Jiazhu Pan[a,b,1] / Yali He[b,2] /

# Tail Behaviours of Multiple-Regime Threshold AR Models with Heavy-Tailed Innovations

[a] Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XH, UK

[b] School of Mathematics and Statistics, Yangtze Normal University, Chongqing 408100, China

**Abstract**

This paper studies the tail behaviours of the stationary distribution of multiple-regime threshold AR models with multiple heavy-tailed innovations. It is shown that the marginal tail probability has the same order as that of the innovation with the heaviest tail. Other new results in this paper include the geometric ergodicity and the tail dependence of TAR models with multiple heavy-tailed innovations.

**Key words.** Heavy-tailed distribution, tail probability, threshold AR model, ergodicity.

**JEL classification:** C22, C32.

# 1 Introduction

Tong's (1978) threshold autoregressive (TAR) model and its many extensions are one standard class of nonlinear time series models. The probabilistic structures of these models were studied by many authors. Examples are Chan, Petruccelli, Tong, and Woolford (1985), Chan and Tong (1985), Tong (1990), Chen and Tsay (1991), Brockwell, Liu, and Tweedie (1992), Liu and Susko (1992), An and Huang (1996), An and Chen (1997), Liu, Li, and Li (1997), Ling (1999), Hansen (2011), and Tsay and Chen (2019) among others. The impact of TAR model in the fields of econometrics and economics was reviewed by Hansen (2011). In contrast, however, it seems that the tail behaviours of the stationary distributions of these threshold time series models have not been discussed in literature.

Tail properties of stationary time series models are very important in applications, particularly in financial risk analysis, see Embrechts, Kluppelberg, and Mikosch (1997). Heavy tail is a well known empirical feature in financial time series. Good evidences show that Gaussian or light-tailed innovations can not describe the tail behaviors of these time series, see Mandelbrot (1963b), Mandelbrot (1963a),

[1]Corresponding email address: jiazhu.pan@strath.ac.uk

[2]Corresponding email address: heyalihy0702@163.com

This is a peer-reviewed, accepted author manuscript of the following article: Pan, J., & He, Y. (Accepted/In press). Tail behaviours of multiple-regime threshold AR models with heavy-tailed innovations. Studies in Nonlinear Dynamics and Econometrics.

Fama (2013) and Adler, Feldaman, and Taqqu (1997). More and more people are interested in time series models with heavy-tailed innovations. The tail behaviours of linear models and bilinear models were studied by Resnick (1997) and Davis and Resnick (1996), respectively. The tail of ARCH and GARCH models were studied by DeHaan, Resnick, Rootzen, and De Vries (1989), Mikosch and Starica (2000), Borkovec and Kluppelberg (2001) and Mittnik, Paolella, and Rachev (2002). Pan, Yu, and Pang (2004) discussed the relation between the marginal tail probability and the innovation's tail probability for three important types of time series models: infinite order moving averages, bilinear time series and solutions of stochastic difference equations. Jin and An (2005) discussed the tail probability of nonparametric AR models with an additive heavy-tailed innovation and Pan and Wu (2005) studied the tail probability of nonparametric AR models with varying conditional variances. Asimita, Gerrarda, Houb, and Peng (2016) used tail dependence to examine financial extreme co-movements. Heavy-tailed time series has been extended to multivariate cases in Xie (2017).

This paper considers the tail behaviours of the general form of TAR model:

$$x_t = \sum_{i=1}^{s}\{\varphi_{i0} + \sum_{j=1}^{p}\varphi_{ij}x_{t-1} + \varepsilon_{it}\}I(r_{i-1} < x_{t-d} \leq r_i), \tag{1}$$

where $(\varepsilon_{1t},\cdots,\varepsilon_{st})'$ is independent of $\{x_h, h<t\}$, $I(\cdot)$ is the indicator function, $d$ is a positive integer and $-\infty = r_0 < r_1 < \cdots < r_s = \infty$. The following assumption is used:

**Assumption** *A*. *For each $i = 1,\cdots,s$, $\varepsilon_{it}$ has a continuous positive probability density function over the real line $R^1$, and*

$$\lim_{x\to\infty}\frac{P(|\varepsilon_{it}|>x)}{x^{-\alpha_i}L(x)} = 1$$

*for some $\alpha_i > 0$, where $L(x) > 0$ is a slowly varying function.*

We say that $\varepsilon_{it}$ has a heavy-tailed distribution if Assumption $A$ holds (sometimes it is said that $\varepsilon_{it}$ follows power-law, and some references called them heavy-tailed only when $\alpha_i < 2$) and $\alpha_i$ is called its tail index. Under Assumption $A$, model (1.1) is driven by multiple heavy-tailed innovations and hence it is a generalization of the traditional TAR in Chan et al. (1985) and Chen and Tsay (1991). These innovations may be different in different regimes. They may not be independent and may not have finite means or finite variances. Since this model is new in nature, we first discuss its geometric ergodicity in Section 2, which guarantees the existence and uniqueness of a stationary distribution for this model. Furthermore, we investigate the tail behaviours of the stationary distribution of model (1.1) in sections 3 and 4 for the following two cases:

**Case** 1. $\{(\varepsilon_{1t}, \cdots, \varepsilon_{st})'\}$ is an iid sequence of $s$-dimensional random vectors, and

$$\lim_{x\to\infty} \frac{P(|\varepsilon_{it}| > x, |\varepsilon_{jt}| > x)}{\min\{P(|\varepsilon_{it}| > x), P(|\varepsilon_{jt}| > x)\}} = 0, \quad 1 \le i < j \le s. \tag{2}$$

**Case** 2. $\varepsilon_{1t} = \cdots = \varepsilon_{st} = \varepsilon_t$, and $\{\varepsilon_t\}$ is an iid sequence of random variables.

In case 2, the innovations are the same for all regimes of $x_{t-d}$. But in case 1, different regimes could have different innovations. When $\varepsilon_{1t}, \cdots, \varepsilon_{st}$ are independent-in-tail, particularly independent, random variables, (2) is satisfied (see Asimita et al. (2016)).

Because the conditions for TAR(1) can be much weaker than those for TAR(p) ($p > 1$), we establish the results for TAR(1) and TAR(p) separately in different sections. For case 1, it is shown that the marginal tail probability has the same order as that of the innovation with the heaviest tail. Section 5 gives an upper bound for the tail conditional probability of model (1.1).

# 2 Geometric ergodicity of multiple-regime TAR models with heavy-tailed innovations

When $p = 1$, $d = 1$, $s = 2$, and $\varepsilon_{1t} = \varepsilon_{2t} = \varepsilon_t$, model (1.1) was studied by Petruccelli and Woolford (1984). They showed that $\{x_t\}$ is geometrically ergodic iff

$$\varphi_{11} < 1, \varphi_{21} < 1, \varphi_{11}\varphi_{21} < 1.$$

This result was extended to the case with $p = 1$, $d > 1, s = 2$ by Chen and Tsay (1991) and it was showed that $\{x_t\}$ is geometrically ergodic iff

$$\varphi_{11} < 1, \varphi_{21} < 1, \varphi_{11}\varphi_{21} < 1, \quad \text{and} \quad \varphi_{11}^{s(d)}\varphi_{21}^{t(d)} < 1, \varphi_{11}^{t(d)}\varphi_{21}^{s(d)} < 1,$$

where $s(d)$ and $t(d)$ are nonnegative integers depending on $d$, and $s(d)$ and $t(d)$ are odd and even numbers, respectively. When $p = 1$, $d = 1$, model (1.1) was studied by Chan et al. (1985). Cline and Pu (1999) discussed the stability of nonlinear AR(1) with delay under very general conditions. But their results can not be applied to model (1.1). In this section, we provide a simple sufficient condition for the geometric ergodicity of model (1.1) with $p=1$ and $p>1$ respectively.

We first introduce the following two lemmas.

**Lemma 2.1.** *Assume that $\{X_t\}$ is an aperiodic $\phi$- irreducible Markov chain, and let g be a nonnegative measurable function. Then $\{X_t\}$ is geometrically ergodic if there exist a small set C and constants $\lambda_1>0$, $\lambda_2>0$, $0<\rho<1$ such that*

$$\begin{array}{ll}(i) & E\{g(X_t)|X_{t-1}=x\}\le \rho g(x)-\lambda_1, \quad \textit{for any} \quad x\notin C;\\ (ii) & E\{g(X_t)|X_{t-1}=x\}\le \lambda_2, \quad \textit{for any} \quad x\in C.\end{array}$$

This lemma is called drift-criteria for the geometric ergodicity of a Markov chain, which comes from Tweedie (1975) (see also Nummelin (1984)). Before given another lemma, we needs the following notations:

$$\begin{aligned} X_t &= (x_t,\cdots,x_{t-m+1})', \\ \varphi(x_{t-1},\cdots,x_{t-m}) &= \sum_{j=1}^{s}(\varphi_{j0}+\varphi_{j1}x_{t-1}+\cdots+\varphi_{jp}x_{t-p})I(r_{j-1}<x_{t-d}\le r_j), \\ \Phi(X_{t-1}) &= (\varphi(x_{t-1},\cdots,x_{t-m}),x_{t-1},\cdots,x_{t-m+1}), \end{aligned}$$

where $m=max(p,d)$. We rewrite model (1.1) in a vector form as follows:

$$X_t=\Phi(X_{t-1})+\varepsilon(x_{t-d})e, \tag{3}$$

where $\varepsilon(x_{t-d})=\sum_{i=1}^{s}I(r_{i-1}<x_{t-d}\le r_i)\varepsilon_{it}$ and $e=(1,0,...,0)'$ is the $m-$dimensional unit vector.

**Lemma 2.2.** *Let $\mu_m$ be the Lebesgue measure on $R^m$. Under Assumption A, $\{X_t\}$ defined by (3) is an aperiodic $\mu_m$- irreducible Markov chain, and every bounded compact set with positive Lebesgue measure is a small set.*

*Proof.* We only give the proof of this lemma for the case with $p=1$ *and* $d=2$. When $p>1$ *or* $d>2$, the proof is similar. Firstly, take $A=[a_3,b_3)\times[a_2,b_2)$. Denote $z=(z_1,z_0)'$. Then, 2-step transition probability is

$$\begin{aligned} P_2(z,A) &= P\{X_3\in A\big|X_1=z\} \\ &= P\{x_3\in[a_3,b_3),x_2\in[a_2,b_2)\big|x_1=z_1,x_0=z_0\}. \end{aligned}$$

Suppose that, for example, $z_0\in(r_{j_0-1},r_{j_0}],z_1\in(r_{j_1-1},r_{j_1}]$. Then, by the definition of the model, we have

$$x_2 \;=\; \varphi_{j_0,0}+\varphi_{j_0,1}x_1+\varepsilon_{j_0,2}=\varphi_{j_0,0}+\varphi_{j_0,1}z_1+\varepsilon_{j_0,2}$$

$$\begin{aligned} x_3 &= \varphi_{j_1,0} + \varphi_{j_1,1}x_2 + \varepsilon_{j_1,3} \\ &= \varphi_{j_1,0} + \varphi_{j_1,1}\varphi_{j_0,0} + \varphi_{j_1,1}\varphi_{j_0,1}z_1 + \varphi_{j_1,1}\varepsilon_{j_0,2} + \varepsilon_{j_1,3}. \end{aligned}$$

Since $\varepsilon_{j_0,2}$ and $\varepsilon_{j_1,3}$ are independent, it follows that

$$\begin{aligned} P_2(z,A) &= P\{\varphi_{j_1,0} + \varphi_{j_1,1}\varphi_{j_0,0} + \varphi_{j_1,1}\varphi_{j_0,1}z_1 + \varphi_{j_1,1}\varepsilon_{j_0,2} + \varepsilon_{j_1,3} \in [a_3,b_3), \\ &\quad \varphi_{j_0,0} + \varphi_{j_0,1}z_1 + \varepsilon_{j_0,2} \in [a_2,b_2)\} \\ &\geq \int_{a_2-\varphi_{j_0,0}-\varphi_{j_0,1}z_1}^{b_2-\varphi_{j_0,0}-\varphi_{j_0,1}z_1} P\{a_3 - \varphi_{j_1,0} - \varphi_{j_1,1}\varphi_{j_0,0} - \varphi_{j_1,1}\varphi_{j_0,1}z_1 - \varphi_{j_1,1}u \\ &\quad < \varepsilon_{j_1,3} \leq b_3 - \varphi_{j_1,0} - \varphi_{j_1,1}\varphi_{j_0,0} - \varphi_{j_1,1}\varphi_{j_0,1}z_1 - \varphi_{j_1,1}u\} f_{j_0}(u)du \\ &= \int_{a_2}^{b_2}\int_{a_3}^{b_3} f_{j_0}(u - \varphi_{j_0,0} - \varphi_{j_0,1}z_1) \\ &\quad f_{j_1}(v - \varphi_{j_1,0} - \varphi_{j_1,1}\varphi_{j_0,0} - \varphi_{j_1,1}\varphi_{j_0,1}z_1 - \varphi_{j_1,1}u)dudv \\ &> 0, \end{aligned}$$

where $f_j(u)$ denotes the density function of $\varepsilon_{jt}$ and it is positive everywhere. In the same way, we can get a similar but more complicated inequality for $P_3(z,A)$. By Assumption $A$, for any $z \in R^2$ and any Borel measurable subset $A$ of $R^2$ with $\mu_2(A) > 0$, we have

$$P_2(z,A) > 0, \quad \text{and} \quad P_3(z,A) > 0.$$

Thus, the Markov chain $\{X_t\}$ is $\mu_2$- irreducible and aperiodic.

Finally, let $A$ be an arbitrary bounded compact set with $\mu_2(A) > 0$. Then, by Assumption $A$,

$$\begin{aligned} \nu = \inf_{(u,v)\in A, z\in A} \quad & f_{j_0}(u - \varphi_{j_0,0} - \varphi_{j_0,1}z_1) \\ & \cdot f_{j_1}(v - \varphi_{j_1,0} - \varphi_{j_1,1}\varphi_{j_0,0} - \varphi_{j_1,1}\varphi_{j_0,1}z_1 - \varphi_{j_1,1}u) > 0. \end{aligned}$$

Therefore, for any Borel subset $B \subset A$ with $\mu_2(B) > 0$, $\inf_{z\in A} P_2(z,B) \geq \nu\mu_2(B) > 0$. This implies that $A$ is a small set. This completes the proof. □

Now we are ready to state the following theorem for TAR(1) with general delay $d$.

**Theorem 2.1.** *Suppose Assumption A holds and* $\max\{|\varphi_{11}|, |\varphi_{s1}|\} < 1$. *Then*

*(a) the TAR(1) process $\{x_t\}$, defined by (1.1) with p=1, is geometrically ergodic, and hence it has a unique stationary distribution and is strong mixing with a geometric rate,*

*(b) the density function of the stationary distribution of $\{x_t\}$ is positive everywhere in $R^1$.*

*Proof.* (a) Let $\rho = \max\{|\varphi_{11}|, |\varphi_{s1}|\}$, $c_0 = \max_{1\le j\le s} |\varphi_{j0}|$, and $c_1 = \max_{1\le j\le s} |\varphi_{j1}|$. By the definition of $\varphi(\cdot)$, it follows that

$$\begin{aligned}
&|\varphi(x_{t-1},\cdots,x_{t-d})| \\
&\quad \le c_0 + \sum_{j=1}^{s} |\varphi_{j1}||x_{t-1}| I(r_{j-1} < x_{t-d} \le r_j) \\
&\quad \le c_0 + \rho|x_{t-1}|[I(-\infty < x_{t-d} \le r_1) + I(x_{t-d} > r_{s-1})] \\
&\qquad + \sum_{j=2}^{s-1} |\varphi_{j1}||x_{t-1}| I(r_{j-1} < x_{t-d} \le r_j).
\end{aligned} \tag{4}$$

Let $r = \max_{1\le r\le s-1} |r_j|$. Then $\sum_{j=2}^{s-1} I(r_{j-1} < x_{t-d} \le r_j) \le I(-r < x_{t-d} \le r)$. It is straightforward to show that when $-r < x_{t-d} < r$,

$$\begin{aligned}
|x_{t-d+1}| &= |\sum_{j=1}^{s} (\varphi_{j0} + \varphi_{j1} x_{t-d}) I(r_{j-1} < x_{t-2d+1} < r_j) + \varepsilon(x_{t-2d+1})| \\
&\le c_0 + c_1|x_{t-d}| + |\varepsilon(x_{t-2d+1})| \le c_0 + c_1 r + |\varepsilon(x_{t-2d+1})| \\
&\cdots \\
|x_{t-d+k}| &\le M_k + \sum_{i=1}^{k} c_1^{k-i} |\varepsilon(x_{t-2d+i})|, \\
&\cdots \\
|x_{t-1}| &\le M_{d-1} + \sum_{i=1}^{d-1} c_1^{d-1-i} |\varepsilon(x_{t-2d+i})|,
\end{aligned} \tag{5}$$

where $M_k = c_0 + c_0 c_1 + c_0 c_1^2 + \cdots + c_0 c_1^k + c_1^k r$. Let $c = c_0 + c_1 M_{d-1}$. Then, combining (4) and (5), we have

$$|\varphi(x_{t-1},\cdots,x_{t-d})| \le \rho|x_{t-1}| + c + \sum_{i=1}^{d-1} c_1^{d-i} |\varepsilon(x_{t-2d+i})|, \tag{6}$$

and

$$\begin{aligned}
|x_t| &\le |\phi(x_{t-1},\cdots,x_{t-d})| + |\varepsilon(x_{t-d})| \\
&\le \rho|x_{t-1}| + c + \sum_{i=1}^{d} c_1^{d-i} |\varepsilon(x_{t-2d+i})|.
\end{aligned} \tag{7}$$

Define a norm by

$$\|X\|_v = |x_1| + \rho|x_2| + \cdots + \rho^{d-1}|x_d| \; for \; X = (x_1,\cdots,x_d)' \in R^d.$$

By (6) and (7), it follows that

$$\|\Phi(X_{t-1})\|_v = |\varphi(x_{t-1},\cdots,x_{t-d})| + \rho|x_{t-1}| + \cdots + \rho^{d-1}|x_{t-d+1}|$$

$$
\begin{aligned}
&\leq \rho|x_{t-1}|+c+\sum_{i=1}^{d}c_1^{d-i}|\varepsilon(x_{t-2d+i})| \\
&\quad +\rho^2|x_{t-2}|+\rho c+\cdots+\rho\sum_{i=1}^{d}c_1^{d-i}|\varepsilon(x_{t-2d+i-1})| \\
&\quad +\cdots+\rho^d|x_{t-d}|+\rho^{d-1}c+\rho^{d-1}\sum_{i=1}^{d}c_1^{d-i}|\varepsilon(x_{t-2d+i-d+1})| \\
&= \rho\|X_{t-1}\|_v+c_\rho+H_{t-d}, \qquad (8)
\end{aligned}
$$

where $c_\rho = c+c\rho+\cdots+c\rho^{d-1}$, and

$$
\begin{aligned}
H_{t-d} &= \sum_{k=0}^{d-1}\rho^k\sum_{i=0}^{d-1}c_1^{d-i-1}|\varepsilon(x_{t-2d+i-k+1})| \\
&= \sum_{k=0}^{d-1}\sum_{h=0}^{d-1}\rho^k c_1^h|\varepsilon(x_{t-d-(k+h)})| \\
&= \sum_{m=0}^{2d-2}\Big(\sum_{0\leq h,k\leq d-1,k+h=m}\rho^k c_1^h\Big)|\varepsilon(x_{t-d-m})| \\
&= \sum_{l=1}^{2d-1}a_l|\varepsilon(x_{t-d-l+1})|,
\end{aligned}
$$

where $a_l=\sum_{0\leq h,k\leq d-1,k+h=l-1}\rho^k c_1^h$. Furthermore, since $|\varepsilon(x_{t-d-l+1})|\leq\sum_{j=1}^{s}|\varepsilon_{j,t-l+1}|$, it follows that

$$
H_{t-d}\leq R_{t-d}=:\sum_{l=1}^{2d-1}a_l\sum_{j=1}^{s}|\varepsilon_{j,t-l+1}|. \qquad (9)
$$

Firstly, by Lemma 2.2, $\{X_t\}$ is an aperiodic and irreducible Markov chain. Pick a constant $\delta\in(0,\min\{1,\alpha_1,\cdots,\alpha_s\})$. By Assumption A, it follows that

$$
\begin{aligned}
E|\varepsilon(x_{t-d})|^\delta &= E|\sum_{j=1}^{s}\varepsilon_{jt}I(r_{j-1}<x_{t-d}<r_j)|^\delta \\
&\leq \sum_{j=1}^{s}E|\varepsilon_{jt}|^\delta =: \Gamma_\delta<\infty,
\end{aligned}
$$

and

$$
ER_{t-d}^\delta \leq \sum_{l=1}^{2d-1}a_l^\delta\Big(\sum_{j=1}^{s}E|\varepsilon_{j,t-l+1}|\Big)^\delta
$$

$$\leq \sum_{l=1}^{2d-1} a_l^\delta \sum_{j=1}^{s} E|\varepsilon_{j,t-l+1}|^\delta =: \Lambda_\delta < \infty.$$

Now we choose the test function as

$$g(X) = 1 + \|X\|_v^\delta, \quad \text{for} \quad X = (x_1, \cdots, x_d) \in R^d.$$

From (3), we have

$$\begin{aligned} E\{g(X_t)|X_{t-1} = X\} &= E\{1 + \|\Phi(X) + e\varepsilon(x_d)\|_v^\delta\} \\ &\leq 1 + E[\|\Phi(X)\|_v + \|e\varepsilon(x_d)\|_v]^\delta \\ &\leq 1 + E\{\|\Phi(X)\|_v^\delta + (|\varepsilon(x_d)|\|e\|_v)^\delta\} \\ &\leq \rho^\delta \|X\|_v^\delta + ER_{t-d}^\delta + c_\rho^\delta + E|\varepsilon(x_d)|^\delta + 1 \\ &\leq \rho^\delta \|X\|_v^\delta + \Lambda_\delta + c_\rho^\delta + \Gamma_\delta + 1. \end{aligned}$$

Taking $\lambda$ and $M$ such that $0 < \rho^\delta < \lambda < 1$ and

$$M > \frac{\Lambda_\delta + c_\rho^\delta + \Gamma_\delta + 1}{\lambda - \rho^\delta},$$

we get, when $\|X\|_v > M$,

$$E\{g(X_t)|X_{t-1} = X\} \leq \lambda g(X) - \lambda_1,$$

where $\lambda_1 = (\lambda - \rho^\delta)M^\delta - \Lambda_\delta - c_\rho^\delta - \Gamma_\delta - 1$. Denote $C = \{X : \|X\|_v \leq M\}$. By Lemma 2.2, $C$ is a small set. For any $X \in C$,

$$E\{g(X_t)|X_{t-1} = X\} \leq \lambda_2$$

where $\lambda_2 = \rho^\delta M^\delta + \Lambda_\delta + c_\rho^\delta + \Gamma_\delta + 1$. Then, from Lemma 2.1, $\{X_t\}$ is geometrically ergodic. This implies that $\{x_t\}$ is geometrically ergodic, and hence it has a unique stationary distribution and satisfies the strong mixing condition: the mixing coefficient

$$\alpha(l) = \sup_k \sup_{A \in \mathscr{F}_{-\infty}^k, B \in \mathscr{F}_{k+l}^\infty} |P(A \cap B) - P(A)P(B)|$$

goes to zero exponentially fast as $l \to \infty$, where $\mathscr{F}_{-\infty}^k$ and $\mathscr{F}_{k+l}^\infty$ are the $\sigma$-algebras generated by $\{x_t\}_{t=-\infty}^k$ and $\{x_t\}_{t=k+l}^\infty$ respectively.

(b) Let $F(x)$ be the stationary distribution function of $x_t$, and denote the density function of $\varepsilon_{it}$ by $f_i$, $i = 1, \cdots, s$. Note that, in model (1.1), these innovations' densities are assumed to be continuous and positive over the whole line $R^1$, and $(\varepsilon_{1t}, \cdots, \varepsilon_{st})$ is independent of $\{x_s, s < t\}$. Then

$$\begin{aligned} F(x) &= P\{x_t \le x\} \\ &= \sum_{j=1}^{s} P\{x_t \le x \big| r_{j-1} < x_{t-d} \le r_j\} P\{r_{j-1} < x_{t-d} \le r_j\} \\ &= \sum_{j=1}^{s} [F(r_j) - F(r_{j-1})] \int_{-\infty}^{+\infty} P\{x_t \le x \big| r_{j-1} < x_{t-d} \le r_j, x_{t-1} = y\} dF(y) \\ &= \sum_{j=1}^{s} [F(r_j) - F(r_{j-1})] \int_{-\infty}^{+\infty} P\{\varepsilon_{jt} \le x - [\varphi_{j0} + \varphi_{j1} y]\} dF(y) \\ &= \sum_{j=1}^{s} [F(r_j) - F(r_{j-1})] \int_{-\infty}^{+\infty} [\int_{-\infty}^{x} f_j(z - [\varphi_{i0} + \varphi_{j1} y]) dz] dF(y) \\ &= \int_{-\infty}^{x} \Big\{ \sum_{j=1}^{s} [F(r_j) - F(r_{j-1})] \int_{-\infty}^{+\infty} f_j(z - [\varphi_{j0} + \varphi_{j1} y]) dF(y) \Big\} dz, \end{aligned}$$

where the last step is guaranteed by the Fubini Theorem. Hence, $F(x)$ has a density function

$$f(x) = \sum_{j=1}^{s} [F(r_j) - F(r_{j-1})] \int_{-\infty}^{+\infty} f_j(x - [\varphi_{j0} + \varphi_{j1} y]) dF(y).$$

Furthermore, since $f_j > 0, j = 1, \cdots, s$, $f(x)$ is strictly positive everywhere. This completes the proof. □

**Remark 2.1**. Comparing the above theorem with those in Petruccelli and Woolford (1984), Chan et al. (1985) and Chen and Tsay (1991), one can see that

$$\max\{|\varphi_{11}|, |\varphi_{s1}|\} < 1$$

is a stronger condition. But it keeps an important feature, as in Chan et al (1985), that the coefficients in middle regimes are irrelevant to the ergodicity of model (1.1) with $p = 1$. It is difficult to get a necessary and sufficient condition for the ergodicity of model (1.1) with $p = 1$. This remains as an open problem for the heavy-tailed case.

For the geometric ergodicity of model (1.1) with $p > 1$, we have the following theorem.

**Theorem 2.2.** *Under Assumption A and the condition*

$$\lambda := \max_{1\le i\le s}\sum_{j=1}^{p}|\varphi_{ij}| < 1, \tag{10}$$

*(1) model (1.1) is geometrically ergodic. Hence, it has a unique stationary distribution with positive density function in $R^1$ and is strong mixing with a geometric rate.*

*(2) the density function of the stationary distribution of model (1.1) is positive everywhere in $R^1$.*

*Proof.* (1) Under Assumption $A$, $\{X_t\}$ is an aperiodic irreducible Markov chain. If condition (10) holds, then

$$\begin{aligned}|\phi(x_{t-1},\cdots,x_{t-m})| &= |\sum_{i=1}^{s} I(r_{i-1}<x_{t-d}\le r_i)[\varphi_{i0}+\sum_{j=1}^{p}\phi_{ij}x_{t-j}]| \\ &\le \lambda\max\{|x_{t-1}|,\cdots,|x_{t-p}|\}+c_0 \\ &\le \lambda\max\{|x_{t-1}|,\cdots,|x_{t-m}|\}+c_0\end{aligned}$$

where $m=\max\{p,d\}$ and $c_0=\max|\varphi_{i0}|$. For $0<\lambda<1$, we can take positive constants $b_1>b_2>\cdots>b_m>0$ and $0<\rho<1$ such that

$$\lambda<\frac{b_1\lambda}{b_i}<\rho,\quad i=2,\cdots,m,\quad \frac{b_{i+1}}{b_i}<\rho, i=1,\cdots,i=1,\cdots,m-1.$$

Consider a norm defined by

$$\|X\|_b=\max\{b_1|x_1|,\cdots,b_m|x_m|\}\quad \text{for}\quad X=(x_1,\cdots,x_m)'\in R^m.$$

Then

$$\begin{aligned}\|\Phi(X_{t-1})\|_b &= \max\{b_1\varphi(x_{t-1},\cdots,x_{t-m}),b_2|x_{t-1}|,\cdots,b_m|x_{t-m+1}|\} \\ &\le \max\{b_1\lambda\max\{|x_{t-1}|,\cdots,|x_{t-m}|\}+b_1c_0,b_2|x_{t-1}|,\cdots,b_m|x_{t-m+1}|\} \\ &\le \max\{\max\{b_1\lambda|x_{t-1}|,\frac{b_1\lambda}{b_2}b_2|x_{t-2}|,\cdots,\frac{b_1\lambda}{b_m}b_m|x_{t-m}|\}, \\ &\quad \frac{b_2\lambda}{b_1}b_1|x_{t-1}|,\cdots,\frac{b_m}{b_{m-1}}b_{m-1}|x_{t-m+1}|\}+b_1c_0 \\ &\le \max\{\rho\max\{b_1|x_{t-1}|,b_2|x_{t-2}|,\cdots,b_m|x_{t-m}|\},\rho b_1|x_{t-1}|, \\ &\quad \cdots,\rho b_{m-1}|x_{t-m+1}|\}+b_1c_0 \\ &\le \rho\max\{b_1|x_{t-1}|,\cdots,b_m|x_{t-m}|\}+b_1c_0\end{aligned}$$

$$= \quad \rho\|X_{t-1}\|_b + C_\rho.$$

Therefore, taking the test function $g(X) = 1 + \|X\|_b^\delta$, where $0 < \delta < \min\{1, \alpha_1, ..., \alpha_s\}$, we can easily prove the geometric ergodicity of $\{X_t\}$ by the drift criteria.

(2) Note that

$$\begin{aligned}
F(x) &= P\{x_t \le x\} \\
&= \sum_{j=1}^{s} [F(r_j) - F(r_{j-1})] \\
&\quad \cdot \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} [\int_{-\infty}^{x} f_j(z - [\varphi_{i0} + \varphi_{j1}y_1 + \cdots + \varphi_{jp}y_p])dz] dF(y_1, ..., y_p) \\
&= \int_{-\infty}^{x} \{ \sum_{j=1}^{s} [F(r_j) - F(r_{j-1})] \\
&\quad \cdot \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_j(z - [\varphi_{j0} + \varphi_{j1}y_1 + \cdots + \varphi_{jp}y_p]) dF(y_1, ..., y_p) \} dz.
\end{aligned}$$

Because $f_j$ is positive everywhere, $F$ has a density function which is positive everywhere. This completes the proof.

□

# 3 Tail behaviours of stationary distribution for multiple TAR(1) with heavy-tailed innovations

This section gives the tail shape of the marginal stationary distribution of a TAR(1) model. We first give the following notations:

$$\begin{aligned}
a_l &= \sum_{m=\max(0,l-d)}^{\min(d-1,l-1)} \rho^{l-1-m} c_1^m, \\
c_1 &= \max_{1 \le j \le s} |\varphi_{j1}|, \\
\rho &= \max\{|\varphi_{11}|, |\varphi_{s1}|\} < 1, \\
\beta &= \min(\alpha_1, \cdots, \alpha_s), \\
N &= \text{the number of the elements in} \quad \{i : 1 \le i \le s, \alpha_i = \beta\}, \\
\tau &= \min_{1 \le j \le s} P\{r_{j-1} < x_{t-d} \le r_j\} > 0.
\end{aligned}$$

Our main result is as follows.

**Theorem 3.1.** *Suppose Assumption A holds and* $\max\{|\varphi_{11}|,|\varphi_{s1}|\} < 1$. *Then*

$$K_1 \le \liminf_{x\to\infty} \frac{P\{|x_t| > x\}}{x^{-\beta}L(x)} \le \limsup_{x\to\infty} \frac{P\{|x_t| > x\}}{x^{-\beta}L(x)} \le K_2$$

*for Case 1 with* $K_1 = \tau N(1+c_1)^{-\beta}$,

$$K_2 = \begin{cases} \frac{N2^\beta}{1-\rho^\beta}, & \text{if} \quad d = 1, \\ \frac{N(2^\beta + \sum_{l=2}^{2d-1} a_l^\beta)}{1-\rho^\beta}, & \text{if} \quad d > 1; \end{cases}$$

*for* <u>*Case 2*</u> *with* $K_1 = (1+c_1)^{-\beta}$,

$$K_2 = \begin{cases} \frac{2^\beta}{1-\rho^\beta}, & \text{if} \quad d = 1, \\ \frac{2^\beta + \sum_{l=2}^{2d-1} a_l^\beta}{1-\rho^\beta}, & \text{if} \quad d > 1. \end{cases}$$

To prove this theorem, we need the following two lemmas. They have general and independent interest for heavy-tailed random variables.

**Lemma 3.1.** *Under Assumptions A and condition (2), it follows that*

$$P\{\sum_{i=1}^{s} |\varepsilon_{it}| > x\} \sim \sum_{i=1}^{s} P\{|\varepsilon_{it}| > x\} \sim Nx^{-\beta}L(x),$$

*as* $x \to \infty$.

*Proof.* Let $G$ be a distribution function such that

$$1 - G(x) \sim x^{-\beta}L(x), \quad \text{as} \quad x \to \infty.$$

Then,

$$\frac{P\{|\varepsilon_{kt}| > x\}}{1-G(x)} \to c_k = \begin{cases} 0, & \text{if} \quad \alpha_k > \beta, \\ 1, & \text{if} \quad \alpha_k = \beta, \end{cases} \quad k = 1, \cdots, s,$$

as $x \to \infty$. Take a sequence of positive constant $\{u_n\}$ such that

$$n[1 - G(u_n)] \to 1, \qquad \text{as} \quad n \to \infty, \tag{11}$$

and define

$$\nu_{n,k}(A) = nP\{u_n^{-1}|\varepsilon_{kt}| \in A\} \qquad \text{for} \quad A \in \mathscr{B}(0,\infty],$$

where $\mathscr{B}(0,\infty]$ denotes the collection of all Borel measurable subsets of $(0,+\infty]$. Note that when $A=(x,+\infty)$, $x>0$,

$$\nu_{n,k}(A) = nP\{|\varepsilon_{kt}| \geq u_n x\} \to c_k x^{-\beta},$$

as $n\to\infty$. We have

$$\nu_{n,k} \xrightarrow{v} \nu_k \tag{12}$$

in $(0,\infty]$ for any fixed $k=1,\cdots,s$ as $n\to\infty$, where

$$\nu_k(dx) = c_k \beta x^{-\beta-1}dx, x>0, \tag{13}$$

and $\xrightarrow{v}$ denotes vague convergence of Radon measures on the space $(0,+\infty]$ (see Resnick (1987) or Kallenberg (1983)).

Under condition (2), for $\{A_k \in \mathscr{B}(0,\infty], k=1,\cdots,s\}$ such that there are at least two sets, say $A_i$ and $A_j$,$i\neq j$, with the form $(x,+\infty]$ and $x>0$, we obtain

$$\begin{aligned}
& nP\{u_n^{-1}(|\varepsilon_{1t}|,\cdots,|\varepsilon_{st}|) \in A_1\times\cdots\times A_s\} \\
& \leq nP\{u_n^{-1}(|\varepsilon_{it}|,|\varepsilon_{jt}|) \in A_i \times A_j\} \\
& = nP\{|\varepsilon_{it}| > u_n x, |\varepsilon_{jt}| > u_n x\} \\
& = \frac{P\{|\varepsilon_{it}| > u_n x, |\varepsilon_{jt}| > u_n x\}}{1-G(u_n)}(1+o(1)) \\
& \leq \frac{P\{|\varepsilon_{it}| > u_n x, |\varepsilon_{jt}| > u_n x\}}{\min\{P\{|\varepsilon_{it}| > u_n x\}, P\{|\varepsilon_{jt}| > u_n x\}\}} \cdot \frac{1-G(u_n x)}{1-G(u_n)}(1+o(1)) \\
& \to 0\cdot x^{-\beta} = 0,
\end{aligned}$$

as $n\to\infty$. Define

$$\theta_n(\mathbf{A}) = nP\{u_n^{-1}(|\varepsilon_{1t}|,\cdots,|\varepsilon_{st}|) \in \mathbf{A}\} \quad \text{for} \quad \mathbf{A}\in\mathscr{B}([0,\infty]^s\setminus\{0\}).$$

We have

$$\theta_n \xrightarrow{v} \theta \tag{14}$$

on $[0,+\infty]^s\setminus\{0\}$, where $\theta$ is a measure on $[0,+\infty]^s\setminus\{0\}$ such that

$$\theta\{y\mathbf{e_k} : y>x\} = \nu_k\{y : y>x\} = c_k x^{-\beta}, \quad x>0,$$

and $\theta\{\cap_{k=1}^s \{y\mathbf{e_k} : y\neq 0\}^c\} = 0$, where $\mathbf{e_k}\in R^s$ is the basis element with $k$th component equal to one and the rest zero. It is easy to see that for any $\mathbf{A}\in\mathscr{B}([0,\infty]^s\setminus\{0\})$,

$$\theta(\mathbf{A}) = \sum_{k=1}^{s} \nu_k\{y\in[0,+\infty] : ye_k \in A\}. \tag{15}$$

Denote

$$\mathbf{Y^{(s)}} = (|\varepsilon_{1t}|,\cdots,|\varepsilon_{st}|)',$$

$$\mathbf{1}^{(\mathbf{s})} = (1, \cdots, 1)',$$
$$\mathbf{B} = \{(y_1, \cdots, y_s)' \in [0, +\infty]^s \setminus \{0\} : \sum_{k=1}^{s} y_k > 1\}.$$

Then,

$$\begin{aligned}\frac{P(\sum_{k=1}^{s} |\varepsilon_{kt}| > u_n)}{1 - G(u_n)} =& nP(\mathbf{1}^{(\mathbf{s})'}\mathbf{Y}^{(\mathbf{s})} > u_n)(1 + o(1)) \\ =& nP(u_n^{-1}\mathbf{Y}^{(\mathbf{s})} \in \mathbf{B})(1 + o(1)) = \theta_n(\mathbf{B})(1 + o(1)) \to \theta(\mathbf{B}),\end{aligned}$$

as $n \to \infty$. But,

$$\begin{aligned}\theta(B) &= \sum_{k=1}^{s} \nu_k\{y \in [0, +\infty] : ye_k \in B\} \\ &= \sum_{k=1}^{s} \nu_k\{y \in [0, +\infty] : y > 1\} \\ &= \sum_{k=1}^{s} \int_{\{y>1\}} \nu_k(dy) = \sum_{k=1}^{s} c_k = N.\end{aligned}$$

It is easily seen that the above equality holds when $u_n$ is replaced by $x$ as $x \to \infty$. Therefore,

$$P\{\sum_{i=1}^{s} |\varepsilon_{it}| > x\} \quad \sim \quad Nx^{-\beta}L(x).$$

It is obvious that$\sum_{i=1}^{s} P\{|\varepsilon_{it}| > x\} \sim Nx^{-\beta}L(x)$. This completes the proof. □

**Lemma 3.2.** *For a slowly varying function $L(x)$, it follows that*

$$\frac{L(ax + b)}{L(x)} \to 1,$$

*as $x \to \infty$ for any constants $a > 0$ and $b \in R^1$.*

*Proof.* The result comes directly from the Karamata representative (see Embrechts et al. (1997)) of a slowly varying function:

$$L(x) = c(x)\exp\{\int_{z}^{x} \frac{\delta(u)}{u} du\}, \quad x \geq z,$$

for some $z > 0$, where $c$ and $\delta$ are measurable functions satisfying $c(x) \to c_0 \in (0, +\infty)$ and $\delta(x) \to 0$ as $x \to \infty$. This completes the proof. □

Now we are ready to prove Theorem 3.1.

**Proof of Theorem 3.1**

Using the notation as in the proof of Theorem 2.1, by (2.6) and (2.7), it follows that

$$\begin{aligned}
\|X_t\|_v &= \|\Phi(X_{t-1})+\varepsilon(x_{t-d})e\|_v \\
&\leq \|\Phi(X_{t-1})\|_v+\|\varepsilon(x_{t-d})e\|_v \\
&\leq \rho\|X_{t-1}\|_v+c_\rho+R_{t-d}+|\varepsilon(x_{t-d})| \\
&\leq \rho^n\|X_{t-n}\|_v+c_\rho\frac{1-\rho^n}{1-\rho}+\sum_{k=0}^{n-1}\rho^k[|\varepsilon(x_{t-d-k})|+R_{t-d-k}] \\
&\leq \frac{c_\rho}{1-\rho}+\sum_{k=0}^{\infty}\rho^k[|\varepsilon(x_{t-d-k})|+R_{t-d-k}] \qquad (\rho^n\|X_{t-n}\|_v\longrightarrow 0) \\
&\leq \frac{c_\rho}{1-\rho}+\sum_{k=0}^{\infty}\rho^k\eta_{t,k}, \qquad (16)
\end{aligned}$$

where

$$R_{t-d} =: \sum_{l=1}^{2d-1} a_l \sum_{j=1}^{s} |\varepsilon_{j,t-l+1}|,$$

and

$$\eta_{t,k} = 2\sum_{j=1}^{s}|\varepsilon_{j,t-k}|+\sum_{l=2}^{2d-1} a_l \sum_{j=1}^{s}|\varepsilon_{j,t-k-l+1}|,$$

because of $a_1=1$, and $\sum_a^b=0$ as $a>b$; For case 2, noting that $\varepsilon(x_{t-d})=\varepsilon_t$ in this case, we have

$$\|X_t\|_v \leq \rho\|X_{t-1}\|_v+c_\rho+H_t+|\varepsilon_t|, \qquad (17)$$

where

$$H_t = \sum_{l=1}^{2d-1} a_l|\varepsilon_{t-l+1}|.$$

Similarly to (16), we get

$$\|X_t\|_v \leq \frac{c_\rho}{1-\rho}+\sum_{k=0}^{\infty}\rho^k\eta'_{t,k},$$

where

$$\eta'_{t,k} = 2|\varepsilon_{t-k}| + \sum_{l=2}^{2d-1} a_l |\varepsilon_{t-k-l+1}|.$$

Hence, for case 1,

$$P\{\|X_t\|_v > x\} \le P\{\frac{c_\rho}{1-\rho} + \sum_{k=0}^{\infty} \rho^k |\eta_{t,k}| > x\}; \tag{18}$$

and for case 2,

$$P\{\|X_t\|_v > x\} \le P\{\frac{c_\rho}{1-\rho} + \sum_{k=0}^{\infty} \rho^k |\eta'_{t,k}| > x\}; \tag{19}$$

Therefore, for case 1, since $(\varepsilon_{1,t_1}, \cdots, \varepsilon_{s,t_1})'$ is independent of $(\varepsilon_{1,t_2}, \cdots, \varepsilon_{s,t_2})'$ when $t_1 \neq t_2$, by lemmas 3.1 and 3.2, we have

$$\begin{aligned} P\{\eta_{t,k} > x\} &\sim P\{2\sum_{j=1}^{s} |\varepsilon_{j,t-k}| > x) + \sum_{l=2}^{2d-1} P\{a_l \sum_{j=1}^{s} |\varepsilon_{j,t-k-l+1}| > x\} \\ &\sim \sum_{j=1}^{s} P\{2|\varepsilon_{j,t-k}| > x) + \sum_{l=2}^{2d-1} \sum_{j=1}^{s} P\{a_l |\varepsilon_{j,t-k-l+1}| > x\} \\ &\sim Nx^{-\beta} L(x)[2^\beta \frac{L(x/2)}{L(x)} + \sum_{l=2}^{2d-1} a_l^\beta \frac{L(x/a_l)}{L(x)}] \\ &\sim Nx^{-\beta} L(x)[2^\beta + \sum_{l=2}^{2d-1} a_l^\beta] \end{aligned}$$

as $x \to \infty$; For case 2,

$$\begin{aligned} P\{\eta'_{t,k} > x\} &\sim P\{2|\varepsilon_{t-k}| > x) + P\{\sum_{l=2}^{2d-1} a_l |\varepsilon_{t-k-l+1}| > x\} \\ &\sim P\{|\varepsilon_{t-k}| > \frac{x}{2}) + \sum_{l=2}^{2d-1} P\{|\varepsilon_{t-k-l+1}| > \frac{x}{a_l}\} \\ &\sim x^{-\beta} 2^\beta L(\frac{x}{2}) + \sum_{l=2}^{2d-1} (\frac{x}{a_l})^{-\beta} L(\frac{x}{a_l}) \\ &\sim x^{-\beta} L(x)[2^\beta + \sum_{l=2}^{2d-1} a_l^\beta] \end{aligned}$$

as $x \to \infty$.

So, for case 1, by lemmas 3.1-3.2, Assumption A, and Cline's result (Cline (1983)),

$$\begin{aligned}\limsup_{x\to\infty}\frac{P\{\|X_t\|_v > x\}}{x^{-\beta}L(x)} &\leq \lim_{x\to\infty}\frac{P\{\sum_{k=0}^{\infty}\rho^k\eta_{t,k} > x-\frac{C_\rho}{1-\rho}\}}{P\{\eta_{t,k} > x-\frac{C_\rho}{1-\rho}\}} \\ &\quad \cdot\lim_{x\to\infty}\frac{N(2^\beta+\sum_{l=2}^{2d-1}a_l^\beta)(x-\frac{c_\rho}{1-\rho})^{-\beta}L(x-\frac{c_\rho}{1-\rho})}{x^{-\beta}L(x)} \\ &= N(2^\beta+\sum_{l=2}^{2d-1}a_l^\beta)\sum_{k=0}^{\infty}\rho^{k\beta} = \frac{N(2^\beta+\sum_{l=2}^{2d-1}a_l^\beta)}{1-\rho^\beta};\end{aligned}$$

For case 2,

$$\limsup_{x\to\infty}\frac{P\{\|X_t\|_v > x\}}{x^{-\beta}L(x)} \leq \frac{2^\beta+\sum_{l=2}^{2d-1}a_l^\beta}{1-\rho^\beta}.$$

Since $|x_t| \leq \|X_t\|_v$, it follows that

$$\limsup_{x\to\infty}\frac{P\{|x_t| > x\}}{x^{-\beta}L(x)} \leq \limsup_{x\to\infty}\frac{P\{\|X_t\|_v > x\}}{x^{-\beta}L(x)} \leq K_2.$$

On the other hand, note that

$$|x_t| \geq |\varepsilon(x_{t-d})| - |\varphi(x_{t-1},x_{t-2},\ldots,x_{t-d})| \geq |\varepsilon(x_{t-d})| - c_1|x_{t-1}| - c_0.$$

Hence

$$\begin{aligned}P\{|x_t| > x\} &\geq P\{|\varepsilon(x_{t-d})| - c_1|x_{t-1}|_v - c_0 > x\} \\ &\geq P\{|\varepsilon(x_{t-d})| - c_1|x_{t-1}| - c_0 > x, |x_{t-1}| \leq x\} \\ &\geq P\{|\varepsilon(x_{t-d})| > (1+c_1)x + c_0, |x_{t-1}| \leq x\}\end{aligned}$$

Therefore, for case 1,

$$\begin{aligned}P\{|x_t| > x\} &= \sum_{j=1}^{s}P\{|\varepsilon_{jt}| > (1+c_1)x+c_0, r_{j-1} < x_{t-d} \leq r_j, |x_{t-1}| \leq x\} \\ &= \sum_{j=1}^{s}P\{|\varepsilon_{jt}| > (1+c_1)x+c_0\}P\{r_{j-1} < x_{t-d} \leq r_j, |x_{t-1}| \leq x\} \\ &\geq \min_{1\leq j\leq s}P\{r_{j-1} < x_{t-d} \leq r_j, |x_{t-1}| \leq x\}\end{aligned}$$

$$\sum_{j=1}^{s} P\{|\varepsilon_{jt}| > (1+c_1)x + c_0\}. \tag{20}$$

Thus, by lemmas 3.1-3.2, we have, for case 1,

$$\begin{aligned}
\liminf_{x\to\infty} \frac{P\{|x_t| > x\}}{x^{-\beta}L(x)} &\geq \lim_{x\to\infty} \frac{\sum_{j=1}^{s} P\{|\varepsilon_{jt}| > (1+c_1)x + c_0\}}{x^{-\beta}L(x)} \\
&\quad \lim_{x\to\infty} \min_{1\leq j\leq s} P\{r_{j-1} < x_{t-d} \leq r_j, |x_{t-1}| < x\} \\
&= N(1+c_1)^{-\beta} \cdot \min_{1\leq j\leq s} P\{r_{j-1} < x_{t-d} \leq r_j\} \\
&= \tau N \cdot (1+c_1)^{-\beta},
\end{aligned}$$

where $\tau = \min_{1\leq j\leq s} P\{r_{j-1} < x_{t-d} \leq r_j\} = \min_{1\leq j\leq s} \int_{r_{j-1}}^{r_j} f(u)du > 0$ by Theorem 2.1; For case 2, noticing that $\varepsilon(x_{t-d}) = \varepsilon_t$, we have

$$\begin{aligned}
P\{|x_t| > x\} &= \sum_{j=1}^{s} P\{|\varepsilon_t| > (1+c_1)x + c_0, |x_{t-1}| \leq x\} \\
&= \sum_{j=1}^{s} P\{|\varepsilon_t| > (1+c_1)x + c_0\} P\{|x_{t-1}| \leq x\}, \qquad (21)
\end{aligned}$$

and then

$$\begin{aligned}
\liminf_{x\to\infty} \frac{P\{|x_t| > x\}}{x^{-\beta}L(x)} &\geq \lim_{x\to\infty} \frac{P\{|\varepsilon_t| > (1+c_1)x + c_0\}}{x^{-\beta}L(x)} \cdot \lim_{x\to\infty} P\{|x_{t-1}| < x\} \\
&= (1+c_1)^{-\beta}.
\end{aligned}$$

This completes the proof. $\square$

**Remark 3.1** Combining (22) and (3.7), we have

$$\begin{aligned}
&\sum_{j=1}^{s} P\{|\varepsilon_{jt}| > (1+c_1)x + c_0\} \min_{1\leq j\leq s} P\{r_{j-1} < x_{t-d} \leq r_j, |x_{t-1}| < x\} \\
&\quad \leq P\{|x_t| > x\} \leq P\{\sum_{j=0}^{\infty} \rho^j |\eta_{t,j}| > x - \frac{c_\rho}{1-\rho}\}.
\end{aligned}$$

This inequality shows that $x_t$ is light-tailed if $\{\varepsilon_t\}$ are i.i.d. normal. So model (1.1) turns light-tailed input into light-tailed output, heavy-tailed input into heavy-tailed output. This feature is similar to that of linear process.

**Remark 3.2** Under the assumption of Theorem 2.1, the coefficients in middle regimes in model (1.1) do not make any contribution to the existence of the stationary distribution. However, Theorem 3.1 shows that the innovations in these regimes may play an important role to the tail index of the stationary distribution. This is a new and interesting finding.

# 4 Tail behaviours of stationary distribution for multiple TAR(p) with heavy-tailed innovations

For the general order case, we can get similar results to the first-order case but in stronger conditions on autoregressive coefficients.

**Theorem 4.1.** *Suppose model (1.1) satisfies Assumption A and* $\max_{1\le i\le s}\sum_{j=1}^{p}|\varphi_{ij}|<1$. *Then, for case 1 and case 2, there exist two constants* $0<K_1<K_2<+\infty$ *such that the stationary distribution of TAR(p) satisfies*

$$K_1 \le \liminf_{x\to\infty}\frac{P\{|x_t|>x\}}{x^{-\beta}L(x)} \le \limsup_{x\to\infty}\frac{P\{|x_t|>x\}}{x^{-\beta}L(x)} \le K_2.$$

*Proof.* By using the notations in the proof of Theorem 2.2, we have

$$\begin{aligned}
\|X_t\|_b &= \|\Phi(X_{t-1})+\varepsilon(x_{t-d})e\|_b \\
&\le \|\Phi(X_{t-1})\|_b+|\varepsilon_t|\|e\|_b \\
&\le \rho\|X_{t-1}\|_b+C_\rho+|\varepsilon_t|C_u \qquad (C_u=\|e\|_b) \\
&\le \rho^n\|X_{t-n}\|_b+C_\rho\frac{1-\rho^n}{1-\rho}+C_u\sum_{j=0}^{n-1}\rho^j|\varepsilon(x_{t-d-j}| \\
&\le \frac{C_\rho}{1-\rho}+C_u\sum_{j=0}^{\infty}\rho^j|\varepsilon(x_{t-d-j})| \qquad (\rho^n\|X_{t-n}\|_b\longrightarrow 0) \\
&\le \frac{C_\rho}{1-\rho}+C_u\sum_{j=0}^{\infty}\rho^j[\sum_{i=1}^{s}|\varepsilon_{i,t-j})|].
\end{aligned}$$

Then,

$$\begin{aligned}
P\{\|X_t\|_b>x\} &\le P\{\frac{C_\rho}{1-\rho}+C_u\sum_{j=0}^{\infty}\rho^j[\sum_{i=1}^{s}|\varepsilon_{i,t-j})|]>x\} \\
&= P\{\sum_{j=0}^{\infty}\rho^j[\sum_{i=1}^{s}|\varepsilon_{i,t-j})|]>\frac{x-\frac{C_\rho}{1-\rho}}{c_u}\}. \qquad (22)
\end{aligned}$$

But $|x_t|\le b_1^{-1}\|X_t\|_b$, and then $P\{|x_t|>x\}\le P\{\|X_t\|_b>b_1x\}$.

On the other hand, note that

$$|x_t| \geq |\varepsilon(x_{t-d})| - |\varphi(x_{t-1}, x_{t-2}, ..., x_{t-m})|$$

and

$$|\varphi(x_{t-1}, x_{t-2}, ..., x_{t-m})| \leq b_1^{-1} \|\Phi(X_{t-1})\|_b \leq b_1^{-1} \rho \|X_{t-1}\|_b + b_1^{-1} C_\rho.$$

Then, the result of Theorem 4.1 can be obtained by the same way as that in the proof of Theorem 3.1.

□

**Remark 4.1.** From the proofs of Theorem 3.1 and Theorem 4.1, it can be seen that $x_t$ is light-tailed if $\{\varepsilon_{it}\}$ are normal. So model (1.1) turns light-tailed input into light-tailed output, heavy-tailed input into heavy-tailed output. This feature is similar to that of linear process.

# 5 The auto-tail-dependence of multiple TAR models with heavy-tailed innovations

The tail conditional probability

$$P\{|x_{t_2}| > x \big| |x_{t_1}| > x\} \quad \text{for large} \quad x > 0$$

can describe the possibility that an extreme event will occur again at time $t_2$ when such an extreme event has already occurred at time $t_1$. It is very useful in practice, particularly in financial risk analysis, see Poon, Rockingger, and Tawn (2003). We call this tail conditional probability the auto-tail-dependence, which is a measure of possibility that an extreme event causes another extreme event in a time series. Pan (2002) has discussed tail dependence of ARCH and heavy-tailed Bilinear models. In this section, based on the main results in section 3 and section 4, we can get the following upper bound of the tail conditional probability of model (1.1).

**Theorem 5.1.** *Assume the conditions for TAR(1) in Theorem 3.1 and for AR(p) in Theorem 4.1 are satisfied. Then, the following upper bound for the auto-tail-dependence of $\{x_t\}$ defined by model (1.1) is true: for any two time points $t_1 < t_2$,*

$$\limsup_{x \to \infty} P\{|x_{t_2}| > x \mid |x_{t_1}| > x\} \leq 1 - \frac{K_1}{K_2} \tag{23}$$

*holds for TAR(1) and TAR(p), where $K_1$ and $K_2$ are defined as in Theorem 3.1 for case 1 and case 2, and in Theorem 4.1, respectively.*

*Proof.* We give the proof for TAR(1) first. Note that, for $t_1 < t_2$,

$$P\{|x_{t_2}| > x, |x_{t_1}| > x\} = P\{|x_{t_2}| > x\} - P\{|x_{t_2}| > x, |x_{t_1}| \le x\}.$$

Since

$$|x_{t_2}| \ge |\varepsilon(x_{t_2-d})| - |\varphi(x_{t_2-1}, x_{t_2-2}, \ldots, x_{t_2-d})| \ge |\varepsilon(x_{t_2-d})| - c_1|x_{t_2-1}| - c_0,$$

we have, for case 1,

$$\begin{aligned}
&P\{|x_{t_2}| > x, |x_{t_1}| \le x\} \\
&\ge P\{|\varepsilon(x_{t_2-d}| - c_1|x_{t_2-1}| - c_0 > x, |x_{t_1}| \le x\} \\
&\ge P\{|\varepsilon(x_{t_2-d}| - c_1|x_{t_2-1}| - c_0 > x, |x_{t_2-1}| \le x, |x_{t_1}| \le x\} \\
&\ge P\{|\varepsilon(x_{t_2-d})| > (1+c_1)x + c_0, |x_{t_2-1}| \le x, |x_{t_1}| \le x\} \\
&= \sum_{j=1}^{s} P\{|\varepsilon_{jt_2}| > (1+c_1)x + c_0, r_{j-1} < x_{t_2-d} \le r_j, |x_{t_2-1}| \le x, |x_{t_1}| \le x\} \\
&= \sum_{j=1}^{s} P\{|\varepsilon_{jt_2}| > (1+c_1)x + c_0\} P\{r_{j-1} < x_{t_2-d} \le r_j, |x_{t_2-1}| \le x, |x_{t_1}| \le x\} \\
&\ge \min_{1\le j\le s} P\{r_{j-1} < x_{t_2-d} \le r_j, |x_{t_2-1}| \le x, |x_{t_1}| \le x\} \sum_{j=1}^{s} P\{|\varepsilon_{jt_2}| > (1+c_1)x + c_0\}.
\end{aligned}$$

Furthermore, by the stationarity of $\{x_t\}$, it follows that

$$\begin{aligned}
&\frac{P\{|x_{t_2}| > x, |x_{t_1}| > x\}}{P\{|x_{t_1}| > x\}} \\
\le\ & 1 - \Big[\frac{\sum_{j=1}^{s} P\{|\varepsilon_{jt_2}| > (1+c_1)x + c_0\}}{P\{|x_{t_1}| > x\}} \\
&\cdot \min_{1\le j\le s} P\{r_{j-1} < x_{t_2-d} \le r_j, |x_{t_2-1}| \le x, |x_{t_1}| \le x\}\Big].
\end{aligned}$$

But, by Lemma 3.1 and Theorem 3.1,

$$\begin{aligned}
&\liminf_{x\to\infty} \frac{\sum_{j=1}^{s} P\{|\varepsilon_{jt_2}| > (1+c_1)x + c_0\}}{P\{|x_{t_1}| > x\}} \\
&\quad \ge \lim_{x\to\infty} \frac{\sum_{j=1}^{s} P\{|\varepsilon_{jt_2}| > (1+c_1)x + c_0\}}{N[(1+c_1)x + c_0]^{-\beta} L((1+c_1)x + c_0)} \\
&\qquad \cdot \lim_{x\to\infty} \frac{N[(1+c_1)x + c_0]^{-\beta} L((1+c_1)x + c_0)}{Nx^{-\beta} L(x)} \cdot \liminf_{x\to\infty} \frac{Nx^{-\beta} L(x)}{P\{|x_{t_1}| > x\}}
\end{aligned}$$

$$\geq (1+c_1)^{-\beta} N[\limsup_{x\to\infty} \frac{P\{|x_{t_1}|>x\}}{x^{-\beta}L(x)}]^{-1}$$
$$\geq (1+c_1)^{-\beta} N K_2^{-1} = \frac{K_1}{\tau K_2}.$$

It follows that

$$\begin{aligned}
&\limsup_{x\to\infty} P\{|x_{t_2}|>x|\ |x_{t_1}|>x\} \\
&\quad = \limsup_{x\to\infty} \frac{P\{|x_{t_2}|>x, |x_{t_1}|>x\}}{P\{|x_{t_1}|>x\}} \\
&\quad \leq 1 - \liminf_{x\to\infty} \frac{\sum_{j=1}^{s} P\{|\varepsilon_{jt_2}|>(1+c_1)x+c_0\}}{P\{|x_{t_1}|>x\}} \\
&\qquad \cdot \lim_{x\to\infty} \min_{1\leq j\leq s} P\{r_{j-1}<x_{t_2-d}\leq r_j, |x_{t_2-1}|<x, |x_{t_1}|\leq x\} \\
&\quad \leq 1-\frac{K_1}{K_2}.
\end{aligned}$$

This completes the proof for case 1.

Now turn to the proof for case 2. For case 2,

$$\begin{aligned}
&P\{|x_{t_2}|>x, |x_{t_1}|\leq x\} \\
&\geq P\{|\varepsilon_{t_2}|>(1+c_1)x+c_0, |x_{t_2-1}|\leq x, |x_{t_1}|\leq x\} \\
&= \sum_{j=1}^{s} P\{|\varepsilon_{t_2}|>(1+c_1)x+c_0\}\cdot P\{|x_{t_2-1}|\leq x, |x_{t_1}|\leq x\}
\end{aligned}$$

Then

$$\begin{aligned}
&\frac{P\{|x_{t_2}|>x, |x_{t_1}|>x\}}{P\{|x_{t_1}|>x\}} \\
\leq\quad & 1-\frac{P\{|\varepsilon_{t_2}|>(1+c_1)x+c_0\}}{P\{|x_{t_1}|>x\}}\cdot P\{|x_{t_2-1}|\leq x, |x_{t_1}|\leq x\}\big].
\end{aligned}$$

But

$$\begin{aligned}
&\liminf_{x\to\infty} \frac{P\{|\varepsilon_{t_2}|>(1+c_1)x+c_0\}}{P\{|x_{t_1}|>x\}} \\
&\quad \geq \lim_{x\to\infty} \frac{P\{|\varepsilon_{t_2}|>(1+c_1)x+c_0\}}{[(1+c_1)x+c_0]^{-\beta}L((1+c_1)x+c_0)} \\
&\qquad \cdot \lim_{x\to\infty} \frac{[(1+c_1)x+c_0]^{-\beta}L((1+c_1)x+c_0)}{x^{-\beta}L(x)} \cdot \liminf_{x\to\infty} \frac{x^{-\beta}L(x)}{P\{|x_{t_1}|>x\}}
\end{aligned}$$

$$\geq (1+c_1)^{-\beta} K_2^{-1} = \frac{K_1}{K_2}.$$

Therefore

$$\limsup_{x\to\infty} P\{|x_{t_2}| > x|\ |x_{t_1}| > x\}$$
$$\leq 1 - \liminf_{x\to\infty} \frac{P\{|\varepsilon_{t_2}| > (1+c_1)x + c_0\}}{P\{|x_{t_1}| > x\}} \cdot \lim_{x\to\infty} P\{|x_{t_2-1}| \leq x, |x_{t_1}| \leq x\} \leq 1 - \frac{K_1}{K_2}.$$

Finally, for TAR(p) with general delay $d$, we can have, for case 1,

$$\begin{aligned} &P\{|x_{t_2}| > x, |x_{t_1}| \leq x\} \\ &\geq \min_{1\leq j\leq s} P\{r_{j-1} < x_{t_2-d} \leq r_j, |x_{t_2-1}| \leq x, \cdots, |x_{t_1+1}| \leq x, |x_{t_1}| \leq x\} \\ &\quad \cdot \sum_{j=1}^{s} P\{|\varepsilon_{jt_2}| > (1+c_1+\cdot+c_p)x + c_0\}; \end{aligned}$$

for case 2,

$$\begin{aligned} P\{|x_{t_2}| > x, |x_{t_1}| \leq x\} \quad \geq \quad & P\{|x_{t_2-1}| \leq x, \cdots, |x_{t_1+1}| \leq x, |x_{t_1}| \leq x\} \\ & \cdot P\{|\varepsilon_{t_2}| > (1+c_1+\cdot+c_p)x + c_0\}, \end{aligned}$$

where $c_i = \max_{1\leq j\leq s} |\varphi_{ji}|, i = 1, ..., s$. Then, by the same way as above for TAR(1), we obtain (5.1) for TAR(p). □

# 6 Concluding remarks

Threshold AR models have been widely used in applications. However, their tail behaviours have rarely been studied. This paper tries to fill in this gap. In this paper, the tail probability and tail dependence of this type of model are studied. There are important questions remaining open: How to estimate beta in Theorem 3.1 and Theorem 4.1? How sharp is the upper bound $1 - K_1/K_2$ in Theorem 5.1? These questions are worthy of further research.

# Acknowledgments

We thank the Editor and the Referee(s) for their insightful comments and suggestions that make us improve our paper significantly. The first author wishes to thank Professor Hongzhi An and Professor Howell Tong for stimulating discussions on TAR models. This work was partially supported by the National Natural Science Foundation of China (Grant No.12171161).

# References

Adler, R. J., R. Feldaman, and M. Taqqu (1997): *A User's Guide to Heavy Tails: Statistical Techniques for Analyzing Heavy Tailed Distribution and Processes*, (eds.) Pages 91-96, Brirkhauser, Boston.

An, H. Z. and S. G. Chen (1997): "A note on the ergodicity of non-linear autoregressive model," *Statist. Probab. Lett.*, 34, 365–372.

An, H. Z. and F. Huang (1996): "The geometrical ergodicity of nonlinear autoreggressive models," *Statist. Sinica*, 6, 943–956.

Asimita, A., R. Gerrarda, Y. Houb, and L. Peng (2016): "Tail dependence measure for examining financial extreme co-movements," *Journal of Econometrics*, 194, 330–348.

Borkovec, M. and C. Kluppelberg (2001): "The tail of the stationary distribution of an autoregressive process with arch(1) errors," *Ann. Appl. Probab.*, 11, 1220–1241.

Brockwell, P., J. Liu, and R. L. Tweedie (1992): "On the existence of stationary threshold autoregressive moving-average processes," *J. Time Ser. Anal.*, 13, 95–107.

Chan, K. S., J. D. Petruccelli, H. Tong, and S. W. Woolford (1985): "A multiple-threshold ar(1) model," *J. Appl. Probab.*, 22, 267–279.

Chan, K. S. and H. Tong (1985): "On the use of the deterministic lyapunov function for the ergodicity of stochastic difference equations," *Adv. Appl. Probab.*, 17, 666–678.

Chen, R. and R. S. Tsay (1991): "On the ergodicity of tar(1) processes," *Ann. Appl. Probab.*, 1, 613–635.

Cline, D. (1983): *Estimation and linear prediction for regression, autoregression and ARMA with infinite variance data*, Ph.D. dissertation, Dept. Statistics, Colorado State Univ.

Cline, D. and H. Pu (1999): "Stability of nonlinear ar(1) time series with delay," *Stochastic Processes and Their Applications*, 82, 307–333.

Davis, R. and S. Resnick (1996): "Limit theory for bilinear processes with heavy-tailed noises," *Ann. Appl. Prob.*, 6, 1191–1210.

DeHaan, L., S. I. Resnick, H. Rootzen, and G. G. De Vries (1989): "Extremal behavior of solutions to a stochastic difference equation with applications to arch processes," *Stochastic Processes and Their Applications*, 32, 213–224.

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

Fama, E. (2013): "The behavior of stock market prices," *J. Business*, 38, 34–105.

Hansen, B. (2011): "Threshold autoregression in economics," *Statistics and Its Interfaces*, 4, 123–127.

Jin, Y. and H. Z. An (2005): "Nonlinear autoregressive models with heavy-tailed innovation," *Science in China (A)*, 48, 333–340.
Kallenberg, O. (1983): *Random Measures (3rd ed.)*, Berlin: Akademie-Verlag.
Ling, S. (1999): "On probability properties of a double threshold arma conditional heteroskedasticity model," *Journal of Applied Probability*, 36, 688–705.
Liu, J., W. Li, and C. Li (1997): "On a threshold autoregression with conditionnal heteroscedastic variances," *J. Statist. Plan. Inference*, 62, 279–300.
Liu, J. and E. Susko (1992): "On strict stationarity and ergodicity of a nonlinear arma model," *J. Appl. Probab.*, 29, 363–373.
Mandelbrot, B. (1963a): "New methods in statistical economics," *Journal of Political Economy*, 71, 411–440.
Mandelbrot, B. (1963b): "The variation of certain speculative prices," *Journal of Busines*, 36, 394–419.
Mikosch, T. and C. Starica (2000): "Limit theory for the sample autocorrelations and extremes of a garch(1,1) process," *Ann. Statist.*, 28, 1427–1451.
Mittnik, S., M. Paolella, and S. Rachev (2002): "Stationarity of stable power-garch processes," *Journal of Econometrics*, 106, 97–107.
Nummelin, E. (1984): *General Irreducible Markov chain and Non-negative Operators. Cambridge Tracts in Math. 83*, Cambridge Univ. Press.
Pan, J. Z. (2002): "Tail dependence of random variables from arch and heavy-tailed bilinear models," *Science in China (A)*, 45, 749–760.
Pan, J. Z. and G. X. Wu (2005): "On tail behavior of nonlinear autoregressive functional conditional heteroscedastic model with heavy-tailed innovations," *Science in China (A)*, 48, 1169–1181.
Pan, J. Z., B. W. T. Yu, and W. K. Pang (2004): "How does innovation's tail risk determine marginal tail risk of a stationary financial time series?" *Science in China (A)*, 47, 321–338.
Petruccelli, J. and S. W. Woolford (1984): "A threshold ar(1) model," *J. Appl. Probab.*, 21, 270–286.
Poon, S. H., M. Rockingger, and J. Tawn (2003): "Modelling extreme-value dependence in international stock markets," *Statistica Sinica*, 13, 929–953.
Resnick, S. (1987): *Extreme Values, Regular Variation, and Point Processes*, New York: Springer.
Resnick, S. (1997): "Heavy tail modeling and teletraffic data. with discussion and a rejoinder by the author," *Ann. Statist.*, 25, 1805–1869.
Tong, H. (1990): *Non-linear Time Series. A dynamical system approach*, Oxford University Press, New York.
Tsay, R. and R. Chen (2019): *Non-linear Time Series*, John Wiley & Sons, Hoboken, NJ.
Tweedie, R. L. (1975): "Sufficient conditions for ergodicity and recurrence of

markov chain on a general state space," *Stochastic Processes and Their Applications*, 3, 385–403.

Xie, X. (2017): *Analysis of Heavy-Tailed Time Series*, PhD Thesis, Department of Mathematical Sciences, University of Copenhagen.