12-2018

Mild-explosive and local-to-mild-explosive autoregressions with serially correlated errors

Yiu Lim LUI
Singapore Management University, yl.lui.2015@phdecons.smu.edu.sg

Weilin XIAO
Zhejiang University

Jun YU
Singapore Management University, yujun@smu.edu.sg

SMU ECONOMICS & STATISTICS

Mild-explosive and Local-to-mild-explosive Autoregressions with Serially Correlated Errors

Yiu Lim Lui, Weilin Xiao, Jun Yu

December 2018

Paper No. 22-2018

Mild-explosive and Local-to-mild-explosive Autoregressions with Serially Correlated Errors*

Yiu Lim Lui
Singapore Management University

Weilin Xiao
Zhejiang University

Jun Yu
Singapore Management University

December 15, 2018

Abstract

This paper firstly extends the results of Phillips and Magdalinos (2007a) by allowing for anti-persistent errors in mildly explosive autoregressive models. It is shown that the Cauchy asymptotic theory remains valid for the least squares (LS) estimator. The paper then extends the results of Phillips, Magdalinos and Giraitis (2010) by allowing for serially correlated errors of various forms in local-to-mild-explosive autoregressive models. It is shown that the result of smooth transition in the limit theory between local-to-unity and mild-explosiveness remains valid for the LS estimator. Finally, the limit theory for autoregression with intercept is developed.

JEL classification: C22
Keywords: Anti-persistent, unit root, mildly explosive, limit theory, bubble, fractional integration, Young integral

1 Introduction

The autoregressive (AR) model with a mildly explosive root was first studied in Phillips and Magdalinos (2007a) (PM_a hereafter). It allows for the development of an invariance principle for the least squares (LS) estimator of the AR parameter. The limit distribution is Cauchy which is the same as that developed in White (1958) and Anderson (1959) for the pure explosive AR models with independent and identically distributed (i.i.d.) Gaussian errors and the zero initial condition.

The AR model with a mildly explosive root considered in PM_a takes the form of

$$y_t = \left(1 + \frac{c}{n^\alpha}\right) y_{t-1} + u_t, \tag{1}$$

*Yiu Lim Lui, School of Economics, Singapore Management University, 90 Stamford Rd, Singapore 178903, Email: yl.lui.2015@smu.edu.sg. Weilin Xiao, School of Management, Zhejiang University, Hangzhou, 310058, China. Email: wlxiao@zju.edu.cn. Jun Yu, School of Economics and Lee Kong Chian School of Business, Singapore Management University, 90 Stamford Rd, Singapore 178903. Email for Jun Yu: yujun@smu.edu.sg. URL: http://www.mysmu.edu/faculty/yujun/.

where $y_0 = o_p(n^{\alpha/2})$, $u_t \overset{iid}{\sim} (0, \sigma^2)$, $t = 1, ..., n$, $c > 0$ and $\alpha \in (0, 1)$. Let $\rho_n = 1 + \frac{c}{n^\alpha}$. PM_a showed that as $n \to \infty$,

$$\frac{n^\alpha}{2c} \rho_n^n \left(\hat{\rho}_n - \rho_n\right) \Rightarrow C, \tag{2}$$

where $\hat{\rho}_n$ denotes the LS estimator of ρ_n and C is a standard Cauchy variate. The model and the asymptotic theory have been used extensively in the literature on identifying rational bubbles in asset prices; see Phillips et al. (2011), Phillips and Yu (2011), Phillips et al. (2015a, 2015b).

Considerable efforts have been made in the literature to extend the results in PM_a to dependent errors with the following a linear structure

$$u_t = \sum_{j=0}^{\infty} c_j \epsilon_{t-j}. \tag{3}$$

For example, Phillips and Magdalinos (2007b) (PM_b hereafter) considered the case where ϵ_t is an i.i.d. sequence and $\sum_{j=0}^{\infty} j\,|c_j| < \infty$ which imply a weakly dependent error process. Magdalinos (2012) assumed that ϵ_t is a martingale difference sequence (MDS) and considered a general class of weakly dependent errors with a weaker summability condition on (c_j), that is,

$$\sum_{j=0}^{\infty} |c_j| < \infty, \text{ and } \sum_{j=0}^{\infty} c_j \neq 0. \tag{4}$$

He also considered the error process where

$$c_j = L(j) j^{-1+d}, \tag{5}$$

for some memory parameter $d \in (0, 0.5)$ with $L(j)$ being a slow-varying function at infinity. If we let $\psi(k)$ be the k^{th} order autocovariance for u_t, it can be shown that u_t does not have absolutely summable autocovariances (i.e. $\sum_{k=1}^{\infty} \psi(k) = \infty$) when $d \in (0, 0.5)$. This range of d therefore corresponds to the long-range-dependent (or long-memory) behavior and covers stationary AFRIMA processes. Both PM_b and Magdalinos (2012) showed that the asymptotic result in (2) remains valid. However, it is unknown if the asymptotic result in (2) remains valid when $d \in (-0.5, 0)$, that is, the error process is anti-persistent.

It is interesting to note that the rate of convergence in (2) bridges that of the local-to-unity model and that of the pure explosive model. However, there is a discontinuity in the form of the limit distributions when the root transits from the local-to-unity to the pure explosive root. Phillips, Magdalinos and Giraitis (2010) (PMG hereafter) showed that the mildly explosive model of PM_a is strongly linked to the local-to-unity model, if one partitions the data with sample size n to m blocks containing K observations, and replacing ρ_n in model (1) with $\rho_{n,m} = 1 + \frac{cm}{n}, c > 0$, that is

$$y_t = \left(1 + \frac{cm}{n}\right) y_{t-1} + u_t, \tag{6}$$

where $u_t \overset{iid}{\sim} (0, \sigma^2)$.[1] Under a sequential-asymptotic approach by first letting $n \to \infty$ and then letting $m \to \infty$, one obtains $\frac{1}{2c}\frac{n}{m}e^{cm}(\hat{\rho}_n - \rho_{n,m}) \Rightarrow C$. By letting $n \to \infty$ with m fixed, one obtains the local-to-unity asymptotic distribution as in Chan and Wei (1988) and Phillips (1987) with $n(\hat{\rho}_n - \rho_{n,m}) \Rightarrow \int_0^1 J_{cm} dW(s) / \int_0^1 J_{cm}^2(s) ds$ where $J_{cm}(t) = \int_0^t \exp(-cm(t-s)) dW(s)$ and W is a standard Brownian motion. Therefore, a smooth transition from the local-to-unity distribution to the Cauchy distribution is achieved. However, it is unknown if this smooth transition continues to hold under serially correlated errors.

Fei (2018) considered the mildly explosive AR model with intercept and i.i.d. errors. He showed that the asymptotic distribution of the LS estimator of the intercept is Gaussian, and somewhat surprisingly, that the asymptotic distribution of the LS estimator of the AR coefficient is also Gaussian. It is unknown if the limit theory remains valid under serially correlated errors.

This paper contributes to this burgeoning literature in three aspects. First, we show that the asymptotic Cauchy theory developed in (2) remains valid when $d \in (-0.5, 0)$ in (5). Second, we show that the smooth transition result of PMG continues to hold when the error process is weakly dependent, or long-range-dependent, or anti-persistent. Third, we show that the limit theory of Fei (2018) does not necessarily hold under serially correlated errors.

Long-range dependence is widely found in economic and financial time series; see Cheung (1993) and Baillie et al. (1996). Empirical relevance of anti-persistent processes in financial time series was recently documented in Gatheral et al. (2018) and Xiao et al. (2018). Hence, it is important to generalize the results of PM_a, PMG and Fei (2018) by allowing for long-range-dependent or anti-persistent errors.

The paper is organized as follows. Section 2 presents our main assumptions and briefly reviews several forms of serially dependent error processes. Section 3 introduces the mildly explosive AR model with anti-persistent errors and derives the asymptotic theory. Section 4 develops asymptotic theory in the PMG model with serially correlated errors using a sequential-asymptotic approach. Section 5 obtains the asymptotic theory for a model with intercept. Proofs of the main results in the paper are given in the Appendix.

We use the following notations throughout the article: $\overset{p}{\to}$, $\overset{a.s.}{\to}$, $\Rightarrow$, $\overset{a}{\sim}$ and $\overset{iid}{\sim}$ denote convergence in probability, convergence almost surely, weak convergence, asymptotic equivalence, and independent and identical distribution, respectively.

2 Assumptions of Errors

As our paper aims to extend model in PM_a, PMG and Fei (2018) with a serially dependent error process, to fix ideas and facilitate discussions, we impose the following 3 distinct assumptions on $\{u_t\}_{t=1}^n$ within the linear process (LP) as in (3). These assumptions induce weak-dependent (WD), long-memory (LM) or anti-persistent (AP) property to the

[1] If $c < 0$, this model is the exactly same as the weak unit root model of Park (2003).

error process.

Assumption 1 (LP) *We assume* $\epsilon_t \overset{iid}{\sim} (0, \sigma^2)$, $c_0 = 1$, *and we impose the following three different assumptions to the coefficients* $\{c_j\}$*:*

- ***(WD)*** $\sum_{j=0}^{\infty} |c_j| < \infty$, $\sum_{j=0}^{\infty} c_j \neq 0$.
- ***(LM)*** *for* $j \geq 1, c_j = L(j)j^{-1+d}, d \in (0, 0.5)$.[2]
- ***(AP)*** *for* $j \geq 1, c_j = L(j)j^{-1+d}, d \in (-0.5, 0)$ *and* $\sum_{j=0}^{\infty} c_j = 0$.
  where $L(\cdot)$ *is a slow-varying function at infinity.*

The autocovariance function of $\{u_t\}$ is very different under these three assumptions. Under LP-WD, one can show that $\{u_t\}$ has absolutely summable autocovariances and the summation is non-zero, i.e., $\sum_{k=1}^{\infty} |\psi(k)| \in (0, \infty)$. Under LP-AP, $\{u_t\}$ also has absolutely summable autocovariances but the summation is zero, i.e., $\sum_{k=1}^{\infty} \psi(k) = 0$. Under LP-LM, the autocovariances are not absolutely summable, i.e., $\sum_{k=1}^{\infty} |\psi(k)| = \infty$, indicating a slow decaying autocovariance function, demonstrating the long-memory property of $\{u_t\}$. For an anti-persistent process, although $\sum_{j=0}^{\infty} |c_j| < \infty$ when $d \in (-0.5, 0)$, $\{u_t\}$ is not weakly dependent due to the restriction $\sum_{j=0}^{\infty} c_j = 0$, violating LP-WD. Moreover, as the autocovariance function of $\{u_t\}$ asymptotically has the same sign as d. A negative value for d implies that the jth autocovariance is negative, suggesting the anti-persistent property of $\{u_t\}$, (see Giraitis et al. (2012), Proposition 3.2.1 (3), p. 39).

Assumption LP is also general enough to include stationary ARFIMA(p, d, q) processes where $u_t = (1 - L)^{-d}\phi(L)^{-1}\theta(L)\epsilon_t = \sum_{j=0}^{\infty} c_j \epsilon_{t-j}$.With $d = 0$, a stationary ARMA(p, q) process has an absolutely summable autocovariance function. Thus, it is covered by LP-WD. With $d \in (-0.5, 0) \cup (0, 0.5)$, LP-AP or LP-LM is applicable, it can be shown that for $j \geq 1$, c_j can be asymptotically approximated by $\frac{\theta(1)}{\phi(1)\Gamma(d)} j^{-1+d}$. When $d \in (-0.5, 0)$, the stationary ARFIMA process has zero sum linear coefficients, i.e., $\sum_{j=0}^{\infty} c_j = 0$. It is well-known that u_t corresponds to a fractional Brownian motion (fBM) with Hurst parameter $H = 1/2 + d$; see Giraitis et al . (2012). An fBM with $H \in (0, 0.5)$ has a rough sample path and is anti-persistent, while $H \in (0.5, 1)$ corresponds to an fBM with a smooth sample path.

3 Mildly Explosive Model

This section studies the mildly explosive model with an anti-persistent error process. Consider the AR model given by (1) and (3) with $\{c_j\}$ satisfying LP-AP.

We are now in a position to develop the asymptotic theory for the centered LS estimator, that is

$$\hat{\rho}_n - \rho_n = \frac{\sum_{t=1}^{n} y_{t-1} u_t}{\sum_{t=1}^{n} y_{t-1}^2}, \tag{7}$$

[2]Note that LP-WD and LP-LM are the same as Assumption LP(i) and LP(ii) in Magdalinos (2012).

where $\hat{\rho}_n = (\sum_{t=1}^n y_{t-1}y_t)(\sum_{t=1}^n y_{t-1}^2)^{-1}$. Following Magdalinos (2012), we define the following two terms $Y_n(d)$ and $Z_n(d)$,

$$Y_n(d) = \frac{1}{n^{\left(\frac{1}{2}+d\right)\alpha}} \sum_{t=1}^{\tau_n(\beta)} \rho_n^{-t} u_{n+1-t}, \tag{8}$$

$$Z_n(d) = \frac{1}{n^{\left(\frac{1}{2}+d\right)\alpha}} \sum_{t=1}^{\tau_n(\beta)} \rho_n^{-t} u_t, \tag{9}$$

where $\tau_n(\beta) = \lfloor n^\beta/2 \rfloor$, and $\beta \in (\alpha, \min\{3\alpha/2, 1\})$. By construction, these two terms have the same variance. We now introduce the following lemma and theorem. The lemma obtains the asymptotic variance of $Z_n(d)$ (and thus that of $Y_n(d)$) and the joint convergence of $Y_n(d)$ and $Z_n(d)$.

Lemma 3.1 *Let $y_0 = o_p\left(n^{\left(\frac{1}{2}+d\right)\alpha} L(n^\alpha)\right)$ and $\{c_j\}$ satisfy Assumption LP-AP. As $n \to \infty$,*

1. $\frac{1}{L(n^\alpha)^2} E\left[Z_n(d)^2\right] \to \sigma^2 c^{-(1+2d)} \frac{\Gamma(d)^2}{2\cos(\pi d)}$;
2. $\frac{1}{L(n^\alpha)}[Y_n(d), Z_n(d)] \Rightarrow [Y_d, Z_d]$,

where Y_d and Z_d are independent $N(0, V_d)$ random variable with $V_d = \sigma^2 c^{-(1+2d)} \frac{\Gamma(d)^2}{2\cos(\pi d)}$.

Theorem 3.1 *Under the same set of assumptions as in Lemma 3.1, as $n \to \infty$, we have*

$$\frac{n^\alpha}{2c} \rho_n^n \left(\hat{\rho}_n - \rho_n\right) \Rightarrow C, \tag{10}$$

where C is a standard Cauchy variable.

Remark 3.1 *If we replace the i.i.d. assumption for $\{\epsilon_t\}_{t=1}^n$ in Assumption LP with the martingale difference sequence $\{\epsilon_t, \mathcal{F}_t\}_{t=1}^n$, where $\mathcal{F}_t$ is the natural filtration and $E[\epsilon_t^2|\mathcal{F}_{t-1}] = \sigma^2$, Lemma 3.1 and Theorem 3.1 remain valid.*

Remark 3.2 *Lemma 3.1 and Theorem 3.1 extend Lemma 1 (ii), Lemma 3, and Theorem 1 of Magdalinos (2012) from the case when $d \in (0, 0.5)$ to that when $d \in (-0.5, 0)$. Note that the convergence rate of the LS estimator does not depend on d.*

4 Local-to-mild-explosive Model

Now we consider the model given by (6) which is a local-to-mild-explosive model. As suggested in PMG (2010), one way of thinking of the model specification is that the total number of observations (n) is partitioned into m blocks with K samples so that $n = m \times K$. Thus, the chronological time for y_t becomes $t = \lfloor Kj \rfloor + k$, for $k \in \{1, ..., K\}$ and $j \in \{0, 1, ...m-1\}$. When $u_t \overset{iid}{\sim} (0, \sigma^2)$, it is easy to see that as $n \to \infty$ with fixed m,

this is a local-to-unity model and hence the standard local-to-unity asymptotic theory is applicable. However, if one assumes $n \to \infty$ followed by $m \to \infty$, (6) is a mildly explosive AR model as the root is in a larger neighborhood of unity compared to the local-to-unity; see Park (2003) and PMG for detailed discussions.

To be more specific, with fixed m, we have

$$n(\hat{\rho}_n - \rho_{n,m}) \Rightarrow \int_0^1 J_{cm}(s)dW(s) / \int_0^1 J_{cm}^2(s)ds.$$

With the sequential asymptotics, we have

$$\begin{aligned} \frac{1}{2c}\frac{n}{m}e^{cm}\left(\hat{\rho}_n - \rho_{n,m}\right) &\Rightarrow \frac{e^{-cm}\int_0^1 J_{cm}(s)dW(s)}{2ce^{-2cm}\int_0^1 J_{cm}^2(s)ds}, \text{as } n \to \infty \text{ with fixed } m \\ &= \frac{e^{-cm}\int_0^m \widetilde{J}_c(s)d\widetilde{W}(s)}{2ce^{-2cm}\int_0^m \widetilde{J}_c^2(s)ds} \Rightarrow C, \text{as } m \to \infty, \end{aligned} \tag{11}$$

where $\widetilde{W}(t) = \sqrt{m}W(t/m)$ and $\widetilde{J}_c(t) = \int_0^t e^{c(t-s)}d\widetilde{W}(s)$. To see the link between this sequential-asymptotic result in (11) and the asymptotic result in (10), we can replace n^α by $\frac{n}{m}$ and note that $e^{cm} = \exp\left(\frac{cm}{n}\right)^n \approx \rho_{n,m}^n$.

Before we develop our limit theory, we first review the functional central limit theorem due to Giraitis et al. (2012) which extends Donsker's theorem.

Lemma 4.1 (Corollary 4.4.1 in Giraitis et al. (2012)) *Suppose* $u_t = \sum_{j=0}^\infty c_j\epsilon_{t-j}$, *and* $\epsilon_t \overset{iid}{\sim} (0, \sigma^2)$. *Assume* $c_j \overset{a}{\sim} \gamma j^{-1+d}$ *with* γ *being a constant, as* $j \to \infty$, *and furthermore, either one of the following conditions is satisfied,*

1. $d \in (0, 0.5)$;
2. $d \in (-0.5, 0)$, $E|\epsilon_t|^p < \infty$ *with* $p > (0.5 + d)^{-1}$ *and* $\sum_{j=0}^\infty c_j = 0$.

Then we have

$$n^{-\left(\frac{1}{2}+d\right)}\sum_{t=1}^{\lfloor nr \rfloor} u_t \Rightarrow \varsigma B^H(r), \tag{12}$$

in $\mathcal{D}[0,1]$ *with the uniform metric, where* $H = \frac{1}{2} + d$, $\varsigma = \sqrt{\sigma^2\gamma^2\frac{\mathcal{B}(d,1-2d)}{d(1+2d)}}$ *with* $\mathcal{B}(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$, $B^H(r)$ *being an fBM with Hurst parameter* H.

An fBM with Hurst parameter $H \in (0,1)$ is a Gaussian process with zero mean and the following covariance,

$$E(B^H(r)B^H(s)) = \frac{1}{2}\left(|r|^{2H} + |s|^{2H} - |r-s|^{2H}\right).$$

Clearly, if $H = 1/2$, $B^H(t)$ becomes a standard Brownian motion, $W(t)$. Unlike $W(t)$, fBM is not a semi-martingale whenever $H \neq 1/2$. Therefore, we cannot interpret the

stochastic integral with respect to fBM as an Itô integral. In this paper, following El Machkouri et al. (2016) and Xiao and Yu (2018a, 2018b), we interpret the stochastic integral with respect to fBM as a Young integral when we study the asymptotic theory for the error process under LP-LM or LP-AP. This interpretation is in contrast to PMG where $\widetilde{J}_c(t) = \int_0^t e^{c(t-s)} d\widetilde{W}(s)$ is viewed as an Itô integral. Moreover, we need a different asymptotic theory to obtain the sequential limit.

We are now ready to extend the result of PMG with a serially dependent error process. We first study the error process with weak dependence, then we consider the case with long memory and anti-persistence.

4.1 Weakly dependent errors

Lemma 4.2 *In model given by (6) and (3) with $\{c_j\}$ satisfying LP-WD. Let $y_0 = o_p(n^{1/2})$ and $E|\epsilon_t|^{\beta+\varepsilon} < \infty$ for some $\beta > 2$ and $\varepsilon > 0$. As $n \to \infty$, with fixed m, we have:*

$$n\left(\hat{\rho}_{n,m} - \rho_{n,m}\right) \Rightarrow \frac{\int_0^1 J_{cm}(r) dW(r) + \frac{1}{2}\left(1 - \frac{\upsilon}{\lambda^2}\right)}{\int_0^1 \left(J_{cm}(r)\right)^2 dr},$$

where $\upsilon = \sigma^2 \sum_{j=0}^{\infty} c_j^2$ and $\lambda = \sigma \sum_{j=0}^{\infty} c_j$.

Note that the above result can be directly obtained from Theorem 1 of Phillips (1987).

Theorem 4.1 *Under the same set of assumptions as in Lemma 4.2, as $n \to \infty$ followed by $m \to \infty$, we have:*

$$\begin{aligned} \frac{1}{2c}\frac{n}{m} e^{cm}\left(\hat{\rho}_n - \rho_{n,m}\right) &\Rightarrow \frac{e^{-cm}\int_0^1 J_{cm}(s) dW(s) + e^{-cm}\frac{1}{2}\left(1 - \frac{\upsilon}{\lambda^2}\right)}{2ce^{-2cm}\int_0^1 J_{cm}^2(s) ds}, \text{ as } n \to \infty \text{ with fixed } m \\ &= \frac{e^{-cm}\int_0^m \widetilde{J}_c(s) d\widetilde{W}(s)}{2ce^{-2cm}\int_0^m \widetilde{J}_c^2(s) ds} + O_p(e^{-cm}) \Rightarrow C, \text{ as } m \to \infty. \end{aligned} \tag{13}$$

Note that the difference between (13) and (11) is the extra term $e^{-cm}\frac{1}{2}\left(1 - \frac{\upsilon}{\lambda^2}\right)$ in (13). This term vanishes when $m \to \infty$.

Remark 4.1 *The limit theory given in Theorem 4.1 is the same as that in (11). Hence a smooth transition between the local-to-unity theory and the mild-explosive theory holds under weakly dependent errors.*

4.2 Long-range-dependent errors

Lemma 4.3 *In model given by (6) and (3) with $\{c_j\}$ satisfying LP-LM. Let $y_0 = o_p\left(n^{1/2+d}\right)$. As $n \to \infty$ with fixed m, we have:*

1. $\frac{1}{n^H} y_{\lfloor nr \rfloor} \Rightarrow \varsigma J_{cm}^H(r)$;
2. $\frac{1}{n^{H+1}} \sum_{t=1}^{n} y_t \Rightarrow \varsigma \int_0^1 J_{cm}^H(r) dr$;

3. $\frac{1}{n^{2H+1}}\sum_{t=1}^{n} y_t^2 \Rightarrow \varsigma^2 \int_0^1 (J_{cm}^H(r))^2 dr$;

4. $\frac{1}{n^{2H}}\sum_{t=1}^{n} y_{t-1}u_t \Rightarrow \varsigma^2 \left(cmZ(1)\int_0^1 e^{cms} dB^H(s) + R(1) \right)$.

where

$$\begin{aligned} J_{cm}^H(r) &= \int_0^r e^{cm(r-s)} dB^H(s),\ \varsigma = \sqrt{\sigma^2 L^2(\infty)\frac{\mathcal{B}(d,1-2d)}{d(1+2d)}}, \\ L(\infty) &= \lim_{j\to\infty} L(j),\ Z(1) = \int_0^1 e^{-cms} B^H(s) ds, \\ R(1) &= \frac{1}{2}\left[B^H(1)\right]^2 - cm\int_0^1 (B^H(s))^2 ds + (cm)^2 \int_0^1 \int_0^s e^{cm(r-s)} B^H(r) B^H(s) dr ds. \end{aligned}$$

Since $B^H(s)$ is not a semimartingale, we treat $J_{cm}^H(r)$ as a Young Integral in the present paper. By the self-similarity property of fBM, we have $B^H\left(\frac{t}{m}\right) = \left(\frac{1}{m}\right)^H B^H(t)$. Let $\widetilde{B}^H(t) := m^H B^H\left(\frac{t}{m}\right)$, we now introduce the following corollary.

Corollary 4.2 *Under the same set of assumptions as in Lemma 4.3 , when $n \to \infty$ with fixed m, we have:*

1. $\frac{1}{n^{2H+1}}\sum_{t=1}^{n} y_t^2 \Rightarrow \frac{\varsigma^2}{m^{2H+1}} \int_0^m \left(\widetilde{J}_c^H(s)\right)^2 ds$;

2. $\frac{1}{n^{2H}}\sum_{t=1}^{n} y_{t-1}u_t \Rightarrow \frac{\varsigma^2}{m^{2H}} \left(c\int_0^m e^{-cr}\widetilde{B}^H(r) dr \int_0^m e^{cs} d\widetilde{B}^H(s) + \widetilde{R}(m) \right)$;

where

$$\begin{aligned} \widetilde{J}_c^H(r) &= \int_0^r e^{c(r-s)} d\widetilde{B}^H(s), \\ \widetilde{R}(m) &= \frac{1}{2}\left[\left(\widetilde{B}^H(m)\right)^2 - 2\int_0^m \left(\widetilde{B}^H(s)\right)^2 ds + c^2 \int_0^m \int_0^r e^{c(r-s)} \widetilde{B}^H(r)\widetilde{B}^H(s) dr ds \right]. \end{aligned}$$

To build the link between the mild-explosive model and the local-to-mild-explosive model, we allow $m \to \infty$ and have the following lemma.

Lemma 4.4 *In model (6), under the same set of assumptions as in Lemma 4.3 , when $m \to \infty$, we have*

1. $e^{-2cm}\int_0^m \left(\widetilde{J}_c^H(s)\right)^2 ds \Rightarrow \frac{c}{2}\left(\int_0^\infty e^{-cs}\widetilde{B}^H(s) ds\right)^2$;

2. $e^{-cm} c\int_0^m e^{-cr}\widetilde{B}^H(r) dr \int_0^m e^{cs} d\tilde{B}^H(s) \Rightarrow c\left(\int_0^\infty e^{-cs}\widetilde{B}^H(s) ds\right)\sqrt{\frac{H\Gamma(2H)}{c^{2H}}}\eta$;

3. $e^{-cm}\widetilde{R}(m) \to 0$,

where η follows a standard normal distribution.

Theorem 4.3 *Under the same set of assumptions as in Lemma 4.3, when $n \to \infty$ followed by $m \to \infty$, we have*

$$\frac{1}{2c}\frac{n}{m}e^{cm}\left(\hat{\rho}_n - \rho_{n,m}\right) \Rightarrow C\,. \tag{14}$$

Remark 4.2 *The limit theory given in Theorem 4.3 is the same as that in (11). Hence a smooth transition between the local-to-unity theory and the mild-explosive theory continues to hold under long-range-dependent errors.*

4.3 Anti-persistent errors

Lemma 4.5 *In model given by (6) and (3) with $\{c_j\}$ satisfying LP-AP. Let $y_0 = o_p\left(n^{1/2+d}\right)$ and $E|\epsilon_t|^p < \infty$ with $p > (0.5+d)^{-1}$. As $n \to \infty$ with fixed m, we have*

1. *The first three results in Lemma 4.3 remain valid.*
2. $\frac{1}{n^{2H}}\sum_{t=1}^{n} y_{t-1}u_t \overset{a}{\sim} \varsigma^2\left[cmZ(1)\int_0^1 e^{cms}dB^H(s) + R(1)\right] + \frac{n^{-2d}}{2}E(u_t^2)$.

Comparing the second result in Lemma 4.5 with the fourth result in Lemma 4.3, we can see that there is an extra term $\frac{n^{-2d}}{2}E(u_t^2)$ which is due to a strong convergence result, $\frac{1}{n}\sum_{t=1}^{n} u_t^2 \overset{a.s.}{\to} E(u_t^2)$ and $n^{-1-2d}\sum_{t=1}^{n} u_t^2 = n^{-2d}\frac{1}{n}\sum_{t=1}^{n} u_t^2$. To obtain a smooth transition between the local-to-unity theory and the mild-explosive theory, we have to strengthen the assumption and make the extra term asymptotically negligible as $m \to \infty$.

Theorem 4.4 *Under the same set of assumptions as in Lemma 4.5, if $n \to \infty$ following by $m \to \infty$ and $\frac{n^{1-2H}}{\exp(\delta m)} \to 0$ for some $\delta \in (0, c)$, then all the results in Lemma 4.4 and Theorem 4.3 continue to hold.*

Remark 4.3 *The smooth transition between the local-to-unity theory and the mild-explosive theory continues to hold under anti-persistent errors.*

5 Mild-explosive Model with Intercept

While PM_a and Magdalinos (2012) showed that the LS estimator enjoys a Cauchy limit theory under the mildly explosive model, Fei (2018) showed that when an intercept is added to a mildly explosive AR(1) model the LS estimator is asymptotically normal. Considering the following model:

$$y_t = \mu + \rho_n y_{t-1} + u_t, \;\; \rho_n = \left(1 + \frac{c}{n^\alpha}\right), \alpha \in (0,1), \mu \neq 0, \tag{15}$$

where $u_t \overset{iid}{\sim} (0, \sigma^2)$, and $y_0 = o(n^{\alpha/2})$. Under this model, Fei (2018) proved that the centered LS estimator $\hat{\mu}$ and $\hat{\rho}_n$ converge to a normal distribution: $n^{1/2}(\hat{\mu} - \mu) \Rightarrow N(0, \sigma^2)$ and $\rho_n^n(\rho_n - 1)^{-3/2}(\hat{\rho}_n - \rho_n) \Rightarrow N(0, 2\sigma^2/\mu^2)$. Now we introduce the three forms of serially correlated errors to model (15).

5.1 Weak-dependent errors

The results are summarized in the following lemma and theorem.

Lemma 5.1 *In model given by (15), (3) with $\{c_j\}$ satisfying LP-WD. Let $y_0 = o_p(n^{\alpha/2})$. As $n \to \infty$, we have*

1. $\frac{\rho_n^{-n}}{n^\alpha} y_n \xrightarrow{p} \frac{\mu}{c}$;
2. $\frac{\rho_n^{-n}}{n^{2\alpha}} \sum_{t=1}^n y_{t-1} \xrightarrow{p} \frac{\mu}{c^2}$;
3. $(\rho_n^2 - 1)\frac{\rho_n^{-2n}}{n^{2\alpha}} \sum_{t=1}^n y_{t-1}^2 \xrightarrow{p} \frac{\mu^2}{c^2}$;
4. $\frac{\rho_n^{-n}}{L(n^\alpha)n^{3\alpha/2}} \sum_{t=1}^n y_{t-1}u_t \Rightarrow \left(\frac{\mu}{c}\right) Y_0$,
   where Y_0 is a $N(0, \lambda^2/2c)$ variate and $\lambda = \sigma \sum_{j=0}^\infty c_j$.

Theorem 5.1 *Under the same set of assumptions as in Lemma 5.1, as $n \to \infty$, we have*

$$n^{1/2}(\hat{\mu} - \mu) \Rightarrow N(0, \lambda^2), \tag{16}$$

$$\frac{\rho_n^n n^{\alpha/2}}{(\rho_n^2 - 1)}(\hat{\rho}_n - \rho_n) \Rightarrow N\left(0, \frac{c\lambda^2}{2\mu^2}\right). \tag{17}$$

Remark 5.1 *Theorem 5.1 implies*

$$\rho_n^n(\rho_n - 1)^{-3/2}(\hat{\rho}_n - \rho_n) \Rightarrow N\left(0, \frac{2\lambda^2}{\mu^2}\right) \tag{18}$$

Remark 5.2 *Since we have weakly dependent errors in model (15), we need to estimate the long run variance λ^2. One can apply the Newey-West estimator of Newey and West (1987) to the LS residual $\hat{u}_t$ to obtain a consistent estimate of the long run variance, denoted by $\hat{\lambda}^2$. A feasible 100(1-a)% confidence interval can be constructed as:*

$$\hat{\rho}_n \pm Z_a \times \frac{\sqrt{2}\hat{\lambda}(\hat{\rho}_n - 1)^{3/2}}{\hat{\rho}_n^n \hat{\mu}},$$

where $Z_a = \Phi^{-1}\left(1 - \frac{a}{2}\right)$, and Φ is the CDF of the standard normal distribution.

Remark 5.3 *If $c_j = 0$ for $j \geq 1$, then $u_t = \epsilon_t$. In this case $\lambda = \sigma$, and Theorem 5.1 is the same as Theorem 2.7 in Fei (2018).*

5.2 Long-range-dependent errors

We now move on to study model (15) with a long-memory error term, and introduce the following lemma.

Lemma 5.2 *In model given by (15) and (3) with $\{c_j\}$ satisfying LP-LM. Let $y_0 = o_p\left(n^{(1/2+d)\alpha} L(n^\alpha)\right)$. As $n \to \infty$,*

1. *the first three results in Lemma 5.1 remain valid.*

2. $\frac{\rho_n^{-n}}{L(n^\alpha)n^{(3/2+d)\alpha}}\sum_{t=1}^n y_{t-1}u_t \Rightarrow \frac{\mu}{c}Y_d$,

where Y_d is defined in Lemma 3.1.

Theorem 5.2 *Under the same set of assumptions as in Lemma 5.2, as $n \to \infty$, we have*

$$n^{1/2-d}(\hat{\mu} - \mu) \Rightarrow \varsigma B^H(1), \tag{19}$$

$$\frac{\rho_n^n n^{(1/2-d)\alpha}}{L(n^\alpha)(\rho_n^2 - 1)}(\hat{\rho}_n - \rho_n) \Rightarrow N\left(0, \frac{\sigma^2 c^{1-2d}}{\mu^2}\frac{\Gamma(d)^2}{2\cos(\pi d)}\right). \tag{20}$$

where ς is defined in Lemma 4.3.

5.3 Anti-persistent errors

Theorem 5.3 *In model given by (15) and (3) with $\{c_j\}$ satisfying LP-AP. Let $y_0 = o_p\left(n^{(1/2+d)\alpha}L(n^\alpha)\right)$. As $n \to \infty$, the results in Lemma 5.2 and Theorem 5.2 continue to hold.*

Without intercept, whether the error process is i.i.d., long-memory or anti-persistent, the convergence speed of the LS estimator and its limit distribution are the same. However, when a non-zero intercept μ is added to the model, as shown in Theorem 5.2 and 5.3, the convergence rates for $\hat{\mu}$ and $\hat{\rho}_n$ and their asymptotic variances depend on d explicitly.

Note that in model (15) and under LP, there is no smooth transition between the local-to-unity asymptotics and the mild-explosive asymptotics. Suppose that in model (15), we replace ρ_n by $\rho_{n,m} = 1 + \frac{cm}{n}$. Following Wang and Yu (2015), we can write $y_n = \frac{\mu}{cm}n(\rho_{n,m}^n - 1) + \rho_{n,m}^n y_0 + \sum_{i=1}^n \rho_{n,m}^{n-i}u_i$. Under LP, one can easily show that if we let $n \to \infty$ with m being fixed, and $y_0 = o_p(n^\vartheta)$, we have

$$\begin{aligned}
\frac{1}{n}y_n &= \frac{\mu}{cm}(\rho_{n,m}^n - 1) + \frac{1}{n}n^\vartheta\left(\frac{1}{n^\vartheta}(\rho_{n,m}^n y_0 + \sum_{j=1}^n \rho_{n,m}^{n-j}u_j)\right) \\
&= \frac{\mu}{cm}(cm) + n^{\vartheta-1}\left(\frac{\exp(cm)o_p(n^\vartheta)}{n^\vartheta} + \frac{1}{n^\vartheta}\sum_{j=1}^n \rho_{n,m}^{n-j}u_j\right) + o(1) \\
&= \mu + n^{\vartheta-1}G(1) + o_p(1) = \mu + o_p(1)\,.
\end{aligned}$$

It is straightforward to see that $G(1) = \lambda J_{cm}(1)$ and $\vartheta = 1/2$ under LP-WD whereas $G(1) = \varsigma J_{cm}^H(1)$ and $\vartheta = 1/2 + d$ under LP-LM or LP-AP. As $\frac{1}{n}y_n$ converges in probability to μ, it does not depend on m and therefore we do not need the sequential asymptotic. This explains the difference between the model with intercept and that without intercept.

However, if we let μ in model (15) be $\mu_n = o_p(n^{\vartheta-1})$, the intercept is asymptotically negligible. From Lemma 1(a) in Phillips (1987), Lemma 4.3 or 4.5, we obtain $\frac{1}{n^\vartheta}y_n \Rightarrow G(1)$. And the smooth transition as in Theorem 4.1, Theorem 4.3 or Theorem 4.4 can be recovered.

Remark 5.4 *If $u_t = \epsilon_t$, then $s = \sigma$, $B^H(1) = W(1) \sim N(0,1)$ and the asymptotic theory of $\hat{\mu}$ becomes that of Fei (2018).*

Remark 5.5 *Theorem 5.2 and Theorem 5.3 imply that*

$$\frac{\rho_n^n}{L(n^\alpha)}(\rho_n - 1)^{-3/2+d}(\hat{\rho}_n - \rho_n) \Rightarrow N\left(0, \frac{\sigma^2}{\mu^2}\frac{2\Gamma(d)^2}{\cos(\pi d)}\right). \tag{21}$$

If d is known, a confidence interval for ρ_n can be constructed based on (21). If d is unknown, a two-step approach can be introduced below to construct a feasible confidence interval for ρ_n.

Remark 5.6 (Feasible Confidence Interval) *For model (6), as the LS estimator converges to a standard Cauchy variable and the convergence rate does not depend on d, one can construct a confidence interval based on Cauchy distribution (see Phillips et al. (2011) for discussion). In model (15), as d and σ appear in equation (21), constructing a confidence interval for ρ_n based on this equation is infeasible as these two parameters are unknown. In the case when $(1-L)^d u_t = \epsilon_t$, we can utilize a two-step approach to obtaining a feasible confidence interval. In the first step, we approximate the error term $\{u_t\}_{t=0}^n$ by the LS residuals $\{\hat{u}_t\}_{t=0}^n$. We can show that*

$$\begin{aligned}
\hat{u}_t - u_t &= y_t - \hat{\mu} - \hat{\rho}_n y_{t-1} - (y_t - \mu - \rho_n y_{t-1}) \\
&= (\hat{\mu} - \mu) + (\rho_n - \hat{\rho}_n)\, y_{t-1} \\
&= O_p(n^{d-1/2}) + O_p\left(\rho_n^{-(n-t)}\frac{t^\alpha}{n^{1-da}}\right) = o_p(1).
\end{aligned}$$

In the second step we can estimate d using the local Whittle (LW) method of Robinson (1994). Denote the LW estimator of d by $\hat{d}_{LW}$. For σ, since u_t is ergodic and stationary with variance $\varphi = E[u_t^2] = \sigma^2\frac{\Gamma(1-2d)}{(\Gamma(1-d))^2}$. This implies $\sigma^2 = \varphi \times \frac{(\Gamma(1-d))^2}{\Gamma(1-2d)}$. Denote $\hat{\varphi} = \frac{1}{n}\sum_{t=1}^n \hat{u}_t^2$. A natural estimator of σ is $\hat{\sigma} = \sqrt{\hat{\varphi} \times \frac{\left(\Gamma(1-\hat{d}_{LW})\right)^2}{\Gamma(1-2\hat{d}_{LW})}}$. Therefore, a feasible 100(1-a)% confidence interval of ρ_n is:

$$\hat{\rho}_n \pm Z_a \times \frac{(\hat{\rho}_n - 1)^{3/2-\hat{d}_{LW}}}{\hat{\rho}_n^n}\frac{\hat{\sigma}}{\hat{\mu}}\sqrt{\frac{2\Gamma(\hat{d}_{LW})^2}{\cos(\pi\hat{d}_{LW})}}. \tag{22}$$

6 Conclusion

In this paper, we have filled several gaps in the rapidly growing literature on explosive time series. First, we show that the limit Cauchy theory of PM_a is applicable to the mildly explosive AR(1) model with an anti-persistent error process. Hence, our result complements that of Magdalinos (2012) where it was shown that the limit Cauchy theory is applicable to the mildly explosive AR(1) model with a weakly dependent or a long-range-dependent error process. The empirical relevance of anti-persistent process was established recently in Gatheral et al. (2018) and Xiao et al. (2018).

Second, we derive the asymptotic distribution of the LS estimator under a local-to-mild-explosive set up with either weak-dependent, or long-range-dependent or anti-persistent errors. Two asymptotic schemes are considered. In the first scheme, only $n \to \infty$. Whereas in the second scheme, $n \to \infty$ is followed by $m \to \infty$. With the sequential asymptotic scheme, we have obtained the same Cauchy asymptotic distribution as in the mildly explosive model. We demonstrate a smooth transition between the asymptotics of a local-to-mild-explosive model and those of a mild-explosive model. Hence, our results extend those of PMG from i.i.d. errors to serially correlated errors.

Finally, we study the mildly explosive model with intercept. It is shown that the convergence rate of intercept depends explicitly on the memory parameter of the error process, and the AR coefficient has a asymptotic normal distribution. Finally, we discuss how to obtain a feasible confidence interval for the AR coefficient.

A Appendix

Throughout the appendix, we follow the notations of Magdalinos (2012) by letting $\kappa = 1-d$ and utilize the following lemmas.

Lemma A.1 (Lemma A.2(i) in Magdalinos (2012)) *As $n \to \infty$, we have*

$$\sup_{1 \le t \le \tau_n(\beta)} \left| \rho_n^{-t} - e^{-\frac{c}{n^\alpha} t} \right| = O\left(n^{-\frac{n^\alpha}{2}} \right).$$

Lemma A.2 (Lemma 2.3 in El Machkouri et al. (2016)) *Suppose we have the following stochastic differential equation:*

$$dX(t) = cX(t)dt + dG(t), X(0) = X_0 = 0,$$

where $G(t)$ is a Gaussian process and $c > 0$. Further assuming the following two assumptions hold for $G = (G(t), t \ge 0)$.

1. *The process G has Hölder continuous paths of order $\delta \in (0, 1]$;*
2. *For every $t \ge 0$, $E(G^2(t)) \le ct^{2\gamma}$ for some positive constants c and γ.*

Then, for every $t \ge 0$, we have

$$\frac{1}{2}X^2(t) = c\int_0^t X^2(s)ds + cZ(t)\int_0^t e^{cs}dG(s) + R(t),$$

where

$$Z(t) = \int_0^t e^{-cs}G(s)ds,$$

$$R(t) = \frac{1}{2}G^2(t) - c\int_0^t G^2(s)ds + c^2\int_0^t \int_0^s e^{-c(s-r)}G(s)G(r)drds.$$

Proof of Lemma 3.1.1

To avoid confusion, note that $Z_n(d)$ and $Y_n(d)$ now become $Z_n(\kappa)$ and $Y_n(\kappa)$. We can write the variance of $Z_n(\kappa)$ as

$$Var(Z_n(\kappa)) = \frac{1}{n^{(3-2\kappa)\alpha}} \left(\sum_{t=1}^{\tau_n(\beta)} \rho_n^{-2t} \gamma_u(0) + 2 \sum_{t=1}^{\tau_n(\beta)} \rho_n^{-2t} \sum_{h=1}^{\tau_n(\beta)-t} \rho_n^{-h} \psi(h) \right), \tag{23}$$

where $\psi(h) = \sigma^2 \sum_{j=0}^{\infty} c_j c_{j+h}$ is the auto-covariance function for the error term.

Note that we can have an asymptotic approximation for $Var(Z_n(\kappa))$. For any positive and finite integer K, equation (23) can be rewritten as a truncated version:

$$Var(Z_n(\kappa)) = \frac{1}{n^{(3-2\kappa)\alpha}} \left(\sum_{t=1}^{\tau_n(\beta)} \rho_n^{-2t} \gamma_u(0) + 2 \sum_{t=K}^{\tau_n(\beta)} \rho_n^{-2t} \sum_{h=K}^{\tau_n(\beta)-t} \rho_n^{-h} \sigma^2 \sum_{j=K}^{\tau_n(\beta)} c_j c_{j+h} \right) + o(1). \tag{24}$$

Now following Magdalinos (2012) (Equations (6), (13) and (14)), and letting $\|.\|_r$ be the L_r norm ($\|x\|_r = (E|x|^r)^{\frac{1}{r}}$), we can show

$$\left\| \frac{Z_n(\kappa)}{L(n^\alpha)} \right\|_2^2 = \frac{\sigma^2}{c} \frac{1}{\lambda_h^2} \sum_{h=K}^{\tau_n(\beta)} \rho_n^{-h} \sum_{j=K}^{\tau_n(\beta)} c_j c_{j+h} + o(1), \tag{25}$$

where $\lambda_n = n^{(1-\kappa)\alpha} L(n^\alpha)$. With $c_{j+h} = L(j+h)(j+h)^{-\kappa}$ and Lemma A.1, we can rewrite the right hand side of (25) as

$$\frac{\sigma^2}{c} \frac{1}{\lambda_n^2} \sum_{h=K}^{\tau_n(\beta)} e^{-\frac{c}{n^\alpha}h} \sum_{j=K}^{\tau_n(\beta)} L(j) j^{-\kappa} L(j+h)(j+h)^{-\kappa} + o(1). \tag{26}$$

Note that $e^{-\frac{c}{n^\alpha}h} \sum_{j=1}^{\tau_n(\beta)} L(j) j^{-\kappa} L(j+h)(j+h)^{-\kappa}$ is a decreasing function in h. We can have the following inequality

$$LB \leq \left\| \frac{Z_n(\kappa)}{L(n^\alpha)} \right\|_2^2 \leq UB,$$

where

$$LB := \frac{\sigma^2}{c} \frac{1}{\lambda_n^2} \int_K^{\tau_n(\beta)+1} e^{-\frac{c}{n^\alpha}x} \int_K^{\tau_n(\beta)+1} L(y) y^{-\kappa} L(y+x)(y+x)^{-\kappa} dy dx + o(1), \tag{27}$$

$$UB := \frac{\sigma^2}{c} \frac{1}{\lambda_n^2} \int_{K-1}^{\tau_n(\beta)} e^{-\frac{c}{n^\alpha}x} \int_{K-1}^{\tau_n(\beta)} L(y) y^{-\kappa} L(y+x)(y+x)^{-\kappa} dy dx + o(1). \tag{28}$$

We now work on the lower bound (27). Letting $m_n = \frac{c(\tau_n(\beta)+1)}{n^\alpha}$, $u = \frac{cx}{n^\alpha}$ and $z = \frac{cy}{n^\alpha}$, we obtain:

$$\frac{\sigma^2}{c} \frac{1}{\lambda_n^2} \int_K^{\tau_n(\beta)+1} e^{-\frac{c}{n^\alpha}x} \int_K^{\tau_n(\beta)+1} L(y) y^{-\kappa} L(y+x)(y+x)^{-\kappa} dy dx$$

$$
\begin{aligned}
&= \frac{\sigma^2}{c}\frac{1}{n^{2\alpha(1-\kappa)}}\int_{\frac{cK}{n^\alpha}}^{m_n} e^{-u}\left[I_{n1}(u)+I_{n2}(u)\right]\left(\frac{n^{2\alpha(1-\kappa)}}{c^{2-2\kappa}}\right)du \\
&= \sigma^2 c^{2\kappa-3}\int_{\frac{cK}{n^\alpha}}^{m_n} e^{-u}\left[I_{n1}(u)+I_{n2}(u)\right]du, \qquad (29)
\end{aligned}
$$

where $I_{n1}(u)=\int_{cK/n^\alpha}^{1} g_n(u,z)dz$, $I_{n2}(u)=\int_1^{m_n} g_n(u,z)dz$ and

$$
g_n(u,z)=L(n^\alpha)^{-2}L\left(\frac{n^\alpha}{c}z\right)L\left(\frac{n^\alpha}{c}(z+u)\right)z^{-\kappa}(u+z)^{-\kappa}.
$$

Note that we can rewrite $I_{n2}(u)$ as

$$
\begin{aligned}
I_{n2}(u) &= \int_1^{m_n} g_n(u,z)dz \\
&= \int_1^{m_n} L(n^\alpha)^{-2}L\left(\frac{n^\alpha}{c}z\right)L\left(\frac{n^\alpha}{c}(z+u)\right)z^{-\kappa}(u+z)^{-\kappa}dz \\
&= \frac{1}{L(n^\alpha)}\int_1^{m_n}\left(\frac{L\left(\frac{n^\alpha}{c}z\right)}{L(n^\alpha)}-1+1\right)L\left(\frac{n^\alpha}{c}(z+u)\right)z^{-\kappa}(z+u)^{-\kappa}dz \\
&= \frac{1}{L(n^\alpha)}\int_1^{m_n}\left(\frac{L\left(\frac{n^\alpha}{c}z\right)}{L(n^\alpha)}-1\right)L\left(\frac{n^\alpha}{c}(z+u)\right)z^{-\kappa}(z+u)^{-\kappa}dz \\
&\quad +\frac{1}{L(n^\alpha)}\int_1^{m_n} L\left(\frac{n^\alpha}{c}(z+u)\right)z^{-\kappa}(z+u)^{-\kappa}dz. \qquad (30)
\end{aligned}
$$

As $n\to\infty$, the first term in the above expression is bounded by

$$
\sup_{z\in[1,\infty)}\left|\frac{L\left(\frac{n^\alpha}{c}z\right)}{L(n^\alpha)}-1\right|\sup_{x\in[1,\infty)}\left|\frac{L\left(\frac{n^\alpha}{c}x\right)}{L(n^\alpha)}\right|\int_1^\infty z^{-2\kappa}dz=o(1), \qquad (31)
$$

uniformly in u. Therefore, as $n\to\infty$, we have

$$
\sup_{u>0}\left|I_{n2}(u)-\int_1^{m_n}\frac{L\left(\frac{n^\alpha}{c}(z+u)\right)}{L(n^\alpha)}z^{-\kappa}(z+u)^{-\kappa}dz\right|=o(1). \qquad (32)
$$

Since $L(.)$ is a slow-varying function, it can be seen that as $n\to\infty$

$$
\int_{cK/n^\alpha}^{m_n} e^{-u}I_{n2}(u)du\to\int_0^\infty e^{-u}\int_1^\infty z^{-\kappa}(u+z)^{-\kappa}dzdu=\int_1^\infty e^z z^{-\kappa}\int_z^\infty e^{-x}x^{-\kappa}dxdz. \qquad (33)
$$

Now we go back to $I_{n1}(u)$. By using the substitution $x=z+u$, we have

$$
\begin{aligned}
&\int_{cK/n^\alpha}^{m_n} e^{-u}I_{n1}(u)du \\
&= \int_{cK/n^\alpha}^{m_n} e^{-u}\int_{cK/n^\alpha}^{1} L(n^\alpha)^2 L\left(\frac{n^\alpha}{c}z\right)L\left(\frac{n^\alpha}{c}(z+u)\right)z^{-\kappa}(u+z)^{-\kappa}dzdu
\end{aligned}
$$

$$= \int_{cK/n^\alpha}^{1} e^z \frac{L\left(\frac{n^\alpha z}{c}\right)}{L(n^\alpha)} z^{-\kappa} \int_{z+cK/n^\alpha}^{z+m_n} e^{-x} \frac{L\left(\frac{n^\alpha x}{c}\right)}{L(n^\alpha)} x^{-\kappa} dxdz$$

$$\to \int_0^1 e^z z^{-\kappa} \int_z^\infty e^{-x} x^{-\kappa} dxdz. \tag{34}$$

The last result is justified by the property of slow-varying function and the dominated convergence theorem.

From (29), (33) and (34), we obtain the asymptotic expression of the lower bound (27), which is $\sigma^2 c^{2\kappa-3} \int_0^\infty e^z z^{-\kappa} \Gamma(1-\kappa, z) dz$, where $\Gamma(\kappa, z) = \int_z^\infty u^{\kappa-1} e^{-u} du$ is the upper incomplete gamma function. Since Magdalinos (2012) showed that $\int_0^\infty e^z z^{-\kappa} \Gamma(1-\kappa, z) dz = \frac{\Gamma(1-\kappa)^2}{2\cos\{\pi(1-\kappa)\}}$, we have obtained the limit of lower bound. As K is arbitrary, the result holds for $K-1$ as well. Therefore, by obtaining the limit of LB, we can also obtain the same limit of UB. Consequently, by applying the squeeze theorem, we obtain the result of the lemma.

Proof of Lemma 3.1.2

Following Magdalinos (2012), we decompose $Z_n(\kappa)$ and $Y_n(\kappa)$ as a sum of two uncorrelated components, such that $Z_n(\kappa) = Z_n^{(1)}(\kappa) + Z_n^{(2)}(\kappa)$, and $Y_n(\kappa) = Y_n^{(1)}(\kappa) + Y_n^{(2)}(\kappa)$, which are defined as

$$Z_n^{(1)}(\kappa) = n^{-\left(\frac{3}{2}-\kappa\right)\alpha} \sum_{t=1}^{\tau_n(\beta)} \rho_n^{-t} \sum_{j=0}^{t} c_j \epsilon_{t-j},$$

$$Z_n^{(2)}(\kappa) = \sum_{j=1}^{\infty} B_{nj} \epsilon_{-j}, B_{nj} = n^{-\left(\frac{3}{2}-\kappa\right)\alpha} \sum_{t=1}^{\tau_n(\beta)} \rho_n^{-t} c_{t+j},$$

$$Y_n^{(1)}(\kappa) = \sum_{k=1}^{\tau_n(\beta)} C_{nk} \epsilon_{n+1+k}, C_{nk} = n^{-\left(\frac{3}{2}-\kappa\right)\alpha} \sum_{t=1}^{k} \rho_n^{-t} c_{k-t},$$

$$Y_n^{(2)}(\kappa) = n^{-\left(\frac{3}{2}-\kappa\right)\alpha} \sum_{k>\tau_n(\beta)} \sum_{t=1}^{\tau_n(\beta)} \rho_n^{-t} \epsilon_{n+1+k}\,. \tag{35}$$

Using the truncation argument as in (24) and the squeeze theorem, we first show the limit of $\left\| \frac{1}{L(n^\alpha)} Z_n^{(1)}(\kappa) \right\|_2^2$. By virtue of the truncation argument and changing the order of summation, we can express $Z_n^{(\kappa)}(1)$ as

$$Z_n^{(1)}(\kappa) = \frac{1}{n^{\left(\frac{3}{2}-\kappa\right)\alpha}} \sum_{j=K}^{\tau_n(\beta)} c_j \rho_n^{-j} \sum_{k=K}^{\tau_n(\beta)-j} \rho_n^{-\kappa} \epsilon_k.$$

Magdalinos (2012) showed that

$$\left\| \frac{1}{L(n^\alpha)} Z_n^{(1)}(\kappa) \right\|_2^2 \overset{a}{\sim} \frac{1}{2c} \left(\frac{1}{\lambda_n} \sum_{j=K}^{\tau_n(\beta)} c_j \rho_n^{-j} \right)^2. \tag{36}$$

Similar to proving Lemma 3.1, we write the upper bound and the lower bound for the right hand side of equation (36) as

$$\frac{1}{\lambda_n}\int_{K-1}^{\tau_n(\beta)} L(t)t^{-\kappa}e^{-\frac{c}{n^\alpha}t}dt \le \frac{1}{\lambda_n}\sum_{j=K}^{\tau_n(\beta)} c_j\rho_n^{-j} \le \frac{1}{\lambda_n}\int_{K}^{\tau_n(\beta)+1} L(t)t^{-\kappa}e^{-\frac{c}{n^\alpha}t}dt. \tag{37}$$

By changing variable and the dominated convergence, we obtain

$$\begin{aligned}
\frac{1}{\lambda_n}\int_{K}^{\tau_n(\beta)+1} L(t)t^{-\kappa}e^{-\frac{c}{n^\alpha}t}dt &= \frac{1}{L(n^\alpha)n^{\alpha(1-\kappa)}}\int_{\frac{cK}{n^\alpha}}^{\frac{c(\tau_n(\beta)+1)}{n^\alpha}} L\left(\frac{n^\alpha}{c}u\right)\left(u\frac{n^\alpha}{c}\right)^{-\kappa}e^{-u}\left(\frac{c}{n^\alpha}\right)du \\
&= \frac{1}{L(n^\alpha)n^{\alpha(1-\kappa)}}\int_{\frac{cK}{n^\alpha}}^{m_n} e^{-u}u^{-\kappa}L\left(\frac{n^\alpha}{c}u\right)\left(\frac{n^{\alpha(1-\kappa)}}{c^{1-\kappa}}\right)du \\
&= c^{\kappa-1}\int_{\frac{cK}{n^\alpha}}^{m_n} e^{-u}u^{-\kappa}\frac{L\left(\frac{n^\alpha}{c}u\right)}{L(n^\alpha)}du \\
&\to c^{\kappa-1}\int_0^\infty e^{-u}u^{-\kappa}du = c^{\kappa-1}\Gamma(1-\kappa).
\end{aligned}$$

By the squeeze theorem, we obtain the limit of $\frac{1}{\lambda_n}\sum_{j=K}^{\tau_n(\beta)} c_j\rho_n^{-j}$. This result combined with (36) implies that

$$\left\|\frac{1}{L(n^\alpha)}Z_n^{(1)}(\kappa)\right\|_2^2 \to \frac{c^{2\kappa-3}}{2}\Gamma(1-\kappa)^2.$$

After pinning down the asymptotic result for $Z_n^{(1)}(\kappa)$, since $Z_n(\kappa) = Z_n^{(1)}(\kappa) + Z_n^{(2)}(\kappa)$, we can show the limit of $Z_n^{(2)}(\kappa)$:

$$\left\|\frac{1}{L(n^\alpha)}Z_n^{(2)}(\kappa)\right\|_2^2 \to c^{2\kappa-3}\Gamma(1-\kappa)^2\{[2\cos\{\pi(1-\kappa)\}]^{-1} - 1/2\}.$$

Now we show $Y_n^{(2)} \to_p 0$. Letting $i = \kappa - \tau_n(\beta)$ and $s = \tau_n(\beta) - t$ in (35), we have

$$\begin{aligned}
\left\|Y_n^{(2)}(\kappa)\right\|_2^2 &= \sigma^2 n^{-(3-2\kappa)\alpha}\sum_{s,t=0}^{\tau_n(\beta)-1}\rho_n^{-(\tau_n(\beta)-s)}\rho_n^{-(\tau_n(\beta)-t)}\sum_{i=1}^{\infty}c_{i+s}c_{i+t} \\
&\le \sigma^2 n^{-(3-2\kappa)\alpha}\sum_{s,t=0}^{\tau_n(\beta)-1}\rho_n^{-(\tau_n(\beta)-s)}\rho_n^{-(\tau_n(\beta)-t)}\left(\sum_{i=1}^{\infty}c_{i+s}^2\right)^{1/2}\left(\sum_{i=1}^{\infty}c_{i+t}^2\right)^{1/2}.
\end{aligned}$$

Note that $c_i = L(i)i^{-\kappa}$. This implies the inequality: $\sum_{i=1}^{\infty}c_{i+s}^2 \le \sup_i L(i)^2\sum_{i>s} i^{-2\kappa}$. And for any $p > 1$, $\sum_{j=n}^{\infty}\frac{1}{j^p} = O(n^{1-p})$ as $n \to \infty$. Therefore, for some constant C, we have

$$\left\|Y_n^{(2)}(\kappa)\right\|_2^2 \le Cn^{-(3-2\kappa)\alpha}\left(\sum_{s=0}^{\tau_n(\beta)}\rho_n^{-(\tau_n(\beta)-s)}s^{(1-2\kappa)/2}\right)^2$$

$$
\begin{aligned}
&\leq Cn^{-(3-2\kappa)\alpha}\left(\sum_{s=0}^{\tau_n(\beta)} e^{-\frac{c}{n^\alpha}(\tau_n(\beta)-s)} s^{(1-2\kappa)/2}\right)^2 \\
&\quad +2Cn^{-(3-2\kappa)\alpha}\left(\sum_{s=0}^{\tau_n(\beta)} e^{-\frac{c}{n^\alpha}(\tau_n(\beta)-s)} s^{(1-2\kappa)/2}\right)\left(\sum_{j=0}^{\tau_n(\beta)} n^{-\frac{\alpha}{2}} j^{\frac{1-2\kappa}{2}}\right) \\
&\quad +Cn^{-(3-2\kappa)\alpha}\left(\sum_{s=0}^{\tau_n(\beta)} n^{-\frac{\alpha}{2}} s^{\frac{1-2\kappa}{2}}\right)^2 \\
&= Cn^{-(3-2\kappa)\alpha}\left(\sum_{s=0}^{\tau_n(\beta)} e^{-\frac{c}{n^\alpha}(\tau_n(\beta)-s)} s^{(1-2\kappa)/2}\right)^2 + O\left(\frac{n^{(3-2\kappa)(\beta-\alpha)}}{n^\alpha}\right) \\
&\leq Cn^{-(3-2\kappa)\alpha}\left(\int_0^{\tau_n(\beta)} e^{-\frac{c}{n^\alpha}t}(\tau_n(\beta)-t)^{(1-2\kappa)/2}\right)^2 + o(1) \\
&= O\left(n^{-(2\kappa-1)(\beta-\alpha)}\right) = o(1).
\end{aligned}
$$

We apply Lemma A.1 to obtain the second inequality. Since $\beta < \frac{3\alpha}{2}$ and $3-2\kappa \in (0,1)$, we have $(3-2\kappa)(\beta-\alpha) < (3-2\kappa)(\frac{3\alpha}{2}-\alpha) = (3-2\kappa)\frac{\alpha}{2} < \alpha$. To show the order of the integrand in the last expression, we can apply the arguments in proving proposition 3.2.3 as in Magdalinos (2012). Since $\left\|Y_n^{(2)}(\kappa)\right\|_2^2 = o(1)$, so we can obtain the result that $Var(Y_n(\kappa)) = Var(Y_n^{(1)}(\kappa))$. Since $Y_n(\kappa)$, and $Z_n(\kappa)$ share the same variance, Lemma 3.1 is applicable to $Y_n(\kappa)$.

After deriving the limit of the variance of $Y_n^{(1)}(\kappa)$, $Z_n^{(1)}(\kappa)$ and $Z_n^{(2)}(\kappa)$, we apply the proof in proposition 3.2.4 in Magdalinos (2012) to obtain the desired result.

Proof of Theorem 3.1

The proof of Theorem 3.1 is omitted due to similarity to Lemma 5 in Magdalinos (2012).

Proof of Lemma 4.3 and 4.5

Note that by backward substitutions, we can obtain

$$
\begin{aligned}
y_t &= \rho_{n,m}^t y_0 + \sum_{j=1}^t \rho_{n,m}^{t-j} u_j \\
&= \rho_{n,m}^t o_p(n^{1/2+d}) + \sum_{j=1}^t \rho_{n,m}^{t-j}(S_j - S_{j-1}), \text{ where } S_j = \sum_{i=1}^j u_i.
\end{aligned}
$$

Following the approach in Phillips (1987) and noting that $\exp\left(\frac{cm}{n}\right) = \rho_{n,m} + O(n^{-2})$, after applying Lemma 4.1**,** we have

$$
\frac{1}{n^H} y_{\lfloor nr \rfloor} = \rho_{n,m}^{-1}\left\{\frac{1}{n^H} S_{\lfloor nr \rfloor} + cm \int_{1/n}^{\lfloor nr \rfloor/n} \exp\left((\lfloor nr \rfloor - \lfloor ns \rfloor)\frac{cm}{n}\right)\frac{1}{n^H} S_{\lfloor nr \rfloor} ds\right\} + o_p(1)
$$

$$\Rightarrow \quad \varsigma\left(B^H(r) + cm\int_0^r \exp(cm(r-s))B^H(s)ds\right) = \varsigma J_{cm}^H(r).$$

Note that the second to the fourth claims are direct results after applying the continuous mapping theorem (Billingsley, 1968, p. 30). For the last result, after squaring and summing the process y_t, we have

$$\sum_{t=1}^n y_t^2 = \left(1 + \frac{2cm}{n} + \frac{(cm)^2}{n^2}\right)\sum_{t=1}^n y_{t-1}^2 + 2\left(1+\frac{cm}{n}\right)\sum_{t=1}^n y_{t-1}u_t + \sum_{t=1}^n u_t^2,$$

which leads to

$$y_n^2 = \frac{2cm}{n}\sum_{t=1}^n y_{t-1}^2 + 2\sum_{t=1}^n y_{t-1}u_t + \sum_{t=1}^n u_t^2 + \frac{(cm)^2}{n^2}\sum_{t=1}^n y_{t-1}^2 + \frac{2cm}{n}\sum_{t=1}^n y_{t-1}u_t.$$

Hence, we can write

$$\begin{aligned}
\frac{2}{n^{2H}}\sum_{t=1}^n y_{t-1}u_t &= \frac{1}{n^{2H}}y_n^2 - \frac{2cm}{n^{2H+1}}\sum_{t=1}^n y_{t-1}^2 - \frac{1}{n^{2H-1}}\frac{1}{n}\sum_{t=1}^n u_t^2 + o_p(1) \\
&= \varsigma^2\left[\left(J_{cm}^H(1)\right)^2 - 2cm\int_0^1 \left(J_{cm}^H(r)\right)^2 dr\right] - n^{1-2H}E(u_t^2) + o_p(1).
\end{aligned}$$

This implies

$$\begin{aligned}
\frac{1}{n^{2H}}\sum_{t=1}^n y_{t-1}u_t &= \varsigma^2\left[\frac{1}{2}\left(J_{cm}^H(1)\right)^2 - cm\int_0^1 (J_{cm}^H(r))^2 dr\right] - \frac{n^{1-2H}}{2}E(u_t^2) + o_p(1) \\
&= \varsigma^2\left[cmZ(1)\int_0^1 e^{cms}d\widetilde{B}^H(t) + R(1)\right] - \frac{n^{1-2H}}{2}E(u_t^2) + o_p(1),
\end{aligned}$$

where for the last equality we have applied Lemma A.2.

Note that if $d \in (0, 0.5), 1-2H < 0$ and $n^{1-2H}E(u_t^2) = o_{as}(1)$. If $d \in (-0.5, 0)$, we have an extra term $\frac{n^{1-2H}}{2}E(u_t^2)$. This explain the difference between Lemma 4.3 and Lemma 4.5.

Proof of Corollary 4.2

We only need to show the following results are correct:

1. $Z(1) = \int_0^1 e^{-cms}B^H(s)ds = \frac{1}{m^{H+1}}\int_0^m e^{-cs}\widetilde{B}^H(s)ds$;
2. $\int_0^1 \left(B^H(s)\right)^2 ds = \frac{1}{m^{2H}}\int_0^m \left(\widetilde{B}^H(s)\right)^2 ds$;
3. $\int_0^1 (J_{cm}^H(r))^2 dr = \frac{1}{m^{2H+1}}\int_0^m \left(\widetilde{J}_c^H(s)\right)^2 ds$;
4. $m^2\int_0^1\int_0^s e^{cm(r-s)}B^H(r)B^H(s)drds = \frac{1}{m^{2H}}\int_0^m\int_0^s e^{c(r-s)}\widetilde{B}^H(r)\widetilde{B}^H(s)drds$.

As the steps to prove the above results are similar, we shall only prove the last two claims. For the third result, we have

$$\begin{aligned}
\int_0^1 \left(J_{cm}^H(r)\right)^2 dr &= \int_0^1 \left(\int_0^r e^{cm(r-s)} dB^H(s)\right)^2 dr \\
&= \int_0^1 e^{2cmr} \left(\int_0^r e^{-cms} dB^H(s)\right)^2 dr \\
&= \int_0^1 e^{2cmr} \left(\int_0^{mr} e^{-cv} dB^H\left(\frac{v}{m}\right)\right)^2 dr \\
&= \frac{1}{m^{2H}} \int_0^1 e^{2cms} \int_0^{mr} e^{-cv} d\left(m^H B^H\left(\frac{v}{m}\right)\right)^2 dr \\
&= \frac{1}{m^{2H}} \int_0^m e^{2cu} \left(\int_0^u e^{-cv} d\widetilde{B}^H(v)\right)^2 d\left(\frac{u}{m}\right) \\
&= \frac{1}{m^{2H+1}} \int_0^m \left(\int_0^u e^{c(u-v)} d\widetilde{B}^H(v)\right)^2 du \\
&= \frac{1}{m^{2H+1}} \int_0^m \left(\widetilde{J}_c^H(u)\right)^2 du.
\end{aligned}$$

For the fourth result, we have

$$\begin{aligned}
m^2 \int_0^1 \int_0^s e^{cm(r-s)} B^H(r) B^H(s) dr ds &= m^2 \int_0^1 e^{-cms} \left(\int_0^s e^{cmr} B^H(r) dr\right) B^H(s) ds \\
&= m^2 \int_0^1 e^{-cms} \left(\int_0^{ms} e^{cr} B^H\left(\frac{r}{m}\right) d\left(\frac{r}{m}\right)\right) B^H(s) ds \\
&= \frac{m}{m^H} \int_0^m e^{-cv} \left(\int_0^{ms} e^{cr} \widetilde{B}^H(r) dr\right) B^H\left(\frac{v}{m}\right) d\left(\frac{v}{m}\right) \\
&= \frac{1}{m^{2H}} \int_0^m e^{-cv} \left(\int_0^v e^{cr} \widetilde{B}^H(r) dr\right) \widetilde{B}^H(v) dv \\
&= \frac{1}{m^{2H}} \int_0^m \int_0^v e^{c(r-v)} \widetilde{B}^H(r) \widetilde{B}^H(v) dr dv.
\end{aligned}$$

Proof of Lemma 4.4 The proof of the first three items in this Lemma can be found in Lemma 2.2, the proof of Theorem 2.2 and expression (3.17) in El Machikouri et al. (2016).

Proof of Theorem 4.3 and Theorem 4.4

To avoid confusion, we now refer $n \to \infty$ with m fixed as "fix-m asymptotics", and $n \to \infty$ followed by $m \to \infty$ as "sequential asymptotics". Note that the fix-m asymptotics lead us to the following expression:

$$\begin{aligned}
& ne^{cm} \left(\hat{\rho}_n - \rho_{n,m}\right) \\
= \; & ne^{cm} \frac{\sum_{t=1}^n y_{t-1} u_t}{\sum_{t=1}^n y_{t-1}^2} = e^{cm} \frac{\frac{1}{n^{2H}} \sum_{t=1}^n y_{t-1} u_t}{\frac{1}{n^{2H+1}} \sum_{t=1}^n y_{t-1}^2}
\end{aligned}$$

$$
\begin{aligned}
&\overset{a}{\sim} \quad e^{cm}\frac{\frac{1}{m^{2H}}\left(c\int_0^m e^{-cr}\widetilde{B}^H(r)dr\int_0^m e^{-cs}d\widetilde{B}^H(s)+\widetilde{R}(m)\right)+\frac{n^{1-2H}}{2}E(u_t^2)}{\frac{1}{m^{2H+1}}\int_0^m\left(\widetilde{J}_c^H(s)\right)^2ds}\\
&= \quad \frac{\frac{1}{m^{2H}}e^{-cm}\left(c\int_0^m e^{-cr}\widetilde{B}^H(r)dr\int_0^m e^{-cs}d\widetilde{B}^H(s)+\widetilde{R}(m)\right)+e^{-cm}\frac{n^{1-2H}}{2}E(u_t^2)}{\frac{1}{m^{2H+1}}e^{-2cm}\int_0^m\left(\widetilde{J}_c^H(s)\right)^2ds}\\
&= \quad \frac{\frac{1}{m^{2H}}e^{-cm}\left(c\int_0^m e^{-cr}\widetilde{B}^H(r)dr\int_0^m e^{-cs}d\widetilde{B}^H(s)+\widetilde{R}(m)\right)}{\frac{1}{m^{2H+1}}e^{-2cm}\int_0^m\left(\widetilde{J}_c^H(s)\right)^2ds}\\
&\quad +\frac{e^{-cm}\frac{n^{1-2H}}{2}E(u_t^2)}{\frac{1}{m^{2H+1}}e^{-2cm}\int_0^m\left(\widetilde{J}_c^H(s)\right)^2ds},
\end{aligned}
\tag{38}
$$

where we have applied Corollary 4.2 to obtain the equality. As in Lemma 4.3, we do not have the second term $\frac{e^{-cm}\frac{n^{1-2H}}{2}E(u_t^2)}{\frac{1}{m^{2H+1}}e^{-2cm}\int_0^m\left(\widetilde{J}_c^H(s)\right)^2ds}$ when $d\in(0,0.5)$ which implies $1-2H<0$. The second term only shows up when $d\in(-0.5,0)$.

For the first term in (38), we utilize Lemma 2.2 and Lemma 2.4 in El Machkouri et al. (2016) to obtain the following three results as $m\to\infty$:

1. $e^{-2cm}\int_0^m\left(\widetilde{J}_c^H(s)\right)^2ds \Rightarrow \frac{c}{2}\left(\int_0^\infty e^{-cs}\widetilde{B}^H(s)ds\right)^2$;
2. $e^{-cm}c\int_0^m e^{-cr}\widetilde{B}^H(r)dr\int_0^m e^{cs}d\widetilde{B}^H(s) \Rightarrow c\left(\int_0^\infty e^{-cs}\widetilde{B}^H(s)ds\right)\sqrt{\frac{H\Gamma(2H)}{c^{2H}}}\eta$;
3. $e^{-cm}\widetilde{R}(m)\to 0$.

For the second term in (38), we have

$$
\begin{aligned}
\frac{e^{-cm}\frac{n^{1-2H}}{2}E(u_t^2)}{\frac{1}{m^{2H+1}}e^{-2cm}\int_0^m\left(\widetilde{J}_c^H(s)\right)^2ds} &= \frac{e^{-cm}n^{1-2H}}{\frac{1}{m^{2H+1}}}\frac{\frac{1}{2}E(u_t^2)}{e^{-2cm}\int_0^m\left(\widetilde{J}_c^H(s)\right)^2ds}\\
&= m\frac{n^{1-2H}}{\exp(\delta m)}\frac{m^{2H}}{\exp((c-\delta)m)}\frac{\frac{1}{2}E(u_t^2)}{e^{-2cm}\int_0^m\left(\widetilde{J}_c^H(s)\right)^2ds},
\end{aligned}
$$

where $\delta\in(0,c)$.

Under the assumption that $\frac{n^{1-2H}}{\exp(\delta m)}\to 0$ as $m\to\infty$, we have

$$
\begin{aligned}
\frac{1}{2c}\frac{n}{m}e^{cm}\left(\hat{\rho}_n-\rho_{n,m}\right) &\Rightarrow \frac{1}{2c}\frac{c\int_0^\infty e^{-cs}\widetilde{B}^H(s)ds\sqrt{\frac{H\Gamma(2H)}{c^{2H}}}\eta}{\frac{c}{2}\left(\int_0^\infty e^{-cs}\widetilde{B}^H(s)ds\right)^2}\\
&= \frac{1}{c}\frac{\sqrt{\frac{H\Gamma(2H)}{c^{2H}}}\eta}{\int_0^\infty e^{-cs}\widetilde{B}^H(s)ds}=\frac{\sqrt{\frac{H\Gamma(2H)}{c^{2H}}}\eta}{\sqrt{\frac{H\Gamma(2H)}{c^{2H}}}\omega}=C,
\end{aligned}
$$

where ω and η are two independent standard normal random variables.

Proof of Lemma 5.1 Note that the proofs under a long-memory or an anti-persistent error process are similar. The former one utilizes Lemma 1 in Magdalinos (2012) while the later utilizes Lemma 3.1. Therefore, we only prove the claims under anti-persistence.

The results of this Lemma are similar to those in Theorem 2.6 of Fei (2018). However, since the assumptions on $\{u_t\}_{t=1}^n$ and y_0 are different, the proofs are different. To show the first result, note that from equation (4) in Fei (2018), we can write $y_n = \frac{\mu}{c} n^\alpha (\rho_n^n - 1) + \tilde{y}_n$, where $\tilde{y}_n = \rho_n^n y_0 + \sum_{j=0}^{n-1} \rho_n^j u_{n-j}$. Therefore, we have

$$\frac{\rho_n^{-n}}{n^{(3/2-\kappa)\alpha} L(n^\alpha)} \tilde{y}_n = \frac{y_0}{n^{(3/2-\kappa)\alpha} L(n^\alpha)} + \frac{1}{n^{(3/2-\kappa)\alpha} L(n^\alpha)} \sum_{j=1}^{n} \rho_n^{-j} u_j = o_p(1) + \frac{Z_n(\kappa)}{L(n^\alpha)} = O_p(1).$$

Consequently, we have

$$\frac{\rho_n^{-n} y_n}{n^\alpha} = \frac{\mu}{c}\left(1 - \frac{1}{\rho_n^n}\right) + O_p\left(L(n^\alpha) n^{(3/2-1-\kappa)\alpha}\right) = \frac{\mu}{c} + o_p(1). \tag{39}$$

For the result of $\frac{\rho_n^{-n}}{n^{2\alpha}} \sum_{t=1}^n y_{t-1}$, we can write

$$\begin{aligned} \frac{c}{n^\alpha} \sum_{t=1}^{n} y_{t-1} &= y_n - y_0 - \mu n - \sum_{t=1}^{n} u_t \\ &= O_p(\rho_n^n n^\alpha) - o_p\left(L(n^\alpha) n^{(3/2-\kappa)\alpha}\right) - O(n) - O_p\left(n^{3/2-\kappa}\right) = O_p(\rho_n^n n^\alpha). \end{aligned}$$

To obtain the second equality above, we apply the first claim in this Lemma and Proposition 4.4.4 in Giraitis et al. (2012) to obtain that $y_n = O_p(\rho_n^n n^\alpha)$ and $\sum_{t=1}^n u_t = O_p\left(n^{3/2-\kappa}\right)$. It is clear the first term plays a dominant role asymptotically. Combined with (39), we obtain $\frac{\rho_n^{-n} c}{n^{2\alpha}} \sum_{t=1}^n y_{t-1} = \frac{\mu}{c} + o_p(1)$ and therefore establish the result.

For the next term,

$$\begin{aligned} & \frac{\rho_n^{-n}}{L(n^\alpha) n^{(5/2-\kappa)\alpha}} \sum_{t=1}^{n} y_{t-1} u_t \\ = & \frac{\rho_n^{-n}}{L(n^\alpha) n^{(5/2-\kappa)\alpha}} \left[\sum_{t=1}^{n} \left(\frac{\mu}{c} n^\alpha (\rho_n^{t-1} - 1) + \tilde{y}_{t-1}\right) u_t\right] \\ = & \frac{\rho_n^{-n}}{L(n^\alpha) n^{(5/2-\kappa)\alpha}} \left[\frac{\mu}{c} n^\alpha \sum_{t=1}^{n} \rho_n^{t-1} u_t - \frac{\mu}{c} n^\alpha \sum_{t=1}^{n} u_t + \sum_{t=1}^{n} \tilde{y}_{t-1} u_t\right] \\ = & \frac{\mu}{c} \frac{1}{L(n^\alpha) n^{(3/2-\kappa)\alpha}} \sum_{t=1}^{n} \rho_n^{-t} u_{n+1-t} - \frac{\mu}{c} \frac{\rho_n^{-n} n^{(3/2-\kappa)(1-\alpha)}}{L(n^\alpha)} \frac{1}{n^{3/2-\kappa}} \sum_{t=1}^{n} u_t \\ & + \frac{L(n^\alpha) n^{(3/2-\kappa)\alpha}}{n^\alpha} \frac{\rho_n^{-n}}{L(n^\alpha)^2 n^{(3-2\kappa)\alpha}} \sum_{t=1}^{n} \tilde{y}_{t-1} u_t \\ = & \frac{\mu}{c} \frac{Y_n(\kappa)}{L(n^\alpha)} - O_p\left(\frac{\rho_n^{-n} n^{(3/2-\kappa)(1-\alpha)}}{L(n^\alpha)}\right) + \frac{L(n^\alpha) n^{(3/2-\kappa)\alpha}}{n^\alpha} \left(\frac{Y_n(\kappa)}{L(n^\alpha)}\right) \left(\frac{Z_n(\kappa)}{L(n^\alpha)}\right) + o_p(1) \end{aligned}$$

$$\Rightarrow \quad \frac{\mu}{c} Y_d.$$

Note that we apply Lemma 3.1 to establish the last result. For the final claim, we have

$$\left(\rho_n^2 - 1\right) \sum_{t=1}^{n} y_{t-1}^2 = y_n^2 - y_0^2 - \mu^2 n - \sum_{t=1}^{n} u_t^2 - 2\mu\rho_n \sum_{t=1}^{n} y_{t-1} - 2d \sum_{t=1}^{n} u_t - 2\rho_n \sum_{t=1}^{n} y_{t-1} u_t.$$

As the first term has the highest order, we obtain the desired result.

Proof of Theorem 5.2 and Theorem 5.3

To analyze the asymptotic behavior of the LS estimator, we can express the centered LS estimator in matrix form:

$$\begin{aligned}
\begin{bmatrix} n^{(1/2-d)}(\hat{\mu} - \mu) \\ \frac{\rho_n^n n^{(1/2-d)\alpha}}{L(n^\alpha)(\rho_n^2 - 1)}(\hat{\rho}_n - \rho_n) \end{bmatrix} &= \begin{bmatrix} 1 & \frac{L(n^\alpha)\rho_n^{-n}(\rho_n^2-1)}{n^{(3/2-d)\alpha+1/2+d}} \sum_{t=1}^{n} y_{t-1} \\ \frac{\rho_n^{-n}}{L(n^\alpha) n^{(3/2+d)\alpha+1/2-d}} \sum_{t=1}^{n} y_{t-1} & \frac{\rho_n^{-2n}(\rho_n^2-1)}{n^{2\alpha}} \sum_{t=1}^{n} y_{t-1}^2 \end{bmatrix}^{-1} \\
&\quad \times \begin{bmatrix} n^{-(1/2+d)} \sum_{t=1}^{n} u_t \\ \frac{\rho_n^{-n}}{L(n^\alpha) n^{(5/2-d)\alpha}} \sum_{t=1}^{n} y_{t-1} u_t \end{bmatrix} \\
&\Rightarrow \begin{bmatrix} 1 & o_p(1) \\ o_p(1) & \frac{\mu^2}{c^2} \end{bmatrix}^{-1} \begin{bmatrix} \varsigma B^H(1) \\ \left(\frac{\mu}{c}\right) Y_d \end{bmatrix}.
\end{aligned}$$

Note that when $d < 0.5$, $(3/2+d)\alpha+1/2-d > 2\alpha$; when $d > -0.5$, $(3/2-d)\alpha+1/2+d > 2\alpha$. Based on Lemma 5.2, the two off-diagonal elements of the inverse matrix converge in probability to 0 as $n \to \infty$, and we have the result for the limiting distribution of $\hat{\mu}$ and $\hat{\rho}_n$.

Proof of Remark 5.5

We can directly obtain the result of Remark 5.5 based on the result of Theorem 5.2 or 5.3. Note that

$$\frac{\rho_n^n n^{(1/2-d)\alpha}}{L(n^\alpha)(\rho_n^2 - 1)}(\hat{\rho}_n - \rho_n) \Rightarrow N\left(0, \frac{\sigma^2 c^{1-2d}}{\mu^2} \frac{\Gamma(d)^2}{2\cos(\pi d)}\right),$$

which implies

$$\frac{\rho_n^n n^{(3/2-d)\alpha}}{L(n^\alpha) 2c}(\hat{\rho}_n - \rho_n) \Rightarrow c^{1/2-d} N\left(0, \frac{\sigma^2}{\mu^2} \frac{\Gamma(d)^2}{2\cos(\pi d)}\right),$$

which in turn implies

$$\frac{\rho_n^n}{L(n^\alpha)} \left(\rho_n - 1\right)^{-3/2+d} (\hat{\rho}_n - \rho_n) \Rightarrow N\left(0, \frac{\sigma^2}{\mu^2} \frac{2\Gamma(d)^2}{\cos(\pi d)}\right).$$

References

Anderson, T.W., (1959). On asymptotic distribution of estimates of parameters of stochastic difference equations. Annals of Mathematical Statistics, 30(3), 676–687.

Baillie, R. T., Bollerslev, T. & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 74(1), 3–30.

Billingsley, P. (1968). Convergence of probability measures. John Wiley & Sons.

Chan, N.H. & Wei, C.Z. (1988). Limiting distributions of least squares estimates of unstable autoregressive processes. Annals of Statistics, 367-401.

Cheung, Y W. (1993). Long memory in foreign-exchange rates. Journal of Business and Economic Statistics, 11(1): 93–101.

Fei, Y. (2018). Limit theory for mildly integrated process with intercept. Economics Letters, 163, 98–101.

El Machkouri, M., Es-Sebaiy, K., & Ouknine, Y. (2016). Least squares estimator for non-ergodic Ornstein–Uhlenbeck processes driven by Gaussian processes. Journal of the Korean Statistical Society, 45(3), 329–341.

Gatheral, J., Jaisson, T. & Rosenbaum, M. (2018). Volatility is rough. Quantitative Finance, 18(6), 933–949.

Giraitis, L., Koul, H. L., & Surgailis, D. (2012). Large sample inference for long memory processes. World Scientific Publishing Company.

Magdalinos, T. (2012). Mildly explosive autoregression under weak and strong dependence. Journal of Econometrics, 169(2), 179–187.

Newey, W. K., West, K. D. (1987). A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica, 55(3): 703–708.

Park, J. (2003). Weak unit roots. Working Paper, Rice University.

Phillips, P. C. B. (1987). Towards a unified asymptotic theory for autoregression. Biometrika, 74(3), 535–547

Phillips, P. C. B., & Magdalinos, T. (2007a). Limit theory for moderate deviations from a unit root. Journal of Econometrics, 136(1), 115–130.

Phillips, P. C. B., & Magdalinos, T. (2007b).. Limit theory for moderate deviations from a unit root under weak dependence. In: Phillips, G.D.A., Tzavalis, E. (Eds.), The Refinement of Econometric Estimation and Test Procedures. Cambridge University Press.

Phillips P. C. B., Magdalinos T, Giraitis L. (2010) Smoothing local-to-moderate unit root theory. Journal of Econometrics, 158(2), 274–279.

Phillips, P. C. B., Shi, S., Yu, J., (2015a). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the s&p 500. International Economic Review, 56(4), 1043–1078.

Phillips, P. C. B., Shi, S., Yu, J., (2015b). Testing for multiple bubbles: Limit theory of real-time detectors. International Economic Review 56 (4), 1079-1134.

Phillips, P. C. B., Wu, Y., Yu, J., (2011). Explosive behavior in the 1990s nasdaq: When did exuberance escalate asset values? International Economic Review 52 (1), 201–226.

Phillips, P. C. B., & Yu, J., (2011). Dating the Timeline of Financial Bubbles During the Subprime Crisis, Quantitative Economics, 2(3), 455–491.

Robinson, P. M. (1994). Semiparametric analysis of long-memory time series. Annals of Statistics, 22(1): 515–539.

Wang, X., & Yu, J., (2015). Limit theory for an explosive autoregressive process. Economics Letters 126, 176–180.

White, J., (1958). The limiting distribution of the serial correlation coefficient in the explosive case. Annals of Mathematical Statistics, 29(4), 1188–1197.

Xiao, W., Wang, X., Yu, J., (2018). Estimation and Inference in Fractional Ornstein-Uhlenbeck Model with Discrete-Sampled Data. Manuscript.

Xiao, W., & Yu, J., (2018a). Asymptotic Theory for Estimating Drift Parameters in the Fractional Vasicek Model. Econometric Theory. Cambridge University Press, 1–34.

Xiao, W., & Yu, J., (2018b). Asymptotic Theory for Rough Fractional Vasicek Models. Working paper.