbb{R}} \setminus \{0\})) \to \overline{\mathbb{R}}, \quad \sum_{k=1}^{\infty} \varepsilon_{(t_k, j_k)} \mapsto \sup\{j_k : t_k \le 1\},$$

it follows from Theorem 3.5 that

$$\begin{aligned} a_n^{-1} M_n \stackrel{w}{\Longrightarrow} Y \;&:=\; \sup\{b_1^{(j)} P_k^+ + b_2^{(j)} P_k^- : s_k \le 1,\ j \in \mathbb{N}_0\} \qquad (n \to \infty) \\ &=\; \sup\{P_k^+, \beta_1^- P_k^- : s_k \le 1\} = T_1(\overline{N}), \end{aligned}$$

where $\overline{N} := \sum_{k=1}^{\infty} \varepsilon_{(s_k, P_k^+)} + \varepsilon_{(s_k, \beta_1^- P_k^-)}$. Here, we used that $\sup\{0, b_1^{(j)} : j \in \mathbb{N}_0\} = 1$ and $\sup\{0, b_2^{(j)} : j \in \mathbb{N}_0\} = \beta_1^-$. Since $\overline{N}$ is PRM$(\overline{\vartheta})$ with mean measure

$$\overline{\vartheta}(dt \times dx) = dt \times \kappa\theta(p^+ + p^-(\beta_1^-)^\kappa)x^{-\kappa-1}\mathbf{1}_{(0,\infty)}(x)\, dx,$$

we conclude that for fixed $x > 0$

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}(a_n^{-1}M_n \le x) &= \mathbb{P}(Y \le x) = \mathbb{P}\left(\overline{N}((0,1] \times (x,\infty)) = 0\right) \\ &= \exp\{-\theta(p^+ + p^-(\beta_1^-)^\kappa)x^{-\kappa}\}.\end{aligned}$$

The extremal index $\rho$ of $(X_k)_{k\in\mathbb{N}_0}$ is then given by

$$\rho = \frac{p^+ + p^-(\beta_1^-)^\kappa}{\widetilde{p}^+ + \widetilde{p}^-} \Big/ \frac{\widetilde{p}^+}{\widetilde{p}^+ + \widetilde{p}^-} = 1 - (\beta_2^+)^\kappa - (\beta_1^-)^\kappa(\beta_2^-)^\kappa,$$

which is the claim. □

We can also compute the asymptotic cluster probabilities $\mathbb{P}(\zeta_1 = j)$ of exceedances $X_k > a_n x$ of length $j \in \mathbb{N}$ for fixed $x > 0$ explicitly, considering the limit behavior of the rescaled times $k/n$ for which $X_k > a_n x$. This is done using the same arguments as in Davis and Resnick [8], Section 3.D. We omit the proof.

**Corollary 3.7** *Let the assumptions of Theorem 3.5 hold with $(a_n)_{n\in\mathbb{N}}$ as defined in (3.42), and suppose that $\widetilde{p}^+ > 0$, i. e. $p^+ + p^-(\beta_1^-)^\kappa > 0$. Let $(\widetilde{s}_k)_{k\in\mathbb{N}}$ be the jump times of a Poisson process with intensity $\theta(p^+ + p^-(\beta_1^-)^\kappa)x^{-\kappa}$, $x > 0$ fixed. Let $(\zeta_k)_{k\in\mathbb{N}}$ be an $\mathbb{N}$-valued i. i. d. sequence, independent of $(\widetilde{s}_k)_{k\in\mathbb{N}}$, with distribution*

$$\mathbb{P}(\zeta_1 = j) = \frac{p^+((\widetilde{b}_1^{(j-1)})^\kappa - (\widetilde{b}_1^{(j)})^\kappa) + p^-((\widetilde{b}_2^{(j-1)})^\kappa - (\widetilde{b}_2^{(j)})^\kappa)}{p^+ + p^-(\beta_1^-)^\kappa} \quad \textit{for } j \in \mathbb{N},$$

*where $1 = \widetilde{b}_1^{(0)} \ge \widetilde{b}_1^{(1)} \ge \ldots$ is the order statistic of the sequence $(\max\{0, b_1^{(j)}\})_{j\in\mathbb{N}_0}$ and $\beta_1^- = \widetilde{b}_2^{(0)} \ge \widetilde{b}_2^{(1)} \ge \ldots$ is the order statistic of the sequence $(\max\{0, b_2^{(j)}\})_{j\in\mathbb{N}_0}$. Then*

$$\sum_{k=1}^\infty \varepsilon_{(k/n, a_n^{-1}X_k)}(\cdot \times (x,\infty)) \overset{w}{\Longrightarrow} \sum_{k=1}^\infty \zeta_k \varepsilon_{\widetilde{s}_k} \quad \textit{in } M_P([0,\infty)) \quad \textit{as } n \to \infty.$$

**Remark 3.8**

*(i)* If $\beta_1 = \beta_2 = \ldots = \beta_S = \beta \in (-1,1)$ and $\alpha_1 = \ldots = \alpha_S = 0$, then the process (3.9) is a causal AR(1) process $X_k = \beta X_{k-1} + Z_k$, which can be written as an infinite moving average process $X_k = \sum_{j=0}^\infty \beta^j Z_{k-j}$. It is then easy to see that the results obtained in this paper are in line with those of Davis and Resnick [8] for infinite moving average processes with regularly varying noise.

*(ii)* Provided $\widetilde{p}^+ > 0$, the extremal index $\rho$ of Corollary 3.6 is strictly less than 1 if and only if $\beta_2 > 0$ or $\beta_1\beta_2 > 0$. In these cases, the TAR$(S,1)$ model can model clusters.

*(iii)* The value $\theta$ is the extremal index of $(|X_k|)_{k\in\mathbb{N}_0}$.

# 4 Noise with at most exponentially decreasing tail

In this section we shall be interested in noise sequences which are lighter tailed than regularly varying functions, but whose tail is still not too light. Since distributions which are not regularly varying and do not have a finite right endpoint can only be in the maximum domain of attraction of the Gumbel distribution if they are in the maximum domain of attraction of some extreme value distribution at all, we shall concentrate on specific i. i. d. noise sequences $(Z_k)_{k\in\mathbb{N}_0}$ for which (3.41) holds with $G(x) = \Lambda(x) = \exp(-\mathrm{e}^{-x})$ and $\xi_k = Z_k$, which we denote by $Z_1 \in \mathrm{MDA}(\Lambda)$. Besides the tail balance condition on $Z_1$, we shall look at distributions which are either subexponential or have a tail which is close to that of an exponential distribution. More precisely, we shall look at distribution functions $F$ in the class $\mathcal{L}(\gamma) \geq 0$ with $\gamma \in [0,\infty)$, i. e. distribution functions which satisfy $F(x) < 1$ for all $x \in \mathbb{R}$ and for which

$$\lim_{x\to\infty} \overline{F}(x+y)/\overline{F}(x) = \exp(-\gamma y) \tag{4.1}$$

holds locally uniformly in $y \in \mathbb{R}$. For $\gamma > 0$ this means that $F$ has a tail which is close to that of an exponential distribution. For $\gamma = 0$ we need a further assumption: a distribution function $F$ (defined on $\mathbb{R}$) is called *subexponential*, if $F \in \mathcal{L}(0)$ and $\lim_{x\to\infty} \overline{F*F}(x)/\overline{F}(x) = 2$. The class of subexponential distributions will be denoted by $\mathcal{S}$. For a random variable $Z$ with distribution function $F$ we simply write $X \in \mathcal{S}$ or $Z \in \mathcal{L}(\gamma)$ if $F$ has the corresponding property. Subexponential distributions and those which are in $\mathcal{L}(\gamma)$ for $\gamma > 0$ are handy classes of (semi-)heavy tailed distributions, and the extremal and the tail behavior of infinite moving average processes associated with such noise sequences have been studied by Davis and Resnick [9] and Rootzén [26], respectively. Examples of distributions in $\mathrm{MDA}(\Lambda) \cap \mathcal{S}$ include tails of the form $\overline{F}(x) \sim \exp\{-x/(\log x)^a\}$, $a > 0$, or $\overline{F}(x) \sim Kx^b\mathrm{e}^{-x^p}$ where $p \in (0,1)$, $K > 0$ and $b \in \mathbb{R}$ $(x \to \infty)$, or the lognormal distribution. The class of subexponential distributions includes also those which have regularly varying tails, but these are not in $\mathrm{MDA}(\Lambda)$. Examples in $\mathcal{L}(\gamma)$ with $\gamma > 0$ include distribution functions with tails of the form $\overline{F}(x) \sim Kx^b\mathrm{e}^{-\gamma x}$ $(x \to \infty)$ with $K > 0$, $b \in \mathbb{R}$, or certain generalized inverse Gaussian distributions. Observe that $\mathcal{L}(\gamma) \subset \mathrm{MDA}(\Lambda)$ for $\gamma > 0$ since by (4.1) equation (3.41) holds with $a_n = \gamma^{-1}$ and $b_n = \overline{F}^{\leftarrow}(1/n)$. Also observe that whether a distribution function is in $\mathcal{S} \cap \mathrm{MDA}(\Lambda)$ or in $\mathcal{L}(\gamma)$ with $\gamma > 0$, respectively, is completely determined by its tail behavior. More precisely, if $\overline{G}(x) \sim c\overline{F}(x)$ $(x \to \infty)$ for some $c \in (0,\infty)$, then $F \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$ or $F \in \mathcal{L}(\gamma)$ implies the same for $G$; see Pakes [22], Lemma 2.4, for the subexponential case (the case $\mathcal{L}(\gamma)$ follows from the definition of $\mathcal{L}(\gamma)$).

As in Section 3 we shall first present the tail behavior of the TAR model if $Z$ has the described noise. But unlike there we do not have to restrict to specific $\mathrm{TAR}(S,1)$ models to remain in the same noise class, but can handle the general TAR$(S,q)$ model.

**Proposition 4.1 (Tail behavior)** *Suppose that conditions DC and TB hold with $p^+ > 0$, and let $(X_k)_{k\in\mathbb{N}_0}$ be a stationary version of the TAR$(S,q)$ process as given in (1.1). Suppose further that $Z_1 \in \mathcal{L}(\gamma)$ with $\gamma > 0$, or that $Z_1 \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$, in which case we put $\gamma := 0$. Then $\mathbb{E}\mathrm{e}^{\gamma(X_1-Z_1)}$ is finite, and*

$$\mathbb{P}(X_1 > x) \sim \mathbb{E}\mathrm{e}^{\gamma(X_1-Z_1)}\mathbb{P}(Z_1 > x) \quad \text{as } x \to \infty. \tag{4.2}$$

*In particular, if $Z_1 \in \mathcal{L}(\gamma)$ or $Z_1 \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$, respectively, then so is $X_1$.*

**Proof.** Similarly to and with the same notations as in Lemma 2.2 we obtain

$$\mathbb{P}(|X_k - Z_k| > x) \le \mathbb{P}(|\widetilde{X}_k + \alpha - \widetilde{Z}_k| > x) = \mathbb{P}\left(\sum_{j=1}^{\infty} \psi_j \widetilde{Z}_j + \alpha > x\right), \tag{4.3}$$

so that

$$\limsup_{x\to\infty} \frac{\mathbb{P}(|X_k - Z_k| > x)}{\mathbb{P}(Z_1 > x)} \le \frac{1}{p^+} \limsup_{x\to\infty} \frac{\mathbb{P}\left(\sum_{j=1}^{\infty} \psi_j \widetilde{Z}_j + \alpha > x\right)}{\mathbb{P}(\widetilde{Z}_1 > x)} \tag{4.4}$$

by condition TB. Since $0 \le \psi_j < 1$ for $j \in \mathbb{N}$ and $(\psi_j)$ are exponentially decreasing, it follows from Proposition 1.3 of Davis and Resnick [9] that the right hand side of (4.4) is 0 if $Z_1 \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$ and TB is valid, and hence that (4.2) holds in the case $Z_1 \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$ ($\gamma = 0$) by Embrechts et al. [11], Proposition 1 (for subexponentials on $\mathbb{R}$ see also Pakes [22], Lemma 5.1).

Now, suppose that $\gamma > 0$ and that $Z_1 \in \mathcal{L}(\gamma)$. Together with TB this implies that $\psi\widetilde{Z}_1 \in \mathcal{L}(\gamma/\psi)$ for $0 < \psi < 1$, so that $\mathbb{E}\mathrm{e}^{\delta\widetilde{Z}_1} < \infty$ for $0 \le \delta < \gamma$. Observe that $0 \le \psi_j < 1$ for $j \in \mathbb{N}$ and that $\sum_{j=1}^{\infty} \psi_j < \infty$. Choosing $0 < \varepsilon < \gamma$ such that $(\gamma+\varepsilon)^{-1}(\gamma-\varepsilon) > \max_{j\in\mathbb{N}}\{\psi_j\}$, it follows from (4.3) and Jensen's inequality that

$$\begin{aligned}
\mathbb{E}\mathrm{e}^{(\gamma+\varepsilon)|X_1-Z_1|} &\le \mathrm{e}^{(\gamma+\varepsilon)\alpha} \prod_{j=1}^{\infty} \mathbb{E}\mathrm{e}^{(\gamma+\varepsilon)\psi_j\widetilde{Z}_j} \\
&\le \mathrm{e}^{(\gamma+\varepsilon)\alpha} \prod_{j=1}^{\infty} \left(\mathbb{E}\mathrm{e}^{(\gamma-\varepsilon)\widetilde{Z}_j}\right)^{\frac{\gamma+\varepsilon}{\gamma-\varepsilon}\psi_j} < \infty.
\end{aligned}$$

But since $Z_1 \in \mathcal{L}(\gamma)$ if and only if $\mathrm{e}^{Z_1} \in \mathcal{R}_{-\gamma}$, and since $\mathbb{E}\mathrm{e}^{|X_1-Z_1|(\gamma+\varepsilon)} < \infty$, it follows from Breiman's [4] result on products of regularly varying distributions that

$$\mathbb{P}\left(\mathrm{e}^{X_1} > x\right) = \mathbb{P}\left(\mathrm{e}^{Z_1}\,\mathrm{e}^{X_1-Z_1} > x\right) \sim \mathbb{E}\mathrm{e}^{\gamma(X_1-Z_1)}\,\mathbb{P}\left(\mathrm{e}^{Z_1} > x\right), \quad \text{as} \quad x \to \infty,$$

which is Equation (4.2) (the latter equation can also be derived from Proposition 1.1 of Davis and Resnick [9] and Pakes [22], Lemma 2.1). $\square$

Similarly to the regularly varying TAR$(S, 1)$ model, the tail of the TAR process is equivalent to the tail of the noise. Next, analog to the regularly varying case we show the convergence of a sequence of point processes. In contrast to Theorem 3.5 we obtain the convergence to a Poisson random measure. Thus, this model cannot exhibit extremal clusters.

**Theorem 4.2 (Point process behavior)** *Suppose that conditions DC and TB hold with $p^+ > 0$, and let $(X_k)_{k\in\mathbb{N}_0}$ be a stationary version of the TAR$(S, q)$ process as given in (1.1). Suppose further that $Z_1 \in \mathcal{L}(\gamma)$ with $\gamma > 0$, or that $Z_1 \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$. Let $a_n > 0$ and $b_n \in \mathbb{R}$ be sequences of constants such that*

$$\lim_{n\to\infty} n\mathbb{P}(X_0 > a_n x + b_n) = \exp(-x) \qquad \text{for } x \in \mathbb{R}.$$

*Then as $n \to \infty$,*

$$\sum_{k=1}^{\infty} \varepsilon_{(k/n, a_n^{-1}(X_k - b_n))} \overset{w}{\Longrightarrow} \sum_{k=1}^{\infty} \varepsilon_{(s_k, P_k)} \qquad \text{in } M_P([0,\infty) \times \mathbb{R}),$$

*where $\sum_{k=1}^{\infty} \varepsilon_{(s_k,P_k)}$ is a $\mathrm{PRM}(dt \times \mathrm{e}^{-x}dx)$.*

**Proof.** First, we investigate the case $Z_1 \in \mathcal{L}(\gamma)$ with $\gamma > 0$. Let $u_n = a_n x + b_n$, $n \in \mathbb{N}$. Since $(X_k)_{k\in\mathbb{N}_0}$ is geometrically strongly mixing by Lemma 2.1, the mixing condition $D_r(u_n)$ of Leadbetter et al. [20], Theorem 5.5.1, holds for $(X_k)_{k\in\mathbb{N}_0}$. It remains to show the anti-cluster condition $D'(u_n)$, i. e.

$$\lim_{k\to\infty} \lim_{n\to\infty} n \sum_{j=2}^{\lfloor n/k \rfloor} \mathbb{P}(X_j > u_n, X_1 > u_n) = 0.$$

By Lemma 2.2 we have

$$n \sum_{j=2}^{\lfloor n/k \rfloor} \mathbb{P}(X_j > u_n, X_1 > u_n) \le n \sum_{j=2}^{\lfloor n/k \rfloor} \mathbb{P}(\widetilde{X}_j > u_n, \widetilde{X}_1 > u_n).$$

Furthermore, $\mathbb{P}(\widetilde{X}_1 > u_n) \sim C\mathbb{P}(X_1 > u_n)$ as $n \to \infty$ and some constant $C > 0$ by Proposition 4.1. Analog to the proof of Theorem 7.4 in Rootzén [26] (cf. Fasen et al. [16], proof of Lemma 2) we have that the MA process $(\widetilde{X}_k)_{k\in\mathbb{N}_0}$ satisfies the $D'(u_n)$ condition, and, hence, also $(X_k)_{k\in\mathbb{N}_0}$ satisfies the $D'(u_n)$ condition. The conclusion then follows by Leadbetter et al. [20], Theorem 5.5.1.

In the remaining case $Z_1 \in \mathcal{S} \cap \mathrm{MDA}(\Lambda)$ the conclusion follows from Fasen [15], Proposition 9, and (4.3)-(4.4), which converges to 0 as $x \to \infty$. □

We can now obtain the behavior of the running maxima of the stationary sequence and hence identify the extremal index to be equal to 1, which implies that the model cannot exhibit clusters on high levels in this case. We omit the proof, which is along the lines of that of Corollary 3.6.

**Corollary 4.3** *Let the assumptions of Theorem 4.2 hold and* $M_n = \max_{k=1,\dots,n} X_k$ *for* $n \in \mathbb{N}$. *Then*

$$\lim_{n\to\infty} \mathbb{P}(a_n^{-1}(M_n - b_n) \le x) = \exp(-\mathrm{e}^{-x}) \quad \text{for } x \in \mathbb{R}.$$

*In particular, the extremal index of* $(X_k)_{k\in\mathbb{N}_0}$ *is equal to 1.*

# 5 Conclusion

We have shown that stationary TAR models with noise in $\mathcal{S}\cap\mathrm{MDA}(\Lambda)$ or $\mathcal{L}(\gamma)$ with $\gamma > 0$ are tail equivalent to their noise sequence, and that they have extremal index equal to one, hence cannot cluster. On the other hand, if the noise sequence is regularly varying, then the tail is in general only O-regularly varying, but for $\mathrm{TAR}(S,1)$ models with intervals as partitions it is tail equivalent to its noise (in particular it is regularly varying). Moreover, in this case the extremal index is less than 1 in many cases depending on the coefficients of the TAR model. In those cases the $\mathrm{TAR}(S,1)$ model can exhibit cluster behavior.

It would be interesting to obtain similar results for noise sequences which are super-exponential, such as distribution functions $F$ with tails like $\overline{F}(x) = Kx^b \exp(-x^p)$ for $p \in (1,\infty)$. However, already the analysis for infinite moving average processes with such noise sequences is very involved and has been carried out by Rootzén [26, 27]. See also Klüppelberg and Lindner [19] for such moving average processes. But for the TAR model, due to the nonlinear regime switch, it seems an open problem how to determine the precise tail behavior of the stationary TAR model even for Gaussian noise, apart from simple situations such as symmetric TAR models.

## Acknowledgement

The authors are indebted to a referee for pointing out the relation between $\mathcal{L}(\gamma)$ for $\gamma > 0$ and $\mathrm{MDA}(\Lambda)$.

# References


[1] An, H. Z. and Huang, F. C. (1996). The geometric ergodicity of nonlinear autoregressive models. *Statist. Sinica* **6**, 943–956.

[2] Basrak, B., Davis, R. A., and Mikosch, T. (2002). Regular variation of GARCH processes. *Stochastic Process. Appl.* **99**, 95–115.

[3] Brachner, C. (2004). Tailverhalten autoregressiver Thresholdmodelle. Diplomarbeit, Technische Universität München.

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

[5] Chan, K. and Tong, H. (1985). On the use of the deterministic Lyapunov function for the ergodicity of stochastic difference equations. *Adv. Appl. Probab.* **17**, 666–678.

[6] Chen, R. and Tsay, R. (1991). On the ergodicity of TAR(1) processes. *Ann. Appl. Probab.* **1**, 613–634.

[7] Davis, R. and Hsing, T. (1995). Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Probab.* **23**, 879–917.

[8] Davis, R. and Resnick, S. (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.* **13**, 179–195.

[9] Davis, R. and Resnick, S. (1988). Extremes of moving averages of random variables from the domain of attraction of the double exponential distribution. *Stochastic Process. Appl.* **30**, 41–68.

[10] Diop, A. and Guegan, D. (2004). Tail behavior of a threshold autoregressive stochastic volatility model. *Extremes* **7**, 369–377.

[11] Embrechts, P., Goldie, C. M., and Veraverbeke, N. (1979). Subexponentiality and infinite divisibility. *Z. Wahrsch. Verw. Gebiete* **49**, 335–347.

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

[13] Fan, J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer, New York.

[14] Fasen, V. (2005). Extremes of regularly varying mixed moving average processes. *Adv. in Appl. Probab.* **37**, 993–1014.

[15] Fasen, V. (2006). Extremes of subexponential Lévy driven moving average processes. *Stochastic Process. Appl.* **116**, 1066–1087.

[16] Fasen, V., Klüppelberg, C., and Lindner, A. (2006). Extremal behavior of stochastic volatility models. In: A. Shiryaev, M. d. R. Grossinho, P. Oliviera, and M. Esquivel (Eds.), *Stochastic Finance*, pp. 107–155. Springer, New York.

[17] Feller, W. (1971). *An Introduction to Probability Theory and its Applications*, vol. 2. 2nd edn. Wiley, New York.

[18] Hsing, T. (1993). On some estimates based on sample behavior near high level excursions. *Probab. Theory Relat. Fields* **95**, 331–356.

[19] Klüppelberg, C., and Lindner, A. (2005). Extreme value theory for moving average processes with light–tailed innovations. *Bernoulli* **11**, 381–410.

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

[21] Meyn, S. and Tweedie, R. (1993). *Markov Chains and Stochastic Stability*. Springer, London.

[22] Pakes, A. G. (2004). Convolution equivalence and infinite divisibility. *J. Appl. Probab.* **41**, 407–424. With correction in *J. Appl. Prob.* **44** (2007), 295-305.

[23] Resnick, S. I. (1986). Point processes, regular variation and weak convergence. *Adv. Appl. Probability* **18**, 66–138.

[24] Resnick, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer, New York.

[25] Resnick, S. I. (2006). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, New York.

[26] Rootzén, H. (1986). Extreme value theory for moving average processes. *Ann. Probab.* **14**, 612–652.

[27] Rootzén, H. (1987). A ratio limit theorem for the tails of weighted sums. *Ann. Probab.* **15**, 728–747.

[28] Tong, H. (1977). Contribution to the discussion of the paper entitled "Stochastic modelling of riverflow time series" by A.J. Lawrance and N.T. Kottegoda. *J. Roy. Statist. Soc. Ser. A* **140**, 34–35.

[29] Tong, H. (1990). *Non-linear Time Series*. Oxford University Press, Oxford.

[30] Tong, H. and Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *J. Roy. Statist. Soc. Ser. B* **42**, 245–292.