Statistica Sinica **19** (2009), 925-947

INFERENCE FOR NEARLY NONSTATIONARY PROCESSES UNDER STRONG DEPENDENCE WITH INFINITE VARIANCE

Ngai Hang Chan and Rong-Mao Zhang

Chinese University of Hong Kong and Zhejiang University

Abstract: Limit distributions of the least squares estimate of the autoregressive coefficient of a nearly nonstationary autoregressive model with strong dependent and infinite variance innovations are established in this paper. It is shown that under some regularity conditions, the ordinary least squares estimator of the autoregressive parameter converges to a functional of a fractional Ornstein-Uhlenbeck stable process. This paper not only generalizes the recent results of Buchmann and Chan to models with long-memory finite variance innovations, but also demonstrates the subtlety involved in the asymptotics when jumps are present. To this end, some newly established weak convergence theory involving so-called M_1 convergence is employed to handle these subtleties. Results of this paper work toward a better understanding of inference for jump processes that are commonly encountered in finance and related fields.

Key words and phrases: Autoregressive process, infinite variance, least squares, fractional Ornstein-Uhlenbeck processes, long-range dependence, nearly nonstationary processes, stochastic integrals, unit-root problem.

1. Introduction

The asymptotic theory of autoregressive time series with roots on or near the unit circle has been actively pursued by statisticians and econometricians. As of today, a relatively complete theory has been established for inference for time series with unit root or near unit root when the variance is finite, see for example the recent survey articles of Chan (2009) and the references therein.

However, large number of empirical studies, ranging from signal processing and network traffic to insurance, indicate that time series with heavy tails provide better models for these kinds of data. For background information on heavy-tailed time series and their applications, see Finkenstädt and Rootzén (2004) for a survey of important theories and applications of extreme values in the areas of finance, insurance, the environment and telecommunications. In financial econometrics, there has also been increasing interest in modeling financial phenomena by time series driven by heavy-tailed innovation. For example, Fama

(1965) and Mandelbrot (1963, 1967) argued that distributions of commodity and stock returns are often heavy-tailed with possible infinite variance, Rachev and Mittnik (2000) considered stable paretian models in finance, Lux and Marchesi (2000) studied agent-based models with heavy tails, and Bayraktar, Horst and Sircar (2003) studied financial market model where order flows have heavy-tailed and long-memory durations.

Although relatively complete theories for the inference for finite variance, nonstationary ARMA models are readily available, see for example the recent monograph of Andersen, Davis, Kreiss and Mikosch (2009), less can be said for the infinite variance counterpart. For a unit-root autoregressive process with heavy-tailed noise, Knight (1989, 1991) considered the asymptotic distribution of M-estimation and a least absolute deviation estimate; Hasan (2001) considered a rank test for the unit-root hypothesis. For nearly non-stationary process with heavy-tailed noise, Chan (1990) and Chan, Peng and Qi (2006) considered the limit distribution of the LSE and quantile inference. Long-range dependence with finite variance noise was considered by Wu (2006). Recently, Buchmann and Chan (2007) establish the asymptotic theory of the LSE for nearly nonstationary processes when innovations are strongly dependent with finite variance.

Due to the intricacy of the asymptotic theory involved in an infinite variance model, even less is known when both long-range dependence and infinite variance structure are exhibited in the time series. On the other hand, time series with long-range dependence and infinite variance phenomenon do exist in financial data, see for example Cont and Tankov (2004), where many convincing examples are given. One of the main purposes of this paper is to establish a unified theory for nearly nonstationary AR(1) model when the noise is a strongly dependent and an infinite variance process. For more information and applications concerning strong dependent and infinite variance processes, we refer the readers to Doukhan, Oppenheim and Taqqu (2003) and Samorodnitsky and Taqqu (1994), and the references therein.

Consider a nearly nonstationary first-order autoregressive (AR(1)) model

$$Y_t = \mu_n + \beta_n Y_{t-1} + \varepsilon_t, \quad t = 1, \ldots, n, \tag{1.1}$$

where μ_n and β_n are two unknown parameters. When $\mu_n = 0$ and $\beta_n = 1$, (1.1) reduces to the traditional random walk model. When μ_n is unknown and $\beta_n = 1$, (1.1) is sometimes known as a differenced-stationary model. In this paper, the limiting behavior of the least squares estimator of β_n is studied when β_n is close to one. In particular, we show that when $\lim_{n\to} n(1-\beta_n) = \gamma$, where γ is a constant, then under some regularity conditions the limit distribution of the least squares estimator (LSE) of β_n is a functional of fractional Ornstein-Uhlenbeck (O-U) stable processes. An important reason to study processes like (1.1) is due to the non standard asymptotic behavior of the LSE between $\beta_n = 1$ and β_n close to one. As indicated in Buchmann and Chan (2007), a key consideration is the kind of approximation that should be used for statistics constructed from the LSE when (1.1) incorporates both long-range dependence and infinite variance.

Although the main result of this paper bears some formal analogy with Buchmann and Chan (2007), it offers a number of important new implications. First, it extends the recent work of Buchmann and Chan (2007) to the case when $\{\varepsilon_t\}$ is heavy-tailed, thereby extending the inference for LSE to the nearly nonstationary infinite variance model under strong dependence. This result can then be used to shed light on finite sample analysis or to conduct test based on local alternatives. Due to the existence of random jumps exhibited in infinite variance models, traditional weak convergence is no longer sufficient. To this end, we rely on a weaker convergence involving the M_1 topology (see Whitt (2002)) and the associated results of fractional O-U stable processes. Second, we show that even in the infinite variance case, the order of β_n is still n. This is somewhat intriguing as this order was originally motivated by the consideration of the magnitude of the observed Fisher's information number in the finite variance case, see Chan and Wei (1987). Third, by considering (1.1) with different rates of convergence for the drift term μ_n, we exhaust all possible scenarios for both differenced-stationary and trend-stationary models. These results reflect the subtlety between differenced stationarity and trend stationarity under strong dependence and infinite variance; this has not been dealt with previously.

The paper is organized as follows. In Section 2, we give the main results of this paper. In Section 3 some elementary lemmas are given, while Section 4 consists of proofs of the main theorems. Simulations are reported on in Section 5, and Section 6 concludes. Throughout the paper, the symbol C is used to denote an unspecified positive and finite constant, which may vary in each appearance.

2. Distributions of Least Squares Estimators

Assume that $Y_0 = 0$ and $Y_1, \ldots Y_n$ are observed. To estimate β_n, consider the statistics $\widehat{\beta}_n$ and $\widehat{\beta}_{\mu n}$ based on the least squares regression of Y_t on Y_{t-1} in (1.1), where for $\mu_n = 0$,

$$\widehat{\beta}_n = \frac{\sum_{t=1}^n Y_{t-1}Y_t}{\sum_{t=1}^n Y_{t-1}^2},$$

and for μ_n unknown,

$$\widehat{\beta}_{\mu n} = \frac{\sum_{t=1}^n Y_{t-1}Y_t - \overline{Y}\sum_{t=1}^n Y_t}{\sum_{t=1}^n Y_{t-1}^2 - n(\overline{Y})^2},$$

where $\overline{Y} = \sum_{t=1}^{n} Y_{t-1}/n$. Define

$$\widehat{\tau}_n = \Big(\sum_{t=1}^{n} Y_{t-1}^2\Big)^{1/2}\Big(\widehat{\beta}_n - \beta_n\Big) = \Big(\sum_{t=1}^{n} Y_{t-1}^2\Big)^{-1/2}\Big(\sum_{t=1}^{n} Y_{t-1}\varepsilon_t\Big),$$

$$\widehat{\rho}_n = n(\widehat{\beta}_n - \beta_n) = \Big(\frac{1}{n}\sum_{t=1}^{n} Y_{t-1}^2\Big)^{-1}\Big(\sum_{t=1}^{n} Y_{t-1}\varepsilon_t\Big),$$

$$\widehat{\tau}_{\mu n} = \Big(\sum_{t=1}^{n} Y_{t-1}^2 - n(\overline{Y})^2\Big)^{1/2}\Big(\widehat{\beta}_{\mu n} - \beta_n\Big)$$
$$= \Big(\sum_{t=1}^{n} Y_{t-1}^2 - n(\overline{Y})^2\Big)^{-1/2}\Big(\sum_{t=1}^{n} Y_{t-1}\varepsilon_t - \overline{Y}\sum_{t=1}^{n}\varepsilon_t\Big),$$

$$\widehat{\rho}_{\mu n} = n(\widehat{\beta}_{\mu n} - \beta_n) = \Big(\frac{1}{n}\sum_{t=1}^{n} Y_{t-1}^2 - (\overline{Y})^2\Big)^{-1}\Big(\sum_{t=1}^{n} Y_{t-1}\varepsilon_t - \overline{Y}\sum_{t=1}^{n}\varepsilon_t\Big).$$

By elementary decompositions (see Section 4), the limit distributions of $\widehat{\tau}_n, \widehat{\rho}_n, \widehat{\tau}_{\mu n}$ and $\widehat{\rho}_{\mu n}$ can be established by studying the limit behaviors of

$$\frac{1}{n}\sum_{i=1}^{n}\Big(\sum_{t=1}^{i}\frac{\beta_n^{i-t}\varepsilon_t}{b_n}\Big)^2, \ \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{i}\frac{\beta_n^{i-t}\varepsilon_t}{b_n} \quad \text{and} \quad \frac{1}{d_n}\sum_{i=1}^{n}\varepsilon_i^2 \tag{2.1}$$

for suitably chosen sequences $\{a_n\}$ and $\{b_n\}$. In many situations, the limit distribution of $\sum_{i=1}^{n}\varepsilon_i^2/d_n$ can be easily obtained. If we can show that the partial sum process $S_{[nt]} = b_n^{-1}\sum_{i=1}^{[nt]}\varepsilon_i$ converges weakly to some process $Z(t)$ on the space of càdlàg functions $D[0,1]$ with the Skorokhod (J_1) topology, then by the Continuous Mapping Theorem, we can deduce the limit distributions in (2.1). Unfortunately, when dependence and heavy-tailed structures are present in the innovation, it is not always true that $S_{[nt]} \overset{J_1}{\Longrightarrow} Z(t)$ in $D(0,1)$ for some $Z(t)$, see Avram and Taqqu (1992). In considering possible dependent noise sequences $\{\varepsilon_t\}$ with infinite variance, a weaker topology is required so that $S_{[nt]}$ converges weakly to $Z(t)$. It turns out that the M_1 topology satisfies this requirement, though there are few results on weak convergence with the M_1 topology. In particular, no result concerning the limit distribution of nearly nonstationary process under the M_1 topology is available. For more information about the definitions of the J_1 and M_1 topologies, see Skorokhod (1956), Avram and Taqqu (2000), and Whitt (2002). Throughout this paper, assume that $\lim_n n(1-\beta_n) = \gamma$ and define the Ornstein-Uhlenbeck (O-U) process $Z_\gamma = (Z_\gamma(t), t \in [0,1])$ driven by $Z = (Z(t))$ via

$$Z_\gamma(t) = Z(t) - \gamma\int_0^t e^{-\gamma(t-s)}Z(s)\,ds, \quad Z_\gamma(0) = 0, \quad t \in [0,1]. \tag{2.2}$$

By Samorodnitsky and Taqqu (1994), the fractional O-U stable process given by (2.2) is well-defined. Let $\xrightarrow{M_1}$ and $\Longrightarrow^{J_1}$ denote weak convergence under the M_1 and J_1 topologies, respectively, and $\xrightarrow{d}$ denote convergence in distribution. With a slight abuse of notation, let a and b denote two generic random variables whose exact meaning may be different from line to line. Our main results can be stated as follows.

Theorem 2.1. *Let Y_t follow model* (1.1) *with $\mu_n = 0$. Suppose that there exist two sequences of constant $\{b_n\}$ and $\{d_n\}$ such that*

$$\Big(\frac{1}{b_n}\sum_{i=1}^{[nt]}\varepsilon_i, \frac{1}{d_n}\sum_{i=1}^{n}\varepsilon_i^2\Big)\xrightarrow{M_1}(Z(t), Z) \quad in \quad D[0,1] \tag{2.3}$$

for some process $Z(t)$ and certain random variable Z with $\lim_{n\to\infty} d_n/b_n^2 = c \in [0,\infty)$. Then

$$\Big(\frac{n^{1/2}\widehat{\tau}_n}{b_n}, \widehat{\rho}_n\Big)\xrightarrow{d}\Big(\frac{a}{\sqrt{2b}}, \frac{a}{b}\Big), \tag{2.4}$$

where $a = 2\gamma\int_0^1 Z_\gamma^2(s)\,ds + Z_\gamma^2(1) - cZ$ and $b = 2\int_0^1 Z_\gamma^2(s)\,ds$.

Theorem 2.2. *Let Y_t follow model* (1.1) *with μ_n unknown. Under the conditions of Theorem* 2.1.

(i) *For $\gamma = 0$ ($\beta_n = 1$) and $\lim_{n\to\infty} n\mu_n/b_n = \nu \in [0,\infty)$,*

$$\Big(\frac{\sqrt{n}\widehat{\tau}_{\mu n}}{b_n}, \widehat{\rho}_{\mu n}\Big)\xrightarrow{d}\Big(\frac{a}{\sqrt{2b}}, \frac{a}{b}\Big), \tag{2.5}$$

where $a = \nu(Z(1) - 2\int_0^1 Z(t)\,dt) + Z(1)^2 - cZ - 2Z(1)\int_0^1 Z(t)\,dt$ and $b = 2(\nu^2/12 + \nu\int_0^1(2t-1)Z(t)\,dt + \int_0^1 Z^2(t)\,dt - (\int_0^1 Z(t)\,dt)^2)$.

(ii) *For $\gamma = 0$ ($\beta_n = 1$) and $\lim_{n\to\infty} n\mu_n/b_n = \infty$,*

$$\Big(\frac{n^{1/2}\widehat{\tau}_{\mu n}}{b_n}, \frac{n\mu_n}{b_n}\widehat{\rho}_{\mu n}\Big) \xrightarrow{d} \Big(\frac{a}{\sqrt{2b}}, \frac{a}{b}\Big), \tag{2.6}$$

where $a = Z(1) - 2\int_0^1 Z(t)\,dt$ and $b = 1/6$.

(iii) *For $\gamma \neq 0$ ($\beta_n \neq 1$) and $\lim_{n\to\infty} n\mu_n/b_n = \nu \in [0,\infty)$,*

$$\Big(\frac{n^{1/2}\widehat{\tau}_{\mu n}}{b_n}, \widehat{\rho}_{\mu n}\Big)\xrightarrow{d}\Big(\frac{a}{\sqrt{2b}}, \frac{a}{b}\Big), \tag{2.7}$$

where $a = 2\gamma\int_0^1(Z_\gamma(s) - \nu e^{-\gamma s}/\gamma)^2\,ds + (Z_\gamma(1) - \nu e^{-\gamma}/\gamma)^2 - Z(1)\int_0^1(Z_\gamma(t) - \nu e^{-\gamma s}/\gamma)\,ds - cZ$ and $b = 2\int_0^1(Z_\gamma(s) - \nu e^{-\gamma s}/\gamma)^2\,ds$.

(iv) *For* $\gamma \neq 0$ $(\beta_n \neq 1)$ *and* $\lim_{n\to\infty} n\mu_n/b_n = \infty$,

$$\Big(\frac{\sqrt{n}\widehat{\tau}_{\mu n}}{b_n}, \frac{n\mu_n}{b_n}\widehat{\rho}_{\mu n}\Big) \xrightarrow{d} \Big(\frac{a}{\sqrt{2b}}, \frac{a}{b}\Big), \tag{2.8}$$

where $a = 2(1-(\gamma+1)e^{-\gamma})Z(1)/\gamma^2 - 2\int_0^1 e^{-\gamma t}Z(t)\,dt$ *and* $b = 2(4e^{-\gamma} - 3e^{-2\gamma} - 1)/\gamma^2$.

In what follows, Theorems 2.1 and 2.2 are applied to the case when $\{\varepsilon_t\}$ is a heavy-tailed dependent process defined by

$$\varepsilon_t = \sum_{j=0}^{\infty} c_j \eta_{t-j}, \quad t = 1, 2, \ldots, \tag{2.9}$$

where $\eta_j, j = 0, \pm 1, \pm 2, \ldots$ are i.i.d. random variables belonging to the domain of attraction α $(DA(\alpha))$ for $\alpha \in (0, 2]$, that is, $\sum_{i=1}^{[nt]} \eta_i/a_n \xrightarrow{M_1} Z_\alpha(t)$ in $D[0,\infty)$, with $a_n = \inf\{y : P\{|\eta_i| > y\} \leq 1/n\} = n^{1/\alpha} l_0(n)$ for some slowly varying function $l_0(x)$, and where $Z_\alpha(t)$ is a stable process with index α. For the coefficients $\{c_i\}$, we require the following conditions.

H1. If $1 < \alpha \leq 2$, then $c_j = j^{-\theta} l(j)$ for some $\theta > 1/\alpha$, where $l(x)$ is a slowly varying function. If $\theta = 1$, then $\lim_{n\to\infty} (\ln n)^{2+1/\alpha+\delta} l(n) = 0$ for some $\delta > 0$.

H2. If $0 < \alpha \leq 1$, then $\sum_{j=0}^{\infty} |c_j|^{\varsigma} < \infty$ for some $\varsigma < \alpha$ with $c_i \geq 0$.

Let $b_n = a_n n^{1-\theta} l(n) = n^{1-\theta+1/\alpha} l_0(n) l(n)$, $\omega_1 = \sum_{i=0}^{\infty} c_i, \omega_2 = \sum_{i=0}^{\infty} c_i^2$, and let $Z_{\alpha\theta}(t)$ be an integrated stable process defined by

$$Z_{\alpha\theta}(t) = \int_{-\infty}^{\infty} \int_0^t (u-s)_+^{-\theta}\,du\,dZ_\alpha(s),$$

where $Z_{\alpha\theta,\gamma}, Z_{\alpha,\gamma}$ are O-U processes defined in (2.2) driven by the processes $\{Z_{\alpha\theta}(t)\}$ and $\{Z_\alpha(t)\}$, respectively.

Theorem 2.3. *If conditions H*1 *and H*2 *are satisfied, the following assertions hold.*

1. *For* $1 < \alpha \leq 2$ *and* $1/\alpha < \theta < 1$, *we have*

   (a)
$$\Big(\frac{n^{1/2}\widehat{\tau}_n}{b_n}, \widehat{\rho}_n\Big) \xrightarrow{d} \Big(\frac{a}{\sqrt{2b}}, \frac{a}{b}\Big), \tag{2.10}$$

   where $a = 2\gamma\int_0^1 Z_{\alpha\theta,\gamma}^2(s)\,ds + Z_{\alpha\theta,\gamma}^2(1)$ *and* $b = 2\int_0^1 Z_{\alpha\theta,\gamma}^2(s)\,ds$;

(b) $\gamma = 0$ ($\beta_n = 1$) *and* $\mu_n = O(n^{-\theta+1/\alpha-\delta})$ *for some* $\delta > 0$,

$$\Big(\frac{n^{1/2}\widehat{\tau}_{\mu n}}{b_n}, \widehat{\rho}_{\mu n}\Big) \xrightarrow{d} \Big(\frac{a}{\sqrt{2b}}, \frac{a}{b}\Big), \tag{2.11}$$

where $a = Z_{\alpha\theta}(1)^2 - 2Z_{\alpha\theta}(1)\int_0^1 Z_{\alpha\theta}(t)dt$ *and* $b = 2(\int_0^1 Z_{\alpha\theta}^2(t)dt - (\int_0^1 Z_{\alpha\theta}(t) dt)^2)$;

(c) $\gamma \neq 0$ ($\beta_n \neq 1$) *and* $\mu_n = O(n^{-\theta+1/\alpha-\delta})$ *for some* $\delta > 0$,

$$\Big(\frac{n^{1/2}\widehat{\tau}_{\mu n}}{b_n}, \widehat{\rho}_{\mu n}\Big) \xrightarrow{d} \Big(\frac{a}{\sqrt{2b}}, \frac{a}{b}\Big), \tag{2.12}$$

where $a = 2\gamma\int_0^1 Z_{\alpha\theta,\gamma}^2(s)\,ds + Z_{\alpha\theta,\gamma}^2(1) - Z_\alpha(1)\int_0^1 Z_{\alpha\theta,\gamma}(t)\,ds$ *and* $b = 2\int_0^1 Z_{\alpha\theta,\gamma}^2(s)\,ds$;

(d) $\gamma = 0$ ($\beta_n = 1$) *and* $\mu_n = Cn^{-\theta+1/\alpha+\delta}$ *for some* $\delta > 0$,

$$\Big(\frac{n^{1/2}\widehat{\tau}_{\mu n}}{b_n}, \frac{n\mu_n}{b_n}\widehat{\rho}_{\mu n}\Big) \xrightarrow{d} \Big(\frac{a}{\sqrt{2b}}, \frac{a}{b}\Big), \tag{2.13}$$

where $a = Z_{\alpha\theta}(1) - 2\int_0^1 Z_{\alpha\theta}(t)\,dt$ *and* $b = 1/6$;

(e) $\gamma \neq 0$ ($\beta_n \neq 1$) *and* $\mu_n = Cn^{-\theta+1/\alpha+\delta}$ *for some* $\delta > 0$,

$$\Big(\frac{n^{1/2}\widehat{\tau}_{\mu n}}{b_n^*}, \frac{n\mu_n}{b_n}\widehat{\rho}_{\mu n}\Big) \xrightarrow{d} \Big(\frac{a}{\sqrt{2b}}, \frac{a}{b}\Big), \tag{2.14}$$

where $a = 2(1-(\gamma+1)e^{-\gamma})Z_{\alpha\theta}(1)/\gamma^2 - 2\int_0^1 e^{-\gamma t} Z_{\alpha\theta}(t)\,dt$ *and* $b = 2(4e^{-\gamma} - 3e^{-2\gamma} - 1)/\gamma^2$.

2. *For* $0 < \alpha \leq 1$ *or* $1 < \alpha < 2$ *and* $\theta \geq 1$, *we have*

(a)

$$\Big(n^{1/2}\frac{\widehat{\tau}_n}{\omega_1 a_n}, \widehat{\rho}_n\Big) \xrightarrow{d} \Big(\frac{a}{\sqrt{2b}}, \frac{a}{b}\Big), \tag{2.15}$$

where $a = 2\gamma\int_0^1 Z_{\alpha,\gamma}^2(s)\,ds + Z_{\alpha,\gamma}^2(1) - \omega_2 Z_{\alpha/2}/\omega_1^2$ *and* $b = 2\int_0^1 Z_{\alpha,\gamma}^2(s)\,ds$;

(b) $\gamma = 0$ ($\beta_n = 1$) *and* $\mu_n = O(n^{-1+1/\alpha-\delta})$ *for some* $\delta > 0$,

$$\Big(\frac{n^{1/2}\widehat{\tau}_{\mu n}}{\omega_1 a_n}, \widehat{\rho}_{\mu n}\Big) \xrightarrow{d} \Big(\frac{a}{\sqrt{2b}}, \frac{a}{b}\Big), \tag{2.16}$$

where $a = Z_\alpha(1)^2 - 2Z_\alpha(1)\int_0^1 Z_\alpha(t)\,dt - \omega_2 Z_{\alpha/2}/\omega_1^2$ *and* $b = 2\int_0^1 Z_\alpha^2(t)\,dt - 2(\int_0^1 Z_\alpha(t)\,dt)^2$;

(c) $\gamma \neq 0$ ($\beta_n \neq 1$) *and* $\mu_n = O(n^{-1+1/\alpha-\delta})$ *for some* $\delta > 0$,

$$\Big(\frac{n^{1/2}\widehat{\tau}_{\mu n}}{\omega_1 a_n}, \widehat{\rho}_{\mu n}\Big) \xrightarrow{d} \Big(\frac{a}{\sqrt{2b}}, \frac{a}{b}\Big), \tag{2.17}$$

where $a = 2\gamma \int_0^1 Z^2_{\alpha,\gamma}(s)\,ds + Z^2_{\alpha,\gamma}(1) - Z_\alpha(1)\int_0^1 Z_{\alpha,\gamma}(t)\,ds - (\omega_2/\omega_1^2)Z_{\alpha/2}$, $b = 2\int_0^1 Z^2_{\alpha,\gamma}(s)\,ds$ *and* $Z_{\alpha/2}$ *is a stable random variable with index* $\alpha/2$ *that will be defined in the next section;*

(d) $\gamma = 0$ ($\beta_n = 1$) *and* $\mu_n = Cn^{-1+1/\alpha+\delta}$ *for some* $\delta > 0$,

$$\Big(\frac{n^{1/2}\widehat{\tau}_{\mu n}}{\omega_1 a_n}, \frac{n\mu_n}{\omega_1 a_n}\widehat{\rho}_{\mu n}\Big) \xrightarrow{d} \Big(\frac{a}{\sqrt{2b}}, \frac{a}{b}\Big), \tag{2.18}$$

where $a = Z_\alpha(1) - 2\int_0^1 Z_\alpha(t)\,dt$ *and* $b = 1/6$;

(e) $\gamma \neq 0$ ($\beta_n \neq 1$) *and* $\mu_n = Cn^{-1+1/\alpha+\delta}$ *for some* $\delta > 0$,

$$\Big(\frac{n^{1/2}\widehat{\tau}_{\mu n}}{\omega_1 a_n}, \frac{n\mu_n}{\omega_1 a_n}\widehat{\rho}_{\mu n}\Big) \xrightarrow{d} \Big(\frac{a}{\sqrt{2b}}, \frac{a}{b}\Big), \tag{2.19}$$

where $a = 2(1-(\gamma+1)e^{-\gamma})Z_\alpha(1)/\gamma^2 - 2\int_0^1 e^{-\gamma t}Z_\alpha(t)\,dt$ *and* $b = 2(4e^{-\gamma} - 3e^{-2\gamma} - 1)/\gamma^2$.

Remark 2.1. Although Theorems 2.2−2.3 may appear cumbersome, together they exhaust all commonly encountered scenarios. For example, consider the special case that $\mu_n = \mu$ and $\gamma = 0$ so that $Y_t = \mu + Y_{t-1} + \varepsilon_t$. Here $\{\varepsilon_t\}$ is a sequence of i.i.d. symmetric stable random variables with index $\alpha \in (0, 2]$. Then $b_n = Cn^{1/\alpha}$ for some $C > 0$ and

$$\lim_{n\to\infty} \frac{n\mu_n}{b_n} = \begin{cases} \infty \text{ if } \alpha > 1, \\ \frac{\mu}{C} \text{ if } \alpha = 1, \\ 0 \text{ if } \alpha < 1. \end{cases}$$

As a result, we need to separately consider the different parts in Theorems 2.2−2.3 to cover the different limit behaviors of the quantity $n\mu_n/b_n$ for different α's. As an illustration, consider Theorem 2.3, part 2 (d). When $\lim_{n\to\infty} n\mu_n/b_n = \infty$, by (2.18) we have

$$\rho_{\mu n} := n(\widehat{\beta}_{\mu n} - \beta_n) \xrightarrow{d} 6Z_\alpha(1) - 12\int_0^1 Z_\alpha(t)\,dt\,.$$

When $\lim_{n\to\infty} n\mu_n/b_n = 0$, by (2.16) of Theorem 2.3, part 2 (b),

$$\rho_{\mu n} \xrightarrow{d} \frac{\Big[Z_\alpha(1)^2 - Z_{\alpha/2} - 2Z_\alpha(1)\int_0^1 Z_\alpha(t)\,dt\Big]}{2\Big[\int_0^1 Z^2_\alpha(t)\,dt - (\int_0^1 Z_\alpha(t)\,dt)^2\Big]}.$$

When $\lim_{n\to\infty} n\mu_n/b_n = \mu/C$, according to Theorem 2.2 part (i) the limit distribution of $\rho_{\mu n}$ has a complex form with $\nu = \mu/C, c = 1, Z(t) = Z_\alpha(t)$ and $Z = Z_{\alpha/2}$ in (2.5) of Theorem 2.2.

Remark 2.2. When $\{\varepsilon_t\}$ is long-range dependent with $1/\alpha < \theta < 1$ and $\mu_n = \mu \neq 0$, by Theorem 2.3, we have that the limit distributions of $\widehat{\tau}_{\mu n}$ and $\widehat{\rho}_{\mu n}$ are functionals of fractional stable processes (see (2.13) and (2.14)). Consider the special case $\alpha = 2$ and $\gamma = 0$. It follows from (2.13) that the limit distribution of the LSE is a functional of a fractional Brownian motion. When $\{\varepsilon_t\}$ is a short-memory process, however, it follows from Theorem 2.3 that the limit distribution of the LSE of a unit root AR(1) with a shift is a functional of a stable process that is independent of θ (see (2.18) and (2.19)). To the best of our knowledge, there seems to be no result concerning nearly nonstationary processes with drifts, not even for the case when $\{\varepsilon_t\}$ is short-range dependent, has finite variance, and μ_n is unknown.

Remark 2.3. It is well-known that for $\{\varepsilon_t\}$ in (2.9) to be well-defined, the coefficients $\{c_i\}$ must satisfy $\sum_{i=0}^{\infty} |c_i|^\varsigma < \infty$ for some $\varsigma < \alpha$. In comparison, assumptions (H_1) and (H_2) are relatively mild. When $\alpha = 2$ and $\mu = 0$, Theorem 2.1 recovers the main result (Theorem 2.1) of Buchmann and Chan (2007). Similar results to Theorem 2.2 are obtained by Buchmann and Chan (2007) for $\alpha = 2$ and $\mu_n = 0$, see also Proposition 2.1 in their paper.

3. Supplementary Lemmas

To prove Theorems 2.1−2.3, we need the following supplementary results.

Lemma 3.1. *Let* $S_{[nt]} = \sum_{i=1}^{[nt]} \varepsilon_i$. *If*

$$\frac{S_{[nt]}}{b_n} \xrightarrow{M_1} Z(t) \text{ in } D[0,1], \tag{3.1}$$

then $Y_{[nt]} := \sum_{i=1}^{[nt]} \beta_n^{[nt]-i} \varepsilon_i / b_n \xrightarrow{M_1} Z_\gamma(t)$, *where* Z_γ *is an O-U process defined by* (2.2).

Proof. Note that $S_0 = Y_0 = 0$ and

$$\begin{aligned} Y_{[nt]} &= \sum_{i=1}^{[nt]} \beta_n^{[nt]-i} \frac{S_i - S_{i-1}}{b_n} \\ &= \frac{S_{[nt]}}{b_n} - \sum_{i=1}^{[nt]} (\beta_n^{[nt]-i} - \beta_n^{[nt]-(i-1)}) \frac{S_{i-1}}{b_n} \\ &= \frac{S_{[nt]}}{b_n} - \frac{1}{n} \sum_{i=1}^{[nt]} \beta_n^{[nt]-i} (1-\beta_n) n \frac{S_{i-1}}{b_n} \end{aligned}$$

$$= \frac{S_{[nt]}}{b_n} - \frac{\gamma}{n}\sum_{i=1}^{[nt]} e^{-\gamma(t-i/n)}\frac{S_{i-1}}{b_n} + \frac{1}{n}\sum_{i=1}^{[nt]}(\gamma e^{-\gamma(t-i/n)} - \beta_n^{[nt]-i}(1-\beta_n)n)\frac{S_{i-1}}{b_n}$$
$$= \frac{S_{[nt]}}{b_n} - \gamma\int_0^t e^{-\gamma(t-s)}\frac{S_{[ns]}}{b_n}d\,s + R_n(t). \tag{3.2}$$

Since $\sup_{0\le t\le 1}|X(t)|$ is continuous on $D[0,1]$ in the M_1 topology, by (3.1) and the Continuous Mapping Theorem in the M_1 topology, we have $\sup_{0\le t\le 1}|S_{[nt]}/b_n| = O_p(1)$. This implies

$$\sup_{0\le t\le 1}|R_n(t)| \le \sup_{0\le t\le 1}|\frac{S_{[nt]}}{b_n}|\cdot \sup_{0\le t\le 1}\frac{1}{n}\sum_{i=1}^{[nt]}|\gamma e^{-c(t-i/n)} - \beta_n^{[nt]-i}(1-\beta_n)n)|$$
$$= O_p(1)o(1) = o_p(1). \tag{3.3}$$

Let $f(X(t)) = \int_0^t X(s)e^{-\gamma(t-s)}\,ds$. By Theorem 11.5.1 of Whitt (2002), we see that f is also a continuous function in D with the M_1 topology. Therefore, by (3.1) and the Continuous Mapping Theorem, we have

$$\int_0^t e^{-\gamma(t-s)}\frac{S_{[ns]}}{b_n}ds \xrightarrow{M_1} \int_0^t e^{-\gamma(t-s)}Z(s)\,ds. \tag{3.4}$$

From (3.1), (3.2) (3.3) and (3.4), it follows that

$$Y_{[nt]} := \sum_{i=1}^{[nt]}\beta_n^{[nt]-i}\frac{\varepsilon_i}{b_n}\xrightarrow{M_1} Z_\gamma(t).$$

This completes the proof.

Let $\varepsilon_i = \sum_{j=0}^\infty c_j\eta_{i-j}$ be defined as in Theorem 2.2. For $v \le 1 < \alpha$, set

$$s(\alpha-\eta) = \Big(\sum_{i=n}^\infty |c_i|^v\Big)\Big(\sum_{i=n}^\infty |c_i|\Big)^{\alpha-\eta-v}.$$

The following lemma is a consequence of Theorem 2 of Avram and Taqqu (1992).

Lemma 3.2. *Suppose that* $\sum_{i=0}^\infty c_i < \infty$ *with* $c_i \ge 0$, *and that either* (i) $\alpha \le 1$ *or* (ii) $\alpha > 1$ *and* $\lim_{n\to\infty}(\log n)^{1+\alpha+\eta}s(\alpha-\eta) = 0$ *for some* $0 < \eta \le \alpha - 1$. *Then*

$$\sum_{i=1}^{[nt]}\frac{\varepsilon_i}{a_n}\xrightarrow{M_1}(\sum_{i=0}^\infty c_i)Z_\alpha(t) \quad in \quad D[0,1].$$

Lemma 3.3. *Under the condition of Theorem* 2.3, *we have*

(i) $1 < \alpha < 2$ *with* $1/\alpha < \theta < 1$,

$$\Big(\sum_{j=1}^{[nt]} \frac{\varepsilon_j}{a_n n^{1-\theta} l(n)}, \sum_{j=1}^{n} \frac{\varepsilon_j^2}{a_n^2}\Big) \xrightarrow{M_1} \Big(Z_{\alpha\theta}(t), (\sum_{i=0}^{\infty} c_i^2) Z_{\frac{\alpha}{2}}\Big)\,; \tag{3.5}$$

(ii) $\alpha = 2$ *with* $1/2 < \theta < 1$,

$$\Big(\sum_{j=1}^{[nt]} \frac{\varepsilon_j}{b_n}, \sum_{j=1}^{n} \frac{\varepsilon_j^2}{b_n^2}\Big) \xrightarrow{M_1} \Big(Z_{2\theta}(t), 0\Big)\,, \tag{3.6}$$

where $b_n = n^{3/2-\theta} l_0(n) l(n)$ *and* $Z_{2\theta}$ *is a fractional Brownian motion with index* θ*;*

(iii) $0 < \alpha \le 1$ *and* $\alpha \in (1, 2)$ *with* $\theta \ge 1$,

$$\Big(\sum_{j=1}^{[nt]} \frac{\varepsilon_j}{a_n}, \sum_{j=1}^{n} \frac{\varepsilon_j^2}{a_n^2}\Big) \xrightarrow{M_1} \Big((\sum_{j=0}^{\infty} c_i) Z_{\alpha}(t), (\sum_{i=0}^{\infty} c_i^2) Z_{\frac{\alpha}{2}}\Big)\,, \tag{3.7}$$

where $Z_{\alpha/2}$ *is a stable process with index* $\alpha/2$.

Proof. The weak convergence of the first component in (3.5) and (3.6) can be found in Maejima (1983), and the asymptotic distribution of the second component in (3.5) and (3.7) are given by Astrauskas (1983), see also Avram and Taqqu (2000). For the second component of (3.6), by $\theta < 1$, we have $(2-\theta)/(3-2\theta) < 1$, so $E|\varepsilon_1|^{2(2-\theta)/(3-2\theta)} < \infty$. This implies

$$E\Big|\sum_{j=1}^{n} \frac{\varepsilon_j^2}{b_n^2}\Big|^{(2-\theta)/(3-2\theta)} \to 0.$$

Hence, $\sum_{j=1}^{n} \varepsilon_j^2 / b_n^2 \xrightarrow{p} 0$. This concludes (3.6). Further, the joint convergence of the finite dimension distribution of the first component and the second component in (3.5) and (3.7) can be found in Avram and Taqqu (2000). It is therefore sufficient to show the weak convergence of the first component in (3.7).

Case one: $\theta > 1$. For $1/\theta < v \le 1$, we have

$$\begin{aligned} s(\alpha-\eta) &= \Big(\sum_{i=n}^{\infty} (i^{-\theta} l(i))^v\Big)\Big(\sum_{i=n}^{\infty} i^{-\theta} l(i)\Big)^{\alpha-\eta-v} \\ &= n^{-\theta v+1} l^v(n) (n^{-\theta+1} l(n))^{\alpha-\eta-v} \\ &= n^{1-v-(\alpha-\eta)(\theta-1)} (l(n))^{\alpha-\eta}. \end{aligned} \tag{3.8}$$

As $1/\theta < v \le 1$, $1 - v - (\alpha-\eta)(\theta-1) < 0$. Thus, $\lim_{n\to\infty} (\ln n)^{1+\alpha+\eta} s(\alpha-\eta) = 0$ and by Lemma 3.2, we have (3.7).

Case two: $\theta = 1$. By the condition $\lim_{n\to\infty}(\log n)^{2+1/\alpha+\delta} l(n) = 0$, we have

$$\sum_{i=n}^{\infty} i^{-1} l(i) = (\ln n) l(n) = 0.$$

This implies that $\sum_{i=0}^{\infty} c_i < \infty$. By taking $v=1$, we have $s(\alpha-\eta) = [(\ln n) l(n)]^{\alpha-\eta}$. Combining this with the condition $\lim_{n\to\infty}(\ln n)^{2+1/\alpha+\delta} l(n) = 0$, we have

$$\lim_{n\to\infty} (\log n)^{1+\alpha+\eta} s(\alpha-\eta) = \lim_{n\to\infty} (\ln n)^{1+2\alpha} (l(n))^{\alpha-\eta} = 0$$

for $0 < \eta < \delta/(2+\delta+1/\alpha)$. By Lemma 2.2, we also have (3.7). The proof is complete.

Lemma 3.4. *Under the conditions of Theorem* 2.3*,*

(i) *For* $1 < \alpha \le 2$ *with* $1/\alpha < \theta < 1$*,*

$$\Big(\sum_{j=1}^{[nt]} \frac{\beta_n^{[nt]-j}\varepsilon_j}{a_n n^{1-\theta} l(n)}, \sum_{j=1}^{n} \frac{\varepsilon_j^2}{a_n^2}\Big) \xrightarrow{M_1} \Big(Z_{\alpha\theta,\gamma}(t), (\sum_{i=0}^{\infty} c_i^2) Z_{\alpha/2}\Big). \tag{3.9}$$

(ii) *For* $\alpha = 2$ *with* $1/2 < \theta < 1$*,*

$$\Big(\sum_{j=1}^{[nt]} \frac{\beta_n^{[nt]-j}\varepsilon_j}{b_n}, \sum_{j=1}^{n} \frac{\varepsilon_j^2}{b_n^2}\Big) \xrightarrow{M_1} \Big(Z_{2\theta,\gamma}(t), 0\Big), \tag{3.10}$$

where $b_n, Z_{2\theta}$ *is defined as that in Lemma* 3.3*.*

(iii) *For* $0 < \alpha \le 1$ *and* $\alpha \in (1,2]$ *with* $\theta \ge 1$*,*

$$\Big(\sum_{j=1}^{[nt]} \frac{\beta_n^{[nt]-j}\varepsilon_j}{a_n}, \sum_{j=1}^{n} \frac{\varepsilon_j^2}{a_n^2}\Big) \xrightarrow{M_1} \Big((\sum_{j=0}^{\infty} c_i) Z_{\alpha,\gamma}(t), (\sum_{i=0}^{\infty} c_i^2) Z_{\alpha/2}\Big). \tag{3.11}$$

Remark 3.1. For $1 < \alpha \le 2$ with $1/\alpha < \theta < 1$, the weak convergence in Lemmas 3.3 can be replaced by the J_1 topology.

4. Proofs of Theorems

In this section, proofs of Theorems 2.1−2.3 are presented. The main idea of showing these theorems is to decompose the statistics $\widehat{\tau}_n, \widehat{\tau}_{\mu n}, \widehat{\rho}_n$ and $\widehat{\rho}_{\mu n}$.

For model (1.1) with $\mu_n = 0$, we have

$$Y_i = \beta_n^i Y_0 + \sum_{j=1}^{i} \beta_n^{i-j} \varepsilon_j\,,$$

$$\begin{aligned}\sum_{i=1}^{n} Y_{i-1}\varepsilon_i &= \frac{1}{\beta_n}\sum_{i=1}^{n}\beta_n Y_{i-1}\varepsilon_i\\ &= \frac{1}{2\beta_n}\Big(\sum_{i=1}^{n}(\beta_n Y_{i-1}+\varepsilon_i)^2-\sum_{i=1}^{n}\beta_n^2 Y_{i-1}^2-\sum_{i=1}^{n}\varepsilon_i^2\Big)\\ &= \frac{1}{2\beta_n}\Big((1-\beta_n^2)\sum_{i=1}^{n}Y_{i-1}^2+Y_n^2-\sum_{i=1}^{n}\varepsilon_i^2\Big). \qquad (4.1)\end{aligned}$$

This yields

$$\begin{aligned}\frac{n^{1/2}\widehat{\tau}_n}{b_n} &= \Big(\frac{\sqrt{n}}{b_n}\Big)\cdot\frac{(1-\beta_n^2)\sum_{i=1}^{n}Y_{i-1}^2+Y_n^2-\sum_{i=1}^{n}\varepsilon_i^2}{2\beta_n\sqrt{\sum_{i=1}^{n}Y_{i-1}^2}}\\ &= \Big(\frac{\sqrt{n}}{b_n}\Big)\cdot\frac{(2\gamma/n)\sum_{i=1}^{n}Y_{i-1}^2+Y_n^2-\sum_{i=1}^{n}\varepsilon_i^2}{2\sqrt{\sum_{i=1}^{n}Y_{i-1}^2}}+o_p(1)\\ &= \frac{(2\gamma/n)\sum_{i=1}^{n}(Y_{i-1}/b_n)^2+(Y_n/b_n)^2-\sum_{i=1}^{n}(\varepsilon_i/b_n)^2}{2\sqrt{n^{-1}\sum_{i=1}^{n}(Y_{i-1}/b_n)^2}}+o_p(1), \quad (4.2)\end{aligned}$$

$$\begin{aligned}\widehat{\rho}_n &= \frac{(1-\beta_n^2)\sum_{i=1}^{n}Y_{i-1}^2+Y_n^2-\sum_{i=1}^{n}\varepsilon_i^2}{2\beta_n\frac{1}{n}\sum_{i=1}^{n}Y_{i-1}^2}\\ &= \frac{(2\gamma/n)\sum_{i=1}^{n}(Y_{i-1}/b_n)^2+(Y_n/b_n)^2-\sum_{i=1}^{n}(\varepsilon_i/b_n)^2}{(2/n)\sum_{i=1}^{n}(Y_{i-1}/b_n)^2}+o_p(1). \qquad (4.3)\end{aligned}$$

For model (1.1) with μ_n unknown and $\beta_n \neq 1$, let $d_n = \mu_n/(1-\beta_n)$, $X_i = Y_i - d_n$. Then

$$X_i = \beta_n X_{i-1} + \varepsilon_i,\ i = 1, 2, \ldots, n,$$

$$\begin{aligned}\frac{X_{[nt]}}{b_n} &= \beta_n^{[nt]}\frac{X_0-d_n}{b_n}+\sum_{i=1}^{[nt]}\beta_n^{[nt]-i}\frac{\varepsilon_i}{b_n}\\ &= \frac{\beta_n^{[nt]}Y_0}{b_n}-\frac{\beta_n^{[nt]}n\mu_n}{b_n n(1-\beta_n)}+\sum_{i=1}^{[nt]}\frac{\beta_n^{[nt]-i}\varepsilon_i}{b_n}\\ &=: I_1(t)-I_2(t)+I_3(t). \qquad (4.4)\end{aligned}$$

Furthermore,

$$\begin{aligned}\widehat{\beta}_{\mu n}-\beta_{\mu n} &= \frac{\sum_{i=1}^{n}Y_{i-1}\varepsilon_i-(1/n)\sum_{i=1}^{n}Y_{i-1}\sum_{i=1}^{n}\varepsilon_i}{\sum_{i=1}^{n}Y_{i-1}^2-n(\overline{Y})^2}\\ &= \frac{\sum_{i=1}^{n}(Y_{i-1}-d_n)\varepsilon_i-(1/n)\sum_{i=1}^{n}(Y_{i-1}-d_n)\sum_{i=1}^{n}\varepsilon_i}{\sum_{i=1}^{n}(Y_{i-1}-d_n)^2-n(\overline{Y}-d_n)^2}\end{aligned}$$

$$= \frac{\sum_{i=1}^n X_{i-1}\varepsilon_i - (1/n)\sum_{i=1}^n X_{i-1}\sum_{i=1}^n \varepsilon_i}{\sum_{i=1}^n X_{i-1}^2 - n(\overline{X})^2}, \tag{4.5}$$

where $\overline{X} = \sum_{i=1}^n X_{i-1}/n$. Similar to the argument of $\mu_n = 0$, we have

$$\frac{n^{1/2}\widehat{\tau}_{\mu n}}{b_n} = \frac{\frac{2\gamma}{n}\sum\limits_{i=1}^n(\frac{X_{i-1}}{b_n})^2 + (\frac{X_n}{b_n})^2 - \sum\limits_{i=1}^n(\frac{\varepsilon_i}{b_n})^2 - \frac{1}{n}\sum\limits_{i=1}^n\frac{X_{i-1}}{b_n}\sum\limits_{i=1}^n\frac{\varepsilon_i}{b_n}}{2\Big[\frac{1}{n}\sum\limits_{i=1}^n(\frac{X_{i-1}}{b_n})^2 - \Big(\frac{1}{n}\sum\limits_{i=1}^n\frac{X_{i-1}}{b_n}\Big)^2\Big]^{1/2}} + o_p(1), \tag{4.6}$$

$$\widehat{\rho}_{\mu n} = \frac{\frac{2\gamma}{n}\sum\limits_{i=1}^n(\frac{X_{i-1}}{b_n})^2 + (\frac{X_n}{b_n})^2 - \sum\limits_{i=1}^n(\frac{\varepsilon_i}{b_n})^2 - \frac{1}{n}\sum_{i=1}^n\frac{X_{i-1}}{b_n}\sum\limits_{i=1}^n\frac{\varepsilon_i}{b_n}}{\frac{2}{n}\sum\limits_{i=1}^n(\frac{X_{i-1}}{b_n})^2 - \Big(\frac{1}{n}\sum\limits_{i=1}^n\frac{X_{i-1}}{b_n}\Big)^2} + o_p(1)\,. \tag{4.7}$$

For model (1.1) with μ_n unknown and $\beta_n = 1$, we have $Y_i = i\mu_n + \sum_{j=1}^i \varepsilon_j$. Then,

$$\begin{aligned}
&\sum_{i=1}^n Y_{i-1}\varepsilon_i - \frac{1}{n}\sum_{i=1}^n Y_{i-1}\sum_{i=1}^n\varepsilon_i \\
&= \mu_n\sum_{i=1}^n((i-1) - \frac{1}{n}\sum_{i=1}^n(i-1))\varepsilon_i + \sum_{i=1}^n(\sum_{j=1}^{i-1}\varepsilon_j)\varepsilon_i - \frac{1}{n}\sum_{i=1}^n(\sum_{j=1}^{i-1}\varepsilon_j)\sum_{i=1}^n\varepsilon_i \\
&= \frac{\mu_n(n-1)}{2}\sum_{i=1}^n\varepsilon_i - \frac{n\mu_n}{n}\sum_{i=1}^n\sum_{j=1}^{i-1}\varepsilon_j + \sum_{i=1}^n(\sum_{j=1}^{i-1}\varepsilon_j)\varepsilon_i - \frac{1}{n}\sum_{i=1}^n(\sum_{j=1}^{i-1}\varepsilon_j)\sum_{i=1}^n\varepsilon_i \\
&= \frac{\mu_n(n-1)}{2}\sum_{i=1}^n\varepsilon_i - \frac{n\mu_n}{n}\sum_{i=1}^n\sum_{j=1}^{i-1}\varepsilon_j + \frac{1}{2}\Big(\Big(\sum_{i=1}^n\varepsilon_i\Big)^2 - \sum_{i=1}^n\varepsilon_i^2\Big) - \frac{1}{n}\sum_{i=1}^n(\sum_{j=1}^{i-1}\varepsilon_j)\sum_{i=1}^n\varepsilon_i \\
&=: H_1 - H_2 + H_3 - H_4,
\end{aligned} \tag{4.8}$$

$$\begin{aligned}
&\sum_{i=1}^n Y_{i-1}^2 - n(\overline{Y})^2 = \sum_{i=1}^n(Y_{i-1} - \frac{1}{n}\sum_{i=1}^n Y_{i-1})^2 \\
&\quad = \sum_{i=1}^n\Big((i-1)\mu_n - \frac{1}{n}\sum_{i=1}^n(i-1)\mu_n + \sum_{j=1}^{i-1}\varepsilon_j - \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^{i-1}\varepsilon_j\Big)^2 \\
&\quad = \mu_n\sum_{i=1}^n\Big((i-1) - \frac{1}{n}\sum_{i=1}^n(i-1)\Big)^2 + 2\mu_n\sum_{i=1}^n\Big((i-1) - \frac{1}{n}\sum_{i=1}^n(i-1)\Big) \\
&\qquad \Big(\sum_{j=1}^{i-1}\varepsilon_j - \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^{i-1}\varepsilon_j\Big) + \sum_{i=1}^n\Big(\sum_{j=1}^{i-1}\varepsilon_j - \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^{i-1}\varepsilon_j\Big)^2
\end{aligned}$$

$$=: J_1 + J_2 + J_3. \tag{4.9}$$

We are now ready to prove Theorems 2.1 and 2.2.

Proof of Theorem 2.1. Since $Y_0 = 0$, by (2.3) and Lemma 3.1, we have

$$\frac{Y_{[nt]}}{b_n} \xrightarrow{M_1} Z_\gamma(t) \quad \text{in} \quad D[0,1].$$

We also have

$$\Big(\frac{Y_{[nt]}}{b_n}, \sum_{i=1}^{n} \frac{\varepsilon_i^2}{d_n}\Big) \xrightarrow{M_1} (Z(t), Z) \quad \text{in} \quad D[0,1].$$

Let $f(X(t), Z) = (\int_0^1 X(t)^2 dt, X(1)^2, Z)$. Then f is a continuous function in $D[0,1]$ under the M_1 topology. By (4.2), (4.3) and the Continuous Mapping Theorem, we have

$$\Big(\frac{n^{1/2}\widehat{\tau}_n}{b_n}, \widehat{\rho}_n\Big) \xrightarrow{d} \Big(\frac{2\gamma\int_0^1 Z_\gamma^2(s)\,ds + Z_\gamma^2(1) - cZ}{2\sqrt{\int_0^1 Z_\gamma^2(s)\,ds}}, \frac{2\gamma\int_0^1 Z_\gamma^2(s)\,ds + Z_\gamma^2(1) - cZ}{2\int_0^1 Z_\gamma^2(s)\,ds}\Big).$$

This completes the proof of Theorem 2.1.

Proof of Theorem 2.2. This theorem is proved by considering the following four cases.

Case one: $\beta_n = 1$ and $\lim_{n\to\infty} n\mu_n/b_n = \nu \in [0,\infty)$. By the Continuous Mapping Theorem under the M_1 topology, we have

$$\frac{1}{b_n^2}(H_1 - H_2 + H_3 - H_4) \xrightarrow{d} \frac{\nu}{2}\Big(Z(1) - 2\int_0^1 Z(t)dt\Big) + \frac{1}{2}(Z(1)^2 - cZ) - Z(1)\int_0^1 Z(t)dt, \tag{4.10}$$

$$\frac{1}{nb_n^2}(J_1 + J_2 + J_3) \xrightarrow{d} \frac{\nu^2}{12} + 2\nu\int_0^1 (t - \frac{1}{2})Z(t)\,dt + \int_0^1 Z(t)^2\,dt - \Big(\int_0^1 Z(t)\,dt\Big)^2. \tag{4.11}$$

By (4.10) and (4.11), we have (2.5) as desired.

Case two: $\beta_n = 1$ and $\lim_{n\to\infty} n\mu_n/b_n = \infty$. Similar to (4.10) and (4.11), we have

$$\frac{1}{nb_n\mu_n}(H_1 - H_2 + H_3 - H_4) \xrightarrow{d} \frac{1}{2}Z(1) - \int_0^1 Z(t)\,dt, \tag{4.12}$$

$$\frac{1}{n^3\mu_n^2}(J_1 + J_2 + J_3) \xrightarrow{p} \frac{1}{12}. \tag{4.13}$$

Then, (2.6) follows from (4.12) and (4.13).

Case three: $\gamma \neq 0$ ($\beta_n \neq 1$) and $\lim_{n\to\infty} n\mu_n/b_n = \nu \in [0,\infty)$. Since $Y_0 = 0$, it follows that for all $t \in [0,1]$, $I_1(t) = 0$. Furthermore, $\lim_{n\to\infty} I_2 = (\nu/\gamma)e^{-\gamma t}$. It follows from Lemma 3.1 that

$$\frac{X_{[nt]}}{b_n} = I_3(t) - I_2(t) \xrightarrow{M_1} Z_\gamma(t) - \frac{\nu}{\gamma}e^{-\gamma t} \quad \text{in} \quad D[0,1].$$

This gives

$$\Big(\frac{X_{[nt]}}{b_n}, \sum_{i=1}^n \frac{\varepsilon_i}{b_n}, \sum_{i=1}^n \frac{\varepsilon_i^2}{d_n}\Big) \xrightarrow{M_1} (Z_\gamma(t) - \frac{\nu}{\gamma}e^{-\gamma t}, Z(1), Z) \quad \text{in} \quad D[0,1]. \tag{4.14}$$

Let $f(X(t), Z(1), Z) = (\int_0^t X(t)^2\,dt, \int_0^t X(t)\,dt, X^2(1), Z(1), Z)$. Then f is continuous and by (4.6), (4.7), (4.14) and the Continuous Mapping Theorem, we have (2.7) as desired.

Case four: $\gamma \neq 0$ ($\beta_n \neq 1$) and $\lim_{n\to\infty} n\mu_n/b_n = \infty$. For all $t \in [0,1]$, we have $\lim_{n\to\infty} I_2(t) = \infty$. On the other hand, by Lemma 3.1, we know that $\sup_{0\le t\le 1} I_3(t) = O_p(1)$. The distribution of $X_{[nt]}/b_n$ is determined by $I_2(t)$. By (4.4) and (4.5), we have

$$\begin{aligned}
&\frac{1}{nb_n\mu_n}\Big(\sum_{i=1}^n Y_{i-1}\varepsilon_i - \frac{1}{n}\sum_{i=1}^n Y_{i-1}\sum_{i=1}^n \varepsilon_i\Big)\\
&= \frac{1}{nb_n\mu_n}\Big(\sum_{i=1}^n X_{i-1}\varepsilon_i - \frac{1}{n}\sum_{i=1}^n X_{i-1}\sum_{i=1}^n \varepsilon_i\Big)\\
&= \frac{1}{nb_n\mu_n}\Big[\mu_n \sum_{i=1}^n \Big(\frac{-\beta_n^{i-1}}{1-\beta_n} + \frac{1}{n}\sum_{i=1}^n \frac{\beta_n^{i-1}}{1-\beta_n}\Big)\varepsilon_i\Big] + o_p(1)\\
&= \frac{1}{\gamma}\sum_{i=1}^n \Big(-\beta_n^{i-1} + \frac{1}{n}\sum_{i=1}^n \beta_n^{i-1}\Big)\frac{\epsilon_i}{b_n} + o_p(1)\\
&= \frac{1}{\gamma}\Big[\Big(\frac{1}{n}\sum_{i=1}^n \beta_n^{i-1}\Big)\sum_{i=1}^n \frac{\epsilon_i}{b_n} - \sum_{i=1}^n \beta_n^{n-1}\frac{\varepsilon_i}{b_n} - \frac{\gamma}{n}\sum_{i=1}^n \beta_n^{i-2}\sum_{j=1}^{i-1}\frac{\varepsilon_j}{b_n}\Big] + o_p(1)\\
&= \Big(\frac{1}{\gamma^2} - \frac{\gamma+1}{\gamma^2}e^{-\gamma}\Big)Z(1) - \int_0^1 e^{-\gamma t}Z(t)\,dt.
\end{aligned} \tag{4.15}$$

On the other hand,

$$\begin{aligned}
\frac{1}{n^3\mu_n^2}\Big(\sum_{i=1}^n Y_{i-1}^2 - n(\overline{Y})^2\Big) &= \frac{1}{n^3\mu_n^2}\Big(\sum_{i=1}^n X_{i-1}^2 - n(\overline{X})^2\Big)\\
&= \frac{1}{n^3\mu_n^2}\sum_{i=1}^n \Big(\frac{1}{n}\sum_{i=1}^n \frac{\mu_n\beta_n^{i-1}}{1-\beta_n} - \frac{\mu_n\beta_n^{i-1}}{1-\beta_n}\Big)^2 + o_p(1)
\end{aligned}$$

$$= \frac{1}{\gamma^2}\Big[\frac{1}{n}\sum_{i=1}^{n} e^{-2\gamma(i-1)/n} - \Big(\frac{1}{n}\sum_{i=1}^{n} e^{-\gamma(i-1)/n}\Big)^2\Big] + o_p(1)$$
$$= \frac{1}{\gamma^2}\Big(\int_0^1 e^{-2\gamma t}\,dt - (\int_0^1 e^{-2\gamma t}dt)^2\Big) + o_p(1)$$
$$= \frac{1}{\gamma^2}\Big(4e^{-\gamma} - 3e^{-2\gamma} - 1\Big) + o_p(1). \tag{4.16}$$

By (4.15) and (4.16), we obtain (2.8). This completes the proof of Theorem 2.2.

Proof of Theorem 2.3. By Theorems 2.1, 2.2 and Lemma 3.3, we have the conclusion of Theorem 2.3.

5. Simulations

In this section, we apply Theorems 2.2−2.3 to calculate the empirical percentiles of the the least squares statistics $\widehat{\rho}_n$, $\sqrt{n}\widehat{\tau}_n/b_n$ and $\widehat{\rho}_{\mu n}$, $\sqrt{n}\widehat{\tau}_{\mu n}/b_n$ of model (1.1). For simplicity, they are represented as ρ_n, τ_n and $\rho_{\mu n}, \tau_{\mu n}$ in the tables. By means of these percentiles, we can conduct inference and testing for model (1.1) under various scenarios. Furthermore, to acquire better understanding of the limiting behaviors of the results, we also plot the simulated probability density functions for several cases. Although it is anticipated that direct simulations are feasible once the limiting forms are established, we have to overcome the difficulty of simulating a long-memory stable process. In particular, the error process $\{\varepsilon_t\}$ in the simulations is drawn from $\varepsilon_t = \sum_{i=1}^{\infty} c_i\eta_{t-i}$, where $\{\eta_t, t = \cdots, -1, 0, 1, \cdots\}$ are i.i.d. stable variables with index α and $c_i = i^{-\theta}$. Note that each ε_t involves an infinite number of terms. It is therefore difficult to directly generate ε_t. Instead, the following algorithm due to Wu, Michailidis and Zhang (2004) is adopted to approximate ε. Define

$$\varepsilon_i' = c_i\eta_0 + c_{i+1}\eta_{-1} + \cdots + c_{m-i}\eta_{-(m-2i)} + c_1\eta_{-(m-2i)-1} + \cdots + c_{i-1}\eta_{-m} + R_{m,i}, \quad i \geq 1, \tag{5.1}$$

where

$$R_{m,i} = \begin{cases} [(m-n)^{1/\alpha-\theta}][(\alpha\theta-1)^{-1/\alpha}]\eta_i & \text{if } \theta < 1, \\ 0 & \text{if } \theta > 1, \end{cases}$$

and n is the sample size. When m is large enough, then ε_i' approximates ε_i reasonably well, see Wu, Michailidis and Zhang (2004).

Let $b_n = n^{1-\theta+1/\alpha}$ for the long-memory case ($\theta < 1$) and $b_n = \sum_{i=1}^{m} c_i n^{1/\alpha}$ for the short-memory case ($\theta > 1$). When $\mu_n = 0$ and $\beta_n = 1 - \gamma/n$, we approximate the limit distributions of $\rho_n := \widehat{\rho}_n = n(\widehat{\beta}_n - \beta_n)$ and $\tau_n := \sqrt{n}\widehat{\tau}_n/b_n =$

$\sqrt{n}(\sum_{i=1}^n Y_{i-1}^2)^{1/2}(\widehat{\beta}_n - \beta_n)/b_n$, respectively, by

$$G_1(\gamma) = n\frac{\left\{\sum_{i=1}^n \sum_{k=1}^i e^{-\gamma(i-k)/n}\varepsilon_k\varepsilon_{i+1}\right\}}{\left\{\sum_{i=1}^n \left[\sum_{k=1}^i e^{-\gamma(i-k)/n}\varepsilon_k\right]^2\right\}}, \tag{5.2}$$

$$G_2(\gamma) = \Big(\frac{\sqrt{n}}{b_n}\Big)\frac{\left\{\sum_{i=1}^n \sum_{k=1}^i e^{-\gamma(i-k)/n}\varepsilon_k\varepsilon_{i+1}\right\}}{\sqrt{\sum_{i=1}^n \left[\sum_{k=1}^i e^{-\gamma(i-k)/n}\varepsilon_k\right]^2}}. \tag{5.3}$$

When μ_n is unknown and $\beta_n = 1 - \gamma/n$ with $\gamma \neq 0$, we approximate the limit distribution of $\rho_{\mu n} := \widehat{\rho}_{\mu n} = n(\widehat{\beta}_{\mu n} - \beta_{\mu n})$ and $\tau_{\mu n} := \sqrt{n}\widehat{\tau}_{\mu n}/b_n = \sqrt{n}(\sum_{i=1}^n Y_{i-1}^2 - n(\overline{Y})^2)^{1/2}(\widehat{\beta}_{\mu,n} - \beta_n)/b_n$, respectively, by

$$G_3(\gamma)=n\frac{\sum\limits_{i=1}^n e^{-\frac{\gamma i}{n}}\Big(\sum\limits_{k=1}^i e^{\frac{\gamma k}{n}}\varepsilon_k - \frac{n\mu}{\gamma}\Big)\varepsilon_{i+1} - \frac{1}{n}\Big[\sum\limits_{i=1}^n e^{-\frac{\gamma i}{n}}\Big(\sum\limits_{k=1}^i e^{\frac{\gamma k}{n}}\varepsilon_k - \frac{n\mu}{\gamma}\Big)\Big]\sum\limits_{i=1}^n \varepsilon_i}{\sum\limits_{i=1}^n \Big[e^{-\frac{\gamma i}{n}}\Big(\sum_{k=1}^i e^{\frac{\gamma k}{n}}\varepsilon_k - \frac{n\mu}{\gamma}\Big)\Big]^2 - \frac{1}{n}\Big[\sum\limits_{i=1}^n e^{-\frac{\gamma i}{n}}\Big(\sum\limits_{k=1}^i e^{\frac{\gamma k}{n}}\varepsilon_k - \frac{n\mu}{\gamma}\Big)\Big]^2}, \tag{5.4}$$

$$G_4(\gamma)=\frac{\sum\limits_{i=1}^n e^{-\frac{\gamma i}{n}}\Big(\sum\limits_{k=1}^i e^{\frac{\gamma k}{n}}\varepsilon_k - \frac{n\mu}{\gamma}\Big)\varepsilon_{i+1} - \frac{1}{n}\Big[\sum\limits_{i=1}^n e^{-\frac{\gamma i}{n}}\Big(\sum\limits_{k=1}^i e^{\frac{\gamma k}{n}}\varepsilon_k - \frac{n\mu}{\gamma}\Big)\Big]\sum\limits_{i=1}^n \varepsilon_i}{\Big\{\sum\limits_{i=1}^n \Big[e^{-\frac{\gamma i}{n}}\Big(\sum\limits_{k=1}^i e^{\frac{\gamma k}{n}}\varepsilon_k - \frac{n\mu}{\gamma}\Big)\Big]^2 - \frac{1}{n}\Big[\sum\limits_{i=1}^n e^{-\frac{\gamma i}{n}}\Big(\sum\limits_{k=1}^i e^{\frac{\gamma k}{n}}\varepsilon_k - \frac{n\mu}{\gamma}\Big)\Big]^2\Big\}^{1/2}} \times(\sqrt{n}b_n^{-1}). \tag{5.5}$$

When μ_n is unknown and $\beta_n = 1$, i.e. $\gamma = 0$, then we approximate the limit distributions of $\rho_{\mu n} = n(\widehat{\beta}_{\mu n} - \beta_{\mu n})$ and $\tau_{\mu n} = \sqrt{n}(\sum_{i=1}^n Y_{i-1}^2 - n(\overline{Y})^2)^{1/2}(\widehat{\beta}_{\mu,n} - \beta_n)/b_n$, respectively, by

$$G_5(\gamma) = n\frac{\sum\limits_{i=1}^n \Big(\sum\limits_{k=1}^i \varepsilon_k + i\mu\Big)\varepsilon_{i+1} - n^{-1}\Big[\sum\limits_{i=1}^n \Big(\sum\limits_{k=1}^i \varepsilon_k + i\mu\Big)\Big]\Big(\sum\limits_{i=1}^n \varepsilon_i\Big)}{\sum\limits_{i=1}^n \Big[\Big(\sum\limits_{k=1}^i \varepsilon_k + i\mu\Big)\Big]^2 - n^{-1}\Big[\sum\limits_{i=1}^n \Big(\sum\limits_{k=1}^i \varepsilon_k + i\mu\Big)\Big]^2}, \tag{5.6}$$

$$G_6(\gamma) = (\sqrt{n}b_n^{-1})\frac{\sum\limits_{i=1}^n \Big(\sum\limits_{k=1}^i \varepsilon_k + i\mu\Big)\varepsilon_{i+1} - n^{-1}\Big[\sum\limits_{i=1}^n \Big(\sum\limits_{k=1}^i \varepsilon_k + i\mu\Big)\Big]\Big[\sum\limits_{i=1}^n \varepsilon_i\Big]}{\Big\{\sum\limits_{i=1}^n \Big[\Big(\sum\limits_{k=1}^i \varepsilon_k + i\mu\Big)\Big]^2 - n^{-1}\Big[\sum\limits_{i=1}^n \Big(\sum\limits_{k=1}^i \varepsilon_k + i\mu\Big)\Big]^2\Big\}^{1/2}}. \tag{5.7}$$

Simulated percentiles in each entry of Tables 1−4 were computed using $m = 5,000,000$ in (5.1) with $n = 1,000$ and $5,000$ repetitions. Table 1 is for the case $\theta = 1.6, \mu = 0.1$ and $\gamma = 10$ for different α's and Table 2 is for the case

Table 1. Empirical percentiles of $\rho_{\mu n}, \tau_{\mu n}$ with $\theta = 1.6, \mu = 0.1$ and $\gamma = 10$.

α	Statistics	Probability of a smaller value							
		0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
2	$\rho_{\mu n}$	0.079	1.411	2.479	3.600	8.451	8.991	9.668	10.468
	$\tau_{\mu n}$	-0.047	0.212	0.442	0.701	2.711	3.037	3.337	3.725
1.5	$\rho_{\mu n}$	0.457	2.021	3.098	4.130	8.455	9.164	10.218	12.178
	$\tau_{\mu n}$	0.049	0.248	0.399	0.611	3.855	5.983	9.890	20.043
0.9	$\rho_{\mu n}$	1.133	2.734	3.870	4.953	8.443	9.662	11.809	19.676
	$\tau_{\mu n}$	0.036	0.161	0.256	0.382	11.498	25.273	53.900	154.581

Table 2. Empirical percentiles of $\rho_{\mu n}, \tau_{\mu n}$ with $\alpha = 1.5, \theta = 0.8$ and $\gamma = 0$.

μ	Statistics	Probability of a smaller value							
		0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
10	$\rho_{\mu n}$	-0.366	-0.281	-0.211	-0.152	0.216	0.352	0.578	1.369
	$\tau_{\mu n}$	-15.917	-10.048	-7.094	-4.814	4.710	6.707	8.957	15.965
5	$\rho_{\mu n}$	-0.439	-0.400	-0.349	-0.263	0.452	0.860	1.387	2.293
	$\tau_{\mu n}$	-15.625	-8.613	-6.426	-4.660	4.142	5.812	7.867	12.719
0	$\rho_{\mu n}$	-0.534	-0.475	-0.445	-0.405	1.301	2.152	3.018	4.412
	$\tau_{\mu n}$	-22.216	-10.898	-6.726	-4.311	3.579	5.798	9.903	18.165

$\alpha = 0.9$ for different μ's. Tables 3 and 4 correspond to $\alpha = 2, \theta = 0.8$ and $\alpha = 1.5, \theta = 1.6$, respectively, for different γ's. By examining the values in Table 1, we observe that the heavy-tailed effects are dominant. The distributions of $G_3(\gamma)$ and $G_4(\gamma)$ have heavier right-tails when α decreases. This phenomena was also noted in Chan (1990) for $\mu = 0$. From Table 2, we see that the bigger is μ, the heavier are the right-tails of $G_5(\gamma)$ and $G_6(\gamma)$.

To gain a further understanding of these phenomena, we plot the probability density functions of $G_3(\gamma)$ and $G_5(\gamma)$ in Figures 1−2. Figure 1 shows that the smaller is α, the heavier is the right-tail of the density $G_3(\gamma)$. Figure 2 shows that the smaller is μ, the heavier is the right-tail of the density $G_5(\gamma)$. Tables 3 and 4 reveal that when γ increases, the values of the empirical percentile of $G_1(\gamma)$ and $G_2(\gamma)$ also increase quickly. In Figure 3, observe that the pdf of $G_1(\gamma)$ shifts to the right quickly as γ increases. This is different from the i.i.d. case reported in Chan (1990). A possible explanation is that the dependence $\{\varepsilon_i\}$ plays an important role in determining the limit distribution of $\widehat{\beta}_n$. When $\{\varepsilon_i\}$ are i.i.d. random variables, there is a term $-Z_{\alpha/2}$ appearing in the limit distribution of $\widehat{\beta}_n$. But when $\{\varepsilon_i\}$ is long-range dependent, this term vanishes (cf., (2.10) and (2.15)). As a result, when γ increases, the percentiles of $G_1(\gamma)$ increase more quickly to the right than in the i.i.d. case.

Table 3. Empirical percentiles of ρ_n, τ_n with $\alpha = 2, \theta = 0.8$.

γ	Statistics	Probability of a smaller value							
		0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
0	ρ_n	-0.257	-0.135	-0.054	0.048	2.416	3.085	3.710	4.443
	τ_n	-0.172	-0.122	-0.068	0.075	7.563	8.925	10.186	12.012
10	ρ_n	8.792	9.069	9.271	9.467	11.264	11.863	12.495	13.288
	τ_n	2.469	2.919	3.286	3.753	10.096	11.540	12.828	14.181
100	ρ_n	90.906	92.376	93.487	94.581	100.947	101.671	102.207	102.798
	τ_n	6.628	7.083	7.494	8.024	12.990	14.148	15.186	16.355

Table 4. Empirical percentiles of ρ_n, τ_n with $\alpha = 0.9, \theta = 1.6$.

γ	Statistics	Probability of a smaller value							
		0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
0	ρ_n	-2.766	-1.731	-1.067	-0.566	2.148	3.243	5.081	9.448
	τ_n	-1.123	-0.614	-0.368	-0.187	5.427	12.219	26.599	73.948
10	ρ_n	1.920	4.029	5.227	6.136	9.195	10.325	12.719	18.590
	τ_n	0.155	0.263	0.380	0.537	13.078	24.933	50.490	142.704
100	ρ_n	53.471	59.874	63.226	66.265	75.228	78.484	83.665	97.763
	τ_n	0.859	1.039	1.237	1.570	29.301	62.832	129.422	361.767

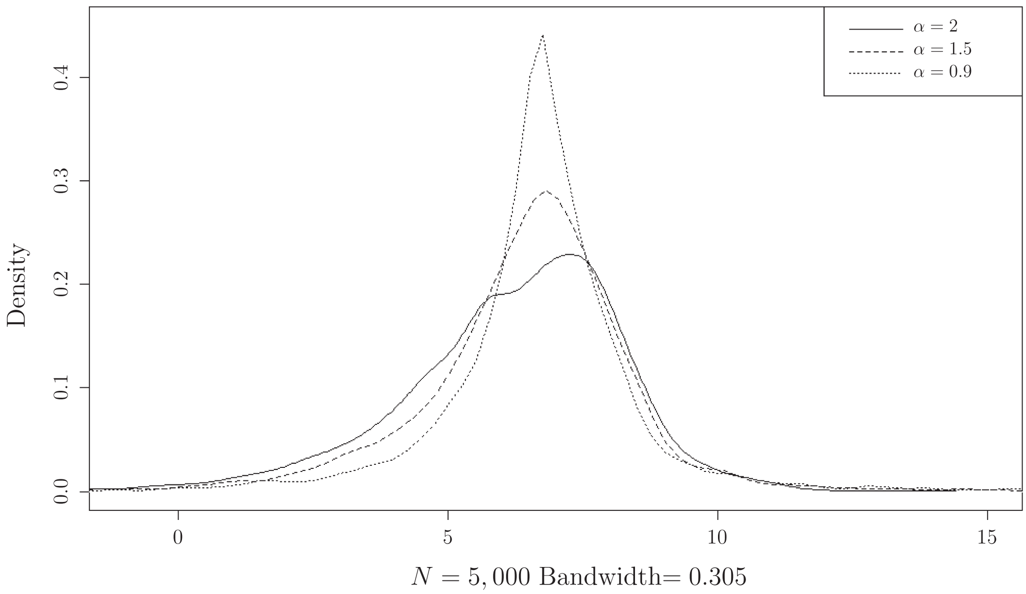

Figure 1. Probability density function of $G_3(10)$, $\theta = 1.6, \mu = 0.1$.

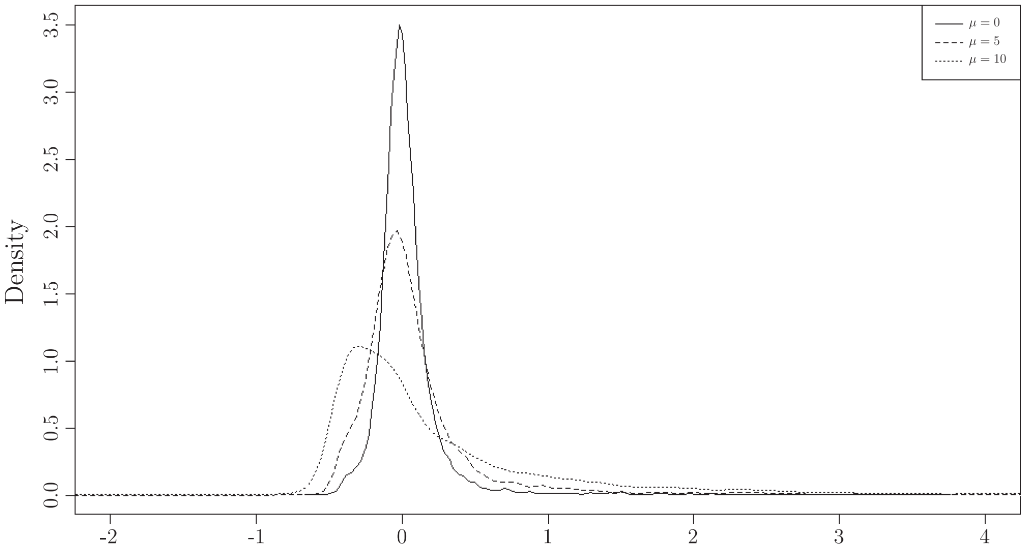

Figure 2. Probability density function of $G_5(\gamma)$, $\alpha = 1.5, \theta = 0.8$.

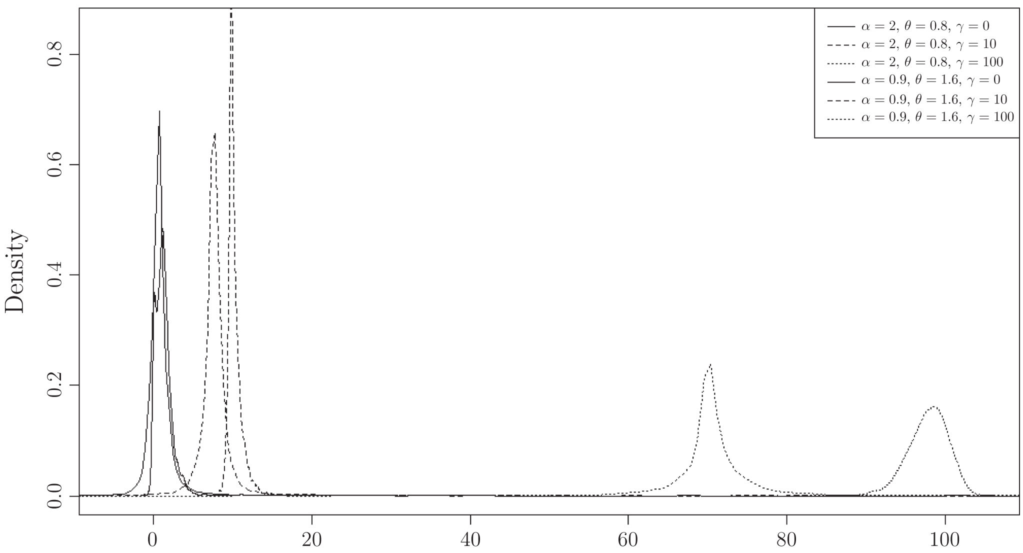

Figure 3. Probability density function of $G_1(\gamma), \alpha = 2, \theta = 0.8$.

6. Conclusion

In this paper, asymptotic distributions of the LSE of a nearly nonstationary AR(1) process with long-memory and infinite variance errors are derived. In particular, we demonstrate the effects of the limit distributions under the presence of known and unknown drifts. It should be pointed out that there exist other methods to deal with infinite variance models and/or long-memory models, notably the M-estimation approach developed by Knight (1991) for infinite variance phenomena, and the semiparametric approach of Robinson (2005) for long-memory phenomena. Although it is arguable that these approaches may offer more efficient estimation procedures than the LSE under infinite variance and/or long-memory, LSE is nevertheless one of the most commonly used procedures in practice. It is useful to study the asymptotic behavior of the LSE under the current setting before tackling the more challenging issues such as quantile inference.

Acknowledgement

We would like to thank the Editors, an associate editor and two anonymous referees for constructive and insightful comments, which lead to a substantial improvement of the paper. Research supported in part by grants from HKSAR-GRF 400306 and 400408, the National Science Foundation of China (10801118), and the Ph.D. Programs Foundation of the Ministry of Education of China (200803351094).

References

Andersen, T. G., Davis, R. A., Kreiss, J. and Mikosch, T. eds. (2009). *Handbook of Financial Time Series*. Springer, New York.

Astrauskas, A. (1983). Limit theorems for quadratic forms of linear processes. *Lithuanian. Math. J.* **23**, 355-361.

Avram, F. and Taqqu, M. S. (1992). Weak convergence of sums of moving averages in the α-stable domain of attraction. *Ann. Probab.* **20**, 483-503.

Avram, F. and Taqqu, M. S. (2000). Robustness of the R/S statistic for fractional stable noises. *Statist. Infer. Stoch. Proc.* **3**, 69-83.

Bayraktar, E., Horst, U. and Sircar, K. R. (2003). A limiting theorem for financial markets with inert investors. Working paper, Princeton University.

Buchmann, B. and Chan, N. H. (2007). Asymptotic theory of the least squares estimators for nearly unstable processes under strong dependence. *Ann. Statist.* **35**, 2001-2017.

Chan, N. H. (1990). Inference for nearly-integrated time series with infinite variance. *J. Amer. Statist. Assoc.* **85**, 1069-1074.

Chan, N. H. (2009). Time series with roots on or near the unit circle. In *Handbook of Financial Time Series*. (Edited by T. G. Andersen, R. A. Davis, J. Kreiss and T. Mikosch), 695-707. Springer, New York.

Chan, N. H., Peng, L. and Qi, Y. C (2006). Quantile inference for near-integrated autoregressive time series with infinite variance. *Statist. Sinica* **16**, 15-28.

Chan, N. H. and Wei, C. Z. (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *Ann. Statist.* **15**, 1050-1063.

Cont, R. and Tankov, P. (2004). *Financial Modeling with Jump Processes.* Chapman and Hall/CRC, New York.

Doukhan, P., Oppenheim, G. and Taqqu, M. (2003). *Theory and Applications of Long-range Dependence.* Birkhäuser, Boston.

Fama, E. (1965). The Behavior of Stock Market Price. *J. Business* **38**, 34-105.

Finkenstädt, B. and Rootzén, H. (2004). *Extreme Values in Finance, Telecommunications, and the Environment.* Chapman and Hall/CRC, New York.

Hasan, M. N. (2001). Rank tests of unit root hypothesis with infinite variance errors. *J. Econometrics* **104**, 49-65.

Knight, K. (1989). Limit theory for autoregressive parameter estimates in an infinite variance random walk. *Canad. J. Statist.* **17**, 261-278.

Knight, K. (1991). Limit theory for M-estimates in an integrated infinite variance processes. *Econometric Theory* **7**, 200-212.

Lux, T. and Marchesi, M. (2000). Volatility clustering in financial markets: a microsimulation of interesting agents. *Int. J. Theor. Appl. Finance.* **3**, 675-702

Maejima, M. (1983). On a class of self-similar processes. *Z. Wahrsch. Verw. Gebiete* **62**, 235-245.

Mandelbrot, B. B. (1963). The variation of certain speculative price. *J. Business* **36**, 394-419.

Mandelbrot, B. B. (1967). The variation of some other speculative price. *J. Business* **40**, 393-413.

Rachev, S. T. and Mittnik, S. (2000). *Stable Paretian Models in Finance.* New York, Wiley.

Robinson, P. M. (2005). Efficiency improvements in inference on stationary and nonstationary fractional time series. *Ann. Statist.* **33**, 1800-1842.

Samorodnitsky, G. and Taqqu, M. S. (1994). *Stable Non-Gaussian Random Processes: Stochastic Models With Infinite Variance.* Chapman and Hall, New York.

Skorokhod, A. V. (1956). Limit theorem for stochastic processes. *Theory Probab. Appl.* **1**, 261-290.

Whitt, W. (2002). *Stochastic-Process Limits— An Introduction to Stochastic-Process Limits and Their Application to Queues.* Springer, New York.

Wu, W. B. (2006). Unit root testing for functionals of linear processes. *Econometric Theory* **22**, 1-14.

Wu, W. B., Michailidis, G. and Zhang, D. L. (2004). Simulating sample paths of linear fractional stable motion. *IEEE Trans. Inform. Theory* **50**, 1086-1096.

Department of Statistics, Room 118, Lady Shaw Building, The Chinese University of Hong Kong Shatin, New Territories, Hong Kong.

E-mail: nhchan@sta.cuhk.edu.hk

Department of Mathematics, Zhejiang University (Yuquan Campus), Hangzhou 310027, China.

E-mail: rmzhang@zju.edu.cn

(Received August 2007; accepted July 2008)