1-1-2003

Asymptotics for general fractionally intergrated processes with applications to unit root tests

Q. Wang
University of Wollongong

Yan-Xia Lin
University of Wollongong, yanxia@uow.edu.au

C. M. Gulati
University of Wollongong, cmg@uow.edu.au

Asymptotics for general fractionally intergrated processes with applications to unit root tests

Abstract

In this paper, functional limit theorems for general fractional processes are established under quite weak conditions+ The results are then used to derive weak convergence of general nonstationary fractionally integrated processes and to characterize unit root distribution in a model with error being a fractional autoregressive moving average process or a nonstationary fractionally integrated process+

Disciplines

Physical Sciences and Mathematics

Publication Details

This article was originally published as Wang, Q, Lin, YX and Gulati, CM, Asymptotics for general fractionally intergrated processes with applications to unit root tests, Econometric Theory, 19, 2003, 143-164. Copyright Cambridge University Press. Original article available here.

Econometric Theory, **19**, 2003, 143–164. Printed in the United States of America.
DOI: 10.1017/S0266466603191062

ASYMPTOTICS FOR GENERAL FRACTIONALLY INTEGRATED PROCESSES WITH APPLICATIONS TO UNIT ROOT TESTS

Qiying Wang
Australian National University and University of Wollongong

Yan-Xia Lin and Chandra M. Gulati
University of Wollongong

In this paper, functional limit theorems for general fractional processes are established under quite weak conditions. The results are then used to derive weak convergence of general nonstationary fractionally integrated processes and to characterize unit root distribution in a model with error being a fractional autoregressive moving average process or a nonstationary fractionally integrated process.

1. INTRODUCTION

Consider a fractionally integrated autoregressive moving average(autoregressive fractionally integrated moving average) (ARFIMA) process $\{X_t\}$ defined by

$$(1-B)^{d+m}X_t = u_t, \qquad \phi(B)u_t = \theta(B)\epsilon_t, \tag{1.1}$$

where $m \geq 0$ is an integer and $d \in (-\frac{1}{2}, \frac{1}{2})$; B is a backshift operator and ϵ_t are independently and identically distributed (i.i.d.) random variables with zero mean and finite variance; $\phi(B)$ and $\theta(B)$ are polynomial functions of B with order p, and q, respectively, and both of them only have roots outside the unit circle, i.e., the ARMA(p,q) process u_t is taken to be stationary and invertible. The fractional difference operator $(1-B)^{\gamma}$ is defined by its Maclaurin series (by its binomial expansion, if γ is an integer):

The authors thank three referees and an associate editor for their detailed reading of this paper and valuable comments, which have led to this much improved version of the paper. Address correspondence to: Qiying Wang, CMA, Mathematical Sciences Institute, ANU, ACT, 0200, Australia; e-mail: qiying@maths.anu.edu.au.

 0266-4666/03 $9.50

$$(1-B)^{\gamma} = \sum_{j=0}^{\infty} \frac{\Gamma(-\gamma+j)}{\Gamma(-\gamma)\Gamma(j+1)} B^{j} \quad \text{where } \Gamma(z) = \begin{cases} \int_0^{\infty} s^{z-1} e^{-z}\, ds & \text{if } z > 0 \\ \infty & \text{if } z = 0. \end{cases} \tag{1.2}$$

If $z < 0$, $\Gamma(z)$ is defined by the recursion formula $z\Gamma(z) = \Gamma(z+1)$.

Since model (1.1) was introduced by Granger and Joyeux (1980) and Hosking (1981), it has become very popular in applications. It nests the usual Box–Jenkins autoregressive integrated moving average (ARIMA) model and has an ability to capture both short-term dynamics and a wide variety of low-frequency behavior at the same time. Also, there is considerable evidence on the success of applying ARFIMA models to describe financial data such as forward premiums, interest rate differentials, and inflation rates. Illustrations can be found in survey and review papers of Robinson (1994) and Baillie (1996).

Because of their applications in economics and finance, ARFIMA processes have been studied quite extensively in recent years. In model (1.1), it is well known that the process is stationary and invertible when $m = 0$ (Hosking, 1981; Odaki, 1993); when $m \geq 1$, the process is nonstationary. In particular, when $d = 0$ and m is an integer, the process becomes a usual unit root process. For estimate of the parameter $d + m$ and other related statistical inferences, because the literature is rather extensive, we here only refer to Hosking (1984), Li and Mcleod (1986), Fox and Taqqu (1986), Dahlhaus (1989), Giraitis and Surgailis (1990), Beran (1995), Beran, Bahnsali, and Ocker (1998), and Tanaka (1999). For more results, see the references cited in these papers and a review book of Beran (1994).

As to the asymptotics of the ARFIMA processes, Sowell (1990) first derived a result that the partial sum process of a simple fractional process (i.e., $m = 0$, $\phi(B) \equiv \theta(B) \equiv 1$ in model (1.1)) converges weakly to a "type I" fractional Brownian motion[1] on $D[0,1]$ instead of a standard Brownian motion. We mention that, combined with the continuous mapping theorem, Sowell's result is quite useful in characterizing limit distributions of the various statistics arising from statistical inference in economic time series such as spurious regression and testing for unit roots. However, Sowell (1990) only provides a weak convergence result on simple fractional processes. This shortcoming limits the applicability of Sowell's result to statistical inference in economic time series.

Motivated by characterizing unit root distribution in a more general model, this paper extends the weak convergence result given by Sowell (1990) to general fractional processes. Instead of assuming that the innovations u_t in model (1.1) are an ARMA(p,q) process, we allow it to be a more general linear process. This means that this paper provides a unified treatment for previous studies on weak convergence for summable linear processes and fractional processes. The weak convergence results for summable linear processes can be found in Hannan (1979), Phillips and Solo (1992), and Chan and Tsay (1996). Our results are then used to derive functional limit theorems for general nonstationary fractionally integrated processes and to derive the limit distribution of the

least square estimate (LSE) of the coefficient for a AR(1) model when the true coefficient is 1 (i.e., the true model has a unit root) and the error process is a general fractional process or a general nonstationary fractionally integrated process. These results improve essentially the related results in the literature (see Section 3 for more details), and they also provide a unified treatment for unit root tests with the error process being a summable linear process, a fractional process, or a nonstationary fractionally integrated process. The main results of this paper are given under quite weak moment conditions for the innovations ϵ_t. For example, weak convergence of general nonstationary fractionally integrated processes and nonstationary unit root distribution are derived whenever the innovations ϵ_t have finite second moment. Such a condition is the best possible moment condition in the literature and it is of interest from a theoretical point of view.

This paper is organized as follows. In the next section, we derive weak convergence of general fractional processes without proofs and compare them to related results in the literature. Applications of these results to general nonstationary fractionally integrated processes and testing for unit roots will be presented in Section 3. Finally in the Appendix, we give the proofs of the main theorems in Section 2.

We end this section with some notation. We denote $\lim_{n\to\infty} a_n/b_n = 1$ by $a_n \sim b_n$; $C, C_1, \dots$ are for positive constants, which may take on different values in different places. The expression $D[0,1]$ denotes the space of functions on $[0,1]$ in which all elements are right continuous and have left-hand limits, endowed with the Skorohod topology (see Billingsley, 1968, p. 111). Convergence in distribution and weak convergence of probability measures on $D[0,1]$ are denoted by $\to_d$ and $\Rightarrow$, respectively. Finally, we define type I fractional Brownian motions with $-\frac{1}{2} < d < \frac{1}{2}$ on $D[0,1]$ as follows:

$$W_d(t) = \frac{1}{A(d)} \int_{-\infty}^{0} [(t-s)^d - (-s)^d]\, dW(s) + \int_0^t (t-s)^d\, dW(s),$$

where $W(s)$ is a standard Brownian motion and

$$A(d) = \left(\frac{1}{2d+1} + \int_0^\infty [(1+s)^d - s^d]^2\, ds \right)^{1/2}.$$

Clearly, $W_d(t)$ is a self-similar Gaussian process with covariance

$$EW_d(s)W_d(t) = \frac{1}{2}\{s^{1+2d} + t^{1+2d} - |s-t|^{1+2d}\}, \quad \text{for } 0 \le s,t \le 1.$$

A more general definition of fractional Brownian motions can be found in Mandelbrot and Van Ness (1968), Samorodnitsky and Taqqu (1994), and Marinucci and Robinson (1999).

2. ASYMPTOTICS OF GENERAL ARFIMA PROCESSES

From here on, we discuss the following general ARFIMA process X_t defined by

$$(1-B)^{d+m}X_t = u_t, \qquad u_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}, \qquad t = 1,2,\ldots, \tag{2.1}$$

where $m \geq 0$ is an integer and $d \in (-\frac{1}{2},\frac{1}{2})$; $(1-B)^d$ is defined by (1.2); $\epsilon_j, j = 0,\pm 1,\ldots$ are i.i.d. random variables with $E\epsilon_0 = 0$, and $\{\psi_j, j \geq 0\}$ is a sequence of real numbers to be specified later.

The two theorems in this section derive results on weak convergence of general stationary fractional processes. They provide a unified treatment for the cases of fractional processes and summable linear processes.

THEOREM 2.1. *Let X_j satisfy (2.1) with $m = 0$ and let $\psi_j, j \geq 0$, satisfy*

$$\sum_{j=0}^{\infty} |\psi_j| < \infty \quad \text{and} \quad b_\psi \equiv \sum_{j=0}^{\infty} \psi_j \neq 0. \tag{2.2}$$

Assume that $E\epsilon_0^2 < \infty$. Then, for $0 \leq d < \frac{1}{2}$,

$$\frac{1}{\kappa(d) n^{1/2+d}} \sum_{j=1}^{[nt]} X_j \Rightarrow W_d(t), \qquad 0 \leq t \leq 1, \tag{2.3}$$

where $\kappa^2(d) = [b_\psi^2 E\epsilon_0^2 \Gamma(1-2d)]/(1+2d)\Gamma(1+d)\Gamma(1-d)$ and $W_d(t)$ is a type I fractional Brownian motion on $D[0,1]$.

If, in addition, $E|\epsilon_0|^{(2+\delta)/(1+2d)} < \infty$, where $\delta > 0$, then (2.3) still holds for $-\frac{1}{2} < d < 0$.

For $0 \leq d < \frac{1}{2}$, Theorem 2.1 gives the result under the best possible moment condition $E\epsilon_0^2 < \infty$. If there is a slightly stronger restriction on ψ_k, the moment condition in Theorem 2.1 for $-\frac{1}{2} < d < 0$ can be weakened to $E|\epsilon_0|^{2/(1+2d)} < \infty$. Explicitly, we have the following theorem.

THEOREM 2.2. *Let X_j satisfy (2.1) with $m = 0$ and let $\psi_j, j \geq 0$, satisfy*

$$\sum_{j=0}^{\infty} j^{1/2-d} |\psi_j| < \infty \quad \text{and} \quad b_\psi \equiv \sum_{j=0}^{\infty} \psi_j \neq 0. \tag{2.4}$$

Assume that $E|\epsilon_0|^{\max\{2,2/(1+2d)\}} < \infty$. Then (2.3) holds for $d \in (-\frac{1}{2},\frac{1}{2})$.

The proofs of Theorems 2.1 and 2.2 are given in the Appendix.

If u_t is a process satisfying $\phi(B)u_t = \theta(B)\epsilon_t$, where polynomials $\phi(B)$ and $\theta(B)$ with order p and q, respectively, have only roots outside the unit circle, Theorem 3.1.1.1 of Brockwell and Davis (1987, p. 85) implies that $u_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$ with $|\psi_k| \leq Ca^{-k}, k \geq 0$, where $a > 1$ and $\sum_{k=0}^{\infty} \psi_k = \theta(1)/\phi(1)$. Therefore, the following corollary is a direct consequence of Theorem 2.2.

COROLLARY 2.1. *Let X_j satisfy (1.1) with $m = 0$. If $E|\epsilon_0|^{\max\{2,2/(1+2d)\}} < \infty$, then (2.3) follows with $b_\psi = \theta(1)/\phi(1)$ for $d \in (-\frac{1}{2}, \frac{1}{2})$.*

We next give some remarks that compare our results to those in the literature.

Remark 2.1. If $\phi(B) \equiv \theta(B) \equiv 1$ in model (1.1), Corollary 2.1 reduces to Theorem 2.2 of Sowell (1990), where the author derives (2.3) provided $E|\epsilon_t|^r < \infty$ for $r \geq \max\{4, -8d/(1+2d)\}$. If $d = 0$, then $X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$ with $\sum_{j=0}^{\infty} |\psi_j| < \infty$. In this case $\kappa^2(0) = b_\psi^2 E\epsilon_0^2$ and $W_0(t)$ is a standard Brownian motion on $D[0,1]$. Thus, Hannan's result (1979) becomes a special case of Theorem 2.1. Theorem 3.1.4 of Phillips and Solo (1992) and Theorem 2.2.5 of Chan and Tsay (1996) also give similar results but impose more restrictions on ψ_j. For further results on weak convergence of linear processes, refer to Wang, Lin, and Gulati (2002).

Remark 2.2. Davydov (1970) and the later work by Gorodetskii (1977), Taqqu (1975), Avram and Taqqu (1987), and Mielniczuk (1997) give results on the functional limit theorem for linear process with square summable weights. Theorems 2.1 and 2.2 derive essentially weak convergence of an infinite weighted function of a linear process, i.e., a process such as $X_t = \sum_{k=-\infty}^{\infty} c_{t-k} u_k$, where $u_k = \sum_{j=1}^{\infty} \psi_j \epsilon_{k-j}$. We note that this infinite weighted function is difficult to rewrite as a linear process with explicit coefficients and therefore the results in the papers cited previously can not be applied directly to our theorems. On the other hand, even in the special case where $\psi_0 = 1$ and $\psi_k = 0, k \geq 1$, Theorems 2.1 and 2.2 are not a direct consequence of the papers cited earlier as their results are held under a higher moment condition for innovations ϵ_k, in particular, in the case that $-\frac{1}{2} < d < 0$.

Remark 2.3. For the results on functional limit theorems for fractional processes defined by (2.1) with other dependent innovations instead of a linear process u_t, refer to Davidson and Jong (2000) and Wang, Lin, and Gulati (2001).

3. APPLICATIONS

In this section, we apply the main results presented in Section 2 to several well-known examples, namely, to derive functional limit theorems for general nonstationary fractionally integrated processes and fractional unit root distributions. These have been studied by various authors in recent years. As will be seen later, the applications of Theorems 2.1 and 2.2 to the related statistics can lead to better results under weaker conditions.

3.1. General nonstationary fractionally integrated processes

In previous research, the functional limit theorem for general nonstationary fractionally integrated processes has been discussed in a very general framework by Chan and Terrin (1995) under the assumption that u_t defined in (2.1) is a class

of stationary Gaussian processes. The result of Chan and Terrin (1995) extends the results of Chan and Wei (1988), Parks and Phillips (1988, 1989), and Sims, Stock, and Watson (1990) from the domain of integer m's (i.e., $d = 0$) to fractional $d + m$'s. More recently, Liu (1998) has derived a functional limit theorem for simple nonstationary fractionally integrated processes (i.e., the process X_t defined by (2.1) with $m \geq 1$ and $u_t = \epsilon_t$) provided $E|\epsilon_t|^r < \infty$ for $r \geq \max\{4, -8d/(1+2d)\}$. Theorem 3.1, which follows gives an essential improvement of the results cited previously.

For convenience, we introduce the following conditions.

Condition A. $\psi_j, j \geq 0$ satisfy (2.2) and $E|\epsilon_0|^p < \infty$, where $p = 2$, for $0 \leq d < \frac{1}{2}$; $p = (2+\delta)/(1+2d) < \infty, \delta > 0$, for $-\frac{1}{2} < d < 0$.

Condition B. $\psi_j, j \geq 0$ satisfy (2.4) and $E|\epsilon_0|^{\max\{2, 2/(1+2d)\}} < \infty$.

THEOREM 3.1. *Let X_j satisfy (2.1) with $m \geq 1$. Let Condition A or Condition B hold. Then, for $0 \leq t \leq 1$,*

$$\frac{1}{\kappa(d) n^{-1/2+d+m}} X_{[nt]} \Rightarrow W_{d,m}(t), \tag{3.1}$$

$$\frac{1}{\kappa(d) n^{1/2+d+m}} \sum_{j=1}^{[nt]} X_j \Rightarrow \int_0^t \int_0^{t_m} \dots \int_0^{t_2} W_d(t_1)\, dt_1\, dt_2 \dots dt_m, \tag{3.2}$$

$$\frac{1}{\kappa^2(d) n^{2(d+m)}} \sum_{j=1}^{[nt]} X_j^2 \Rightarrow \int_0^t [W_{d,m}(s)]^2\, ds, \tag{3.3}$$

where $\kappa(d)$, $W_d(t)$ are defined as in Theorem 2.1 and

$$W_{d,m}(t) = \begin{cases} W_d(t), & \text{if } m = 1, \\ \displaystyle\int_0^t \int_0^{t_{m-1}} \dots \int_0^{t_2} W_d(t_1)\, dt_1\, dt_2 \dots dt_{m-1}, & \text{if } m \geq 2. \end{cases}$$

Proof. We first assume $m = 1$. Put

$$Y_j = X_j - X_{j-1}, \qquad j = 2, 3, \dots. \tag{3.4}$$

By definition of the lag operator, we have that

$$(1-B)^d Y_t = (1-B)^{d+1} X_t = u_t, \qquad t = 2, 3, \dots \tag{3.5}$$

and $X_j = X_1 + \sum_{i=2}^{j} Y_i$ for $j \geq 2$. Now Theorems 2.1 and 2.2 imply that, for $0 \leq t \leq 1$,

$$\frac{1}{\kappa(d) n^{1/2+d}} X_{[nt]} = \frac{1}{\kappa(d) n^{1/2+d}} X_1 + \frac{1}{\kappa(d) n^{1/2+d}} \sum_{i=2}^{[nt]} Y_i \Rightarrow W_d(t),$$

and hence (3.1) holds. By using (3.1) and the continuous mapping theorem, we obtain that (let $\sum_{i=1}^{[ns]} X_i = 0$ if $s < 1/n$)

$$\frac{1}{\kappa(d)n^{3/2+d}} \sum_{j=1}^{[nt]} X_j = \int_0^t \left(\frac{1}{\kappa(d)n^{1/2+d}} X_{[ns]} \right) ds \Rightarrow \int_0^t W_d(t_1)\, dt_1,$$

$$\frac{1}{\kappa^2(d)n^{2+2d}} \sum_{j=1}^{[nt]} X_j^2 = \int_0^t \left(\frac{1}{\kappa(d)n^{1/2+d}} X_{[ns]} \right)^2 ds \Rightarrow \int_0^t W_d^2(t_1)\, dt_1.$$

The relations (3.2) and (3.3) thus hold true for $m = 1$.

For general $m \geq 2$, Theorem 3.1 follows by induction, and details are omitted. This completes the proof of Theorem 3.1. ■

We next consider another application of our main results.

3.2. Fractional unit root distribution

Let $\{y_t\}$ be a stochastic process generated according to

$$y_t = \alpha y_{t-1} + X_t, \qquad t = 1,2,\ldots,n, \tag{3.6}$$

where $y_0 = 0$ and $\{X_t\}$ is a sequence of errors. Denote the LSE of α by $\hat{\alpha}_n$. We have that

$$n(\hat{\alpha}_n - 1) = \left\{ \sum_{t=1}^{n} y_{t-1}(y_t - y_{t-1}) \right\} \Big/ \left\{ n^{-1} \sum_{t=1}^{n} y_{t-1}^2 \right\}. \tag{3.7}$$

For the case where model (3.6) has a unit root (i.e., the null hypothesis $\alpha = 1$ holds), the limit distribution of $n(\hat{\alpha}_n - 1)$ was first considered by Dickey and Fuller (1979) under the assumption that X_t are i.i.d. random variables. Since then, considerable attention has been focused on weakening the i.i.d. assumption. Here we only cite Said and Dickey (1984), Phillips (1987), Hall (1989), and Chan and Tsay (1996). In these papers, the unit root distribution is obtained but only for the situation where the error process is a short memory process, such as an ARMA process. For similar results, more references can be found in Phillips and Xiao (1998), where the authors present a survey of unit root theory with an emphasis on testing principles and recent developments.

On weakening the assumption of i.i.d. errors, another important contribution is made by Sowell (1990). By assuming that the error process is a simple fractional process (i.e., $\psi_0 = 1, \psi_j = 0$, $j \geq 1$, and $m = 0$ in model (2.1)), Sowell (1990) establishes a well-known fractional unit root distribution[2] and points out that the asymptotics in this case significantly differ from those in the case of short memory errors. The results of Sowell (1990) have been extended to nonstationary fractionally integrated processes by Chan and Terrin (1995) and Tanaka (1999). With a Gaussian innovation, Chan and Terrin (1995) study the

general unstable AR unit root test, which extends those in Chan and Wei (1988), Parks and Phillips (1988, 1989), and Sims et al. (1990) to fractional cases.

In Theorem 3.2, which follows, as an application of Theorems 2.1, 2.2, and 3.1, we derive the limit distribution of $n(\hat{\alpha}_n - 1)$ while the error process X_t satisfies (2.1). Under quite weak moment conditions, this result provides a unified treatment of the previously most cited results. In particular, we point out that the limit distribution of $n(\hat{\alpha}_n - 1)$ is free to the choice of the weights ψ_k of the u_t in model (2.1) if $d + m > 0$.

THEOREM 3.2. *Let X_t satisfy (2.1) with $m \geq 0$ and y_t satisfy model (3.6). Let Condition A or Condition B hold.*

(a) If $m \geq 1$ or $m = 0$ and $0 < d < \frac{1}{2}$, then

$$n(\hat{\alpha}_n - 1) \Rightarrow \frac{1}{2}\,[W_{d,m+1}(1)]^2 \Big/ \int_0^1 [W_{d,m+1}(s)]^2\, ds, \tag{3.8}$$

where $W_{d,m}$ is defined as in Theorem 3.1.

(b) If $m = 0$ and $d = 0$, then

$$n(\hat{\alpha}_n - 1) \Rightarrow \frac{1}{2}\,[W(1)^2 - \gamma] \Big/ \int_0^1 [W(s)]^2\, ds, \tag{3.9}$$

where $\gamma = \sum_{k=0}^{\infty} \psi_k^2 / b_\psi^2$.

(c) If $m = 0$ and $d \in (-\frac{1}{2}, 0)$, then

$$n^{1+2d}(\hat{\alpha}_n - 1) \Rightarrow -\frac{M(d,\psi)}{2\kappa^2(d) \displaystyle\int_0^1 [W_d(s)]^2\, ds}, \tag{3.10}$$

where $\kappa^2(d)$ is defined as in Theorem 2.1 and

$$M(d,\psi) = \frac{E\epsilon_0^2}{2\pi} \int_{-\pi}^{\pi} |1 - e^{i\lambda}|^{-2d} |\psi(e^{-i\lambda})|^2\, d\lambda, \quad \textit{with } \psi(e^{-i\lambda}) = \sum_{k=0}^{\infty} \psi_k e^{-ik\lambda}.$$

Proof. By noting

$$\sum_{t=1}^{n} y_{t-1}(y_t - y_{t-1}) = \frac{1}{2} \sum_{t=1}^{n} \{y_t^2 - y_{t-1}^2 - (y_t - y_{t-1})^2\} = \frac{1}{2}\, y_n^2 - \frac{1}{2} \sum_{j=1}^{n} X_j^2,$$

we can rewrite $n(\hat{\alpha}_n - 1)$ as

$$n(\hat{\alpha}_n - 1) = \frac{1}{2\kappa^2(d) n^{1+2(d+m)}} \left\{ y_n^2 - \sum_{j=1}^{n} X_j^2 \right\} \Big/ \left\{ \frac{1}{n} \sum_{t=1}^{n-1} \left(\frac{y_t}{\kappa(d) n^{1/2+d+m}} \right)^2 \right\}. \tag{3.11}$$

Because $y_t = \sum_{j=1}^{t} X_j$, where X_j satisfy (2.1), it follows from Theorem 3.1 that

$$\left(\frac{1}{\kappa(d)n^{1/2+d+m}}\, y_{[nt]}\right)^2 \Rightarrow [W_{d,m+1}(t)]^2, \quad \text{for } m \geq 0; \tag{3.12}$$

$$\frac{1}{n}\sum_{t=1}^{n-1}\left(\frac{y_t}{\kappa(d)n^{1/2+d+m}}\right)^2 \Rightarrow \int_0^1 [W_{d,m+1}(s)]^2\, ds, \quad \text{for } m \geq 0; \tag{3.13}$$

$$\frac{1}{\kappa^2(d)n^{2(d+m)}}\sum_{j=1}^{[nt]} X_j^2 \Rightarrow \int_0^t [W_{d,m}(s)]^2\, ds, \quad \text{for } m \geq 1. \tag{3.14}$$

Because of (3.14), it is clear that if $m \geq 1$, then

$$\frac{1}{n^{1+2(d+m)}}\sum_{j=1}^{n} X_j^2 \to 0, \quad \text{in probability.} \tag{3.15}$$

On the other hand, if $m = 0$ in (2.1), it follows from Lemma 3.3 in the Appendix that $\{X_t, t \geq 1\}$ is a stationary linear process with

$$EX_1^2 = \frac{E\epsilon_0^2}{2\pi}\int_{-\pi}^{\pi} |1 - e^{i\lambda}|^{-2d}|\psi(e^{-i\lambda})|^2\, d\lambda < \infty, \quad \text{where } \psi(e^{-i\lambda}) = \sum_{k=0}^{\infty}\psi_k e^{-ik\lambda}.$$

Because stationary linear processes are ergodic, by using the stationary ergodic theorem (see Stout, 1974, p. 181), we obtain that

$$\frac{1}{n}\sum_{j=1}^{n} X_j^2 \to EX_1^2 = M(d,\psi) < \infty, \quad \text{a.s.} \quad \text{as } n \to \infty. \tag{3.16}$$

In particular, we point out that $EX_1^2 = E\epsilon_0^2 \sum_{k=0}^{\infty}\psi_k^2 < \infty$ if $d = 0$. Because (3.16) implies (3.15) when $m = 0$ and $0 < d < \frac{1}{2}$, the relation (3.8) follows immediately from (3.11)–(3.13), (3.15), (3.16) and the continuous mapping theorem.

Similarly, the relation (3.9) and (3.10) follow easily from (3.11)–(3.13), (3.16), and the fact that $W_{d,1}(s) = W_d(s)$ and $W_0(s) = W(s)$. This completes the proof of Theorem 3.2. ■

NOTES

1. A definition can be found later in this section. For a correction of Sowell's Theorem 2.2, see Theorem 2.1.1 of Liu (1998).

2. The fractional Brownian motion used as limiting process is insufficiently defined in Sowell (1990). For a correction of Sowell's Theorem 3.1, see Marinucci and Robinson (2000). Also, it can be found in Theorem 3.2 which follows.

REFERENCES

Avram, F. & M.S. Taqqu (1987) Noncentral limit theorems and Appell polynomials. *Annals of Probability* 15, 767–775.

Baillie, R.T. (1996) Long memory processes and fractional integration in econometrics. *Journal of Econometrics* 73, 5–59.

Beran, J. (1994) *Statistics for Long Memory Processes*. New York: Chapman and Hall.

Beran, J. (1995) Maximum likelihood estimation of the differencing parameter for invertible short and long memory *ARIMA* models. *Journal of the Royal Statistical Society, Series B* 57, 695–672.

Beran, J., R.J. Bhansali, & D. Ocker (1998) On unified model selection for stationary and nonstationary short and long memory autoregressive processes. *Biometrika* 85, 921–934.

Billingsley, P. (1968) *Convergence of Probability Measures*. New York: Wiley.

Brockwell, P.J. & R.A. Davis (1987) *Time Series: Theory and Methods*. New York: Springer-Verlag.

Chan, N.H. & N. Terrin (1995) Inference for unstable long-memory processes with applications to fractional unit root autoregressions. *Annals of Statistics* 23, 1662–1683.

Chan, N.H. & B.S. Tsay (1996) Asymptotic inference for non-invertible moving-average time series. *Journal of Time Series Analysis* 17, 1–17.

Chan, N.H. & C.Z. Wei (1988) Limiting distributions of least squares estimates of unstable autoregressive processes. *Annals of Statistics* 16, 367–401.

Chow, Y.S. (1965) Local convergence of martingales and the law of large numbers. *Annals of Mathematics and Statistics* 36, 552–558.

Csörgő, M. & J. Horváth (1993) *Weighted Approximation in Probability and Statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Statistics. Chichester: Wiley.

Dahlhaus, R. (1989) Efficient parameter estimation for self-similar processes. *Annals of Statistics* 17, 1749–1766.

Davidson, J. & R.M. Jong (2000) The functional central limit theorem and weak convergence to stochastic integrals, II: Fractionally integrated processes. *Econometric Theory* 16, 643–666.

Davydov, Y.A. (1970) The invariance principle for stationary processes. *Theory of Probability and Its Applications* 15, 487–498.

Dickey, D.A. & W.A. Fuller (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427–431.

Fox, R. & M.S. Taqqu (1986) Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Annals of Statistics* 14, 517–532.

Giraitis, L. & D. Surgailis (1990) A central limit theorem for quadratic forms in strongly dependent linear variables and its applications to the asymptotic normality of Whittle's estimate. *Probability Theory and Related Fields* 86, 87–104.

Gorodetskii, V.V. (1977) On convergence to semi-stable Gaussian process. *Theory of Probability and Its Applications* 22, 498–508.

Granger, C.W.J. & A. Joyeux (1980) An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis* 1, 15–29.

Hall, A.R. (1989) Testing for a unit root in the presence of moving average errors. *Biometrika* 76, 49–56.

Hall, P. & C.C. Heyde (1980) *Martingale Limit Theory and its Applications*. New York: Academic Press.

Hannan, E.J. (1979) The central limit theorem for time series regression. *Stochastic Processes and Their Applications* 9, 281–289.

Hosking, J.R.M. (1981) Fractional differencing. *Biometrika* 68, 165–176.

Hosking, J.R.M. (1984) Modeling persistence in hydrological time series using fractional differencing. *Water Resources Research* 20, 1898–1907.

Kokoszka, P.S. & M.S. Taqqu (1995) Fractional ARIMA with stable innovations. *Stochastic Processes and Their Applications* 60, 19–47.

Komlós, J., P. Major, & G. Tusnády (1975) An approximation of partial sums of independent R.V.'s and sample DF, part I. *Zeitschrift für Wahrscheinlichskeitstheorie und Verwandte Gebiete* 32, 111–131.

Komlós, J., P. Major, & G. Tusnády (1976) An approximation of partial sums of independent R.V.'s and sample DF, part II. *Zeitschrift für Wahrscheinlichskeitstheorie und Verwandte Gebiete* 34, 33–58.

Li, W.K. & A.I. McLeod (1986) Fractional time series modeling. *Biometrika* 73, 217–221.
Liu, M. (1998) Asymptotics of nonstationary fractional integrated series. *Econometric Theory* 14, 641–662.
Mandelbrot, B.B. & J.W. Van Ness (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Review* 10, 423–437.
Marinucci, D. & P.M. Robinson (1999) Alternative forms of fractional Brownian motion. *Journal of Statistical Planning and Inference* 80, 3–13.
Marinucci, D. & P.M. Robinson (2000) Weak convergence of multivariate fractional processes. *Stochastic Processes and Their Applications* 86, 103–120.
Mielniczuk, J. (1997) Long and short-range dependent sums of infinite-order moving averages and regression estimation. *Acta Sci. Math.* (Szeged). 63, 301–316.
Odaki, M. (1993) On the invertibility of fractionally differenced ARIMA processes. *Biometrika* 80, 703–709.
Parks, J.Y. & P.C.B. Phillips (1988) Statistical inference in regressions with integrated processes, part 1. *Econometric Theory* 4, 468–497.
Parks, J.Y. & P.C.B. Phillips (1989) Statistical inference in regressions with integrated processes, part 2. *Econometric Theory* 5, 95–131.
Phillips, P.C.B. (1987) Time series regression with a unit root. *Econometrica* 55, 277–302.
Phillips, P.C.B. & V. Solo (1992) Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.
Phillips, P.C.B. & Z. Xiao (1998) A Primer on Unit Root Testing. Manuscript.
Robinson, P.M. (1994) Time series with strong dependence. In C.A. Sims (ed.), *Advance in Econometrics, Sixth World Congress*. vol. 1, pp. 47–95. Cambridge, UK: Cambridge University Press.
Robinson, P.M. (1997) Large sample inference for nonparametric regression with dependent errors. *Annals of Statistics* 25, 2054–2084.
Said, S.E. & D.A. Dickey (1984) Testing for unit roots in autoregressive moving average models of unknown order. *Biometrika* 71, 599–607.
Samorodnitsky, G. & M.S. Taqqu (1994) *Stable Non-Gaussian Random Processes*. New York: Chapman and Hall.
Sims, C.A., J.H. Stock, & M.W. Watson (1990) Inference in linear time series models with some unit roots. *Econometrica* 58, 113–144.
Sowell, F.B. (1990) The fractional unit root distribution. *Econometrica* 58, 495–505.
Stout, W.F. (1974) *Almost Sure Convergence*. New York: Academic Press.
Tanaka, K. (1999) The nonstationary fractional unit root. *Econometric Theory* 15, 549–582.
Taqqu, M.S. (1975) Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Zeitschrift für Wahrscheinlichskeitstheorie und Verwandte Gebiete* 31, 287–302.
Wang, Q., Y.-X. Lin, & C.M. Gulati (2001) Asymptotics for moving average processes with dependent innovations. *Statistics and Probability Letters* 54, 347–356.
Wang, Q., Y.-X. Lin, & C.M. Gulati (2002) The invariance principle for linear processes with applications. *Econometric Theory* 18, 119–139.

APPENDIX: PROOFS OF THEOREMS 2.1 and 2.2

We first give several preliminary lemmas, which are also interesting in their own right.

Let $\{v_j, j = 0, \pm 1, \dots\}$ be a sequence of random variables, $\{a_{n,k}, k = 0, \pm 1, \pm 2, \dots\}$ be a triangular array of constants, and $A_n^2 = \sum_{k=-\infty}^{\infty} a_{n,k}^2$. For reading convenience, we give the following basic assumptions.

Assumption 1. The sequence $\{v_j^2\}$ is uniformly integrable.

Assumption 2. $E(v_j|\mathcal{F}_{j-1}) = 0$, $E(v_j^2|\mathcal{F}_{j-1}) = \sigma^2$, a.s., for $j = 0, \pm 1, \ldots$, where $\mathcal{F}_j$ is the σ-field of events generated by $\{v_i,\ i \le j\}$.

Assumption 3. $0 < A_n < \infty$ for each fixed $n \ge 1$, and as $n \to \infty$, $A_n \to \infty$ and $\max_k |a_{n,k}|/A_n \to 0$.

Assumption 4. There exists a positive constant C such that

$$\sup_{n\ge 1} \frac{1}{A_n} \sum_{j=-\infty}^{\infty} |a_{n,j} - a_{n,j-1}| \le C. \tag{A.1}$$

LEMMA 3.1. *Let Assumptions 1–4 hold,* $\sigma \sum_{j=0}^{\infty} \psi_j \neq 0$, *and* $\sum_{j=0}^{\infty} |\psi_j| < \infty$. *Then, as* $n \to \infty$,

$$\frac{1}{\sigma_n} \sum_{j=0}^{\infty} \psi_j Y_{nj} \to_d N(0,1) \quad \text{and} \quad \sigma_n^2 \sim A_n^2 b_0^2, \tag{A.2}$$

where $Y_{nj} = \sum_{k=-\infty}^{\infty} a_{nk} v_{k-j}$, $\sigma_n^2 = \mathrm{Var}(\sum_{j=0}^{\infty} \psi_j Y_{nj})$, *and* $b_0 = \sigma \sum_{j=0}^{\infty} \psi_j$.

Proof. Because $\sum_{i=0}^{\infty} |\psi_i| < \infty$, there exists a sequence of positive increasing constants λ_n such that

$$\sum_{j=\lambda_n}^{\infty} |\psi_j| \le A_n^{-3/2}, \qquad n = 1, 2, \ldots. \tag{A.3}$$

For this λ_n, we rewrite

$$\sum_{j=0}^{\infty} \psi_j Y_{nj} = \sum_{j=0}^{\lambda_n} \psi_j Y_{nj} + \sum_{j=\lambda_n+1}^{\infty} \psi_j Y_{nj}.$$

By a simple calculation, to prove (A.2), it suffices to show that

$$E\left(\sum_{j=\lambda_n+1}^{\infty} |\psi_j Y_{nj}| \right)^2 = o(1), \tag{A.4}$$

$$\frac{1}{\sigma_n^*} \sum_{j=0}^{\lambda_n} \psi_j Y_{nj} \to_d N(0,1), \tag{A.5}$$

and

$$\sigma_n^{*2} := \mathrm{Var}\left(\sum_{j=0}^{\lambda_n} \psi_j Y_{nj} \right) \sim A_n^2 b_0^2. \tag{A.6}$$

Note that relation (A.4) also implies that $\sum_{j=0}^{\infty} \psi_j Y_{nj}$ is well defined almost surely.

We next give the proofs of (A.4)–(A.6). It can be easily shown (cf. Chow, 1965) that

$$EY_{nj}^2 = \sigma^2 \sum_{k=-\infty}^{\infty} a_{nk}^2, \quad \text{for } j = 0, 1, 2, \ldots. \tag{A.7}$$

By Hölder's inequality, it follows from (A.3) and (A.7) that

$$E\left(\sum_{j=\lambda_n+1}^{\infty}|\psi_j Y_{nj}|\right)^2 \leq \sum_{j=\lambda_n+1}^{\infty}|\psi_j| \sum_{j=\lambda_n+1}^{\infty}|\psi_j| EY_{nj}^2$$
$$\leq \sigma^2 A_n^2 \left(\sum_{j=\lambda_n+1}^{\infty}|\psi_j|\right)^2 \leq \sigma^2 A_n^{-1}. \tag{A.8}$$

This implies (A.4) because $A_n \to \infty$ from Assumption 3.

To prove (A.5) and (A.6), put

$$B_n^2 = \sum_{k=-\infty}^{\infty} b_{nk}^2, \quad \text{where } b_{nk} = \sum_{j=0}^{\lambda_n} \psi_j a_{n,k+j}.$$

We have that for each fixed $n \geq 1$ (recalling that Y_{nj} are well defined),

$$\sum_{j=0}^{\lambda_n} \psi_j Y_{nj} = \sum_{j=0}^{\lambda_n} \psi_j \sum_{k=-\infty}^{\infty} a_{nk} v_{k-j}$$
$$= \sum_{k=-\infty}^{\infty} v_k \sum_{j=0}^{\lambda_n} \psi_j a_{n,k+j} = \sum_{k=-\infty}^{\infty} v_k b_{nk}, \tag{A.9}$$

$$\max_k |b_{nk}|/B_n \leq \sum_{j=0}^{\infty} |\psi_j| \max_k |a_{n,k}|/B_n \tag{A.10}$$

and similar to (A.7)

$$\sigma_n^{*2} = \operatorname{Var}\left(\sum_{j=0}^{\lambda_n} \psi_j Y_{nj}\right) = \sigma^2 B_n^2. \tag{A.11}$$

Because of (A.9)–(A.11) and Assumptions 1–3, tracing the proof of Lemma 3.1 given in Robinson (1997), (A.5) and (A.6) hold if we prove, as $n \to \infty$,

$$\frac{\sigma^2 B_n^2 - b_0^2 A_n^2}{A_n^2} \to 0, \quad \text{i.e., } \sigma_n^{*2} = \sigma^2 B_n^2 \sim A_n^2 b_0^2. \tag{A.12}$$

Because Hölder's inequality implies that for each $n \geq 1, i, j \geq 0$,

$$\sum_{k=-\infty}^{\infty} |a_{n,k+i} a_{n,k+j}| \leq \sum_{k=-\infty}^{\infty} a_{n,k}^2 < \infty,$$

elementary calculation shows that

$$B_n^2 = \sum_{k=-\infty}^{\infty} \sum_{i,j=0}^{\lambda_n} \psi_i \psi_j a_{n,k+i} a_{n,k+j}$$
$$= \sum_{i,j=0}^{\lambda_n} \psi_i \psi_j \sum_{k=-\infty}^{\infty} a_{n,k+i} a_{n,k+j} = \sum_{i,j=0}^{\lambda_n} \psi_i \psi_j \sum_{k=-\infty}^{\infty} a_{n,k} a_{n,k+j-i}.$$

Write

$$b_0^* = \sum_{i=0}^{\lambda_n} \psi_i \quad \text{and} \quad \eta_n^2 = A_n / \max_k |a_{n,k}|.$$

It follows that

$$\begin{aligned} B_n^2 - b_0^{*2} A_n^2 &= \sum_{i,j=0}^{\lambda_n} \psi_i \psi_j \sum_{k=-\infty}^{\infty} a_{n,k}(a_{n,k+j-i} - a_{n,k}) \\ &= \left(\sum_{|j-i|>\eta_n} + \sum_{|j-i|\le\eta_n} \right) \psi_i \psi_j \sum_{k=-\infty}^{\infty} a_{n,k}(a_{n,k+j-i} - a_{n,k}) \\ &= \Delta_{n1} + \Delta_{n2}, \quad \text{say.} \end{aligned} \tag{A.13}$$

In view of Assumptions 3 and 4, we have $\eta_n \to \infty$ and

$$\begin{aligned} |\Delta_{n1}| &\le \sum_{|j-i|>\eta_n} |\psi_i \psi_j| \left(\sum_{k=-\infty}^{\infty} a_{n,k}^2 \right)^{1/2} \left(\sum_{k=-\infty}^{\infty} |a_{n,k+j-i} - a_{n,k}|^2 \right)^{1/2} \\ &\le 4 \sum_{j=\eta_n}^{\infty} |\psi_j| \sum_{i=0}^{\infty} |\psi_i| \sum_{k=-\infty}^{\infty} a_{n,k}^2 = o(A_n^2). \end{aligned} \tag{A.14}$$

Taking account of the following inequality:

$$\max_{|j-i|\le\eta_n} |a_{n,k+j-i} - a_{n,k}| \le \sum_{t=-\eta_n}^{\eta_n} |a_{n,k+t} - a_{n,k+t-1}|,$$

$\max_k |a_{n,k}| = A_n/\eta_n^2$, and Assumption 4, we have

$$\begin{aligned} |\Delta_{n2}| &\le \sum_{|j-i|\le\eta_n} |\psi_i \psi_j| \sum_{k=-\infty}^{\infty} |a_{n,k}| \sum_{t=-\eta_n}^{\eta_n} |a_{n,k+t} - a_{n,k+t-1}| \\ &\le \left(\sum_{i=0}^{\infty} |\psi_i| \right)^2 \sum_{t=-\eta_n}^{\eta_n} \max_k |a_{n,k}| \sum_{k=-\infty}^{\infty} |a_{n,k+t} - a_{n,k+t-1}| \\ &= 2\eta_n^{-1} \left(\sum_{i=0}^{\infty} |\psi_i| \right)^2 A_n \sum_{k=-\infty}^{\infty} |a_{n,k} - a_{n,k-1}| = o(A_n^2). \end{aligned} \tag{A.15}$$

Therefore, using (A.13)–(A.15), we obtain that

$$\frac{B_n^2 - b_0^{*2} A_n^2}{A_n^2} \to 0, \quad \text{i.e.,}\ \sigma^2 B_n^2 \sim \sigma^2\, b_0^{*2}\, A_n^2. \tag{A.16}$$

Now, (A.12) follows immediately from (A.16) and

$$|b_0^2 - \sigma^2 b_0^{*2}| \le 2\sigma^2 \sum_{j=\lambda_n}^{\infty} |\psi_j| \sum_{j=0}^{\infty} |\psi_j| = o(1).$$

This also completes the proof of Lemma 3.1. ■

LEMMA 3.2. *Let* $c_k = \Gamma(d+k)/\Gamma(d)\Gamma(k+1)$ *for* $k \geq 0$ *and* $c_k = 0$ *for* $k < 0$, *where* $-\frac{1}{2} < d < \frac{1}{2}$. *Then,*

$$c_0 = 1, \qquad c_k \leq Ck^{d-1}, \quad \text{for } k \geq 1; \tag{A.17}$$

$$\max_k \sum_{i=1}^{n} c_{i+k} \leq C_1 \max\{1, n^d\}, \quad \text{for } d \neq 0; \tag{A.18}$$

$$|c_{n+k} - c_k| \leq 3Cnk^{d-2}, \quad \text{for all } 1 \leq n \leq k; \tag{A.19}$$

$$\sum_{k=-\infty}^{\infty} \left(\sum_{j=[ns]+1}^{[nt]} c_{j-k} \right)^2 \sim \frac{n^{1+2d}\Gamma(1-2d)(t-s)^{1+2d}}{(1+2d)\Gamma(1+d)\Gamma(1-d)}, \tag{A.20}$$

for $0 \leq s < t \leq 1$.

Proof. For the proof of (A.17), see Theorem 1 in Hosking (1981). The proof of (A.18) follows easily from (A.17). By noting $\Gamma(z+1) = z\Gamma(z)$ for all z, we have that for $1 \leq n \leq k$ and $d \in (-\frac{1}{2}, \frac{1}{2})$,

$$\begin{aligned} |c_{n+k} - c_k| &= c_k \left(1 - \frac{(k+n+d-1)\ldots(k+d)}{(k+n)\ldots(k+1)} \right) \\ &\leq c_k \left(1 - \frac{(k+d)(k+d+1)}{(k+n)(k+n-1)} \right) \\ &\leq \frac{c_k}{k^2} \{(k+n)^2 - (k+d)^2\} \leq 3Cnk^{d-2}, \end{aligned}$$

which implies (A.19).

To prove (A.20), let $\zeta_k, k = 0, \pm 1, \pm 2, \ldots$, be i.i.d. $N(0,1)$ random variables and $Y_j = \sum_{k=0}^{\infty} c_k \zeta_{j-k}$. Because $c_k = 0$ for $k < 0$ and hence

$$\sum_{j=[ns]+1}^{[nt]} Y_j = \sum_{k=-\infty}^{\infty} \zeta_k \sum_{j=[ns]+1}^{[nt]} c_{j-k},$$

clearly we have that for $0 \leq s < t \leq 1$

$$\sum_{k=-\infty}^{\infty} \left(\sum_{j=[ns]+1}^{[nt]} c_{j-k} \right)^2 = E\left(\sum_{j=[ns]+1}^{[nt]} Y_j \right)^2.$$

By noting that $Y_j, j \geq 1$ are stationary random variables, using Theorem 2.1 of Sowell (1990), we obtain

$$E\left(\sum_{j=[ns]+1}^{[nt]} Y_j \right)^2 = E\left(\sum_{j=1}^{[nt]-[ns]} Y_j \right)^2 \sim \frac{n^{1+2d}\Gamma(1-2d)(t-s)^{1+2d}}{(1+2d)\Gamma(1+d)\Gamma(1-d)},$$

where we use the estimate: $([nt] - [ns])/n \sim t - s$. Thus (A.20) follows. This also completes the proof of Lemma 3.2. ■

LEMMA 3.3. *Let X_t satisfy (2.1) with $m = 0$ and $\psi_j, j \geq 0$ satisfy (2.2). If $E\epsilon_0^2 < \infty$, then*

$$(a)\ X_t = \sum_{k=-\infty}^{\infty} c_{t-k} u_k = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \tag{A.21}$$

where c_k is as in Lemma 3.2 and $Z_t = \sum_{k=0}^{\infty} c_k \epsilon_{t-k}$.

(b) $\{X_t, t \geq 1\}$ is a strictly stationary random sequence with zero mean and

$$EX_1^2 = \frac{E\epsilon_0^2}{2\pi} \int_{-\pi}^{\pi} |1 - e^{i\lambda}|^{-2d} |\psi(e^{-i\lambda})|^2 \, d\lambda < \infty, \tag{A.22}$$

where $\psi(e^{-i\lambda}) = \sum_{k=0}^{\infty} \psi_k e^{-ik\lambda}$, in particular, $EX_1^2 = E\epsilon_0^2 \sum_{k=0}^{\infty} \psi_k^2 < \infty$ if $d = 0$.

Proof of Lemma 3.3. Writing $X_t = \Psi(B)\epsilon_t$, we have $\Psi(z) = (1 - z)^d \psi(z)$, where $\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j$. Since $\sum_{j=0}^{\infty} |\psi_j| < \infty$, similar to the proof of Theorem 2 part (a) given by Hosking (1981), the power series expansion of $\Psi(z)$ converges for all $|z| \leq 1$ when $d < \frac{1}{2}$. Thus, if $-\frac{1}{2} < d < \frac{1}{2}$, we have

$$X_t = (1 - B)^d \psi(B)\epsilon_t = \psi(B)(1 - B)^d \epsilon_t.$$

Now (A.21) follows from the binomial expansion of $(1 - z)^d$ (cf. Hosking, 1981).

As is well-known (e.g., see Hosking, 1981), $\{Z_t, t \geq 1\}$ is a strictly stationary random sequence with 0 mean. It follows from Theorem 3.5.3 of Stout (1974, p. 170) that $X_t = \sum_{j=0}^{\infty} \psi_j Z_{tj}$ has the same properties as Z_t. To prove (A.22), let $f_z(.)$ be a spectral density of $\{Z_t\}$. According to Theorem 12.4.1 in Brockwell and Davis (1987, p. 466),

$$f_Z(\Lambda) = \frac{E\epsilon_0^2}{2\pi} |1 - e^{i\lambda}|^{2d}, \quad \text{for } -\pi \leq \lambda \leq \pi.$$

This, together with the second equality of (A.21) and (4.4.3) in Brockwell and Davis (1987, p. 121), shows that $\{X_t\}$ has a spectral density $f_X(.)$ and $f_X(\lambda) = |\psi(e^{i\lambda})|^2 f_Z(\lambda)$, where $\psi(e^{i\lambda}) = \sum_{k=0}^{\infty} \psi_k e^{ik\lambda}$. In terms of $\sum_{j=0}^{\infty} |\psi_j| < \infty$, we obtain

$$EX_1^2 = \frac{E\epsilon_0^2}{2\pi} \int_{\pi}^{\pi} |1 - e^{i\lambda}|^{2d} |\psi(e^{i\lambda})|^2 \, d\lambda \leq \frac{E\epsilon_0^2}{2\pi} \left(\sum_{j=0}^{\infty} |\psi_j| \right)^2 \int_{\pi}^{\pi} |1 - e^{i\lambda}|^{2d} \, d\lambda < \infty.$$

If $d = 0$, the result is obvious since $X_1 = \sum_{k=0}^{\infty} \psi_k \epsilon_{1k}$. The proof of Lemma 3.3 is complete. ∎

After these preliminaries, we now give the proof of the main results.

Proof of Theorem 2.1. If $d = 0$, X_t reduces to summable linear processes. In this case, (2.3) follows from Hannan (1979). So, we assume that $d \neq 0$ in the sequel. Put

$$V_n(t) = \frac{1}{n^{1/2+d}} \sum_{j=1}^{[nt]} X_j, \qquad B_d(s,t) = \frac{1}{2} \{ s^{1+2d} + t^{1+2d} - |s - t|^{1+2d} \}.$$

In view of part (b) of Lemma 3.3, X_t is a strictly stationary random sequence with zero mean. Now, using the usual method in the proof of weak convergence for stationary random sequence (cf. Taqqu, 1975, Theorem 2.1), it suffices to show that

(i) for each fixed $l \geq 1$ and real constants $0 < t_1 \neq t_2 \neq \cdots \neq t_l \leq 1$,

$$\tau_1 V_n(t_1) + \cdots + \tau_l V_n(t_l) \to_d N(0, \sigma_1^2), \tag{A.23}$$

where $\tau_1, \tau_2, \ldots, \tau_l$ are any real constants and $\sigma_1^2 = \kappa^2(d) \sum_{i,j=1}^{l} \tau_i \tau_j B_d(t_i, t_j)$;

(ii) for some $a > 1/(1+2d)$,

$$ES_n^2 = O(n^{1+2d}) \quad \text{and} \quad E|S_n|^{2a} = O((ES_n^2)^a), \tag{A.24}$$

where $S_n = \sum_{j=1}^{n} X_j$.

We first prove part (i). From part (a) of Lemma 3.3, we have that

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \quad \text{where } Z_t = \sum_{k=0}^{\infty} c_k \epsilon_{t-k} \tag{A.25}$$

and c_k are defined as in Lemma 3.2. Let $m_i = [nt_i], i = 1, \ldots, l$. It follows from (A.25) that

$$\tau_1 V_n(t_1) + \cdots + \tau_l V_n(t_l) = \frac{1}{n^{1/2+d}} \sum_{i=1}^{l} \tau_i \sum_{t=1}^{m_i} X_t = \frac{1}{n^{1/2+d}} \sum_{j=0}^{\infty} \psi_j Y_{nj}, \tag{A.26}$$

where, by an elementary calculation (recalling $c_k = 0$ if $k < 0$),

$$\begin{aligned} Y_{nj} &= \sum_{i=1}^{l} \tau_i \sum_{t=1}^{m_i} Z_{t-j} \\ &= \sum_{i=1}^{l} \tau_i \sum_{t=1}^{m_i} \sum_{k=0}^{\infty} c_k \epsilon_{t-k-j} = \sum_{i=1}^{l} \tau_i \sum_{t=1}^{m_i} \sum_{k=-\infty}^{\infty} c_{t-k} \epsilon_{k-j} \\ &= \sum_{k=-\infty}^{\infty} b_{n,k} \epsilon_{k-j} \end{aligned} \tag{A.27}$$

with $b_{n,k} = \sum_{i=1}^{l} \tau_i \sum_{t=1}^{m_i} c_{t-k}$.

To apply Lemma 3.1 to (A.26), we first show that

$$B_n^2 \equiv \sum_{k=-\infty}^{\infty} b_{n,k}^2 \sim \frac{n^{1+2d} \Gamma(1-2d)}{(1+2d)\Gamma(1+d)\Gamma(1-d)} \sum_{i,j=1}^{l} \tau_i \tau_j B_d(t_i, t_j). \tag{A.28}$$

In fact, (A.28) follows immediately from Lemma 3.2 (see (A.20)) and

$$\begin{aligned} b_{n,k}^2 &= \sum_{i,j=1}^{l} \tau_i \tau_j \left(\sum_{s=1}^{m_i} c_{s-k} \right) \left(\sum_{t=1}^{m_j} c_{t-k} \right) \\ &= \frac{1}{2} \sum_{i,j=1}^{l} \tau_i \tau_j \left\{ \left(\sum_{s=1}^{m_i} c_{s-k} \right)^2 + \left(\sum_{t=1}^{m_j} c_{t-k} \right)^2 - \left(\sum_{s=m_i}^{m_j} c_{s-k} \right)^2 \right\}. \end{aligned}$$

From Lemma 3.2 (i.e., (A.17)–(A.19)), on the other hand, we have that for $d \neq 0$,

$$\max_k |b_{n,k}| \le \sum_{i=1}^{l} |\tau_i| \max_k \sum_{t=0}^{n} |c_{t-k}| \le C \max\{1, n^d\}; \tag{A.29}$$

$$\begin{aligned}\sum_{k=-\infty}^{\infty} |b_{n,k} - b_{n,k-1}| &= \sum_{k=-\infty}^{\infty} \left| \sum_{i=1}^{l} \tau_i \left(\sum_{t=1}^{m_i} c_{t-k} - \sum_{t=1}^{m_i} c_{t-k+1} \right) \right| \\ &\le \sum_{i=1}^{l} |\tau_i| \sum_{k=-\infty}^{\infty} |c_{1-k} - c_{m_i-k+1}| \\ &\le \sum_{i=1}^{l} |\tau_i| \left\{ \sum_{|k|\le n} (|c_{1-k}| + |c_{m_i+1-k}|) + \sum_{k=n+1}^{\infty} |c_{k+1} - c_{m_i+k+1}| \right\} \\ &\le \sum_{i=1}^{l} |\tau_i| \left\{ C \max\{1, n^d\} + C_1 n \sum_{k=n+1}^{\infty} k^{d-2} \right\} \\ &\le C_2 \max\{1, n^d\}. \end{aligned} \tag{A.30}$$

In view of (A.28)–(A.30), conditions of Lemma 3.1 hold for $b_{n,k}$ defined in (A.26) and (A.27). By applying (A.26), (A.28), and Lemma 3.1, we obtain that

$$\frac{1}{B_n^2 b_0^2} \sum_{j=0}^{\infty} \psi_j Y_{nj} \to_d N(0,1), \tag{A.31}$$

$$E\left(\sum_{i=1}^{l} \tau_i \sum_{j=1}^{m_i} X_j \right)^2 \sim B_n^2 b_0^2 \sim n^{1+2d} \kappa^2(d) \sum_{i,j=1}^{l} \tau_i \tau_j B_d(t_i, t_j). \tag{A.32}$$

The relation (A.31), together with (A.26) and (A.32), implies (A.23). This completes the proof of part (i).

We next prove part (ii). By applying (A.26) and (A.27) with $l = \tau_1 = t_1 = 1$, we obtain that

$$S_n = \sum_{j=1}^{n} X_j = \sum_{j=0}^{\infty} \psi_j Y_{nj}, \quad \text{where } Y_{nj} = \sum_{k=-\infty}^{\infty} b_{n,k} \epsilon_{k-j}$$

with $b_{n,k} = \sum_{t=1}^{n} c_{t-k}$. Furthermore, it follows from (A.32) that

$$ES_n^2 \sim b_0^2 \sum_{k=-\infty}^{\infty} b_{n,k}^2 \sim n^{1+2d} \kappa^2(d). \tag{A.33}$$

Thus the first relation of (A.24) holds.

If $0 < d < \frac{1}{2}$, the second relation of (A.24) is obvious by letting $a = 1$. To establish the second relation of (A.24) for $-\frac{1}{2} < d < 0$, we let $2a = (2+\delta)/(1+2d)$. Obviously, we have that $a > 1/(1+2d) > 1$ and $E|\epsilon_0|^{2a} < \infty$ when $d < 0$. By the Marcinkiewicz–Zygmund inequality (also Burkholder's inequality; see Hall and Heyde, 1980, p. 23) and Hölder's inequality, there exists a constant C_a depending only on a such that for all integers j and $s \le h$,

$$
\begin{aligned}
E\left|\sum_{k=s}^{h}\epsilon_k b_{n,k+j}\right|^{2a} &\le C_a E\left(\sum_{k=s}^{h}\epsilon_k^2 b_{n,k+j}^2\right)^a \\
&= C_a E\left(\sum_{k=s}^{h}\epsilon_k^2 |b_{n,k+j}|^{2/a}|b_{n,k+j}|^{(2a-2)/a}\right)^a \\
&\le C_a E\left\{\sum_{k=s}^{h}|\epsilon_k|^{2a} b_{n,k+j}^2\left(\sum_{k=s}^{h} b_{n,k+j}^2\right)^{a-1}\right\} \\
&\le C_a\left(\sum_{k=-\infty}^{\infty} b_{n,k}^2\right)^a E|\epsilon_0|^{2a}. \qquad \textbf{(A.34)}
\end{aligned}
$$

Because of (A.33) and (A.34), it follows from the Fatou lemma that for all j,

$$
\begin{aligned}
E|Y_{nj}|^{2a} &= E\left|\sum_{k=-\infty}^{\infty}\epsilon_k b_{n,k+j}\right|^{2a} \\
&\le C_a E\left|\lim_{h\to\infty}\sum_{k=1}^{h}\epsilon_k b_{n,k+j}\right|^{2a} + C_a E\left|\lim_{s\to\infty}\sum_{k=-s}^{0}\epsilon_k b_{n,k+j}\right|^{2a} \\
&\le C_a \limsup_{h\to\infty} E\left|\sum_{k=1}^{h}\epsilon_k b_{n,k+j}\right|^{2a} + C_a \limsup_{s\to\infty} E\left|\sum_{k=-s}^{0}\epsilon_k b_{n,k+j}\right|^{2a} \\
&\le C\left(\sum_{k=-\infty}^{\infty} b_{nk}^2\right)^a = O((ES_n^2)^a).
\end{aligned}
$$

Hence, by Hölder's inequality again, we obtain that

$$
\begin{aligned}
E|S_n|^{2a} &\le E\left(\sum_{j=0}^{\infty}|\psi_j|^{(2a-1)/(2a)}|\psi_j|^{1/(2a)}|Y_{nj}|\right)^{2a} \\
&\le \left(\sum_{j=0}^{\infty}|\psi_j|\right)^{2a-1}\sum_{j=0}^{\infty}|\psi_j|E|Y_{nj}|^{2a} = O((ES_n^2)^a),
\end{aligned}
$$

which implies the desired result. This also completes the proof of Theorem 2.1. ∎

Proof of Theorem 2.2. We only need to consider the case where $-\frac{1}{2} < d < 0$. In this case, because ϵ_k are i.i.d. random variables with $E\epsilon_0 = 0$ and $E|\epsilon_0|^{2/(1+2d)} < \infty$, by applying Komlós, Major, and Tusnády (1975, 1976) (also see Csörgő and Horváth, 1993 Corollary 1.1), on a suitable probability space $(\Omega, \mathcal{F}, \mathcal{P})$, we can construct $\eta_k, k = 0, \pm 1, \pm 2, \ldots$, which are i.i.d. $N(0, E\epsilon_0^2)$ random variables such that as $n \to \infty$

$$
\max_{1\le m\le n}\left|\sum_{j=1}^{m}\epsilon_j - \sum_{j=1}^{m}\eta_j\right| = o(n^{1/2+d}), \quad \text{a.s.}, \qquad \textbf{(A.35)}
$$

$$
\max_{1\le m\le n}\left|\sum_{j=0}^{m}\epsilon_{-j} - \sum_{j=0}^{m}\eta_{-j}\right| = o(n^{1/2+d}), \quad \text{a.s.} \qquad \textbf{(A.36)}
$$

Let Y_t satisfy

$$(1-B)^d\, Y_t = u_t, \qquad u_t = \sum_{j=0}^{\infty} \psi_j\, \eta_{t-j}, \qquad t = 1,2,\ldots,$$

where $-\frac{1}{2} < d < 0$ and ψ_k satisfies (2.4). Because η_k are i.i.d. $N(0, E\epsilon_0^2)$, by applying Theorem 2.1, we have that, for $0 \le t \le 1$,

$$\frac{1}{\kappa(d) n^{1/2+d}} \sum_{j=1}^{[nt]} Y_j \Rightarrow W_d(t). \tag{A.37}$$

In view of (A.37) and Theorem 1.4.1 of Billingsley (1968), to prove Theorem 2.2, it suffices to show that

$$\sup_{0\le t\le 1} \left| \sum_{j=1}^{[nt]} X_j - \sum_{j=1}^{[nt]} Y_j \right| = o_P(n^{1/2+d}), \tag{A.38}$$

where $o_P(.)$ denotes convergence in probability.

By applying part (a) of Lemma 3.3, for any $m \ge 1$, we can write

$$\begin{aligned}\sum_{j=1}^{m} (X_j - Y_j) &= \sum_{j=1}^{m} \sum_{s=0}^{\infty} \psi_s \sum_{k=0}^{\infty} c_k(\epsilon_{j-k-s} - \eta_{j-k-s}) \\ &= \sum_{j=1}^{m} \sum_{s=0}^{\infty} \psi_s \sum_{k=s}^{\infty} c_{k-s}(\epsilon_{j-k} - \eta_{j-k}) \\ &= A_{1n}^m + A_{2n}^m + A_{3n}^m + A_{4n}^m, \end{aligned} \tag{A.39}$$

where $c_k = \Gamma(d+k)/\Gamma(d)\Gamma(k+1)$ for $k \ge 0$ and $c_k = 0$ for $k < 0$;

$$A_{1n}^m = \sum_{j=1}^{m} \sum_{s=n+1}^{\infty} \psi_s \sum_{k=s}^{\infty} c_{k-s}(\epsilon_{j-k} - \eta_{j-k}),$$

$$A_{2n}^m = \sum_{j=1}^{m} \sum_{s=0}^{n} \psi_s \sum_{k=s}^{2n} c_{k-s}(\epsilon_{j-k} - \eta_{j-k}),$$

$$A_{3n}^m = \sum_{j=1}^{m} \sum_{s=0}^{n} \psi_s \sum_{k=2n+1}^{n^2} c_{k-s}(\epsilon_{j-k} - \eta_{j-k}),$$

$$A_{4n}^m = \sum_{j=1}^{m} \sum_{s=0}^{n} \psi_s \sum_{k=n^2+1}^{\infty} c_{k-s}(\epsilon_{j-k} - \eta_{j-k}).$$

Clearly, (A.38) follows if

$$\max_{1\le m\le n} |A_{jn}^m| = o_P(n^{1/2+d}), \quad \text{for } j = 1,2,3,4. \tag{A.40}$$

We next prove (A.40). Write

$$V_{l,j,s} = \sum_{k=l}^{\infty} c_{k-s}(\epsilon_{j-k} - \eta_{j-k}).$$

By noting $c_0 = 0$ and $|c_k| \le Ck^{d-1}$ for $k \ge 1$ (see Lemma 3.2), it can be easily shown that for all $l \ge s \ge 0$ and $j \ge 1$,

$$E|V_{l,j,s}| \le (EV_{l,j,s}^2)^{1/2} \le C\left(\sum_{k=l}^{\infty} c_{k-s}^2\right)^{1/2} \le C_1 \min\{1, l^{d-1/2}\}.$$

Therefore, it follows that

$$E \max_{1\le m\le n} |A_{1n}^m| \le \sum_{j=1}^{n} \sum_{s=n+1}^{\infty} |\psi_s|\, E|V_{s,j,s}| \le Cn^{1/2+d} \sum_{s=n+1}^{\infty} s^{1/2-d}|\psi_s|. \qquad \textbf{(A.41)}$$

Similarly, we obtain that

$$E \max_{1\le m\le n} |A_{4n}^m| \le \sum_{j=1}^{n} \sum_{s=0}^{n} |\psi_s|\, E|V_{n^2+1,j,s}| \le Cn(n^2)^{d-1/2} \sum_{s=0}^{\infty} |\psi_s| \le C_1 n^{2d}. \qquad \textbf{(A.42)}$$

In view of (2.4), (A.41), (A.42), and $d < 0$, Markov's inequity implies that (A.40) holds for $j = 1$ and $j = 4$.

On the other hand, it follows from (A.35) and (A.36) that

$$\begin{aligned}
\max_{1\le m\le n} |A_{2n}^m| &\le \sum_{s=0}^{n} |\psi_s| \sum_{k=s}^{2n} |c_{k-s}| \max_{\substack{1\le m\le n\\ 0\le k\le 2n}} \left|\sum_{j=1}^{m} (\epsilon_{j-k} - \eta_{j-k})\right| \\
&\le C\left(\max_{1\le m\le n}\left|\sum_{j=1}^{m}(\epsilon_j - \eta_j)\right| + \max_{0\le k\le n}\left|\sum_{j=0}^{2k}(\epsilon_{-j} - \eta_{-j})\right|\right) \\
&= o(n^{1/2+d}), \quad \text{a.s.}
\end{aligned}$$

This implies that (A.40) holds for $j = 2$.

We now prove (A.40) for $j = 3$. For convenience, write $S_k = \sum_{i=0}^{k}(\epsilon_{-i} - \eta_{-i})$. We have that

$$\begin{aligned}
A_{3n}^m &= \sum_{s=0}^{n} \psi_s \sum_{k=2n+1}^{n^2} c_{k-s} \sum_{j=1}^{m} (S_{k-j} - S_{k-j-1}) \\
&= \sum_{s=0}^{n} \psi_s \sum_{k=2n+1}^{n^2} c_{k-s}(S_{k-1} - S_{k-m-1}). \qquad \textbf{(A.43)}
\end{aligned}$$

Clearly, it follows that

$$\begin{aligned}
&\sum_{k=2n+1}^{n^2} c_{k-s}(S_{k-1} - S_{k-m-1}) \\
&\quad = \sum_{k=2n}^{n^2-1} c_{k+1-s}S_k - \sum_{k=2n-m}^{n^2-m-1} c_{k+1+m-s}S_k \\
&\quad = \sum_{k=n^2-m}^{n^2-1} c_{k+1-s}S_k - \sum_{k=2n-m}^{2n-1} c_{k+1+m-s}S_k + \sum_{k=2n}^{n^2-m-1} (c_{k+1-s} - c_{k+1+m-s})S_k \\
&\quad = I_{1ns}^m + I_{2ns}^m + I_{3ns}^m, \quad \text{say.} \qquad \textbf{(A.44)}
\end{aligned}$$

Recalling (A.36), $|c_k| \le Ck^{d-1}$, and $d < 0$, we get that

$$\max_{\substack{1\le m\le n\\0\le s\le n}} |I_{1ns}^m| \le C \sum_{k=n^2-n}^{n^2-1} (k-n)^{d-1} \max_{0\le k\le n^2} |S_k|$$
$$\le Cn(n^2)^{d-1}(n^2)^{1/2+d} = o(n^{1/2+d}), \quad \text{a.s.} \tag{A.45}$$

Similarly, we have

$$\max_{\substack{1\le m\le n\\0\le s\le n}} |I_{2ns}^m| = o(n^{1/2+d}), \quad \text{a.s.} \tag{A.46}$$

On the other hand, by applying Lemma 3.2, we know that (noting $m \le k - s$)

$|c_{k+1-s} - c_{k+1+m-s}| \le Cm(k-s)^{d-2}$.

Therefore, it follows from (A.36) that (recalling $d < 0$)

$$\max_{\substack{1\le m\le n\\0\le s\le n}} |I_{3ns}^m| \le Cn \sum_{k=2n}^{n^2} (k-n)^{d-2} k^{1/2+d} = o(n^{1/2+d}), \quad \text{a.s.} \tag{A.47}$$

In view of (A.43)–(A.47), we have that

$$\max_{1\le m\le n} |A_{3n}^m| \le \max_{\substack{1\le m\le n\\0\le s\le n}} (|I_{1ns}^m| + |I_{2ns}^m| + |I_{3ns}^m|) \sum_{s=0}^{n} |\psi_s| = o(n^{1/2+d}), \quad \text{a.s.}$$

This implies (A.40) for $j = 3$. This also completes the proof of Theorem 2.2. ∎