ORIGINAL RESEARCH

Limit theory for an AR(1) model with intercept and a possible infinite variance

Qing Liu · Chiyu Xia · Xiaohui Liu

Abstract In this paper, we derive the limit distribution of the least squares estimator for an AR(1) model with a non-zero intercept and a possible infinite variance. It turns out that the estimator has a quite different limit for the cases of $|\rho| < 1$, $|\rho| > 1$, and $\rho = 1 + \frac{c}{n^\alpha}$ for some constant $c \in R$ and $\alpha \in (0, 1]$, and whether or not the variance of the model errors is infinite also has a great impact on both the convergence rate and the limit distribution of the estimator.

Keywords Limit distribution · Autoregressive model · Infinite variance

1 Introduction

As a simple but useful tool, the auto-regression (AR) models have been widely used in economics and many other fields. Among them, the simplest one is the autoregressive process with order 1, i.e., an AR(1) model, which is usually defined as

$$y_t = \mu + \rho y_{t-1} + e_t, \quad t = 1, \cdots, n. \tag{1}$$

where y_0 is a constant and e_t's are independent and identically distributed (hereafter, i.i.d.) random errors with mean zero and finite variance. The process $\{y_t\}$ is (i) stationary if $|\rho| < 1$ independent of n, (ii) unit root if $\rho = 1$, (iii) near unit root if $\rho = 1 + c/n$ for some nonzero constant c, iv) explosive if $|\rho| > 1$ independent of n, and v) moderate deviation from a unit root if $\rho = 1 + c/k_n$ for some nonzero constant c and a sequence $\{k_n\}$ satisfying $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$.

When $\mu = 0$ and the error variance in the model (1) is finite, it is well known in the literature that the least squares estimator for ρ has a quite different limit distribution in cases of stationary, unit root and near unit root; see [17]. The convergence rate of the correlation coefficient is $\sqrt{n}$, n for cases (i)–(iii), respectively, and may even be $(1 + c)^n$ in the case of (v) for some $c > 0$ as stated in [19]. More studies on this model can be found in [8], [9], [16], [1,2], [20], [4], [13] and references therein.

When $\mu \neq 0$ with finite variance, [21] and [11] studied the limit theory for the AR(1) for cases of iv) and v), respectively. It is shown that the inclusion of a nonzero intercept may change drastically the large sample

Q. Liu · X. Liu (✉)
School of Statistics and Data Science, Jiangxi University of Finance and Economics, Nanchang 330013, Jiangxi, China
E-mail: liuxiaohui@jxufe.edu.cn

Q. Liu · X. Liu
Key Laboratory of Data Science in Finance and Economics, Jiangxi University of Finance and Economics, Nanchang 330013, Jiangxi, China

Q. Liu
Research Center of Applied Statistics, Jiangxi University of Finance and Economics, Nanchang 330013, Jiangxi, China

properties of the least squares estimator compared to [19]. More recently, [15] studied how to construct a uniform confidence region for (μ, ρ) regardless of (i)–(v) based on the empirical likelihood method.

Observe that e_t may have an infinite variance in practice [5,18], and most of the aforementioned research was focused on the case of e_t having a finite variance. In this paper, we are interested in considering the model (1) when $\mu \neq 0$ and the variance of e_t may possibly be infinite. We will derive the limit distribution of the least squares estimator of (μ, ρ) for the following cases:

- (P1) $|\rho| < 1$ independent of n;
- (P2) $|\rho| > 1$ independent of n;
- (P3) $\rho = 1$;
- (P4) $\rho = 1 + \frac{c}{n}$ for some constant $c \neq 0$;
- (P5) $\rho = 1 + \frac{c}{n^\alpha}$ for some constants $c < 0$ and $\alpha \in (0, 1)$;
- (P6) $\rho = 1 + \frac{c}{n^\alpha}$ for some constants $c > 0$ and $\alpha \in (0, 1)$.

Since the current paper allows for the inclusion of both the intercept and a possible infinite variance, it can be treated as an extension of the existing literature, i.e., [11,14,17,19,21], among others.

We organize the rest of this paper as follows. Section 2 provides the methodology and main limit results. Detailed proofs are put in Section 3.

2 Methodology and main results

Under model (1), by minimizing the sum of squares:

$$\sum_{t=1}^{n}(y_t - \mu - \rho y_{t-1})^2,$$

with respect to $(\mu, \rho)^\top$, we get the least squares estimator for $(\mu, \rho)^\top$ as follows

$$\begin{cases} \hat{\mu} = \dfrac{\sum\limits_{s=1}^{n} y_s \sum\limits_{t=1}^{n} y_{t-1}^2 - \sum\limits_{s=1}^{n} y_{s-1} \sum\limits_{t=1}^{n} y_t y_{t-1}}{n \sum\limits_{t=1}^{n} y_{t-1}^2 - \left(\sum\limits_{t=1}^{n} y_{t-1}\right)^2} \\ \hat{\rho} = \dfrac{n \sum\limits_{t=1}^{n} y_t y_{t-1} - \sum\limits_{s=1}^{n} y_{s-1} \sum\limits_{t=1}^{n} y_t}{n \sum\limits_{t=1}^{n} y_{t-1}^2 - \left(\sum\limits_{t=1}^{n} y_{t-1}\right)^2}. \end{cases} \tag{2}$$

Here $A^\top$ denotes the transpose of the matrix or vector A. In the sequel, we will investigate the limit distribution of $(\hat{\mu} - \mu, \hat{\rho} - \rho)^\top$.

To derive the limit distribution of this least squares estimator, we follow [19] by assuming that

- (C1) The innovations $\{e_t\}$ are i.i.d. with $E[e_t] = 0$;
- (C2) The process is initialized at $y_0 = O_p(1)$.

Observing that the variance of e_t's may not exist, we use the slowly varying function $l(x) = E[e_t^2 I(|e_t| \leq x)]$ instead as did in [14] to characterize the dispersion property of the random errors, which is supposed to satisfy

- C3) $l(nx)/l(n) \to 1$ as $n \to \infty$ for any $x > 0$.

An example of slowly varying function is when $l(x)$ has a limit, say $\lim\limits_{x \to \infty} l(x) = \sigma^2$, which implies $\{e_t\}$ having a finite variance σ^2. Another example is $l(x) = \log(x)$, $x > 1$, which implies that the variance of e_t's does not exist. One known property of $l(x)$ is that $l(x) = o(x^\varepsilon)$ as $x \to \infty$, for any $\varepsilon > 0$. More properties on $l(x)$ can be found in [10]. To deal with the possibly infinite variance, we introduce the following sequence $\{b_k\}_{k=0}^{\infty}$, where

$$b_0 = \inf\{x \geq 1 : l(x) > 0\}$$

and

$$b_j = \inf\left\{s : s \geq b_0 + 1, \frac{l(s)}{s^2} \leq \frac{1}{j}\right\}, \text{ for } j = 1, 2, \ldots, n,$$

which imply directly $nl(b_n) \leq b_n^2$ for all $n \geq 1$; see also [12].

For convenience, in the sequel we still call $|\phi| < 1$ the stationary case, $\rho = 1$ the unit root case, $\phi = 1 + \frac{c}{n}$ for some $c \neq 0$ the near unit root case, $\rho = 1 + \frac{c}{n^\alpha}$ for some $c \neq 0$ the moderate deviation case and $|\rho| > 1$ the explosive case, even when the variance of v_t is infinite. We will divide the theoretical derivations into four separate subsections.

2.1 Limit theory for the stationary case

We first consider the stationary case $|\rho| < 1$, which is independent of n. Observe that

$$y_t = \mu + \rho y_{t-1} + e_t = \frac{\mu}{1-\rho} + \left(y_0 - \frac{\mu}{1-\rho}\right)\rho^t + \sum_{j=1}^{t} \rho^{t-j} e_j.$$

We write $\bar{y}_t = y_t - \frac{\mu}{1-\rho}$, and then have

$$\bar{y}_t = \rho \bar{y}_{t-1} + e_t.$$

To prove the main result for this case, we need the following preliminary lemma.

Lemma 1 *Suppose conditions (C1)–(C3) hold. Under P1, as $n \to \infty$, we have*

$$\frac{1}{n} \sum_{t=1}^{n} y_{t-1} \xrightarrow{p} \frac{\mu}{1-\rho},$$

$$\frac{1}{nl(b_n)} \sum_{t=1}^{n} y_{t-1}^2 \xrightarrow{p} \begin{cases} \frac{1}{1-\rho^2}, & \text{if } \lim_{m\to\infty} l(b_m) = \infty, \\ \frac{1}{1-\rho^2} + \frac{\mu^2}{\sigma^2(1-\rho)^2}, & \text{if } \lim_{m\to\infty} l(b_m) = \sigma^2, \end{cases}$$

and

$$\begin{pmatrix} \frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^{n} e_t \\ \frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^{n} \bar{y}_{t-1} e_t \end{pmatrix} \xrightarrow{d} \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{1-\rho^2} \end{pmatrix}\right). \tag{3}$$

Based on Lemma 1, we can show the following theorem.

Theorem 1 *Under conditions (C1)–(C3), as $n \to \infty$, we have under P1 that*

$$\begin{pmatrix} \sqrt{\frac{n}{l(b_n)}}(\hat{\mu} - \mu) \\ \sqrt{n}(\hat{\rho} - \rho) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix};$$

where $X_1 = W_1 - \frac{\mu(1+\rho)}{\sigma^2} W_2$ and $X_2 = (1-\rho^2)W_2$ if $\lim_{m\to\infty} l(b_n) = \sigma^2$, and $X_1 = W_1$ and $X_2 = (1-\rho^2)W_2$ if $\lim_{m\to\infty} l(b_n) = \infty$.

Remark 1 Theorem 1 indicates that the possible infinite variance may affect the convergence rate of the least squares estimator of the intercept, but has no impact on that of ρ.

Remark 2 When $\lim_{m\to\infty} l(b_m)$ exists and is equal to σ^2, we have $(X_1, X_2)^\top \sim N(0, \Sigma_1)$, where $\Sigma_1 = (\sigma_{ij}^2)_{1\leq i,j\leq 2}$ with $\sigma_{11}^2 = 1 + \frac{\mu^2(1+\rho)}{\sigma^4(1-\rho)}$, $\sigma_{12}^2 = \sigma_{21}^2 = -\frac{\mu(1+\rho)}{\sigma^2}$ and $\sigma_{22}^2 = 1 - \rho^2$. That is, the limit distribution reduces to the ordinary case; see [15] and references therein for details.

2.2 Limit theory for the explosive case

For this case, let $\tilde{y}_t = \sum_{i=1}^{t} \rho^{t-i} e_i + \rho^t y_0$, then

$$y_t = \mu \frac{1-\rho^t}{1-\rho} + \rho^t y_0 + \sum_{i=1}^{t} \rho^{t-i} e_i = \mu \frac{1-\rho^t}{1-\rho} + \tilde{y}_t.$$

Along the same line as Section 2.1, we derive a preliminary lemma first as follows.

Lemma 2 *Suppose conditions (C1)–(C3) hold. Under P2, as $n \to \infty$, we have*

$$\begin{pmatrix} \frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^{n} e_t \\ \frac{1}{\sqrt{l(b_n)}} \sum_{t=1}^{n} \rho^{-(n-t)} e_t \\ \frac{\rho}{\sqrt{l(b_n)}} \sum_{t=1}^{n-1} \rho^{-t} e_t + \rho y_0 \end{pmatrix} \xrightarrow{d} \begin{pmatrix} W_1 \\ U_1 \\ U_2 \end{pmatrix},$$

and

$$\begin{pmatrix} \rho^{-(n-2)} \{l(b_n)\}^{-1} \sum_{t=1}^{n} \tilde{y}_{t-1} e_t \\ (\rho^2 - 1)\rho^{-2(n-1)} \{l(b_n)\}^{-1} \sum_{t=1}^{n} \tilde{y}_{t-1}^2 \end{pmatrix} \xrightarrow{d} \begin{pmatrix} U_1 U_2 \\ U_2^2 \end{pmatrix},$$

where $U_1 \sim \lim_{m\to\infty} \frac{1}{\sqrt{l(b_m)}} \sum_{t=1}^{m} \rho^{-(m-t)} e_t$, $U_2 \sim \rho y_0 + \lim_{m\to\infty} \frac{\rho}{\sqrt{l(b_m)}} \sum_{t=1}^{m-1} \rho^{-t} e_t$, and W_1 are mutually independent random variables. W_1 is specified in Lemma 1.

Using this lemma, we can obtain the following theorem.

Theorem 2 *Under conditions (C1)–(C3), as $n \to \infty$, we have*

$$\begin{pmatrix} \sqrt{\frac{n}{l(b_n)}} (\hat{\mu} - \mu) \\ \rho^n (\hat{\rho} - \rho) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} W_1 \\ (\rho^2 - 1)\frac{U_1}{U_2 + \mu\rho/(\rho-1)} \end{pmatrix},$$

under P2.

Similar to the case of $|\phi| < 1$, Theorem 2 indicates that the possible infinite variance only affects the convergence rate of $\hat{\mu}$. The joint limit distribution reduces to that obtained in [21] if $\lim_{m\to\infty} l(b_m)$ is finite.

2.3 Limit theory for the unit root and near unit root cases

In these two cases, $\rho =: \rho_n = 1 + \frac{c}{n}, c \in \mathbb{R}$. ($c = 0$ corresponds to $\rho = 1$, i.e., the unit root case.) Let $\tilde{y}_t = \sum_{i=1}^{t} \rho^{t-i} e_i$, then

$$y_t = \mu + \rho y_{t-1} + e_t = \mu \left(\sum_{j=0}^{t-1} \rho^j \right) + \rho^t y_0 + \tilde{y}_t.$$

We have the following lemma:

Lemma 3 *Let* $E_n(t) = \dfrac{\sum_{i=1}^{\lfloor ns \rfloor} e_i}{\sqrt{nl(b_n)}}$, $s \in [0, 1]$. *Then*

$$E_n(s) \xrightarrow{D} \bar{W}(s), \ \text{in} \ D[0, 1] \ \text{as} \ n \to \infty,$$

where $\{\bar{W}(s), s \geq 0\}$ *is a standard Brownian process,* $\lfloor \cdot \rfloor$ *is the floor function, and* $\xrightarrow{D}$ *denotes the weak convergence. Moreover, define* $J_c(s) = \lim\limits_{a \to c} \dfrac{1 - e^{as}}{-a}$, *then as* $n \to \infty$, *we have under P3 and P4 that*

$$\begin{cases} n^{-2} \sum_{t=1}^{n} \left(\sum_{j=0}^{t-2} \rho^j \right) \to \int_0^1 J_c(s)\, ds, \\[2ex] n^{-3} \sum_{t=1}^{n} \left(\sum_{j=0}^{t-2} \rho^j \right)^2 \to \int_0^1 J_c^2(s)\, ds, \\[2ex] n^{-3/2} \sum_{t=1}^{n} \left(\sum_{j=0}^{t-2} \rho^j \right) \frac{e_t}{\sqrt{l(b_n)}} \xrightarrow{d} \int_0^1 J_c(s)\, d\bar{W}(s). \end{cases}$$

and in turn

$$\frac{\sum_{k=1}^{n} \tilde{y}_t^2}{n^2 l(b_n)} \xrightarrow{d} \int_0^1 e^{-2c(1-s)} \bar{W}^2(B_c(s))\, ds,$$

$$\frac{\sum_{k=1}^{n} \tilde{y}_t}{n^{3/2} \sqrt{l(b_n)}} \xrightarrow{d} \int_0^1 e^{-c(1-s)} \bar{W}(B_c(s))\, ds,$$

$$\frac{\sum_{k=1}^{n} \tilde{y}_{t-1} e_t}{n l(b_n)} \xrightarrow{d} -c \int_0^1 e^{-2c(1-s)} \bar{W}^2(B_c(s))\, ds + \frac{\bar{W}^2(B_c(1))}{2} - \frac{1}{2},$$

where $B_c(s) = e^{2c}(e^{-2cs} - 1)/(-2c)$.

This lemma can be easily proved using similar techniques as in [6] based on the fact that

$$\frac{1}{n^{3/2} \sqrt{l(b_n)}} \sum_{t=1}^{n} \left(\sum_{j=0}^{t-2} \rho^j \right) e_t^{(2)} \xrightarrow{p} 0, \ n \to \infty.$$

Using this lemma, it is easy to check the following theorem.

Theorem 3 *Under the conditions (C1)–(C3), as* $n \to \infty$, *we have*

$$\begin{pmatrix} \sqrt{\frac{n}{l(b_n)}} (\hat{\mu} - \mu) \\[1ex] \sqrt{\frac{n^3}{l(b_n)}} (\hat{\rho} - \rho) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} Y_1/d \\ Y_2/(\mu d) \end{pmatrix};$$

under P3 and P4, where

$$d = \int_0^1 J_c^2(s)\, ds - \left(\int_0^1 J_c(s)\, ds \right)^2,$$

$$Y_1 = \bar{W}(1) \int_0^1 J_c^2(s)\, ds - \int_0^1 J_c(s)\, ds \int_0^1 J_c(s)\, d\bar{W}(s),$$

$$Y_2 = \int_0^1 J_c(s)\, d\bar{W}(s) - \bar{W}(1) \int_0^1 J_c(s)\, ds.$$

2.4 Limit theory for the moderate deviation cases

As stated in [19], the moderate deviation cases bridge the different convergence rates of cases P1–P4. That is, case P5 bridges the stationary case and the near unit root case, while Case P5 bridges the explosive case and the near unit root case. The derivation of these two cases needs to be handled differently because for the case $c > 0$, the martingale center limit theorem fails to hold. Following [19], we consider them separately.

The following lemma is useful in deriving the limit distribution of the least square estimator under cases P5-P6.

Lemma 4 *Suppose conditions (C1)–(C3) hold.*

- *(i) Under P5, as $n \to \infty$, we have*

$$
\begin{pmatrix}
\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{e_t}{\sqrt{l(b_n)}} \\
\frac{1}{\sqrt{n^\alpha}} \sum_{t=1}^{n} \frac{\rho^{t-1} e_t}{\sqrt{l(b_n)}} \\
\frac{1}{\sqrt{n^\alpha}} \sum_{t=1}^{n} \frac{\rho^{n-t} e_t}{\sqrt{l(b_n)}} \\
\frac{1}{\sqrt{n^{1+\alpha}}} \sum_{t=2}^{n} \frac{e_t}{\sqrt{l(b_n)}} \sum_{j=1}^{t-1} \frac{\rho^{t-1-j} e_j}{\sqrt{l(b_n)}}
\end{pmatrix}
\xrightarrow{d}
\begin{pmatrix} V_{11} \\ V_{12} \\ V_{13} \\ V_{14} \end{pmatrix} \sim N(0, \Sigma_2),
$$

where $\Sigma_2 = diag(1, -\frac{1}{2c}, -\frac{1}{2c}, -\frac{1}{2c})$, which implies that V_{1i}'s are independent;
- *(ii) Under P6, as $n \to \infty$, we have*

$$
\begin{pmatrix}
\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{e_t}{\sqrt{l(b_n)}} \\
\frac{1}{\sqrt{n^\alpha}} \sum_{t=1}^{n} \frac{\rho^{-t} e_t}{\sqrt{l(b_n)}} \\
\frac{1}{\sqrt{n^\alpha}} \sum_{t=1}^{n} \frac{\rho^{t-1-n} e_t}{\sqrt{l(b_n)}}
\end{pmatrix}
\xrightarrow{d}
\begin{pmatrix} V_{21} \\ V_{22} \\ V_{23} \end{pmatrix},
$$

and

$$
\frac{1}{\rho^n n^\alpha} \sum_{t=2}^{n} \left(\sum_{i=1}^{t-1} \frac{\rho^{t-1-i} e_i}{\sqrt{l(b_n)}} \right) \frac{e_t}{\sqrt{l(b_n)}} \xrightarrow{d} V_{22} V_{23},
$$

where $(V_{21}, V_{22}, V_{23})^\top \sim N(0, \Sigma_3)$ with $\Sigma_3 = diag(1, \frac{1}{2c}, \frac{1}{2c})$, which implies that V_{2i}'s are independent.

Theorem 4 *Suppose conditions (C1)–(C3) hold.*

- *(i) Under P5, we have as $n \to \infty$*

$$
\begin{pmatrix} a_n(\hat{\mu} - \mu) \\ a_n n^\alpha(\hat{\rho} - \rho) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \frac{\mu}{cd} \\ \frac{1}{d} \end{pmatrix}^\top Z;
$$

where

$$
a_n = \frac{\max(n^{1-\alpha/2} l(b_n), n^{3\alpha/2})}{\max(n^\alpha \sqrt{l(b_n)}, n^{1/2} l(b_n))},
$$

$$
Z = \frac{\mu}{c} V_{12} I(\alpha > 1/2) + V_{14} I(\alpha \le 1/2),
$$

$$
d = \frac{\mu^2}{-2c^3} I(\alpha > 1/2) + \frac{1}{-2c} I(\alpha \le 1/2),
$$

if $\lim_{n \to \infty} l(b_n) = \infty$, and

$$
a_n = n^{\max(\alpha, 1/2) - \alpha/2},
$$

$$Z = \frac{\mu}{c} V_{12} I(\alpha \geq 1/2) + V_{14} I(\alpha \leq 1/2),$$

$$d = \frac{\mu^2}{-2c^3} I(\alpha \geq 1/2) + \frac{1}{-2c} I(\alpha \leq 1/2),$$

if $\lim_{m \to \infty} l(b_m) = \sigma^2$.

• (ii) Under P6, we have as $n \to \infty$

$$\begin{pmatrix} \sqrt{\frac{n}{l(b_n)}} (\hat{\mu} - \mu) \\ \frac{n^{3\alpha/2} \rho^n}{\sqrt{l(b_n)}} (\hat{\rho} - \rho) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} V_{21} \\ \frac{2c^2}{\mu} V_{23} \end{pmatrix}.$$

Remark 3 Theorem 4 indicates that the possible infinite variance affects both estimators of μ and ρ. Similar to [21], the limit distribution of the least squares estimators under P5 is degenerated, but slightly differently, we obtain the exact limit distribution under this case.

Remark 4 Under some mild conditions, it is possible to extend the current result to the case that $\rho = 1 + \frac{c}{k_n}$ by using similar arguments for some general sequence $\{k_n\}$ such that $k_n = o(n)$ and $k_n/n \to 0$ as $n \to \infty$ as studied in [19].

3 Detailed proofs of the main results

In this section, we provide all detailed proofs of the lemmas and theorems stated in Section 2.

Proof of Lemma 1 To handle the possible infinite variance, we use the truncated random variables. Let

$$\begin{cases} e_t^{(1)} = e_t I(|e_t| \leq b_n) - E[e_t I(|e_t| \leq b_n)], \\ e_t^{(2)} = e_t I(|e_t| > b_n) - E[e_t I(|e_t| > b_n)], \end{cases} \tag{4}$$

where $I(\cdot)$ denotes the indicative function. The key step is to show that the difference of replacing e_t by $e_t^{(1)}$ in the summations is negligible.

Let $\{\bar{y}_t^{(1)}\}$ and $\{\bar{y}_t^{(2)}\}$ be two time series satisfying

$$\bar{y}_t^{(k)} = \rho \bar{y}_{t-1}^{(k)} + e_t^{(k)}, \quad k = 1, 2.$$

Obviously, $\{e_t^{(1)}/\sqrt{l(b_n)} : t \geq 1\}$ are i.i.d., and under P1), it is easy to check that $\{\bar{y}_{t-1} e_t^{(1)}/\sqrt{l(b_n)} : t \geq 1\}$ is a martingale differences sequence with respect to $\mathcal{F}_{t-1} = \sigma(\{e_s : s \leq t - 1\})$ for $t = 1, 2, \cdots, n$, which satisfy the Lindeberg's condition. Hence, by the Cramér-Wold device and the central limit theorem for martingale difference sequences, we have

$$\begin{pmatrix} \frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^{n} e_t^{(1)} \\ \frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^{n} \bar{y}_{t-1}^{(1)} e_t^{(1)} \end{pmatrix} \xrightarrow{d} \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}. \tag{5}$$

Next, under condition C3), it follows from [7] that

$$E[|e_t| I(|e_t| > b_n)] = o(l(b_n) b_n^{-1}), \quad n \to \infty.$$

Then by $nl(b_n) \leq b_n^2$ and the Markov inequality, we have, for any $\varepsilon > 0$,

$$P\left(\left| \sum_{t=1}^{n} e_t^{(2)} \right| \geq \sqrt{nl(b_n)} \varepsilon \right) \leq \frac{\sum_{t=1}^{n} E|e_t^{(2)}|}{\sqrt{nl(b_n)} \varepsilon}$$

$$\leq 2 \frac{\sum_{t=1}^{n} E[|e_t| I(|e_t| > b_n)]}{\sqrt{nl(b_n)} \varepsilon}$$

$$= o\left(\frac{\sqrt{nl(b_n)}}{b_n}\right) = o(1), \quad \text{as } n \to \infty,$$

That is,

$$\frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^{n} e_t^{(2)} = o_p(1), \quad n \to \infty, \tag{6}$$

Furthermore, note that $\bar{y}_{t-1}^{(k)} = \rho^{t-1} \bar{y}_0^{(k)} + \sum_{i=1}^{t-1} \rho^{t-1-i} e_i^{(k)}$, $k = 1, 2$. By the Hölder inequality, we have

$$E\left(\left|\frac{e_t^{(1)}}{\sqrt{l(b_n)}}\right|\right) \le \left\{E\left(\frac{e_t^{(1)}}{\sqrt{l(b_n)}}\right)^2\right\}^{1/2} \le 1.$$

Using the Markov inequality, we have

$$P\left(\left|\sum_{t=1}^{n} \bar{y}_{t-1}^{(1)} e_t^{(2)}\right| \ge \sqrt{nl(b_n)}\varepsilon\right)$$

$$\le \frac{1}{\sqrt{nl(b_n)}\varepsilon} \sum_{t=1}^{n} E(|\bar{y}_{t-1}^{(1)}|) E(|e_t^{(2)}|)$$

$$\le \frac{1}{\sqrt{nl(b_n)}\varepsilon} \frac{1}{1-\rho} \{E(|\bar{y}_0^{(1)}|) + nE(|e_t^{(1)}|)\} E(|e_t^{(2)}|)$$

$$= o(n^{-1/2+\alpha/2} b_n^{-1}) + o\left(\frac{\sqrt{nl(b_n)}}{b_n}\right) = o(1),$$

i.e.,

$$\frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^{n} \bar{y}_{t-1}^{(1)} e_t^{(2)} = o_p(1), \quad \text{as } n \to \infty. \tag{7}$$

Similarly, we can show

$$\frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^{n} \bar{y}_{t-1}^{(2)} e_t^{(j)} = o_p(1), \quad n \to \infty, \ j = 1, 2. \tag{8}$$

This, together with (6)-(8), shows

$$\begin{pmatrix} \frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^{n} e_t \\ \frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^{n} \bar{y}_{t-1} e_t \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^{n} e_t^{(1)} \\ \frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^{n} \bar{y}_{t-1}^{(1)} e_t^{(1)} \end{pmatrix} + o_p(1), \quad \text{as } n \to \infty,$$

while combined with (5) shows (3).

Note that $\bar{y}_t = \rho \bar{y}_{t-1} + e_t$. Multiplying both sides with $\bar{y}_t$ and y_{t-1} respectively, and taking summation, we have

$$\begin{cases} \sum_{t=1}^{n} \bar{y}_t^2 = \rho \sum_{t=1}^{n} \bar{y}_t \bar{y}_{t-1} + \sum_{t=1}^{n} \bar{y}_t e_t, \\ \sum_{t=1}^{n} \bar{y}_t \bar{y}_{t-1} = \rho \sum_{t=1}^{n} \bar{y}_{t-1}^2 + \sum_{t=1}^{n} \bar{y}_{t-1} e_t. \end{cases}$$

Since

$$\frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^{n} \bar{y}_{t-1} e_t \xrightarrow{d} W_2,$$

and

$$\sum_{t=1}^{n} \bar{y}_t e_t = \rho \sum_{t=1}^{n} \bar{y}_{t-1} e_t + \sum_{t=1}^{n} e_t^2,$$

we have

$$\frac{\sum_{t=1}^{n} \bar{y}_t e_t}{nl(b_n)} \xrightarrow{P} 1, \ n \to \infty,$$

by noting that $\frac{\sum_{t=1}^{n} e_t^2}{nl(b_n)} \xrightarrow{P} 1$ (see (3.4) in [12]). Hence,

$$\begin{cases} \dfrac{\sum\limits_{t=1}^{n} \bar{y}_t^2}{nl(b_n)} = \rho \dfrac{\sum\limits_{t=1}^{n} \bar{y}_t \bar{y}_{t-1}}{nl(b_n)} + 1 + o_p(1) \\ \dfrac{\sum\limits_{t=1}^{n} \bar{y}_t \bar{y}_{t-1}}{nl(b_n)} = \rho \dfrac{\sum\limits_{t=1}^{n} \bar{y}_{t-1}^2}{nl(b_n)} + o_p(1), \end{cases}$$

which implies that as $n \to \infty$

$$\begin{cases} \dfrac{\sum\limits_{t=1}^{n} \bar{y}_t^2}{nl(b_n)} \xrightarrow{P} \dfrac{1}{1-\rho^2} \\ \dfrac{\sum\limits_{t=1}^{n} \bar{y}_t \bar{y}_{t-1}}{nl(b_n)} \xrightarrow{P} \dfrac{\rho}{1-\rho^2}. \end{cases}$$

Note that

$$\frac{1}{n} \sum_{t=1}^{n} y_{t-1} = \frac{1}{n} \sum_{t=1}^{n} \bar{y}_{t-1} + \frac{\mu}{1-\rho} = \frac{\mu}{1-\rho} + o_p(1),$$

and

$$\frac{1}{nl(b_n)} \sum_{t=1}^{n} y_{t-1}^2 = \frac{1}{nl(b_n)} \sum_{t=1}^{n} \bar{y}_{t-1}^2 + \frac{2\mu}{1-\rho} \frac{1}{nl(b_n)} \sum_{t=1}^{n} \bar{y}_{t-1} + \frac{1}{l(b_n)} \frac{\mu^2}{(1-\rho)^2}.$$

Using these, the rest proof of this lemma follows directly by the law of large numbers.

Proof of Theorem 1 For the least squares estimator, it is easy to check that

$$\begin{pmatrix} \hat{\mu} - \mu \\ \hat{\rho} - \rho \end{pmatrix} = \begin{pmatrix} \dfrac{\sum\limits_{t=1}^{n} y_{t-1}^2 \sum\limits_{s=1}^{n} e_s - \sum\limits_{s=1}^{n} y_{s-1} \sum\limits_{t=1}^{n} e_t y_{t-1}}{n \sum\limits_{t=1}^{n} y_{t-1}^2 - \left(\sum\limits_{t=1}^{n} y_{t-1}\right)^2} \\ \dfrac{n \sum\limits_{t=1}^{n} y_{t-1} e_t - \sum\limits_{t=1}^{n} y_{t-1} \sum\limits_{s=1}^{n} e_s}{n \sum\limits_{t=1}^{n} y_{t-1}^2 - \left(\sum\limits_{t=1}^{n} y_{t-1}\right)^2} \end{pmatrix}.$$

For convenience, hereafter write

$$\Delta_1 = \sum_{t=1}^{n} y_{t-1}^2 \sum_{s=1}^{n} e_s - \sum_{s=1}^{n} y_{s-1} \sum_{t=1}^{n} y_{t-1} e_t,$$

$$\Delta_2 = n \sum_{t=1}^{n} y_{t-1} e_t - \sum_{t=1}^{n} y_{t-1} \sum_{s=1}^{n} e_s,$$

$$\Delta_3 = n \sum_{t=1}^{n} y_{t-1}^2 - \left(\sum_{t=1}^{n} y_{t-1}\right)^2.$$

Observe that

$$\Delta_1 = \sum_{t=1}^n y_{t-1}^2 \sum_{s=1}^n e_s - \sum_{s=1}^n y_{s-1} \sum_{t=1}^n y_{t-1} e_t$$

$$= \left(\sum_{t=1}^n \bar{y}_{t-1}^2 + \frac{\mu}{1-\rho} \sum_{t=1}^n \bar{y}_{t-1} \right) \sum_{s=1}^n e_s - \sum_{s=1}^n y_{s-1} \sum_{t=1}^n \bar{y}_{t-1} e_t.$$

Hence, by Lemma 1 we have, as $n \to \infty$,

$$\frac{1}{(nl(b_n))^{3/2}} \Delta_1 \xrightarrow{d} \frac{1}{1-\rho^2} W_1 - \frac{\mu}{\sigma^2(1-\rho)} W_2 I \left(\lim_{m\to\infty} l(b_m) = \sigma^2 \right).$$

Next, relying on

$$\Delta_2 = n \sum_{t=1}^n y_{t-1} e_t - \sum_{t=1}^n y_{t-1} \sum_{s=1}^n e_s$$

$$= n \sum_{t=1}^n \bar{y}_{t-1} e_t - \sum_{t=1}^n \bar{y}_{t-1} \sum_{s=1}^n e_s$$

$$= n \left(\sum_{t=1}^n \bar{y}_{t-1} e_t \right) \{1 + o_p(1)\},$$

we obtain

$$\frac{1}{n^{3/2} l(b_n)} \Delta_2 \xrightarrow{d} W_2.$$

Following a similar fashion, we have

$$\frac{1}{n^2 l(b_n)} \Delta_3 = \frac{1}{nl(b_n)} \sum_{t=1}^n y_{t-1}^2 - \frac{1}{l(b_n)} \left(\frac{1}{n} \sum_{t=1}^n y_{t-1} \right)^2 \xrightarrow{p} \frac{1}{1-\rho^2}.$$

Then this theorem follows immediately by using the continuous mapping theorem.

Proof of Lemma 2 For the first part, by following a similar fashion to Lemma 1, we can show that

$$\begin{pmatrix} \frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^n e_t \\ \frac{1}{\sqrt{l(b_n)}} \sum_{t=1}^n \rho^{-(n-t)} e_t \\ \frac{\rho}{\sqrt{l(b_n)}} \sum_{t=1}^{n-1} \rho^{-t} e_t + \rho y_0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^n e_t^{(1)} \\ \frac{1}{\sqrt{l(b_n)}} \sum_{t=1}^n \rho^{-(n-t)} e_t^{(1)} \\ \frac{\rho}{\sqrt{l(b_n)}} \sum_{t=1}^{n-1} \rho^{-t} e_t^{(1)} + \rho y_0 \end{pmatrix} + o_p(1).$$

The rest proof is similar to [3] and [21]. We omit the details.

For the second part, we only prove the case of $\lim_{m\to\infty} l(b_m) = \infty$. Let $\tilde{y}_t^{(k)} = \sum_{i=1}^t \rho^{t-i} e_i^{(k)} + \rho^t y_0$, $k = 1, 2$, $t = 1, 2, \cdots, n$. Similar to Lemma 1, we have

$$\rho^{-n} \{l(b_n)\}^{-1} \sum_{t=1}^n E|\tilde{y}_{t-1}^{(k)} e_t^{(j)}| \xrightarrow{p} 0,$$

$$\rho^{-2n} \{l(b_n)\}^{-1} \sum_{t=1}^n E|\tilde{y}_{t-1}^{(k)} \tilde{y}_{t-1}^{(j)}| \xrightarrow{p} 0,$$

for $(k, j) \in \{(1, 2), (2, 1), (2, 2)\}$, as $n \to \infty$, and in turn, we can obtain that

$$
\begin{pmatrix}
\rho^{-(n-2)}\{l(b_n)\}^{-1} \sum_{t=1}^{n} \tilde{y}_{t-1} e_t \\
(\rho^2 - 1)\rho^{-2(n-1)}\{l(b_n)\}^{-1} \sum_{t=1}^{n} \tilde{y}_{t-1}^2
\end{pmatrix}
=
\begin{pmatrix}
\rho^{-(n-2)}\{l(b_n)\}^{-1} \sum_{t=1}^{n} \tilde{y}_{t-1}^{(1)} e_t^{(1)} \\
(\rho^2 - 1)\rho^{-2(n-1)}\{l(b_n)\}^{-1} \sum_{t=1}^{n} (\tilde{y}_{t-1}^{(1)})^2
\end{pmatrix}
+ o_p(1).
$$

The rest proof is similar to the first part. We omit the details.

Proof of Theorem 2 Using the same arguments as [21], it follows from Lemma 2 that

$$
\begin{cases}
\rho^{-(n-1)}\{l(b_n)\}^{-1/2} y_n \xrightarrow{d} U_2 + \frac{\mu\rho}{\rho-1}, \\
\rho^{-(n-2)}\{l(b_n)\}^{-1} \sum_{t=1}^{n} y_{t-1} e_t \xrightarrow{d} U_1(U_2 + \frac{\mu\rho}{\rho-1}), \\
(\rho - 1)\rho^{-(n-1)}\{l(b_n)\}^{-1/2} \sum_{t=1}^{n} y_{t-1} \xrightarrow{d} U_2 + \frac{\mu\rho}{\rho-1}, \\
(\rho^2 - 1)\rho^{-2(n-1)}\{l(b_n)\}^{-1} \sum_{t=1}^{n} y_{t-1}^2 \xrightarrow{d} (U_2 + \frac{\mu\rho}{\rho-1})^2.
\end{cases}
$$

Then as $n \to \infty$, we have

$$
\begin{aligned}
\frac{1}{n^{1/2}\rho^{2n}\{l(b_n)\}^{3/2}} \Delta_1 &= \rho^{-2n}\{l(b_n)\}^{-1} \sum_{t=1}^{n} y_{t-1}^2 \times \frac{1}{\sqrt{nl(b_n)}} \sum_{t=1}^{n} e_t + o_p(1) \\
&\xrightarrow{d} \frac{1}{\rho^2(\rho^2 - 1)} W_1 \left(U_2 + \frac{\mu\rho}{\rho - 1} \right)^2 \\
\frac{1}{n\rho^n l(b_n)} \Delta_2 &= \rho^{-n}\{l(b_n)\}^{-1} \sum_{t=1}^{n} y_{t-1} e_t + o_p(1) \\
&\xrightarrow{d} \frac{1}{\rho^2} U_1 \left(U_2 + \frac{\mu\rho}{\rho - 1} \right).
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{1}{n\rho^{2n} l(b_n)} \Delta_3 &= \rho^{-2n}\{l(b_n)\}^{-1} \sum_{t=1}^{n} y_{t-1}^2 + o_p(1) \\
&\xrightarrow{d} \frac{1}{\rho^2(\rho^2 - 1)} \left(U_2 + \frac{\mu\rho}{\rho - 1} \right)^2.
\end{aligned}
$$

Then the theorem has been proved.

Proof of Theorem 3 Similar to the proof of Theorem 1, by Lemma 3, we have, as $n \to \infty$,

$$
\begin{aligned}
\Delta_1 &= \sum_{t=1}^{n} y_{t-1}^2 \sum_{t=1}^{n} e_t - \sum_{t=1}^{n} y_{t-1} \sum_{t=1}^{n} y_{t-1} e_t \\
&= \mu^2 \left\{ \sum_{t=1}^{n} \left(\sum_{j=0}^{t-2} \rho^j \right)^2 \sum_{s=1}^{n} e_s - \left(\sum_{s=1}^{n} \sum_{j=0}^{s-2} \rho^j \right) \left(\sum_{t=1}^{n} \sum_{j=0}^{t-2} \rho^j e_t \right) \right\} \{1 + o_p(1)\},
\end{aligned}
$$

which implies

$$
\frac{1}{\sqrt{n^7 l(b_n)}} \Delta_1 \xrightarrow{d} \mu^2 \left(\bar{W}(1) \int_0^1 J_c^2(s)\, ds - \int_0^1 J_c(s)\, ds \int_0^1 J_c(s)\, d\bar{W}(s) \right),
$$

and

$$\Delta_2 = n \sum_{t=1}^{n} y_{t-1} e_t - \sum_{t=1}^{n} y_{t-1} \sum_{t=1}^{n} e_t$$

$$= \mu \left\{ n \left(\sum_{t=1}^{n} \sum_{j=0}^{t-2} \rho^j e_t \right) - \left(\sum_{t=1}^{n} \sum_{j=0}^{t-2} \rho^j \right) \sum_{t=1}^{n} e_t \right\} \{1 + o_p(1)\},$$

which leads

$$\frac{1}{\sqrt{n^5 l(b_n)}} \Delta_2 \xrightarrow{d} \mu \left(\int_0^1 J_c(s) \, d\bar{W}(s) - \bar{W}(1) \int_0^1 J_c(s) \, ds \right),$$

and

$$\Delta_3 = n \sum_{t=1}^{n} y_{t-1}^2 - \left(\sum_{t=1}^{n} y_{t-1} \right)^2$$

$$= \mu^2 n \sum_{t=1}^{n} \left(\sum_{j=0}^{t-2} \rho^j \right)^2 - \mu^2 \left(\sum_{t=1}^{n} \sum_{j=0}^{t-2} \rho^j \right)^2 + o_p(n^4),$$

which results in

$$\frac{1}{n^4} \Delta_3 \to \mu^2 \left(\int_0^1 J_c^2(s) \, ds - \left(\int_0^1 J_c(s) \, ds \right)^2 \right), \quad n \to \infty.$$

Then the theorem has been proved.

Proof of Lemma 4 (i) Similar to Lemma 1, under P5, by the Markov inequality and the fact $nl(b_n) \le b_n^2$, we have for any $\varepsilon > 0$

$$P \left(\left| \sum_{t=1}^{n} \rho^{n-t} e_t^{(2)} \right| \ge \sqrt{n^\alpha l(b_n)} \varepsilon \right) \le \frac{\sum_{t=1}^{n} \rho^{t-1} E |e_t^{(2)}|}{\sqrt{n^\alpha l(b_n)} \varepsilon}$$

$$\le 2 \frac{\sum_{t=1}^{n} \rho^{t-1} E[|e_t| I (|e_t| > b_n)]}{\sqrt{n^\alpha l(b_n)} \varepsilon}$$

$$= o \left(\frac{l(b_n)}{b_n} \right) \frac{1 - \rho^n}{1 - \rho} \frac{1}{\sqrt{n^\alpha l(b_n)} \varepsilon}$$

$$= o \left(\frac{\sqrt{n^\alpha l(b_n)}}{b_n} \right) \to 0, \quad \text{as } n \to \infty.$$

This implies that

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{e_t^{(2)}}{\sqrt{l(b_n)}} \xrightarrow{P} 0 \quad \text{and} \quad \frac{1}{\sqrt{n^\alpha}} \sum_{t=1}^{n} \frac{\rho^{t-1} e_t^{(2)}}{\sqrt{l(b_n)}} \xrightarrow{P} 0 \quad \text{as } n \to \infty.$$

Similarly, we can show that

$$\frac{1}{\sqrt{n^\alpha}} \sum_{t=1}^{n} \frac{\rho^{n-t} e_t^{(2)}}{\sqrt{l(b_n)}} \xrightarrow{P} 0, \quad n \to \infty.$$

Next, if one of i or j equals 2, it follows from Lemma A.2 of [14] that

$$\frac{1}{\sqrt{n^{1+\alpha}}} \sum_{t=2}^{n} \left\{ \sum_{k=1}^{t-1} \frac{\rho^{t-1-k} e_k^{(i)}}{\sqrt{l(b_n)}} \right\} \frac{e_t^{(j)}}{\sqrt{l(b_n)}} \xrightarrow{P} 0, \quad n \to \infty.$$

We actually obtain

$$
\begin{pmatrix}
\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{e_t}{\sqrt{l(b_n)}} \\[2mm]
\frac{1}{\sqrt{n^{\alpha}}} \sum_{t=1}^{n} \frac{\rho^{t-1} e_t}{\sqrt{l(b_n)}} \\[2mm]
\frac{1}{\sqrt{n^{\alpha}}} \sum_{t=1}^{n} \frac{\rho^{n-t} e_t}{\sqrt{l(b_n)}} \\[2mm]
\frac{1}{\sqrt{n^{1+\alpha}}} \sum_{t=2}^{n} \frac{e_t}{\sqrt{l(b_n)}} \sum_{j=1}^{t-1} \frac{\rho^{t-1-j} e_j}{\sqrt{l(b_n)}}
\end{pmatrix}
=
\begin{pmatrix}
\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{e_t^{(1)}}{\sqrt{l(b_n)}} \\[2mm]
\frac{1}{\sqrt{n^{\alpha}}} \sum_{t=1}^{n} \frac{\rho^{t-1} e_t^{(1)}}{\sqrt{l(b_n)}} \\[2mm]
\frac{1}{\sqrt{n^{\alpha}}} \sum_{t=1}^{n} \frac{\rho^{n-t} e_t^{(1)}}{\sqrt{l(b_n)}} \\[2mm]
\frac{1}{\sqrt{n^{1+\alpha}}} \sum_{t=2}^{n} \frac{e_t^{(1)}}{\sqrt{l(b_n)}} \sum_{j=1}^{t-1} \frac{\rho^{t-1-j} e_j^{(1)}}{\sqrt{l(b_n)}}
\end{pmatrix}
+ o_p(1). \tag{9}
$$

Based on the Cramér-Wold device and central limit theorem for martingales differences sequence, Lindeberg's condition for the first part of the right side of (9) can be proved by using the same arguments as [19] and [14]. We omit the details.

(ii) The proof of the case under P6 is similar to that of (i) and [19], thus is omitted.

Proof of Theorem 4 (i) Under P5, observe that $\rho^n = o(n^{-\alpha})$, $y_0 = o(n^{\alpha/2})$, and

$$
y_t = \mu + \rho y_{t-1} + e_t
$$
$$
= \frac{\mu}{1-\rho} + \left(\frac{\mu}{c} n^{\alpha} + y_0 \right) \rho^t + \sum_{i=1}^{t} \rho^{t-i} e_i,
$$

which implies

$$
y_n = \left(\frac{\mu}{1-\rho} + \sum_{i=1}^{n} \rho^{t-i} e_i \right) \{1 + o_p(1)\}.
$$

It is easy to verify that

$$
\sum_{t=1}^{n} y_{t-1} = \frac{1}{1-\rho} \left(n\mu - y_n + \sum_{t=1}^{n} e_t \right),
$$
$$
\sum_{t=1}^{n} y_{t-1} e_t = \left\{ \sum_{t=1}^{n} e_t \left(\frac{\mu}{1-\rho} + \sum_{i=1}^{t-1} \rho^{t-1-i} e_i \right) \right\} \{1 + o_p(1)\},
$$
$$
\sum_{t=1}^{n} y_{t-1}^2 = \frac{1}{1-\rho^2} \left(n\mu^2 - y_n^2 + \sum_{t=1}^{n} e_t^2 + 2\mu\rho \sum_{t=1}^{n} y_{t-1} + 2\mu \sum_{t=1}^{n} e_t + 2\rho \sum_{t=1}^{n} y_{t-1} e_t \right)
$$
$$
= \frac{1}{1-\rho^2} \left(n\mu^2 - y_n^2 + \sum_{t=1}^{n} e_t^2 + 2\mu\rho \sum_{t=1}^{n} y_{t-1} \right) \{1 + o_p(1)\}.
$$

Hence,

$$
\Delta_1 = \sum_{t=1}^{n} y_{t-1}^2 \cdot \sum_{t=1}^{n} e_t - \sum_{t=1}^{n} y_{t-1} \cdot \sum_{t=1}^{n} y_{t-1} e_t
$$
$$
= \left\{ \frac{1}{1-\rho^2} \left(n\mu^2 - y_n^2 + \sum_{t=1}^{n} e_t^2 + 2\mu\rho \sum_{t=1}^{n} y_{t-1} + 2\mu \sum_{t=1}^{n} e_t + 2\rho \sum_{t=1}^{n} y_{t-1} e_t \right) \cdot \sum_{t=1}^{n} e_t \right.
$$
$$
\left. - \frac{1}{1-\rho} \left(n\mu - y_n + \sum_{t=1}^{n} e_t \right) \cdot \sum_{t=1}^{n} y_{t-1} e_t \right\} \{1 + o_p(1)\}
$$
$$
= \frac{1}{1-\rho} \left\{ \left(\frac{n\mu^2 - y_n^2 + \sum_{t=1}^{n} e_t^2}{2} + \mu\rho \sum_{t=1}^{n} y_{t-1} + \mu \sum_{t=1}^{n} e_t \right) \cdot \sum_{t=1}^{n} e_t - \left(n\mu - y_n \right) \right.
$$

$$\times\left[\frac{\mu}{1-\rho}\sum_{t=1}^{n}e_t - \frac{\mu}{1-\rho}\sum_{t=1}^{n}\rho^{t-1}e_t + \sum_{t=1}^{n}\left(e_t\sum_{j=0}^{t-1}\rho^{t-1-j}e_j\right)\right]\Big\}\{1+o_p(1)\}$$

$$=\frac{n\mu}{1-\rho}\left\{\frac{\mu}{1-\rho}\sum_{t=1}^{n}\rho^{t-1}e_t - \sum_{t=1}^{n}\left(e_t\sum_{j=0}^{t-1}\rho^{t-1-j}e_j\right)\right\}\{1+o_p(1)\}$$

$$=\left\{n^{1+5\alpha/2}\sqrt{l(b_n)}\frac{\mu^2}{c^2}\left(\frac{1}{\sqrt{n^\alpha}}\sum_{t=1}^{n}\rho^{t-1}\frac{e_t}{\sqrt{l(b_n)}}\right)\right.$$

$$\left.+n^{3(1+\alpha)/2}l(b_n)\frac{\mu}{c}\left(\frac{1}{\sqrt{n^{1+\alpha}}}\sum_{t=1}^{n}\frac{e_t}{\sqrt{l(b_n)}}\sum_{j=0}^{t-1}\rho^{t-1-j}\frac{e_j}{\sqrt{l(b_n)}}\right)\right\}\{1+o_p(1)\},$$

and

$$\Delta_2 = -\sum_{t=1}^{n}y_{t-1}\cdot\sum_{t=1}^{n}e_t + n\sum_{t=1}^{n}y_{t-1}e_t$$

$$=-\frac{1}{1-\rho}\left(n\mu - y_n + \sum_{t=1}^{n}e_t\right)\cdot\sum_{t=1}^{n}e_t + n\left[\frac{\mu}{1-\rho}\sum_{t=1}^{n}e_t\right.$$

$$\left.-\frac{\mu}{1-\rho}\sum_{t=1}^{n}\rho^{t-1}e_t + \sum_{t=1}^{n}\left(e_t\sum_{j=0}^{t-1}\rho^{t-1-j}e_j\right)\right]$$

$$=\left\{n^{1+3\alpha/2}\sqrt{l(b_n)}\cdot\frac{\mu}{c}\cdot\left(\frac{1}{\sqrt{n^\alpha}}\sum_{t=1}^{n}\rho^{t-1}\frac{e_t}{\sqrt{l(b_n)}}\right)\right.$$

$$\left.+n^{3/2+\alpha/2}l(b_n)\left(\frac{1}{\sqrt{n^{1+\alpha}}}\sum_{t=1}^{n}\frac{e_t}{\sqrt{l(b_n)}}\sum_{j=0}^{t-1}\rho^{t-1-j}\frac{e_j}{\sqrt{l(b_n)}}\right)\right\}\{1+o_p(1)\},$$

and

$$\Delta_3 = n\sum_{t=1}^{n}y_{t-1}^2 - \left(\sum_{t=1}^{n}y_{t-1}\right)^2$$

$$=\frac{n}{1-\rho}\left\{\frac{1}{1+\rho}\left[\frac{2n\mu^2\rho}{1-\rho} + \left(n\mu^2 + \sum_{t=1}^{n}e_t^2\right) - y_n^2 - \frac{2\mu\rho}{1-\rho}y_n + \frac{4\rho\mu}{1-\rho}\sum_{t=1}^{n}e_t\right]\right.$$

$$\left.-\frac{1}{1-\rho}\left[n\mu^2 - 2\mu y_n + 2\mu\sum_{t=1}^{n}e_t\right]\right\}\{1+o_p(1)\}$$

$$=\frac{n}{1-\rho}\left\{\frac{1}{1+\rho}\left[\sum_{t=1}^{n}e_t^2 - y_n^2 - \frac{2\mu\rho}{1-\rho}y_n\right] + \frac{2\mu y_n}{1-\rho}\right\}\{1+o_p(1)\}$$

$$=\left\{\frac{n^{2+\alpha}l(b_n)}{-2c}\left(\frac{1}{n}\sum_{t=1}^{n}\frac{e_t^2}{l(b_n)}\right) + n^{1+3\alpha}\frac{\mu^2}{-2c^3}\right\}\{1+o_p(1)\}.$$

These, together with Lemma 4, lead directly to (i).

(ii) Similarly, we have under P6

$$y_n = \frac{\mu}{c}n^\alpha\rho^n + n^{\alpha/2}\rho^n\sqrt{l(b_n)}V_2\{1+o_p(1)\},$$

as $n\to\infty$, which implies that

$$y_n^2 = \frac{\mu^2}{c^2}n^{2\alpha}\rho^{2n} + 2\frac{\mu}{c}n^{3\alpha/2}\rho^{2n}\sqrt{l(b_n)}V_2\{1+o_p(1)\}$$

and

$$\sum_{t=1}^{n} y_{t-1} = \frac{1}{c} n^{\alpha} y_n - \frac{1}{c} n^{\alpha} y_0 - \frac{\mu}{c} n^{\alpha+1} - \frac{1}{c} n^{\alpha} \sum_{t=1}^{n} e_t$$
$$= \frac{\mu}{c^2} n^{2\alpha} \rho^n + \frac{1}{c} n^{3\alpha/2} \rho^n \sqrt{l(b_n)} V_2 \{1 + o_p(1)\}.$$

Using a similar technique, we obtain, as $n \to \infty$,

$$\sum_{t=1}^{n} y_{t-1} e_t = \sum_{t=1}^{n} \left(-\frac{\mu}{c} n^{\alpha} + \frac{\mu}{c} n^{\alpha} \rho^{t-1} + y_0 \rho^{t-1} + \sum_{i=1}^{t-1} \rho^{t-1-i} e_i \right) e_t$$
$$= -\frac{\mu}{c} n^{\alpha} \sum_{t=1}^{n} e_t + \frac{\mu}{c} n^{\alpha} \sum_{t=1}^{n} \rho^{t-1} e_t + y_0 \sum_{t=1}^{n} \rho^{t-1} e_t + \sum_{t=1}^{n} \left(\sum_{i=1}^{t-1} \rho^{t-1-i} e_i \right) e_t$$
$$= \frac{\mu}{c} n^{3\alpha/2} \rho^n \sqrt{l(b_n)} V_3 \{1 + o_p(1)\},$$

and

$$\sum_{t=1}^{n} y_{t-1}^2 = \frac{\mu^2}{2c^3} n^{3\alpha} \rho^{2n} + \frac{\mu}{c^2} n^{5\alpha/2} \rho^{2n} \sqrt{l(b_n)} V_2 \{1 + o_p(1)\}.$$

Then as $n \to \infty$, we have

$$\Delta_1 = \frac{\mu^2}{2c^3} n^{3\alpha+1/2} \rho^{2n} \sqrt{l(b_n)} V_1 \{1 + o_p(1)\},$$
$$\Delta_2 = \frac{\mu}{c} n^{3\alpha/2+1} \rho^n \sqrt{l(b_n)} V_3 \{1 + o_p(1)\},$$
$$\Delta_3 = \frac{\mu^2}{2c^3} n^{3\alpha+1} \rho^{2n} \{1 + o_p(1)\}$$

Therefore, the result holds.

4 Concluding remarks

In this paper, we investigated the limit distribution of the least squares estimator of (μ, ρ) for the first-order autoregressive model with $\mu \neq 0$. The discussions were taken under the assumption that the error variance may be infinite. The existing results fail to hold under this assumption. Our results show that the possible infinite variance affects the convergence rate of the estimator of the intercept in all cases, but only in some cases for the correlation coefficient; see Sections 3.3 and 3.4 for details. Based on the current results, one could build some testing procedures, e.g., t-statistics. However, their limit distributions may be quite complex because the least squares estimator has a different limit distribution in different cases, and even is degenerated in moderate deviations from a unit root case. Hence, it is interesting to construct some uniform statistical inferential procedures, e.g., confidence region for $(\mu, \rho)^{\top}$, which are robust to all cases above. Nevertheless, this topic is beyond the scope of the current paper and will be pursued in the future.

References

1. Andrews, D.W.K. and Guggenberger, P. (2009). Hybrid and size-corrected subsampling methods. *Econometrica* **77**, 721–762.
2. Andrews, D.W.K. and Guggenberger, P. (2014). A conditional-heteroskedasticity-robust confidence interval for the autoregressive parameter. *The Review of Economics and Statistics* **96**, 376–381.
3. Anderson, T.W. (1959). On asymptotic distributions of estimates of parameters of stochastic difference equations. *The Annals of Mathematical Statistics 30, 676–687.*
4. Chan, N.H., Li, D. and Peng, L. (2012). Toward a unified interval estimation of autoregressions. *Econometric Theory* **28**, 705–717.
5. Cavaliere, G. and Taylor, A.M.R. (2009). Heteroskedastic time series with a unit root. *Econometric Theory* 25: 1228–1276.
6. Chan, N.H. and Wei, C.Z. (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *The Annals of Statistics 15, 1050–1063.*
7. Csörgő, M., Szyszkowicz, B., and Wang, Q.Y., (2003). Donsker's theorem for self-normalized partial sums processes. *The Annals of Probability 31, 1228–1240.*
8. Dickey, D.A. and Fuller, W.A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* **49**, 1057–1072.
9. Dios-Palomares, R. and Roldan, J.A. (2006). A strategy for testing the unit root in AR(1) model with intercept: a Monte Carlo experiment. *Journal of Statistical Planning and Inference* **136**, 2685–2705.
10. Embrechts, P., Klüppelberg, K., and Mikosch, T., (1997). Modelling Extremal Events: for Insurance and Finance. *Springer.*
11. Fei, Y., (2018). Limit theory for mildly interated process with intercept, *Economics Letters* **163** 98–101.
12. Giné, E., Götze, F., and Mason, D.M. (1997). When is the student t-statistic asymptotically standard normal? *The Annals of Probability 25, 1514–1531.*
13. Hill, J.B., Li, D. and Peng, L. (2016). Uniform Interval Estimation for an AR(1) Process with AR Errors. *Statistica Sinica* **26**, 119–136.
14. Huang, S.H., Pang, T.X., and Weng, C. (2014). Limit theory for moderate deviations from a unit root under innovations with a possibly infinite variance. *Methodology & Computing in Applied Probability 16, 187–206.*
15. Liu, X., and Peng, L., (2017). Asymptotic Theory And Uniform Confidence Region For An Autoregressive Model. *Technical report.*
16. Mikusheva, A. (2007). Uniform inference in autoregressive models. *Econometrica* 75, 1411–1452.
17. Phillips, P.C.B. (1987). Towards a unified asymptotic theory for autoregression. *Biometrika* 74, 535-547.
18. Phillips, P.C.B. (1990). Time series regression with a unit root and infinite variance errors. *Econometric Theory* 6: 44–62.
19. Phillips, P.C.B., and Magdalinos, T. (2007). Limit theory for moderate deviations from a unit root. *Journal of Econometrics 136, 115–130.*
20. So, B.S. and Shin, D.W. (1999). Cauchy estimators for autoregressive processes with applications to unit root tests and confidence intervals. *Econometric Theory* **15**, 165–176.
21. Wang, X., and Yu, J., (2015). Limit theory for an explosive autoregressive process. *Economics Letters 126, 176–180.*