Bernoulli **17**(3), 2011, 969–986
DOI: 10.3150/10-BEJ306

On non-stationary threshold autoregressive models

WEIDONG LIU*, SHIQING LING** and QI-MAN SHAO†

Department of Mathematics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China. E-mail: *liuweidong99@gmail.com; **maling@ust.hk; †maqmshao@ust.hk

In this paper we study the limiting distributions of the least-squares estimators for the non-stationary first-order threshold autoregressive (TAR(1)) model. It is proved that the limiting behaviors of the TAR(1) process are very different from those of the classical unit root model and the explosive AR(1).

Keywords: explosive TAR(1) model; least-squares estimator; unit root TAR(1) model

1. Introduction

Since [13], threshold autoregressive (TAR) models have been extensively investigated in the literature. The standard TAR(1) model can be written as follows:

$$Y_t = \begin{cases} \gamma + \alpha Y_{t-1} + \varepsilon_t, & \text{if } Y_{t-1} > r, \\ \delta + \beta Y_{t-1} + \varepsilon_t, & \text{if } Y_{t-1} \le r, \end{cases} \tag{1.1}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables with zero mean and a finite variance $\sigma^2 > 0$. Petrucceli and Woolford [10] and Chan et al. [4] showed that, if ε_n has a strictly positive density, then the necessary and sufficient condition for the strictly stationary and geometrically ergodic solution to model (1.1) when $\gamma = \delta = 0$ is

$$\alpha < 1, \qquad \beta < 1 \quad \text{and} \quad \alpha\beta < 1; \tag{1.2}$$

see also [5,12]. The properties of the least-squares estimator (LSE) of model (1.1) were established when $\{Y_t\}$ is stationary by Chan [3] and later by Chan and Tsay [5] for the continuous case (i.e., $\gamma + r\alpha = \delta + r\beta$). When (α, β) does not lie in the stationary region (1.2), the estimation theory of the LSE of model (1.1) is challenging.

Pham, Chan and Tong [11] were the first to consider the non-stationary case of model (1.1). They focus on the following case:

$$\gamma = \delta \quad \text{and} \quad r = 0 \tag{1.3}$$

and assume that δ is a known parameter. For the LSE of (α, β):

$$\hat{\alpha}_n = \frac{\sum_{t=1}^{n-1} I(Y_t > r) Y_t (Y_{t+1} - \gamma)}{\sum_{t=1}^{n-1} I(Y_t > r) Y_t^2}, \tag{1.4}$$

1350-7265

$$\hat{\beta}_n = \frac{\sum_{t=1}^{n-1} I(Y_t \le r) Y_t (Y_{t+1} - \gamma)}{\sum_{t=1}^{n-1} I(Y_t \le r) Y_t^2},$$

they show that

$$(\hat{\alpha}_n, \hat{\beta}_n) \to (\alpha, \beta) \qquad \text{a.s.}$$

if and only if one of the following conditions holds:

$$\begin{aligned} &\alpha \le 1, \qquad \beta \le 1 \quad \text{and} \quad \gamma = 0, \\ &\alpha < 1, \qquad \beta \le 1 \quad \text{and} \quad \gamma > 0, \\ &\alpha \le 1, \qquad \beta < 1 \quad \text{and} \quad \gamma < 0. \end{aligned} \tag{1.5}$$

They also showed that, when $\alpha\beta = 1$, the estimator of α is strongly consistent. However, the rate of convergence and the limiting distribution of LSE are two open questions when (α, β) lies in the non-ergodic region.

Following [11], in this paper we study the limiting distribution of $(\hat{\alpha}_n, \hat{\beta}_n)$ for the following cases:

$$\begin{aligned} &\text{Case I:} \quad \gamma = \delta = 0, \qquad \alpha = 1 \quad \text{and} \quad \beta < 1, \\ &\text{Case II:} \quad \gamma = \delta = 0, \qquad \alpha > 1 \quad \text{and} \quad \beta \le 1. \end{aligned}$$

For each case, we partially derive the limiting distribution of $(\hat{\alpha}_n, \hat{\beta}_n)$ under some suitable conditions. Case I is related to the unit root problem, which is particularly interesting in economics and finance. One usually tests whether or not a market is efficient via testing a unit root in AR model. Unit root tests have been extensively studied in the literature; see, for example, [6–8]. When Y_t denotes a market index, case I can describe the phenomena that the market moves from efficiency to inefficiency when the index crosses the threshold r and $|\beta| < 1$. Our result may provide a way to test this phenomena. The results for case II can help us to understand the limiting behaviors of the LSE in this complicated and dynamic system. Our proof is based on the limiting behavior of Y_t as $t \to \infty$. The method of the proof is non-standard and may provide some insights for future research in this area.

The paper is organized as follows. The main results are stated in Section 2. The proofs of the main results are given in Sections 3 and 4. This paper also includes consistency of the LSE when $\alpha\beta = 1$ in Section 5, which is of independent interest. Throughout, we let C and $C_{(\cdot)}$ denote positive constants that may be different in every place, $\mathcal{F}_t = \sigma\{Y_0, \varepsilon_1, \dots, \varepsilon_t\}$, and we assume the initial value Y_0 in model (1.1) is a random variable independent with n and $\{\varepsilon_t; t \ge 1\}$.

2. Main results

We consider two different cases.

2.1. Case I

Assume $r \leq 0$ or $r > 0$ with $\alpha = 1$ and $\beta = -1$. The results are stated as follows.

Theorem 2.1. *Assume $\gamma = \delta = 0$ and $\mathsf{E}Y_0^2 < \infty$. If either*

(i) $\alpha = 1$, $\beta < 1$ *and* $r \leq 0$; *or*
(ii) $\alpha = 1$, $\beta = -1$ *and* $r \in R$; *is satisfied, then we have*

$$n(\hat{\alpha}_n - 1) \Rightarrow \frac{B^2(1) - 1}{2\int_0^1 B^2(t)\,\mathrm{d}t}, \tag{2.1}$$

where $B(t)$ is a standard Brownian motion.

Remark 2.1. Unlike the stationary case in [3], the limiting distribution of $\hat{\alpha}_n$ is independent of r and β. (2.1) could be used to test whether (α, β) lies on the boundary $\{\alpha = 1, \beta < 1\}$ if we know r is zero or negative. We note that this test is the same as the Dickey–Fuller test. The limiting distribution of $\hat{\beta}_n$ is still unclear. But when $\alpha < 1$, $\beta = 1$ and $r \geq 0$, from the proof of Theorem 2.1 and Remark 3.1, we have

$$n(\hat{\beta}_n - 1) \Rightarrow \frac{B^2(1) - 1}{2\int_0^1 B^2(t)\,\mathrm{d}t}. \tag{2.2}$$

We should mention that Caner and Hansen [2] developed an asymptotic theory for a TAR model with a unit root, but their model is not the same as model (1.1) since their threshold variable is $Y_{t-1} - Y_{t-2}$.

Remark 2.2. When $\alpha = 1$, $\beta < 1$ and $\gamma = \delta < 0$, Chan *et al.* [4] show that $\{Y_t\}$ is ergodic, and hence is strictly stationary by assuming that Y_0 has its distribution $\pi(\cdot)$ that is the invariant probability distribution of $\{Y_t\}$. For the case $\alpha = 1$, $\beta < 1$, $r \leq 0$ and $\gamma = \delta > 0$, we have

$$Y_n \geq \gamma + \varepsilon_n + Y_{n-1} \geq n\gamma + \sum_{k=1}^{n} \varepsilon_k + Y_0.$$

Hence $Y_n \to \infty$ a.s. It follows that $\max_{1 \leq k \leq n} |Y_k - k\gamma - \sum_{i=1}^{k} \varepsilon_i| = \mathrm{O}(1)$ a.s. By some standard arguments using the martingale central limit theorem (CLT), it is not hard to see $n^{3/2}(\hat{\alpha}_n - 1) \Rightarrow N(0, 3\sigma^2/\gamma^2)$. In this case, $\hat{\beta}_n$ is not a strongly consistent estimator.

2.2. Case II

By (1.5), $(\hat{\alpha}_n, \hat{\beta}_n)$ is not a consistent estimator of (α, β) in this case. However, the following theorem shows that $\hat{\alpha}_n$ is a consistent estimator of α.

Theorem 2.2. *Assume that $\gamma = \delta = 0$ and one of the following conditions holds:*

(H1) $\alpha > 1$, $\beta \le 1$, $r = 0$ *and* $\mathsf{E}Y_0^2 < \infty$;
(H2) $\alpha > 1$, $\beta \le 1$, $r \ne 0$, $\mathsf{E}Y_0^2 < \infty$ *and* $\mathsf{P}(\varepsilon_1 \le x) < 1$ *for any* $x \in R$.

Then we have

$$(\alpha^2 - 1)^{-1}\alpha^n(\hat{\alpha}_n - \alpha) \Rightarrow \eta^*/\xi^*,$$

where η^* *and* ξ^* *are independent random variables*, $\eta^* \stackrel{d}{=} \sum_{t=1}^{\infty}\alpha^{-t}\varepsilon_t$, $\xi^* \stackrel{d}{=} \xi$ *and*

$$\xi = \sum_{k=1}^{\infty}\alpha^{-k+1}\left(\frac{\beta}{\alpha}\right)^{m_k}\varepsilon_k + \left(\frac{\beta}{\alpha}\right)^{m_0}Y_0 > 0 \qquad \text{a.s. for } \beta \le 1 \text{ and } \beta \ne 0; \tag{2.3}$$

$$\xi = \sum_{k=1}^{\infty}\frac{\prod_{t=k}^{\infty}I\{Y_t > r\}}{\alpha^k}\varepsilon_k + \prod_{t=0}^{\infty}I\{Y_t > r\}Y_0 > 0 \qquad \text{a.s. for } \beta = 0, \tag{2.4}$$

where $m_k = \sum_{t=k}^{\infty}I\{Y_t \le r\}$ *is almost surely finite.*

Remark 2.3. In the explosive AR(1) model, $Y_t = \alpha Y_{t-1} + \varepsilon_t$, it is well known that the LSE of α asymptotically follows a Cauchy distribution if ε_t is normal. By Theorem 2.2, this conclusion does not hold any more for model (1.1).

3. Proof of Theorem 2.1

Before proving Theorem 2.1, we first establish the limiting distribution for $\{Y_t\}$ when $t \to \infty$ as follows.

Theorem 3.1. *If either* (i) *or* (ii) *in Theorem* 2.1 *holds, then*

$$\frac{Y_{[nt]}}{\sqrt{n}} \Rightarrow \sigma|B(t)| \qquad \text{on } D[0, 1], \tag{3.1}$$

as $n \to \infty$, *where* $D[0, 1]$ *is the Skorokhod space.*

Remark 3.1. It is interesting to see that the limiting distribution in (3.1) does not depend on β and r. This means that the effect of β and r on Y_t is ignorable when t is long enough. The pattern of Y_t is quite different from the unit root process in the AR(1) model in which $X_{[nt]}/\sqrt{n} \Rightarrow \sigma B(t)$ on $D[0, 1]$, where $X_t = X_{t-1} + \varepsilon_t$. If $\beta = 1$, $\alpha < 1$ and $r \ge 0$, then replacing Y_t by $-Y_t$, we can get $Y_{[nt]}/\sqrt{n} \Rightarrow -\sigma|B(t)|$ on $D[0, 1]$.

Proof of Theorem 3.1 under (ii). We first consider the case when $\alpha = 1$ and $\beta = -1$. Denote Y_n by $Y_n^\star$ in this case. If $r \ge 0$, we have $Y_n^\star = \varepsilon_n + |Y_{n-1}^\star| - 2Y_{n-1}^\star I\{0 \le Y_{n-1}^\star \le r\}$, and if $r < 0$, we have $Y_n^\star = \varepsilon_n + |Y_{n-1}^\star| + 2Y_{n-1}^\star I\{r < Y_{n-1}^\star \le 0\}$. Hence,

$$\max_{1\le k\le n}\bigl|Y_k^\star - |Y_{k-1}^\star|\bigr| \le \max_{1\le k\le n}|\varepsilon_k| + 2|r| = o_{\mathsf{P}}\bigl(\sqrt{n}\bigr).$$

So it is enough to show that $|Y^\star_{[nt]}|/\sqrt{n} \Rightarrow \sigma|B(t)|$ on $D[0,1]$. Note that

$$Y^\star_n = \sum_{k=1}^{n}\prod_{j=k}^{n-1} I_j\varepsilon_k + \prod_{j=0}^{n-1} I_j Y^\star_0,$$

where $I_k = I\{Y^\star_k > r\} - I\{Y^\star_k \le r\}$. It follows that

$$A_n^{-1}Y^\star_n = \sum_{k=1}^{n} A_k^{-1}\varepsilon_k + Y^\star_0,$$

where $A_n = \prod_{k=0}^{n-1} I_k$. Since $\mathsf{E}[A_k^{-2}\varepsilon_k^2|\mathcal{F}_{k-1}] = 1$, we have by the martingale CLT (cf. [1]) that

$$\frac{1}{\sqrt{n}}A^{-1}_{[nt]}Y^\star_{[nt]} \Rightarrow \sigma B(t). \tag{3.2}$$

Now, (3.1) follows from $|A_k| = 1$ and the continuous mapping theorem. □

Proof of Theorem 3.1 under (i). Recall the definition of $\{Y^\star_n\}$ with the initial value $Y^\star_0 = Y_0$. For any $p > 0$, observe that

$$\begin{aligned}|Y_n - Y^\star_n|^p &= |Y_{n-1} - Y^\star_{n-1}|^p I\{Y_{n-1} > r, Y^\star_{n-1} > r\}\\ &\quad + |Y_{n-1} + Y^\star_{n-1}|^p I\{Y_{n-1} > r, Y^\star_{n-1} \le r\}\\ &\quad + |\beta Y_{n-1} - Y^\star_{n-1}|^p I\{Y_{n-1} \le r, Y^\star_{n-1} > r\}\\ &\quad + |\beta Y_{n-1} + Y^\star_{n-1}|^p I\{Y_{n-1} \le r, Y^\star_{n-1} \le r\}.\end{aligned} \tag{3.3}$$

Since $r \le 0$, it follows that

$$\begin{aligned}&|Y_{n-1} + Y^\star_{n-1}|^p I\{Y_{n-1} > r, Y^\star_{n-1} \le r\}\\ &\quad \le |Y_{n-1} - Y^\star_{n-1}|^p I\{Y_{n-1} > 0, Y^\star_{n-1} \le r\}\\ &\qquad + C_p|Y^\star_{n-1}|^p I\{Y^\star_{n-1} \le r\} + C_{p,r} I\{Y_{n-1} \le 0\}.\end{aligned} \tag{3.4}$$

Furthermore, since $\beta \le 1$ and $r \le 0$, we have

$$\begin{aligned}&|\beta Y_{n-1} - Y^\star_{n-1}|^p I\{Y_{n-1} \le r, Y^\star_{n-1} > r, Y^\star_{n-1} \ge \beta Y_{n-1}\}\\ &\quad \le |Y_{n-1} - Y^\star_{n-1}|^p I\{Y_{n-1} \le r, Y^\star_{n-1} > r, Y^\star_{n-1} \ge \beta Y_{n-1}\},\\ &|\beta Y_{n-1} - Y^\star_{n-1}|^p I\{Y_{n-1} \le r, Y^\star_{n-1} > r, Y^\star_{n-1} < \beta Y_{n-1,\beta}\}\\ &\quad \le 2^p|\beta Y_{n-1}|^p I\{Y_{n-1} \le r\} + C_{p,\beta,r} I\{Y_{n-1} \le r\}.\end{aligned} \tag{3.5}$$

It follows from (3.3)–(3.5) that

$$|Y_n - Y^\star_n|^p \le |Y_{n-1} - Y^\star_{n-1}|^p + q_n \le \sum_{k=1}^{n} q_k, \tag{3.6}$$

where

$$q_n = C|Y_{n-1}|^p I\{Y_{n-1} \le r\} + C|Y^\star_{n-1}|^p I\{Y^\star_{n-1} \le r\} + CI\{Y_{n-1} \le 0\}. \tag{3.7}$$

We first have $\mathsf{P}(Y_n \le r) \to 0$ by Lemma 3.2 below and $\mathsf{P}(Y^\star_{n-1} \le r) \to 0$ by (3.1) under $\alpha = 1$ and $\beta = -1$ as $n \to \infty$. Furthermore, applying Lemma 3.1 below with $p = 2$ for Y_t and $Y^\star_t$, we have

$$\begin{aligned}\mathsf{E}q_n &\le C \sup_k \mathsf{E}(\varepsilon_k^2 + Y_0^2) I\{Y_{n-1} \le r\} + C \sup_k \mathsf{E}(\varepsilon_k^2 + Y_0^{\star 2}) I\{Y^\star_{n-1} \le r\} \\ &\quad + C\mathsf{P}(Y_{n-1} \le r) + C\mathsf{P}(Y^\star_{n-1} \le r) \\ &\to 0.\end{aligned}$$

Thus, by (3.3) with $p = 2$ and the previous inequality, we have

$$\mathsf{E} \max_{1 \le k \le n} |Y_k - Y^\star_k|^2 / n \le \frac{1}{n} \sum_{k=1}^{n} \mathsf{E}q_k \to 0$$

as $n \to \infty$. By (i) of Theorem 3.1 and the previous inequality, (ii) of Theorem 3.1 holds. This completes the proof. □

We are now ready to prove Theorem 2.1.

Proof of Theorem 2.1. Without loss of generality, we assume $\sigma = 1$. Note that

$$n(\hat{\alpha}_n - 1) = \frac{n \sum_{t=1}^{n-1} I\{Y_t > r\} Y_t \varepsilon_{t+1}}{\sum_{t=1}^{n-1} I\{Y_t > r\} Y_t^2}$$

and

$$\begin{aligned}\frac{1}{n} \sum_{t=1}^{n-1} I\{Y_t > r\} Y_t \varepsilon_{t+1} &= \frac{1}{2n} \sum_{t=1}^{n-1} I\{Y_t > r\}[(Y_{t+1}^2 - Y_t^2) - \varepsilon_{t+1}^2] \\ &= \frac{1}{2n} \sum_{t=1}^{n-1} (Y_{t+1}^2 - Y_t^2) - \frac{1}{2n} \sum_{t=1}^{n-1} I\{Y_t \le r\}(Y_{t+1}^2 - Y_t^2) \\ &\quad - \frac{1}{2n} \sum_{t=1}^{n-1} I\{Y_t > r\} \varepsilon_{t+1}^2.\end{aligned}$$

Since $\mathsf{P}(y_n \le r) \to \mathsf{P}(|B(1)| \le 0) = 0$ as $n \to \infty$, we have

$$\frac{1}{n} \sum_{t=1}^{n-1} \mathsf{E}I\{Y_t \le r\} \varepsilon_{t+1}^2 = \frac{1}{n} \sum_{t=1}^{n-1} \mathsf{P}(Y_t \le r) \to 0,$$

as $n \to \infty$. Thus,

$$\frac{1}{n}\sum_{t=1}^{n-1} I\{Y_t > r\}\varepsilon_{t+1}^2 \to 1 \tag{3.8}$$

in probability. Furthermore, we have $n^{-1}\sum_{t=1}^{n-1} \mathsf{E}Y_t^2 I\{Y_t \le r\} \to 0$ by Lemmas 3.1 and 3.2 below, and hence

$$\frac{1}{n}\sum_{t=1}^{n-1} I\{Y_t \le r\}(Y_{t+1}^2 - Y_t^2) \to 0 \tag{3.9}$$

in probability. Thus, by (3.8) and (3.9), we have

$$\frac{1}{n}\sum_{t=1}^{n-1} I\{Y_t > r\}Y_t\varepsilon_{t+1} = \frac{1}{2n}(Y_n^2 - Y_1^2) - \frac{1}{2} + o_{\mathsf{P}}(1) \Rightarrow \frac{1}{2}B^2(1) - \frac{1}{2}.$$

Note that

$$\frac{\sum_{t=1}^{n-1} I\{Y_t > r\}Y_t^2}{n^2} = \int_0^1 I\bigl\{Y_{[nt]} > r\bigr\}\frac{Y_{[nt]}^2}{n}\,\mathrm{d}t \Rightarrow \int_0^1 B^2(t)\,\mathrm{d}t.$$

Theorem 2.1 follows from the continuous mapping theorem. □

We now prove Lemmas 3.1 and 3.2, which were used in the proof of Theorem 2.1.

Lemma 3.1. *Suppose that* $\mathsf{E}|Y_0|^p < \infty$ *and* $\mathsf{E}|\varepsilon_0|^p < \infty$ *for some* $p > 0$. *Under the conditions* $\gamma = \delta = 0$, $\alpha = 1$ *and* $\beta < 1$, *for any event* A, *it holds that*

$$\mathsf{E}|Y_n|^p I\{Y_n \le r, A\} \le C\Bigl(\sup_k \mathsf{E}|\varepsilon_k|^p I\{A\} + \mathsf{E}|Y_0|^p I\{A\} + \mathsf{P}(A)\Bigr).$$

Proof. Set $X_n = Y_n - r$ for $n \ge 0$. We can see that $X_n = e_n + X_{n-1}^+ - \beta X_{n-1}^-$, where $e_n = \varepsilon_n + (\beta - 1)rI\{X_{n-1} \le 0\}$. Suppose $\beta \le 0$. The lemma follows from

$$\begin{aligned}
&\mathsf{E}|X_n|^p I\{X_n \le 0, A\} \\
&\quad = \mathsf{E}|X_n|^p I\{e_n \le -(X_{n-1}^+, -\beta X_{n-1}^-), A\} \\
&\quad \le C_p\mathsf{E}|e_n|^p I\{A\} + C_p\mathsf{E}|X_{n-1}^+ - \beta X_{n-1}^-|^p I\{|e_n| \ge X_{n-1}^+ - \beta X_{n-1}^-, A\} \\
&\quad \le 2C_p \sup_k \mathsf{E}|\varepsilon_k|^p I\{A\} + C_{p,\beta,r}\mathsf{P}(A).
\end{aligned} \tag{3.10}$$

Now we prove the lemma when $0 < \beta < 1$. Set the events $A_k = \{X_k \le 0\}$ for $1 \le k \le n$. Note that

$$\mathsf{E}|X_n|^p I\{X_n \le 0, A\} = \sum_{k=0}^{n-1} \mathsf{E}|X_n|^p I\{A_n \cdots A_{n-k}A_{n-k-1}^c, A\} + \mathsf{E}|X_n|^p I\{A_n \cdots A_0, A\}. \tag{3.11}$$

We need to estimate $\mathsf{E}|X_n|^p I\{A_n \cdots A_{n-k} A^c_{n-k-1}\}$. In fact, on $A_n \cdots A_{n-k} A^c_{n-k-1}$, we have

$$X_n = \sum_{j=0}^{k-1} \beta^j e_{n-j} + \beta^k X_{n-k}.$$

Set $\xi = \sum_{j=0}^{n} \beta^j |e_{n-j}|$. It follows that

$$\begin{aligned}
&\mathsf{E}|X_n|^p I\{A_n \cdots A_{n-k} A^c_{n-k-1}, A\} \\
&\quad \le C_{\beta,\delta} \mathsf{E}|\xi|^p I\{A_n \cdots A_{n-k} A^c_{n-k-1}, A\} \\
&\qquad + C_p \beta^{kp} \mathsf{E}|X_{n-k}|^p I\{A_n \cdots A_{n-k} A^c_{n-k-1}, A\} \\
&\quad \le C_{p,\beta} \mathsf{E}|\xi|^p I\{A_n \cdots A_{n-k} A^c_{n-k-1}, A\} \\
&\qquad + C_p \beta^{kp} \Bigl(\sup_k \mathsf{E}|\varepsilon_k|^p I\{A\} + C_{p,\beta,r} \mathsf{P}(A)\Bigr) \\
&\qquad + C_p \beta^{kp} \mathsf{E}|X_{n-k-1}|^p I\{|e_{n-k}| \ge X_{n-k-1}, A^c_{n-k-1}, A\} \\
&\quad \le C_{p,\beta} \mathsf{E}|\xi|^p I\{A_n \cdots A_{n-k} A^c_{n-k-1}, A\} \\
&\qquad + 2C_p \beta^{kp} \Bigl(\sup_k \mathsf{E}|\varepsilon_k|^p I\{A\} + C_{p,\beta,r} \mathsf{P}(A)\Bigr).
\end{aligned} \tag{3.12}$$

Clearly, on $A_n \cdots A_0$, we have $X_n = \sum_{j=0}^{n-1} \beta^j e_{n-j} + \beta^n X_0$ and hence by (3.11) and (3.12), $\mathsf{E}|X_n|^p I\{X_n \le 0, A\} \le C(\sup_k \mathsf{E}|\varepsilon_k|^p I\{A\} + \mathsf{E}|Y_0|^p I\{A\} + \mathsf{P}(A))$. The lemma is now proved. □

Lemma 3.2. *Suppose that* $\mathsf{E}Y_0^2 < \infty$, $\mathsf{E}\varepsilon_0 = 0$ *and* $\mathsf{E}\varepsilon_0^2 < \infty$. *Under the conditions* $\gamma = \delta = 0$, $\alpha = 1$, $\beta < 1$ *and* $r \le 0$, *we have* $Y_n/\sqrt{n} \Rightarrow \sigma|B(1)|$ *as* $n \to \infty$.

Proof. For $K > 0$, set

$$\widetilde{\varepsilon}_k = \varepsilon_k I\{|\varepsilon_k| \le K\} - \mathsf{E}\varepsilon_k I\{|\varepsilon_k| \le K\}, \qquad \hat{\varepsilon}_k = \varepsilon_k - \widetilde{\varepsilon}_k, \qquad k \ge 1.$$
$$\widetilde{Y}_0 = Y_0 I\{|Y_0| \le K\}, \qquad \widehat{Y}_0 = Y_0 - \widetilde{Y}_0.$$

We now construct two TAR(1) processes $\{\widetilde{Y}_t\}$ and $\{\widetilde{Y}^\star_t\}$ as follows:

$$\widetilde{Y}_n = \widetilde{\varepsilon}_n + \widetilde{Y}_{n-1} I\{\widetilde{Y}_{n-1} > r\} + \beta \widetilde{Y}_{n-1} I\{\widetilde{Y}_{n-1} \le r\}, \qquad n \ge 1; \tag{3.13}$$
$$\widetilde{Y}^\star_n = \widetilde{\varepsilon}_n + \widetilde{Y}^\star_{n-1} I\{\widetilde{Y}^\star_{n-1} > r\} - \widetilde{Y}^\star_{n-1} I\{\widetilde{Y}^\star_{n-1} \le r\}, \qquad n \ge 1. \tag{3.14}$$

By Theorem 3.1, when $\alpha = 1$ and $\beta = -1$, we can see that

$$\widetilde{Y}^\star_{[nt]}/\sqrt{n} \Rightarrow \sigma_K |B(t)| \qquad \text{on } D[0,1], \tag{3.15}$$

with $\sigma_K^2 = \mathrm{Var}(\widetilde{\varepsilon}_1)$. Let q_k', $1 \le k \le n$, be defined as q_k in (3.7) by replacing $\{Y_n\}$ and $\{Y_n^\star\}$ with $\{\widetilde{Y}_n\}$ and $\{\widetilde{Y}_n^\star\}$, respectively. Taking $p > 2$, by virtue of (3.6), we have

$$\mathsf{E} \max_{1\le k\le n} \left| \frac{\widetilde{Y}_k - \widetilde{Y}_k^\star}{\sqrt{n}} \right|^p \le \frac{\sum_{k=1}^n \mathsf{E} q_k'}{n^{p/2}}.$$

Furthermore, using Lemma 3.1 with Y_t replaced by $\{\widetilde{Y}_n\}$ and $\{\widetilde{Y}_n^\star\}$, respectively, we know that q_k' is uniformly bounded for all $k \ge 1$. Thus, we have

$$\mathsf{E} \max_{1\le k\le n} \left| \frac{\widetilde{Y}_k - \widetilde{Y}_k^\star}{\sqrt{n}} \right|^p \le C n^{-p/2+1}.$$

By (3.15) and the previous inequality, we have

$$\widetilde{Y}_{[nt]}/\sqrt{n} \Rightarrow \sigma_K |B(t)| \qquad \text{on } D[0, 1]. \tag{3.16}$$

Since $\sigma_K \to \sigma$ as $K \to \infty$, it suffices to show that for any $\delta > 0$,

$$\lim_{K\to\infty} \limsup_{n\to\infty} \mathsf{P}\big(|Y_n - \widetilde{Y}_n| \ge \delta\sqrt{n}\big) = 0. \tag{3.17}$$

By model (1.1) and model (3.13), we have

$$\begin{aligned}
\mathsf{E}(Y_n - \widetilde{Y}_n)^2 &= \mathsf{E}\hat{\varepsilon}_n^2 + \mathsf{E}(Y_{n-1} - \widetilde{Y}_{n-1})^2 I\{Y_{n-1} > r, \widetilde{Y}_{n-1} > r\} \\
&\quad + \mathsf{E}(Y_{n-1} - \beta\widetilde{Y}_{n-1})^2 I\{Y_{n-1} > r, \widetilde{Y}_{n-1} \le r\} \\
&\quad + \mathsf{E}(\beta Y_{n-1} - \widetilde{Y}_{n-1})^2 I\{Y_{n-1} \le r, \widetilde{Y}_{n-1} > r\} \\
&\quad + \mathsf{E}(\beta Y_{n-1} - \beta\widetilde{Y}_{n-1})^2 I\{Y_{n-1} \le r, \widetilde{Y}_{n-1} \le r\}.
\end{aligned}$$

It can be verified that

$$\begin{aligned}
&\mathsf{E}(Y_{n-1} - \beta\widetilde{Y}_{n-1})^2 I\{Y_{n-1} > r, \widetilde{Y}_{n-1} \le r\} \\
&\quad \le \mathsf{E}(Y_{n-1} - \widetilde{Y}_{n-1})^2 I\{Y_{n-1} > r, \widetilde{Y}_{n-1} \le r, Y_{n-1} > \beta\widetilde{Y}_{n-1}\} \\
&\qquad + C\mathsf{E}\widetilde{Y}_{n-1}^2 I\{\widetilde{Y}_{n-1} \le r\} + C\mathsf{P}(\widetilde{Y}_{n-1} \le r).
\end{aligned}$$

Let M be any positive number. Then,

$$\begin{aligned}
&\mathsf{E}(\beta Y_{n-1} - \widetilde{Y}_{n-1})^2 I\{Y_{n-1} \le r, \widetilde{Y}_{n-1} > r\} \\
&\quad \le \mathsf{E}(Y_{n-1} - \widetilde{Y}_{n-1})^2 I\{Y_{n-1} \le r, \widetilde{Y}_{n-1} > r, \widetilde{Y}_{n-1} > \beta Y_{n-1}\} \\
&\qquad + C\mathsf{E}Y_{n-1}^2 I\{Y_{n-1} < -M, \beta Y_{n-1} \ge \widetilde{Y}_{n-1} > r\} + C_{\beta,M,r}\mathsf{P}(\widetilde{Y}_{n-1} \le r + |\beta| M).
\end{aligned}$$

Combining the above inequalities, we can see that

$$\mathsf{E}(Y_n - \widetilde{Y}_n)^2 \le \mathsf{E}\hat{\varepsilon}_n^2 + \mathsf{E}(Y_{n-1} - \widetilde{Y}_{n-1})^2 + \widetilde{q}_n,$$

where

$$\widetilde{q}_n = C\mathsf{E}Y_{n-1}^2 I\{Y_{n-1} \le r, \widetilde{Y}_{n-1} \le r\} + C\mathsf{E}\widetilde{Y}_{n-1}^2 I\{\widetilde{Y}_{n-1} \le r\} + C\mathsf{E}Y_{n-1}^2 I\{Y_{n-1} < -M, \beta Y_{n-1} \ge \widetilde{Y}_{n-1} > r\} + C_{\beta,M,r}\mathsf{P}(\widetilde{Y}_{n-1} \le r + |\beta|M).$$

By induction we have

$$\mathsf{E}(Y_n - \widetilde{Y}_n)^2 \le n\mathsf{E}\hat{\varepsilon}_0^2 + \mathsf{E}\widehat{Y}_0^2 + \sum_{k=1}^n \widetilde{q}_k. \tag{3.18}$$

Since $\widetilde{Y}_n/\sqrt{n} \Rightarrow \sigma_K|B(1)|$, we have $\mathsf{P}(\widetilde{Y}_{n-1} \le r + |\beta|M) \to 0$ as $n \to \infty$. Note that

$$I\{y_{n-1} < -M, \beta Y_{n-1} \ge \widetilde{Y}_{n-1} > r\} \le \begin{cases} I\{\varepsilon_{n-1} < -M + |r|\}, & \text{if } \beta \le 0, \\ I\{\widetilde{Y}_{n-1} \le -\beta M\}, & \text{if } 0 < \beta < 1. \end{cases}$$

By Lemma 3.1,

$$\begin{aligned} &\mathsf{E}Y_{n-1}^2 I\{Y_{n-1} < -M, \beta Y_{n-1} \ge \widetilde{Y}_{n-1} > r\} \\ &\quad \le C\sup_k \mathsf{E}(\varepsilon_k^2 + Y_0^2 + 1)I\{\varepsilon_{n-1} < -M + |r|\} \\ &\qquad + C\sup_k \mathsf{E}(\varepsilon_k^2 + Y_0^2 + 1)I\{\widetilde{Y}_{n-1} \le -\beta M\} \end{aligned}$$

and

$$\mathsf{E}Y_{n-1}^2 I\{Y_{n-1} \le r, \widetilde{Y}_{n-1} \le r\} + \mathsf{E}\widetilde{Y}_{n-1}^2 I\{\widetilde{Y}_{n-1} \le r\} \le C\sup_k \mathsf{E}(\varepsilon_k^2 + Y_0^2 + 1)I\{\widetilde{Y}_{n-1} \le r\}.$$

Since $\lim_{n\to\infty} \mathsf{P}(\widetilde{Y}_n \le x) = 0$ for any $x \in R$, we have

$$\lim_{M\to\infty} \limsup_{n\to\infty} n^{-1} \sum_{k=1}^n \widetilde{q}_k = 0.$$

This, together with $\mathsf{E}\hat{\varepsilon}_1^2 \to 0$ as $K \to \infty$ and (3.18), implies (3.17). □

4. Proof of Theorem 2.2

To prove Theorem 2.2, we need to establish the limiting distribution for $\{Y_t\}$ as $t \to \infty$.

Theorem 4.1. *Let $\gamma = \delta = 0$. Suppose either* (H1) *or* (H2) *in Theorem 2.2 holds. Then we have $\mathsf{E}|Y_n| = \mathrm{O}(\alpha^n)$, $\sum_{t=0}^\infty I\{Y_t \le r\} < \infty$ a.s. and $Y_n/\alpha^n \to \xi > 0$ a.s., where ξ is defined in Theorem 2.2.*

From this theorem, we can see that $\hat{\beta}_n - \beta \to Z$ a.s. for some random variable Z. Thus, $\hat{\beta}_n$ is not a strongly consistent estimator for β. This explains why (1.5) is the necessary and sufficient condition for consistency of $(\hat{\alpha}_n, \hat{\beta}_n)$. To prove Theorem 4.1, we need the following lemma.

Lemma 4.1. *Under the conditions of Theorem* 4.1, *we have*: (i) $\mathsf{E}|Y_n| = \mathrm{O}(\alpha^n)$; (ii) $\lim_{n\to\infty} Y_n/n = \infty$ *a.s. if* $\limsup_{n\to\infty} Y_n = \infty$ *a.s.*

Proof. (i) Following the proof of Lemma 3.1, we can prove that, under the conditions (H1) or (H2) in Theorem 2.2, $\mathsf{E}|Y_n|I\{|Y_n| \le r\} = \mathrm{O}(1)$ if $\beta < 1$ and $\mathsf{E}|Y_n|I\{|Y_n| \le r\} = \mathrm{O}(n)$ if $\beta = 1$. We next show that $\mathsf{E}|Y_n| = \mathrm{O}(\alpha^n)$. Since

$$\begin{aligned}\mathsf{E}|Y_n| &= \mathsf{E}|Y_n|I\{Y_n \le r\}\\ &\quad + \sum_{k=0}^{n-1} \mathsf{E}|Y_n|I\{Y_n > r, Y_{n-1} > r, \ldots, Y_{k+1} > r, Y_k \le r\} \qquad (4.1)\\ &\quad + \mathsf{E}|Y_n|I\{Y_n > r, Y_{n-1} > r, \ldots, Y_0 > r\}\end{aligned}$$

and $\mathsf{E}|Y_n|I\{Y_n > r, Y_{n-1} > r, \ldots, Y_0 > r\} = \mathrm{O}(\alpha^n)$, we only need to estimate the second term on the right-hand side of (4.1). Set $B_k = \{Y_n > r, Y_{n-1} > r, \ldots, Y_{k+1} > r, Y_k \le r\}$. On B_k, we have $Y_n = \sum_{j=k+2}^n \alpha^{n-j}\varepsilon_j + \alpha^{n-k-1}Y_{k+1}$. Thus by noting that $\mathsf{E}|Y_n|I\{B_k\} \le \mathsf{E}\max_{0\le i\le n} |\sum_{j=i+2}^n \alpha^{n-j}\varepsilon_j|I\{B_k\} + \alpha^{n-k-1}\mathsf{E}|Y_{k+1}|I\{B_k\}$ and $\mathsf{E}|Y_n|I\{Y_n \le r\} = \mathrm{O}(n)$, we have

$$\sum_{k=1}^{n-1} \mathsf{E}|Y_n|I\{B_k\} \le \mathsf{E}\max_{0\le i\le n}\left|\sum_{j=i+2}^{n} \alpha^{n-j}\varepsilon_j\right| + \mathrm{O}(1)\sum_{k=0}^{n-1}\alpha^{n-k-1}k = \mathrm{O}(\alpha^n).$$

This together with (4.1) gives $\mathsf{E}|Y_n| = \mathrm{O}(\alpha^n)$.

(ii) For any $M > 1$, define $\mathbf{A}_n = \bigcup_{t=n}^{\infty}\{Y_t \le t^{3/2}\}$. Let $\delta > 0$ and $T > \max(r, 0)$ satisfy $\alpha > 1 + \delta + 8T^{-1/8}$. Define $\tau = \max\{k : Y_{-1} \le T, \ldots, Y_k \le T, Y_{k+1} > T, k \ge -1\}$, $Y_{-1} = 0$. We can see that $\tau < \infty$ a.s. and $\{\tau = k\} = \{Y_{-1} \le T, \ldots, Y_k \le T, Y_{k+1} > T\}$ is $\sigma(Y_0, \varepsilon_1, \ldots, \varepsilon_{k+1})$ measurable. For any $n_0 + 3 < n$, $M > 0$, $T > M$

$$\begin{aligned}\mathsf{P}(\mathbf{A}_n) &\le \mathsf{P}(\tau > n_0) + \sum_{k=-1}^{n_0} \mathsf{P}(\tau = k, \mathbf{A}_n)\\ &\le \mathsf{P}(\tau > n_0) + \sum_{k=-1}^{n_0} \mathsf{P}\left(\tau = k, \mathbf{A}_n, \bigcap_{j=k+2}^{\infty}\big\{|\varepsilon_j| \le \delta\big((j-k-2)^2 + T\big)\big\}\right)\\ &\quad + \sum_{k=-1}^{n_0} \mathsf{P}\left(\tau = k, \mathbf{A}_n, \bigcup_{j=k+2}^{\infty}\big\{|\varepsilon_j| > \delta\big((j-k-2)^2 + T\big)\big\}\right).\end{aligned}$$

Note that on the event

$$\mathbf{B} := \left\{\tau = k, \bigcap_{j=k+2}^{\infty}\big\{|\varepsilon_j| \le \delta\big((j-k-2)^2 + T\big)\big\}\right\},$$

since $\alpha > 1 + \delta + 8T^{-1/8}$ and $\frac{(t-k-1)^2-(t-k-2)^2}{(t-k-2)^2+T} < 8T^{-1/8}$ for $t \geq k+3$, we have

$$
\begin{aligned}
&Y_{k+1} > T > r,\\
&Y_{k+2} = \alpha Y_{k+1} + \varepsilon_{k+2} \geq \alpha T - \delta T > T + 1,\\
&Y_{k+3} = \alpha Y_{k+2} + \varepsilon_{k+3} \geq \alpha(T+1) - \delta(1+T) > T + 2^2,\\
&\quad\vdots\\
&\quad Y_t = \alpha Y_{t-1} + \varepsilon_t > \alpha\big((t-k-2)^2 + T\big) - \delta\big((t-k-2)^2 + T\big)\\
&\qquad > (t-k-1)^2 + T
\end{aligned}
$$

for any $t \geq k+1$. That is, for any t satisfying $t - n_0 - 1 > t^{3/4}$ and $k \leq n_0$, we have $Y_t > t^{3/2}$ on event $\mathbf{B}$. Thus, for n satisfying $n - n_0 - 1 > n^{3/4}$ and $k \leq n_0$, we have

$$
\Bigg\{\tau = k, \mathbf{A}_n, \bigcap_{j=k+2}^{\infty} \big\{|\varepsilon_j| \leq \delta\big((j-k-2)^2 + T\big)\big\}\Bigg\} = \varnothing
$$

and

$$
\begin{aligned}
\mathsf{P}(\mathbf{A}_n) &\leq \mathsf{P}(\tau > n_0) + \sum_{k=-1}^{n_0} \mathsf{P}\Bigg(\tau = k, \bigcup_{j=k+2}^{\infty} \big\{|\varepsilon_j| > \delta\big((j-k-2)^2 + T\big)\big\}\Bigg)\\
&= \mathsf{P}(\tau > n_0) + \sum_{k=-1}^{n_0} \mathsf{P}(\tau = k)\mathsf{P}\Bigg(\bigcup_{j=k+2}^{\infty} \big\{|\varepsilon_j| > \delta\big((j-k-2)^2 + T\big)\big\}\Bigg)\\
&\leq \mathsf{P}(\tau > n_0) + \sum_{k=-1}^{n_0} \mathsf{P}(\tau = k) \sum_{j=k+2}^{\infty} \mathsf{P}\big(|\varepsilon_j| > \delta\big((j-k-2)^2 + T\big)\big).
\end{aligned}
$$

Letting $n \to \infty$ and then $n_0 \to \infty$ implies that for any $M > 0$,

$$
\begin{aligned}
\mathsf{P}\Bigg(\bigcap_{n=1}^{\infty} \mathbf{A}_n\Bigg) &\leq \delta^{-2} \sum_{k=-1}^{\infty} \mathsf{P}(\tau = k) \sum_{j=k+2}^{\infty} \frac{\mathsf{E}|\varepsilon_1|^2}{((j-k-2)^2 + T)^2}\\
&\leq CT^{-1} \sum_{k=1}^{\infty} k^{-2} \to 0 \qquad \text{as } T \to \infty.
\end{aligned}
$$

The lemma is proved. □

Proof of Theorem 4.1. $\mathsf{E}|Y_n| = \mathrm{O}(\alpha^n)$ follows from Lemma 4.1(i). We next give the proofs for the other conclusions.

Proof under (H1). Define X_m by the equations $X_n = \varepsilon_n + X_{n-1}^+ - \beta^+ X_{n-1}^-$, $X_0 = Y_0$. Then we have $Y_n \geq X_n$ for any $n \geq 0$. If $\beta = 1$, then $X_n = \sum_{k=1}^n \varepsilon_k + X_0$ and $\limsup_{n\to\infty} Y_n =$

$\limsup_{n\to\infty} X_n = \infty$ a.s. If $\beta < 1$, then by Theorem 3.1, $X_n \to \infty$ in probability. So $Y_n \to \infty$ in probability, which implies $\limsup_{n\to\infty} Y_n = \infty$ a.s. By Lemma 4.1 we have $Y_n/n \to \infty$ a.s. Hence $\sum_{t=0}^{\infty} I\{Y_t \le 0\} < \infty$ a.s. Thus ξ in (2.3) or (2.4) is well defined and $Y_n/\alpha^n \to \xi$ a.s.

Now we prove $\xi > 0$ a.s. Let $e_n = \varepsilon_n - \beta Y_{n-1}^{-}$. We have $\mathsf{E}|e_n| = \mathrm{O}(n)$ for $\beta \le 1$. Define $m = \sup\{n : \alpha^n/n < M\}$. Then $m \sim \log M/\log\alpha \to \infty$ as $M \to \infty$. By the inequality $(a+b)^+ \ge a - |b|$, we have

$$\frac{Y_n}{\alpha^n} = \frac{e_n}{\alpha^n} + \frac{Y_{n-1}^+}{\alpha^{n-1}} \ge \frac{e_n}{\alpha^n} - \frac{|e_{n-1}|}{\alpha^{n-1}} + \frac{Y_{n-2}^+}{\alpha^{n-2}} \ge \cdots \ge -\sum_{k=m+1}^{n} \frac{|e_k|}{\alpha^k} + \frac{Y_m^+}{\alpha^m}. \tag{4.2}$$

From (4.2) we can get $\xi \ge -\sum_{k=m+1}^{\infty} \frac{|e_k|}{\alpha^k} + \frac{Y_m^+}{\alpha^m}$ a.s. Since $M(m+1)/\alpha^{m+1} \le 1$, we have

$$\sum_{k=m+1}^{\infty} (M\alpha^{-k})\mathsf{E}|e_k| \le C \sum_{k=m+1}^{\infty} \frac{Mk}{\alpha^k} \le C \sum_{k=m+1}^{\infty} \frac{k-m}{\alpha^{k-m-1}} \frac{Mk}{\alpha^{m+1}(k-m)} \le C \sum_{k=0}^{\infty} \frac{k+1}{\alpha^k},$$

where C does not depend on M. So we have

$$\limsup_{M\to\infty} \mathsf{P}\left(M \sum_{k=m+1}^{\infty} \frac{|e_k|}{\alpha^k} \ge \eta \right) \le C\eta^{-1} \to 0 \tag{4.3}$$

as $\eta \to \infty$. It is easy to see that $MY_m^+/\alpha^m \ge Y_m/m \to \infty$ a.s. as $M \to \infty$. Hence, by (4.2) and (4.3), $\mathsf{P}(\xi \le 0) = \mathsf{P}(M\xi \le 0) \le \mathsf{P}(Y_m/m \le \eta) + \mathsf{P}(M \sum_{k=m+1}^{\infty} \frac{|e_k|}{\alpha^k} \ge \eta) \to 0$ by letting $M \to \infty$ first and then $\eta \to \infty$. This proves $\xi > 0$ a.s.

Proof under (H2). We first assume that $\limsup_{n\to\infty} Y_n = \infty$ a.s. Then it follows from Lemma 4.1 that $Y_n/n \to \infty$ a.s. and hence $\sum_{t=1}^{\infty} I\{Y_t \le r\} < \infty$ a.s., $Y_n/\alpha^n \to \xi$ a.s. By writing $Y_n = e_n + \alpha Y_{n-1}^+$, where $e_n = \varepsilon_n + \beta Y_{n-1} I\{Y_{n-1} \le r\} - \alpha Y_{n-1} I\{0 \le Y_{n-1} \le r\}$ if $r > 0$, and $e_n = \varepsilon_n + \beta Y_{n-1} I\{Y_{n-1} \le r\} + \alpha Y_{n-1} I\{r < Y_{n-1} \le 0\}$ if $r \le 0$, we can show that $\xi > 0$ a.s. following the proof of Theorem 4.1 under (H1).

It remains to show that $\limsup_{n\to\infty} Y_n = \infty$ a.s. We claim that if, for all $y \le r$,

$$\mathsf{P}(Y_t < r \text{ for all } t \ge 0 | Y_0 = y) = 0, \tag{4.4}$$

then $\limsup_{n\to\infty} Y_n = \infty$ a.s. The proof is similar to that of [11]. Let $c > |r|$. Since for any $x > 0$, $\mathsf{P}(\varepsilon_1 \le x) < 1$, we have for all $r \le y \le c$,

$$\mathsf{P}(Y_1 \ge c | Y_0 = y) = \mathsf{P}(\alpha y + \varepsilon_1 \ge c) \ge \mathsf{P}\big(\varepsilon_1 \ge c(1+\alpha)\big) > 0,$$

which yields that for any $c > |r|$

$$\inf_{r \le y \le c} \mathsf{P}(Y_t \ge c \text{ for some } t > 0 | Y_0 = y) > 0.$$

Then by Proposition 5.1 in [9], for any initial distribution on Y_0,

$$\{Y_t \in [r, c) \text{ infinitely often}\} \subseteq \{Y_t \in [c, \infty) \text{ infinitely often}\}. \tag{4.5}$$

Using similar arguments in [11], we can see that if for all $y \in R$

$$\mathsf{P}(Y_t \geq r \text{ for some } t \,|\, Y_0 = y) = 1, \tag{4.6}$$

then

$$\mathsf{P}(Y_t \geq r \text{ infinitely often}) = 1$$

and hence by (4.5) we have $\mathsf{P}(Y_t \in [c, \infty) \text{ infinitely often}) = 1$ for any $c > 0$. This yields $\limsup_{n\to\infty} Y_n = \infty$ a.s.

Now it suffices to show that (4.6) or, equivalently, (4.4) holds. Note that (4.4) is a direct consequence of the following results:

$$\lim_{n\to\infty} \mathsf{P}\left(\max_{1\leq k\leq n}\left(\sum_{i=1}^{k} \beta^{k-i}\varepsilon_i + \beta^k y\right) \leq r\right) = 0 \qquad \text{for } y \leq r. \tag{4.7}$$

If $\beta = 1$, then (4.7) holds by the law of iterated logarithm. If $\beta \leq 0$, we have

$$\begin{aligned}
&\left\{\max_{1\leq k\leq n}\left(\sum_{i=1}^{k} \beta^{k-i}\varepsilon_i + \beta^k y\right) \leq r\right\} \\
&\quad\subseteq \left\{\sum_{i=1}^{k} \beta^{k-i}\varepsilon_i + \beta^k y = \varepsilon_k + \beta\left(\sum_{i=1}^{k-1} \beta^{k-1-i}\varepsilon_i + \beta^{k-1} y\right) \leq r, 1 \leq k \leq n\right\} \\
&\quad\subseteq \{\varepsilon_k \leq r + |\beta r|, 1 \leq k \leq n\}.
\end{aligned}$$

Therefore

$$\mathsf{P}\left(\max_{1\leq k\leq n}\left(\sum_{i=1}^{k} \beta^{k-i}\varepsilon_i + \beta^k y\right) \leq r\right) \leq \mathsf{P}\left(\max_{1\leq k\leq n} \varepsilon_k \leq r + |\beta r|\right) \to 0.$$

It remains to prove (4.7) for $0 < \beta < 1$. Set $k_j = jn^{1/2}$ for $1 \leq j \leq n^{1/2}$. Then for any $x > 0$ we have

$$\begin{aligned}
\mathsf{P}\left(\max_{1\leq k\leq n}\left(\sum_{i=1}^{k} \beta^{k-i}\varepsilon_i\right) \leq x\right) &\leq \mathsf{P}\left(\max_{1\leq j\leq n^{1/2}}\left(\sum_{i=1}^{k_j} \beta^{k_j-i}\varepsilon_i\right) \leq x\right) \\
&\leq \mathsf{P}\left(\max_{1\leq j\leq n^{1/2}}\left(\sum_{i=k_j-n^{1/4}}^{k_j} \beta^{k_j-i}\varepsilon_i\right) \leq 2x\right) \\
&\quad + \mathsf{P}\left(\max_{1\leq j\leq n^{1/2}}\left(\sum_{i=1}^{k_j-n^{1/4}-1} \beta^{k_j-i}\varepsilon_i\right) \geq x\right).
\end{aligned}$$

Since $\mathsf{E}|\varepsilon_1| < \infty$, we have

$$\mathsf{P}\left(\max_{1\le j\le n^{1/2}}\left(\sum_{i=1}^{k_j-n^{1/4}} \beta^{k_j-i}\varepsilon_i\right) \ge x\right) \le Cn^{1/2}\sum_{j=n^{1/4}}^{\infty}\beta^j \to 0.$$

By independence, we have

$$\mathsf{P}\left(\max_{1\le j\le n^{1/2}}\left(\sum_{i=k_j-n^{1/4}}^{k_j} \beta^{k_j-i}\varepsilon_i\right) \le 2x\right) = \left(\mathsf{P}\left(\sum_{j=1}^{n^{1/4}+1}\beta^{j-1}\varepsilon_j \le 2x\right)\right)^{n^{1/2}}.$$

Also

$$\mathsf{P}\left(\sum_{j=1}^{n^{1/4}+1}\beta^{j-1}\varepsilon_j \le 2x\right) \le \mathsf{P}\left(\sum_{j=1}^{\infty}\beta^{j-1}\varepsilon_j \le 3x\right) + \Delta_n,$$

where $\Delta_n \le C\sum_{j=n^{1/4}}^{\infty}\beta^j \to 0$ as $n\to\infty$. So it suffices to show that for any $x>0$, $\mathsf{P}(\sum_{j=1}^{\infty}\beta^{j-1}\varepsilon_j \le x) < 1$. In fact, if there exists some $x>0$ such that

$$1 = \mathsf{P}\left(\sum_{j=1}^{\infty}\beta^{j-1}\varepsilon_j \le x\right) = \mathsf{E}F\left(x - \sum_{j=2}^{\infty}\beta^{j-1}\varepsilon_j\right),$$

where $F(\cdot)$ is the distribution function of ε_1, then $F(x-\sum_{j=2}^{\infty}\beta^{j-1}\varepsilon_j) = 1$ a.s. That is, $\sum_{j=2}^{\infty}\beta^{j-1}\varepsilon_j = -\infty$ a.s. This is impossible since $0<\beta<1$ and $\mathsf{E}|\varepsilon_1|<\infty$. □

We are now ready to prove Theorem 2.2.

Proof of Theorem 2.2. Note that

$$\frac{\alpha^n(\hat{\alpha}_n-\alpha)}{\alpha^2-1} = \frac{\alpha^{-n}\sum_{t=1}^{n-1} I(Y_t>r)Y_t\varepsilon_{t+1}}{\alpha^{-2n}(\alpha^2-1)\sum_{t=1}^{n-1} I(Y_t>r)Y_t^2}.$$

Since $Y_n/\alpha^n \to \xi > 0$ a.s., we have $(\alpha^2-1)\alpha^{-2n}\sum_{t=1}^{n-1} Y_t^2 I\{Y_t>r\} \to \xi^2$ a.s. By the fact $\mathsf{E}|Y_n|I\{Y_n\le r\} = \mathrm{O}(n)$ for $\beta\le 1$, we have

$$\alpha^{-n}\left(\sum_{t=1}^{n-1} Y_t\varepsilon_{t+1} - \sum_{t=1}^{n-1} Y_t\varepsilon_{t+1}I\{Y_t>r\}\right) \to 0 \qquad \text{a.s.}$$

We next prove that $\alpha^{-n}(\sum_{t=1}^{n-1} Y_t\varepsilon_{t+1} - \xi\sum_{t=1}^{n-1}\alpha^t\varepsilon_{t+1}) \to 0$ in probability. For $K>0$, let

$$\widetilde{\varepsilon}_t = \varepsilon_t I\{|\varepsilon_t|\le K\}, \qquad 1\le t\le n.$$

We have $\alpha^{-n}\mathsf{E}|\sum_{t=1}^{n-1} Y_t(\varepsilon_{t+1} - \widetilde{\varepsilon}_{t+1})| \le C\mathsf{E}|\varepsilon_0| I\{|\varepsilon_0| > K\} \to 0$ as $K \to \infty$. So it suffices to prove that $\alpha^{-n}(\sum_{t=1}^{n-1} Y_t \widetilde{\varepsilon}_{t+1} - \xi \sum_{t=1}^{n-1} \alpha^t \widetilde{\varepsilon}_{t+1}) \to 0$ in probability, which follows from $Y_n/\alpha^n \to \xi$ a.s. and $|\widetilde{\varepsilon}_t| \le K$. Hence

$$\frac{\alpha^n(\hat{\alpha}_n - \alpha)}{\alpha^2 - 1} - \frac{\alpha^{-n}\sum_{t=1}^{n-1}\alpha^t\varepsilon_{t+1}}{\xi} \to 0 \qquad \text{in probability.}$$

Note that $Y_{[n/2]}/\alpha^{[n/2]} \to \xi$ a.s. and $\alpha^{-n}(\sum_{t=1}^{n-1}\alpha^t\varepsilon_{t+1} - \sum_{t=[n/2]+1}^{n-1}\alpha^t\varepsilon_{t+1}) \to 0$ in probability. We have

$$\frac{\alpha^n(\hat{\alpha}_n - \alpha)}{\alpha^2 - 1} - \frac{\alpha^{-n}\sum_{t=[n/2]+1}^{n-1}\alpha^t\varepsilon_{t+1}}{Y_{[n/2]}/\alpha^{[n/2]}} \to 0 \qquad \text{in probability.}$$

By the independence between $\sum_{t=[n/2]+1}^{n-1}\alpha^t\varepsilon_{t+1}$ and $Y_{[n/2]}$, we see that

$$\left(\alpha^{-n}\sum_{t=[n/2]+1}^{n-1}\alpha^t\varepsilon_{t+1}, Y_{[n/2]}/\alpha^{[n/2]}\right) \Rightarrow (\eta^*, \xi^*),$$

which finishes the proof. □

5. A further result when $\alpha\beta = 1$

We next consider the LSE of (α, β) under the constraints $\alpha\beta = 1$. We estimate α by minimizing $Q_n(x)$, where

$$Q_n(x) = \sum_{t=2}^{n}(Y_t - xY_{t-1}I\{Y_{t-1} < r\} - x^{-1}Y_{t-1}I\{Y_{t-1} \ge r\})^2.$$

Pham, Chan and Tong [11] showed that the estimator $\hat{\alpha}_n$, by minimizing $Q_n(x)$ under $\alpha\beta = 1$ and $\alpha < 0$, is strongly consistent. The following theorem shows that $\hat{\alpha}_n$ is still strongly consistent under $\alpha\beta = 1$ and $\alpha > 0$.

Theorem 5.1. *Let $\gamma = \delta = 0$, $\alpha\beta = 1$ and $0 < \alpha \ne 1$. Assume that $\mathsf{P}(\varepsilon_1 \le x) < 1$ and $\mathsf{P}(\varepsilon_1 \ge x) < 1$ for any $x \in R$. Then $\hat{\alpha}_n$ obtained by minimizing $Q_n(x)$ is strongly consistent.*

Proof. We only prove the theorem for $\alpha > 1$. The proof for the other case $0 < \alpha < 1$ is similar. We have

$$\begin{aligned} &Q_n(x) - Q_n(\alpha) \\ &\quad = (x-\alpha)^2\sum_{t=2}^{n}Y_{t-1}^2 I\{Y_{t-1} > r\} - 2(x-\alpha)\sum_{t=2}^{n}\varepsilon_t Y_{t-1}I\{Y_{t-1} > r\} \end{aligned}$$

$$
\begin{aligned}
&+ (x^{-1} - \alpha^{-1})^2 \sum_{t=2}^{n} Y_{t-1}^2 I\{Y_{t-1} \le r\} \\
&- 2(x^{-1} - \alpha^{-1}) \sum_{t=2}^{n} \varepsilon_t Y_{t-1} I\{Y_{t-1} \le r\} \\
\ge{}& (x-\alpha)^2 \sum_{t=2}^{n} Y_{t-1}^2 I\{Y_{t-1} > r\} - 2(x-\alpha) \sum_{t=2}^{n} \varepsilon_t Y_{t-1} I\{Y_{t-1} > r\} - \sum_{t=2}^{n} \varepsilon_t^2.
\end{aligned}
$$

By Theorem 4.1, we can see that

$$
\begin{aligned}
\frac{1}{\alpha^{2n}} \sum_{t=2}^{n} Y_{t-1}^2 I\{Y_{t-1} > r\} &\to (\alpha^2 - 1)^{-1} \xi \qquad \text{a.s.;} \\
\sum_{t=2}^{n} \varepsilon_t Y_{t-1} I\{Y_{t-1} > r\} &= \mathrm{O}(\alpha^{3n/2}) \qquad \text{a.s;} \\
\sum_{t=2}^{n} Y_{t-1}^2 I\{Y_{t-1} \le r\} &= \mathrm{O}(1) \qquad \text{a.s.;} \\
\sum_{t=2}^{n} \varepsilon_t Y_{t-1} I\{Y_{t-1} \le r\} &= \mathrm{O}(1) \qquad \text{a.s.}
\end{aligned}
$$

Hence for any $\delta > 0$, we have

$$
\lim_{n\to\infty} \inf_{x:|x-\alpha|>\delta} \big(Q_n(x) - Q_n(\alpha)\big) = \infty \qquad \text{a.s.}
$$

Since $Q_n(x)$ is continuous on $[\alpha - \delta, \alpha + \delta]$, it always admits a minimum on this interval. This shows that $\limsup_{n\to\infty} |\hat{\alpha}_n - \alpha| \le \delta$ a.s. for any $\delta > 0$ and completes the proof. □

Acknowledgements

The authors thank the referee and the editor Holger Rootzén for their helpful comments. Shiqing Ling's research supported by HK RGC CERG 601607 and CERG 602609. Qi-Man Shao's research partially supported by Hong Kong RGC CERG 602608.

References

[1] Brown, B.M. (1971). Martingale central limit theorems. *Ann. Math. Statist.* **42** 59–66. MR0290428

[2] Caner, M. and Hansen, B.E. (2001). Threshold autoregression with a unit root. *Econometrica* **69** 1555–1596. MR1865221

[3] Chan, K.S. (1993). Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. *Ann. Statist.* **21** 520–533. MR1212191

[4] Chan, K.S., Petruccelli, J.D., Tong, H. and Woolford, S.W. (1985). A multiple threshold AR(1) model. *J. Appl. Probab.* **22** 267–279. MR0789351

[5] Chan, K.S. and Tsay, R.S. (1998). Limiting properties of the least squares estimator of a continuous threshold autoregressive model. *Biometrika* **85** 413–426. MR1649122

[6] Chan, N.H. and Wei, C.Z. (1988). Limiting distributions of least squares estimates of unstable autoregressive processes. *Ann. Statist.* **16** 367–401. MR0924877

[7] Choi, I. (1999). Testing the random walk hypothesis for real exchange rates. *J. Appl. Econom.* **14** 293–308.

[8] Dickey, D.A. and Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *J. Amer. Statist. Assoc.* **74** 427–431. MR0548036

[9] Orey, S. (1971). *Lecture Notes on Limit Theorems for Markov Chain Transition Probabilities*. London: Van Nostrand and Reinhold Company. MR0324774

[10] Petruccelli, J.D. and Woolford, S.W. (1984). A threshold AR(1) model. *J. Appl. Probab.* **21** 270–286. MR0741130

[11] Pham, D.T., Chan, K.S. and Tong, H. (1991). Strong consistency of the least squares estimator for a nonergodic threshold autoregressive model. *Statist. Sinica* **1** 361–369. MR1130124

[12] Tjøstheim, D. (1990). Nonlinear time series and Markov chains. *Adv. in Appl. Probab.* **22** 587–611. MR1066965

[13] Tong, H. (1978). On a threshold model. In *Pattern Recognition and Signal Processing* (H. Chen, ed.) 575–586. Amsterdam: Sijthoff and Noordhoff.

Received May 2009 and revised June 2010