

# Asymptotic normality for weighted sums of linear processes <sup>1</sup>

K. Abadir, W. Distaso, L. Giraitis, H.L. Koul  
Imperial College London, Imperial College London,  
Queen Mary, University of London, Michigan State University

## Abstract

We establish asymptotic normality of weighted sums of linear processes with general triangular array weights and when the innovations in the linear process are martingale differences. The results are obtained under minimal conditions on the weights and innovations. We also obtain weak convergence of weighted partial sum processes. The results are applicable to linear processes that have short or long memory or exhibit seasonal long memory behavior. In particular, they are applicable to GARCH and ARCH( $\infty$ ) models and to their squares. They are also useful in deriving asymptotic normality of kernel type estimators of a nonparametric regression function with short or long memory moving average errors.

## 1 Introduction

Numerous inference procedures in statistics and econometrics are based on the sums

$$S_n = \sum_{j=1}^n X_j, \quad W_n = \sum_{j=1}^n z_{nj} X_j$$

of a linear process  $\{X_j\}$ , where  $\{z_{nj}, 1 \leq j \leq n\}$  is an array of known real numbers. In this paper we focus on deriving asymptotic distributions of  $S_n$  and  $W_n$ . In addition, the weak convergence property of the corresponding partial sum processes is also discussed. The linear process is assumed to be a moving average with martingale differences innovations, and may exhibit short or long range dependence. It will be shown that  $\{\text{Var}(T_n)\}^{-1/2}(T_n - ET_n)$ , with  $T_n = S_n$  or  $W_n$ , converges weakly to a normal distribution under easily verifiable minimal assumptions on weights and innovations. In particular, these assumptions are valid for nonlinear squared ARCH process where innovations are conditionally heteroscedastic martingale differences. The proofs use central limit theorem for martingale differences.

Numerous testing procedures, e.g. testing for unit root, CUSUM change-point detection and KPSS test for stationarity are based on the weak convergence of the partial sum process  $S_{[nt]}$ ,  $0 \leq t \leq 1$  to a Gaussian process, and, in particular, require verification of the CLT for  $S_n$ . The CLT for weighted sums  $W_n$ , in turn, is used in kernel estimation and in spectral analysis, e.g. obtaining asymptotic normality of the discrete Fourier transforms of  $\{X_j\}$ .

---

<sup>1</sup>Research in part supported by ESRC grant RES062230790 and USA NSF DMS Grant 0704130.

*Keywords:* Linear process, weighted sum, Lindeberg-Feller.

It is thus of interest to provide easy to use CLTs and invariance principles for the above statistics.

The book of Ibragimov and Linnik (1971) contains a number of useful results on classical asymptotic theory of weakly dependent random variables. Davydov (1970) obtained weak convergence result for the partial sum process of linear processes with i.i.d. innovations whereas Phillips and Solo (1992) developed CLT and invariance principles for sums of linear processes based on Beveridge-Nelson decomposition. Peligrad and Utev (2006) extended Ibragimov and Linnik (1971, Theorem 18.6.5) for linear processes with innovations following a more general dependence framework. Gordin (1969) introduced a general method for proving central limit theorems for stationary processes using martingale approximation. In the case of short memory, his method gives the same result as the Beveridge-Nelson decomposition. Wu and Woodroffe (2004) obtained a CLT for the sums of stationary and ergodic sequences using martingale approximation method. Merlevède, Peligrad and Utev (2006) provide a further survey of some recent results on CLT and its weak invariance principle for stationary processes.

Section 2 deals with asymptotic normality of  $S_n$  and  $W_n$ , whereas in Section 3 we discuss weak convergence of the corresponding partial sum processes. Section 4 contains examples and applications, while Section 5 includes simulations. In the sequel, all limits are taken as  $n \rightarrow \infty$ , unless specified otherwise,  $\rightarrow_p$  and  $\rightarrow_D$ , respectively, denote the convergence in probability and in distribution,  $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$ , and  $\mathcal{N}_k(\mu, \Sigma)$  denotes the  $k$ -dimensional normal distribution with mean vector  $\mu$  and covariance matrix  $\Sigma$ ,  $k \geq 1$ . We write  $\mathcal{N}$  for  $\mathcal{N}_1$ . For any two sequences of real numbers,  $a_n b_n$ ,  $a_n \sim b_n$  means  $a_n/b_n \rightarrow 1$ .

## 2 CLT for weighted sums

In this section we consider asymptotic normality of the sums  $S_n$  and  $W_n$  when  $\{X_j\}$  is a moving average process

$$(2.1) \quad X_j = \sum_{k=0}^{\infty} a_k \zeta_{j-k} = \sum_{k=-\infty}^j a_{j-k} \zeta_k, \quad j \in \mathbb{Z}, \quad \sum_{k=0}^{\infty} a_k^2 < \infty.$$

Here,  $(\zeta_j, \mathcal{F}_j)$ ,  $j \in \mathbb{Z}$  is a martingale difference sequence (m.d.s.) with constant variance, where  $\mathcal{F}_j := \sigma\text{-field}\{\zeta_i, i \leq j\}$ ,  $j \in \mathbb{Z}$ , i.e.,  $E(\zeta_j | \mathcal{F}_{j-1}) = 0$ ,  $E\zeta_j^2 = \sigma_\zeta^2 < \infty$ ,  $j \in \mathbb{Z}$ .

The existing literature on asymptotic distributions of the sums of  $X_j$ 's often assumes that  $\{\zeta_j\} \sim \text{IID}(0, \sigma_\zeta^2)$ , i.e., the innovations  $\zeta_j$ ,  $j \in \mathbb{Z}$  are i.i.d. r.v.'s with zero mean and finite and positive variance  $\sigma_\zeta^2$ . Several papers establish a CLT for  $S_n$  under the weaker assumption where  $\{\zeta_j\}$  is a stationary ergodic m.d.s. In some applications even this assumption is too restrictive.

Let  $V_j := E(\zeta_j^2 | \mathcal{F}_{j-1})$  be the conditional variance of  $\{\zeta_j\}$ , and  $\gamma_V(j, k) := \text{Cov}(V_j, V_k)$ ,

$j, k \in \mathbb{Z}$  stand for covariance function of  $V_j$ 's. To allow for broader applications we make the following

**Assumption 2.1**  $\{(\zeta_j, \mathcal{F}_j), j \in \mathbb{Z}\}$  is a m.d.s. of r.v.'s such that  $E\zeta_j^2 = \sigma_\zeta^2$ , for all  $j \in \mathbb{Z}$ ,

- (a)  $\max_j EV_j^2 < \infty$ ,
- (b)  $\Delta_K := \max_{|j-k| \geq K} |\gamma_V(j, k)| \rightarrow 0$ , as  $K \rightarrow \infty$ ,
- (c)  $\max_{j \in \mathbb{Z}} E\zeta_j^2 I(|\zeta_j| > K) \rightarrow 0$ , as  $K \rightarrow \infty$ .

First, we discuss the asymptotic normality of  $S_n$ . Theorem 18.6.5 of Ibragimov and Linnik (1971) gives a CLT for  $S_n$  in the case of i.i.d. innovations under general conditions. Theorem 2.1 below extends it to m.d. innovations and shows that any rate of divergence  $\text{Var}(S_n) \rightarrow \infty$  of the variance guaranties the CLT. Peligrad and Utev (2006, Proposition 4) proved this result when  $\{\zeta_j\}$  is a stationary and ergodic m.d.s. For clarity, we provide a brief proof based on the ideas of these works.

Note that if  $\{X_j\}$  is a zero-mean Gaussian process, then  $(\text{Var}(S_n))^{-1/2} S_n =_D \mathcal{N}(0, 1)$  for all  $n \geq 1$ , and the question about the asymptotic distribution of  $S_n$  reduces to finding the asymptotics of the  $\text{Var}(S_n)$ .

**Theorem 2.1** Suppose  $\{X_j\}$  is a linear process (2.1) where  $\{\zeta_j\}$  is either a stationary and ergodic m.d.s., or satisfies Assumption 2.1. Then

$$\sigma_n^2 := \text{Var}(S_n) \rightarrow \infty,$$

implies

$$(2.2) \quad \sigma_n^{-1} S_n \rightarrow_D \mathcal{N}(0, 1).$$

**Proof.** For simplicity of notation, set  $a_k = 0$ ,  $k = -1, -2, \dots$  in (2.1), and let  $c_{nj} = \sigma_n^{-1} \sum_{k=\max(j,1)}^n a_{k-j} = \sigma_n^{-1} \sum_{k=1}^n a_{k-j}$ ,  $j \in \mathbb{Z}$ . Without loss of generality, assume  $\sigma_\zeta^2 = 1$ . Then,

$$(2.3) \quad \sigma_n^{-1} S_n = \sum_{j=-\infty}^n c_{nj} \zeta_j, \quad \sigma_n^{-2} \text{Var}(S_n) = \sum_{j=-\infty}^n c_{nj}^2 = 1, \quad \forall n \geq 1.$$

Next we show

$$(2.4) \quad \sum_{j=-\infty}^n (c_{nj} - c_{n,j-1})^2 = o(1), \quad c_n := \sup_{j \leq n} |c_{nj}| = o(1),$$

which together with Lemma 2.1 below implies (2.2).

To prove (2.4), note that for any  $s \leq n$ ,  $t = 1, 2, \dots$ ,

$$\begin{aligned} \sum_{j=s-t}^s c_{nj}^2 &= \sum_{j=s-t}^s \left( c_{n,j-1} + (c_{nj} - c_{n,j-1}) \right)^2, \\ c_{n,s}^2 &= c_{n,s-t-1}^2 + 2 \sum_{j=s-t}^s c_{n,j-1} (c_{nj} - c_{n,j-1}) + \sum_{j=s-t}^s (c_{nj} - c_{n,j-1})^2. \end{aligned}$$

Because  $\sum_{j \leq n} c_{n,j}^2 = 1$ , then for every  $s \in \mathbb{Z}$  and  $n \geq 1$ ,  $\lim_{t \rightarrow \infty} c_{n,s-t-1}^2 = 0$ . Since  $t$  is arbitrary, take the limit  $t \rightarrow \infty$  and use the Cauchy-Schwarz inequality, to obtain

$$(2.5) \quad \begin{aligned} c_{n,s}^2 &= 2 \sum_{j=-\infty}^s c_{n,j-1}(c_{nj} - c_{n,j-1}) + \sum_{j=-\infty}^s (c_{nj} - c_{n,j-1})^2 \\ &\leq 2 \left( \sum_{j=-\infty}^n c_{nj}^2 \right)^{1/2} B_n + B_n^2 \leq 2B_n + B_n^2, \end{aligned}$$

where  $B_n := \left\{ \sum_{j \in \mathbb{Z}} (c_{nj} - c_{n,j-1})^2 \right\}^{1/2}$  does not depend on  $s$ . Since by definition  $c_{nj} - c_{n,j-1} = \sigma_n^{-1}(a_{1-j} - a_{n-j+1})$ ,

$$B_n^2 \leq 4\sigma_n^{-2} \sum_{j=0}^{\infty} a_j^2 = o(1),$$

because  $\sigma_n^2 \rightarrow \infty$  and  $\sum_{j=0}^{\infty} a_j^2 < \infty$ , proving (2.4).  $\square$

**Remark 2.1** The approach of Beveridge-Nelson decomposition used in Phillips and Solo (1992, Theorem 3.15) allows one to obtain the CLT for a linear process  $\{X_j\}$  with  $a_k$ 's satisfying  $\sum_{k=0}^{\infty} k^2 a_k^2 < \infty$ , and  $\zeta_j$  being a uniformly integrable m.d.s. and satisfying  $n^{-1} \sum_{k=1}^n E[\zeta_k^2 | \mathcal{F}_{k-1}] \rightarrow \sigma_\zeta^2$  a.s. In contrast, Theorem 2.1 gives the same result under much weaker condition on  $a_k$ 's.

Compared to Proposition 4 of Peligrad and Utev (2006), the above Theorem 2.1 carries more explanatory character demonstrating simple technical tools for the proof of such CLT and relaxing the assumption of stationarity and ergodicity on the m.d.s.  $\{\zeta_j\}$ .

The next lemma provides some sufficient conditions for the CLT of the weighted sums of m.d.s. with the weights being a triangular array of real numbers.

**Lemma 2.1** *Suppose  $S_n = \sum_{j=-\infty}^n d_{nj} \zeta_j$ ,  $n \geq 1$ , where  $\{\zeta_j\}$  is a standardized m.d.s., and  $\{d_{nj}\}$  are such that  $\text{Var}(S_n) = \sum_{j=-\infty}^n d_{nj}^2 = 1$ , for all  $n \geq 1$ .*

(i) *In addition, if  $\{\zeta_j\}$  satisfies Assumption 2.1 and*

$$(2.6) \quad d_n := \sup_{j \leq n} |d_{nj}| = o(1),$$

*then*

$$(2.7) \quad S_n \rightarrow_D \mathcal{N}(0, 1).$$

(ii) *In addition, if  $\{\zeta_j\}$  is stationary ergodic and*

$$(2.8) \quad \sum_{j \leq n} (d_{nj} - d_{n,j-1})^2 = o(1),$$

*then (2.6) and (2.7) hold.*

**Proof.** Since  $\sum_{j=-\infty}^n d_{nj}^2 = 1$ , for a fixed  $n \geq 1$  we can choose  $M = M(n) \rightarrow \infty$  such that

$$(2.9) \quad \sum_{j=-\infty}^{-M-1} d_{nj}^2 \leq 1/\log(n), \quad n \geq 1.$$

Write

$$S_n = \sum_{j=-\infty}^{-M-1} d_{nj} \zeta_j + \sum_{j=-M}^n d_{nj} \zeta_j =: s_{n,1} + s_{n,2}, \quad \text{say.}$$

Then  $Es_{n,1}^2 = \sum_{j=-\infty}^{-M-1} d_{nj}^2 \leq 1/\log(n) \rightarrow 0$  implies  $s_{n,1} = o_p(1)$ . To prove (2.7), it remains to show

$$(2.10) \quad s_{n,2} \rightarrow_D \mathcal{N}(0, 1).$$

We consider cases (i) and (ii) separately.

*Case (i).* To show (2.10), by the CLT for m.d.s., see Hall and Heyde (1980, Corollary 3.1), it suffices to check the following two conditions:

$$(2.11) \quad \sum_{j=-M}^n d_{nj}^2 E[\zeta_j^2 | \mathcal{F}_{j-1}] \rightarrow_p 1,$$

$$(2.12) \quad \sum_{j=-M}^n E[|d_{nj} \zeta_j|^2 I(|d_{nj} \zeta_j| \geq \delta) | \mathcal{F}_{j-1}] \rightarrow_p 0, \quad \forall \delta > 0.$$

Let  $q_n := \sum_{j=-M}^n d_{nj}^2 V_j$  denote the left hand side of (2.11). We shall show that

$$Eq_n \rightarrow 1, \quad \text{Var}(q_n) \rightarrow 0,$$

which yields  $E(q_n - 1)^2 \rightarrow 0$  and, together with the Chebyshev's inequality, will imply (2.11). The first claim follows from  $EV_j = E\zeta_j^2 = 1$ , (2.9) and  $\sum_{j=-\infty}^n d_{nj}^2 = 1$ . To prove the second claim, set  $K = K(n) = \lceil 1/d_n \rceil$ . By the Hölders inequality,

$$\begin{aligned} \text{Var}(q_n) &= \sum_{j,k=-M}^n d_{nj}^2 d_{nk}^2 \gamma_V(j, k) \leq \sum_{j,k=-M: |j-k| > K}^n [\dots] + \sum_{j,k=-M: |j-k| \leq K}^n [\dots] \\ &\leq \Delta_K \left( \sum_{j=-\infty}^n d_{nj}^2 \right)^2 + d_n^2 \sum_{j=-M}^n d_{nj}^2 \sum_{k: |j-k| \leq K} \{|\gamma_V(j, j) \gamma_V(k, k)|\}^{1/2} \\ &\leq \Delta_K + \Delta_0 (2K + 1) d_n^2 \rightarrow 0, \end{aligned}$$

by where  $\Delta_0 := \max_{j \in \mathbb{Z}} EV_j^2 < \infty$ , by Assumption 2.1(a,b) and (2.6), which completes the proof of (2.11).

To prove (2.12), note that the expected value of the l.h.s. of (2.14) is bounded above by

$$\sum_{j=-M}^n d_{nj}^2 E(\zeta_j^2 I(|\zeta_j| \geq d_n^{-1} \delta)) \leq \max_j E(\zeta_j^2 I(|\zeta_j| \geq d_n^{-1} \delta)) \rightarrow 0,$$

by Assumption 2.1(c) and (2.6), which together with the Markov inequality implies (2.12), and also completes the proof of the lemma.

*Case (ii).* The proof of (2.10) combines arguments used in the proofs of Peligrad and Utev (1997, Theorem 2.1; 2006, Proposition 4). Let  $X_{nj} := d_{nj} \zeta_j$ ,  $-M \leq j \leq n$ . By Hall and Heyde (1980, Theorem 3.2) to prove (2.10), it suffices to verify (a)  $\max_{-M \leq j \leq n} |X_{nj}| \rightarrow_p 0$ , (b)  $\sum_{j=-M}^n X_{nj}^2 \rightarrow_p 1$  and (c)  $E \max_{-M \leq j \leq n} X_{nj}^2 = O(1)$ .

First, we show that (2.8) implies (2.6). To see this, use the bound analogous to (2.5), which does not depend on any particular form of  $c_{nj}$ , to obtain  $d_{n,s}^2 \leq 2B_n + B_n^2$ , where now  $B_n := \{ \sum_{j \in \mathbb{Z}} (d_{nj} - d_{n,j-1})^2 \}^{1/2}$ . Hence, (2.8) implies  $d_n \rightarrow 0$ .

Now, claims (a) and (c) follow because for any  $\varepsilon > 0$ ,

$$E \sum_{j=-M}^n (d_{nj} \zeta_j)^2 I(|d_{nj} \zeta_j| \geq \varepsilon) \leq E[\zeta_1^2 I(|\zeta_1| \geq d_n^{-1} \varepsilon)] \sum_{j=-\infty}^n d_{nj}^2 \rightarrow 0.$$

To show (b), we need to verify

$$(2.13) \quad \sum_{j=-M}^n d_{nj}^2 \zeta_j^2 \rightarrow_p 1.$$

Letting  $k \geq 1$ , write

$$\begin{aligned} \sum_{j=-M}^n d_{nj}^2 \zeta_j^2 &= \sum_{j=-M}^n d_{nj}^2 + \sum_{j=-M}^n d_{nj}^2 \left\{ \zeta_j^2 - k^{-1} \sum_{l=1}^k \zeta_{j+l}^2 \right\} \\ &\quad + \sum_{j=-M}^n d_{nj}^2 \left\{ k^{-1} \sum_{l=1}^k \zeta_{j+l}^2 - 1 \right\} =: q_{n1} + q_{n2} + q_{n3}. \end{aligned}$$

In view of (2.9),  $q_{n1} = \sum_{j=-\infty}^n d_{nj}^2 + O(1/\log n) \rightarrow 1$ . Since  $E \zeta_j^2 = 1$ , then

$$\begin{aligned} E|q_{n2}| &\leq k^{-1} E \left| \sum_{j=-M}^n k d_{nj}^2 \zeta_j^2 - \sum_{l=1}^k \sum_{j=-M+l}^{n+l} d_{n,j-l}^2 \zeta_j^2 \right| \\ &\leq k^{-1} \sum_{j=-\infty}^n |k d_{nj}^2 - d_{n,j-1}^2 - \cdots - d_{n,j-k}^2| E[\zeta_j^2] + d_n E[\zeta_{n+1}^2 + \cdots + \zeta_{n+k}^2 + \zeta_{-M}^2] \\ &\leq k \sum_{j=-\infty}^n |d_{nj}^2 - d_{n,j-1}^2| + (k+1) d_n^2 \rightarrow 0, \end{aligned}$$

because  $\sum_{j=-\infty}^n |d_{nj}^2 - d_{n,j-1}^2| \leq (\sum_{j=-\infty}^n (d_{nj} - d_{n,j-1})^2)^{1/2} (2 \sum_{j=-\infty}^n d_{nj}^2)^{1/2} \rightarrow 0$ , by (2.8).

Finally, since  $\{\zeta_j\}$  is stationary and ergodic, by the Ergodic Theorem; see e.g. Stout (1974, Corollary 3.5.2),  $E|k^{-1} \sum_{l=1}^k \zeta_{j+l}^2 - E\zeta_1^2| = E|k^{-1} \sum_{l=1}^k \zeta_l^2 - 1| \rightarrow 0$ ,  $k \rightarrow \infty$ , and thus  $E|q_{n3}| \leq o(1) \sum_{j=-\infty}^n d_{nj}^2 = o(1)$ , which implies (2.13) and completes the proof.  $\square$

**Remark 2.2** Note that the above lemma clearly includes the case of  $\{\zeta_j\} \sim \text{IID}(0, 1)$ . Furthermore, it gives a generalization of Theorem V.1.2.1 of Hájek and Šidák (1967) where  $d_{nj} = 0$ ,  $j \leq 0$  and  $\{\zeta_j\} \sim \text{IID}(0, 1)$ . If  $\zeta_j \sim \text{IID}(0, 1)$ , assumption (2.6) implies the weak Lindeberg condition, see Hall and Heyde (1980, p. 53), and is minimal. Indeed, let  $S_n = \sum_{j=1}^n d_{nj} \zeta_j$ ,  $d_{nj} = (1 + cn^{-\alpha})^{-j} n^{-\alpha}$ ,  $c > 0$ . If  $\alpha > 0$ , then  $\max_{1 \leq j \leq n} |d_{nj}| = o(1)$  and the CLT holds, as also shown in Phillips and Magdalinos (2007). However, if  $\alpha = 0$ , then  $\max_{1 \leq j \leq n} |d_{nj}| = (1 + c)^{-1}$ , (2.6) does not hold and  $S_n \rightarrow_D \sum_{j=1}^{\infty} (1 + c)^{-j} \zeta_j$ , which is Gaussian only if  $\{\zeta_j\}$  is Gaussian, as pointed out in Anderson (1959). Peligrad and Utev (1997, Theorem 2.1) derived Lemma 2.1, requiring instead of Assumption 2.1 that the m.d.s.  $\{\zeta_j\}$  be a pairwise mixing sequence. In their Example 2.1 they showed that (2.6) and stationarity and ergodicity of m.d.s. alone are not sufficient for the CLT.

The next corollary provides a straightforward generalization of Lemma 2.1 to a two-sided moving average of m.d.s., while the result for weighted sums of a stationary ergodic process stated in Proposition 2.1 is useful in various applications.

**Corollary 2.1** *Suppose  $S_n = \sum_{j=-\infty}^{\infty} d_{nj} \zeta_j$ ,  $n \geq 1$ , where  $\{\zeta_j\}$  is a m.d.s., and  $\{d_{nj}\}$  are such that  $\sum_{j \in \mathbb{Z}} d_{nj}^2 = 1$ .*

(i) *In addition, if  $\{\zeta_j\}$  satisfies Assumption 2.1 and  $\max_{j \in \mathbb{Z}} |d_{nj}| = o(1)$ , then  $S_n \rightarrow_D \mathcal{N}(0, \sigma_\zeta^2)$ .*

(ii) *In addition, if  $\{\zeta_j\}$  is stationary ergodic and  $\sum_{j \in \mathbb{Z}} (d_{nj} - d_{n,j-1})^2 = o(1)$ , then  $S_n \rightarrow_D \mathcal{N}(0, \sigma_\zeta^2)$ , and*

$$(2.14) \quad \sum_{j \in \mathbb{Z}} d_{nj}^2 \zeta_j^2 = E\zeta_1^2 + o_p(1).$$

**Proposition 2.1** *Suppose  $S_n = \sum_{j=-n}^n z_{nj} \eta_j$ ,  $n \geq 1$ , where  $\{\eta_j\}$  is a stationary ergodic sequence,  $E|\eta_1| < \infty$ , and  $\{z_{nj}\}$  are such that*

$$(2.15) \quad \sum_{j=2}^n |z_{nj} - z_{n,j-1}| + |z_{n1}| = o\left(\sum_{j=1}^n |z_{nj}|\right).$$

*Then*

$$(2.16) \quad \sum_{j=1}^n z_{nj} \eta_j = E\eta_1 \left( \sum_{j=1}^n z_{nj} \right) + o_p\left(\sum_{j=1}^n |z_{nj}|\right).$$

**Proof.** Since  $|z_{nj}| \leq \sum_{j=2}^j |z_{nj} - z_{n,j-1}| + |z_{n1}|$ ,  $j = 2, \dots, n$ , then by (2.15),  $\max_{1 \leq j \leq n} |z_{nj}| = o(\sum_{j=1}^n |z_{nj}|)$ , and (2.16) follows by the same argument as in the proof of (2.13).  $\square$

**Weighted sums.** We now turn to the asymptotic normality of  $W_n$ . Let  $z_{nj}$ ,  $j, n \geq 1$  be an array of real numbers, and consider the weighted sums

$$(2.17) \quad W_n = \sum_{j=1}^n z_{nj} X_j.$$

The following theorem gives three sufficient (non-equivalent) conditions for the verification of the CLT for  $W_n$  for different types of weights  $z_{nj}$ . Subsequently, they will be strengthened in Proposition 2.2 for easy verification. Here,  $\sigma_n^2 := \text{Var}(W_n)$ .

**Theorem 2.2** *Suppose  $\{X_j\}$  is a linear process (2.1) with m.d. innovations  $\{\zeta_j\}$ . Suppose the weights  $\{z_{nj}\}$  in  $W_n$ ,  $\{a_j\}$  and  $\{\zeta_j\}$  in (2.1) satisfy one of the following three conditions:*

(i) *m.d.s.  $\{\zeta_j\}$  satisfies Assumption 2.1,*

$$\max_{1 \leq j \leq n} |z_{nj}| = o(\sigma_n), \quad \text{and} \quad \sum_{j=1}^n z_{nj}^2 \leq C\sigma_n^2.$$

(ii) *m.d.s.  $\{\zeta_j\}$  satisfies Assumption 2.1,*

$$\max_{1 \leq j \leq n} |z_{nj}| = o(\sigma_n), \quad \text{and} \quad \sum_{j=0}^{\infty} |a_j| < \infty.$$

(iii) *m.d.s.  $\{\zeta_j\}$  is either stationary and ergodic, or satisfies Assumption 2.1, and*

$$|z_{n1}| + |z_{nn}| + \sum_{j=2}^n |z_{nj} - z_{n,j-1}| = o(\sigma_n).$$

Then,

$$(2.18) \quad \sigma_n^{-1} W_n \rightarrow_D \mathcal{N}(0, 1).$$

**Proof.** Similarly as in (2.3), set  $a_k = 0$ ,  $k < 0$ , and let  $d_{nj} := \sigma_n^{-1} \sum_{k=\max(j,1)}^n z_{nk} a_{k-j} \equiv \sigma_n^{-1} \sum_{k=1}^n z_{nk} a_{k-j}$ ,  $j \in \mathbb{Z}$ . Then

$$(2.19) \quad \sigma_n^{-1} W_n = \sum_{j=-\infty}^n d_{nj} \zeta_j.$$

Recall  $\text{Var}(\sigma_n^{-1} W_n) = \sum_{j=-\infty}^n d_{nj}^2 v_j^2 = 1$ . We will verify that in cases (i) and (ii),  $d_{nj}$ 's satisfy (2.6), whereas in case (iii) (2.8) holds, which by Lemma 2.1 proves (2.18).

*Case (i).* Clearly, here  $K_n := \sigma_n / \max_{1 \leq j \leq n} |z_{nj}| \rightarrow \infty$  and

$$\begin{aligned} |d_{nj}| &\leq \sigma_n^{-1} \sum_{k=1}^n |z_{nk} a_{k-j}| I(|k-j| \geq K_n) \\ &\quad + \sigma_n^{-1} \sum_{k=1}^n |z_{nk} a_{k-j}| I(|k-j| < K_n) := q_{n,1j} + q_{n,2j}, \quad \text{say.} \end{aligned}$$



By the Cauchy-Schwarz inequality,

$$\begin{aligned}
q_{n,1j} &\leq \sigma_n^{-1} \left( \sum_{k=1}^n z_{nk}^2 \right)^{1/2} \left( \sum_{k=1}^n a_{k-j}^2 I(|k-j| \geq K_n) \right)^{1/2} \leq C \sum_{i \geq K_n} a_i^2, \\
q_{n,2j} &\leq \sigma_n^{-1} \max_{1 \leq k \leq n} |z_{nk}| \sum_{k=1}^n |a_{k-j}| I(|k-j| < K_n) \\
&\leq K_n^{-1} (2K_n)^{1/2} \left( \sum_{i=0}^{\infty} a_i^2 \right)^{1/2} \leq CK_n^{-1/2}, \quad \forall j \leq n.
\end{aligned}$$

Hence, assumption (i) and  $\sum_{i=0}^{\infty} a_i^2 < \infty$  yield

$$\max_{j \leq n} |d_{nj}| \leq C \left( \sum_{i \geq K_n} a_i^2 + K_n^{-1/2} \right) \rightarrow 0,$$

thereby proving (2.6).

*Case (ii):* Here, (2.6) follows because

$$\max_{1 \leq j \leq n} |d_{nj}| \leq \sigma_n^{-1} \max_{1 \leq j \leq n} |z_{nj}| \sum_{l=0}^{\infty} |a_l| \leq C \sigma_n^{-1} \max_{1 \leq j \leq n} |z_{nj}|.$$

*Case (iii):* To verify (2.8), define for simplicity  $z_{n0} = z_{n,n+1} = 0$ . Then one can write

$$d_{nj} - d_{n,j-1} = \sigma_n^{-1} \sum_{k=1}^n z_{nk} (a_{k-j} - a_{k+1-j}) = \sigma_n^{-1} \sum_{k=1}^{n+1} (z_{nk} - z_{n,k-1}) a_{k-j}.$$

Hence,

$$\begin{aligned}
\sum_{j \in \mathbb{Z}} (d_{nj} - d_{n,j-1})^2 &= \sigma_n^{-2} \sum_{k,s=1}^{n+1} (z_{nk} - z_{n,k-1})(z_{ns} - z_{n,s-1}) \sum_{j \in \mathbb{Z}} a_{k-j} a_{s-j} \\
&\leq \sigma_n^{-2} \sum_{k,s=1}^{n+1} |z_{nk} - z_{n,k-1}| |z_{ns} - z_{n,s-1}| \sum_{j=0}^{\infty} a_j^2 \\
&\leq C \sigma_n^{-2} \left( |z_{n1}| + |z_{nn}| + \sum_{k=2}^n |z_{nk} - z_{n,k-1}| \right)^2 \rightarrow 0,
\end{aligned}$$

by condition (iii) of the proposition. This completes the proof of the theorem.  $\square$

**Remark 2.3** Assumption (ii) of Theorem 2.2 can be applied in case of short and negative memory linear processes  $\{X_j\}$  that satisfy  $\sum_{j=0}^{\infty} |a_j| < \infty$ , whereas condition (i) is useful in the case when  $\{X_j\}$  has long or short memory. Condition on  $z_{nj}$  in (iii) is stronger than in (i) but allows m.d.s.  $\{\zeta_j\}$  to be stationary and ergodic, which is especially tractable.

**Remark 2.4** The CLT results of this paper can be applied to the sum  $W_n$  of a non-stationary process  $\{X_j\}$  for which the first differences  $Y_j := X_j - X_{j-1}$  form a linear process, by rewriting  $W_n = \sum_{j=1}^n z_{nj} X_j = \sum_{k=1}^n \{\sum_{j=k}^n z_{nj}\} Y_k + X_0 \sum_{j=1}^n z_{nj}$ . This reduces the problem to the CLT for the weighed sum  $W_n$  of  $Y_j$ 's with weights  $z'_{nk} := \sum_{j=k}^n z_{nj}$ .

In applications, the verification of conditions for asymptotic normality of  $W_n$  in Theorem 2.2 often reduces to analyzing the asymptotic behavior of the variance  $\sigma_n^2 \equiv \text{Var}(W_n) = \sum_{j,k=1}^n z_{nj} \gamma_X(j-k) z_{nk}$ , since the remaining conditions on weights  $z_{nj}$  are usually easy to verify. In the next proposition we provide stronger sufficient conditions in terms of the spectral density  $f$  of  $\{X_j\}$  and weights  $z_{nj}$  for analyzing the asymptotic behavior of  $\sigma_n^2$  and validating the conditions of Theorem 2.2. In particular, condition (2.20) on weights is mild and satisfied in most of the applications. Part a) requires  $f$  to be only continuous at 0, with no restrictions on  $f$  at higher frequencies. It is satisfied by spectral densities of ARMA and seasonal GARMA models. Part b) allows the spectral density to be unbounded at the origin, i.e. to have long memory. Part c) focuses on the case when  $f$  is bounded away from 0 in the whole spectrum, which includes the case of long memory and seasonal long memory models; see, e.g., Granger and Joyeux (1980), Hosking (1981) and Gray, Zhang and Woodward (1989). Parts (a) and (b) are applicable when m.d.s.  $\{\zeta_j\}$  is stationary and ergodic.

**Proposition 2.2** *Let  $\{X_j\}$ ,  $z_{nj}$ , and  $W_n$  be as in (2.1) and (2.17). Assume that m.d.s.  $\{\zeta_j\}$  is either stationary and ergodic, or satisfies Assumption 2.1. Then the following holds.*

(a) *Suppose  $f(u) \rightarrow f(0)$ ,  $u \rightarrow 0$ ,  $0 < f(0) < \infty$ , and*

$$(2.20) \quad |z_{n1}| + |z_{nn}| + \sum_{j=2}^n |z_{nj} - z_{n,j-1}| = o\left(\left(\sum_{j=1}^n z_{nj}^2\right)^{1/2}\right).$$

*Then*

$$(2.21) \quad \text{Var}(W_n) \sim 2\pi f(0) \sum_{j=1}^n z_{nj}^2.$$

*Moreover, the conditions of Theorem 2.2(iii) are satisfied, and CLT (2.18) holds.*

(b) *If  $f(u) \geq c > 0$ ,  $|u| \leq u_0$ , for some  $c > 0$  and  $u_0 > 0$ , and (2.20) holds, then the conditions of Theorem 2.2(iii) are satisfied, and CLT (2.18) holds.*

(c) *Suppose Assumption 2.1 is satisfied, there exists  $c > 0$  such that  $f(u) \geq c > 0$ ,  $u \in \Pi$ , and*

$$(2.22) \quad \max_{1 \leq j \leq n} |z_{nj}| = o\left(\left(\sum_{j=1}^n z_{nj}^2\right)^{1/2}\right).$$

*Then the conditions of Theorem 2.2(i) are satisfied, and CLT (2.18) holds.*

**Proof.** (a) Let  $G(u) := \sum_{j=1}^n e^{-iju} z_{nj}$ ,  $u \in \Pi$ . Since  $\gamma_X(k) = \int_{\Pi} e^{iku} f(u) du$ , then

$$(2.23) \quad \begin{aligned} \sigma_n^2 &= \sum_{j,k=1}^n z_{nj} \gamma_X(j-k) z_{nk} = \int_{\Pi} f(u) |G(u)|^2 du \\ &= 2\pi f(0) \sum_{j=1}^n z_{nj}^2 + i_n, \quad i_n := \sigma_n^2 - 2\pi f(0) \sum_{j=1}^n z_{nj}^2. \end{aligned}$$

It remains to show that

$$(2.24) \quad i_n = \left| \sigma_n^2 - 2\pi f(0) \sum_{j=1}^n z_{nj}^2 \right| = o\left( \sum_{j=1}^n z_{nj}^2 \right),$$

which proves (2.21), and together with (2.20) verifies condition (iii) Theorem 2.2 and thus (2.18).

Verification of (2.24) is based on two facts

$$(2.25) \quad \int_{\Pi} |G(u)|^2 du = 2\pi \sum_{j=1}^n z_{nj}^2, \quad \sup_{|u| \leq \pi} |uG(u)| = o\left( \sum_{j=1}^n z_{nj}^2 \right),$$

shown below. We proceed as follows. Let  $\epsilon > 0$ . Choose  $\delta > 0$ , such that  $\sup_{0 \leq u \leq \delta} |f(u) - f(0)| \leq \epsilon$ . Then by (2.25),

$$\begin{aligned} i_n &\leq \int_{|u| \leq \pi} |f(u) - f(0)| |G(u)|^2 du \\ &\leq \epsilon \int_{|u| \leq \delta} |G(u)|^2 du + \int_{\delta < |u| \leq \pi} |f(u) - f(0)| \delta^{-2} |uG(u)|^2 du \\ &\leq \epsilon 2\pi \sum_{j=1}^n z_{nj}^2 + o\left( \sum_{j=1}^n z_{nj}^2 \right) \int_{\delta < |u| \leq \pi} |f(u) - f(0)| du \\ &\leq \epsilon 2\pi \sum_{j=1}^n z_{nj}^2 + o\left( \sum_{j=1}^n z_{nj}^2 \right), \end{aligned}$$

since  $\int_{|u| \leq \pi} f(u) du = E\zeta_j^2 = \sigma_\zeta^2$  and  $f(0) < \infty$ , which yields (2.24).

To show the first claim of (2.25), use  $\int_{\Pi} e^{isu} du = 0$  for  $s \neq 0$ , to obtain

$$\int_{\Pi} |G(u)|^2 du = \int_{\Pi} \sum_{j,k=1}^n e^{i(j-k)u} z_{nj} z_{nk} du = 2\pi \sum_{j=1}^n z_{nj}^2.$$

To show the second claim, use summation by parts, to write

$$G(u) = \sum_{j=1}^{n-1} \left( \sum_{l=1}^j e^{-ilu} \right) (z_{nj} - z_{n,j+1}) + z_{nn} \sum_{l=1}^n e^{-ilu}.$$

For  $j = 1, \dots, n$ , one can bound

$$\left| \sum_{l=1}^j e^{ilu} \right| = \left| \frac{\sin(ju/2)}{\sin(u/2)} \right| \leq |\sin(nu/2)|^{-1} \leq \pi u^{-1}, \quad |u| \leq \pi.$$

Therefore,  $|uG(u)| \leq \pi \left( \sum_{j=1}^{n-1} |z_{nj} - z_{n,j+1}| + |z_{nn}| \right) = o\left(\sum_{j=1}^n z_{nj}^2\right)$ , by (2.20), which completes the proof of (2.25) and the part (a) of the proposition.

(b) By assumption of the proposition and (2.25),

$$\begin{aligned} \sigma_n^2 &= \int_{\Pi} f(u) |G(u)|^2 du \geq c \int_{|u| \leq u_0} |G(u)|^2 du \\ &= c \left\{ \int_{|u| \leq u_0} |G(u)|^2 du - \int_{u_0 < |u| \leq \pi} |G(u)|^2 du \right\} = c 2\pi (1 + o(1)) \sum_{j=1}^n z_{nj}^2. \end{aligned}$$

This together with (2.20) verifies (iii) of Theorem 2.2 which implies (2.18).

(c) Assumption  $f(u) \geq c > 0$  thus implies

$$(2.26) \quad \sigma_n^2 \geq c \int_{\Pi} |G(u)|^2 du = 2\pi c \sum_{j=1}^n z_{nj}^2,$$

which together with (2.22) yields (i) of Theorem 2.2 which implies (2.18) and completes the proof of the proposition.  $\square$

**Remark 2.5** Proposition 2.2 verifies the CLT for “smooth” weights  $z_{nj}$  and stationary and ergodic m.d.s.  $\{\zeta_j\}$ , e.g., in Corollary 4.4 below, it is shown that (2.20) holds for kernel weights  $z_{nj} = K((nx-j)/(nb))$  in nonparametric regression setups. As long as  $f$  is continuous at zero, (2.20) yields the asymptotic behavior of the variance (2.21) and the asymptotic normality of  $W_n$ .

If (2.20) does not hold, Theorem 2.2(i, ii) may be applied. For instance, alternating weights  $z_{nj} = (-1)^j = e^{\pm i\pi j}$  do not satisfy (2.20), but Theorem 2.2(i) is applicable. Indeed,  $\max_j |z_{nj}| = 1$  and  $\sum_{j=1}^n z_{nj}^2 = n$ . In addition, if  $f(u) \rightarrow f(\pi) > 0$ ,  $u \rightarrow \pi$ , then, letting  $D_n(u) := \sum_{l=1}^n e^{ilu}$ ,

$$\begin{aligned} \text{Var}(W_n) &= \sum_{j,k=1}^n z_{nj} \gamma_X(j-k) z_{nk} = \int_{\Pi} |D_n(u+\pi)|^2 f(u) du \\ &= \int_0^\pi f(u-\pi) |D_n(u)|^2 du + \int_{-\pi}^0 f(u+\pi) |D_n(u)|^2 du \\ &= f(\pi) \int_{-\pi}^\pi |D_n(u)|^2 du + o(n) = 2\pi f(\pi)n + o(n), \end{aligned}$$

using  $f(-\pi) = f(\pi)$ , similarly as in the proof of (2.21). This shows the applicability of Theorem 2.2(i), and hence the CLT for  $W_n$ . Note that here (2.21) for  $\text{Var}(W_n)$  does not hold.

The following proposition, that is valid for any short memory covariance-stationary process  $\{X_j\}$ , provides an upper bound for  $\text{Var}(W_n)$  and analyzes its asymptotic behavior.

**Proposition 2.3** *Let  $\{X_j\}$  be a covariance-stationary process with zero mean, finite variance and covariance function  $\gamma$  such that*

$$(2.27) \quad \sum_{k \in \mathbb{Z}} |\gamma(k)| < \infty.$$

Let  $W_n$  be as in (2.17) with  $z_{nj}$  satisfy conditions (2.20) and (2.22). Then,

$$(2.28) \quad \begin{aligned} \left( \sum_{j=1}^n z_{nj}^2 \right)^{-1} \text{Var}(W_n) &\leq \sum_{k \in \mathbb{Z}} |\gamma(k)|, \\ \left( \sum_{j=1}^n z_{nj}^2 \right)^{-1} \text{Var}(W_n) &\rightarrow \sigma^2 := \sum_{k \in \mathbb{Z}} \gamma(k). \end{aligned}$$

**Proof.** Under (2.27),

$$\begin{aligned} f(u) &= (2\pi)^{-1} \sum_{k \in \mathbb{Z}} e^{-iku} \gamma(k) \leq (2\pi)^{-1} \sum_{k \in \mathbb{Z}} |\gamma(k)| < \infty, \quad u \in \Pi, \\ f(u) &\rightarrow f(0) = (2\pi)^{-1} \sum_{k \in \mathbb{Z}} \gamma(k) = (2\pi)^{-1} \sigma^2, \quad u \rightarrow 0. \end{aligned}$$

Thus, by (2.23) and (2.25),

$$\begin{aligned} \text{Var}(W_n) &= \int_{\Pi} f(u) |G(u)|^2 du \leq \sup_{u \in \Pi} f(u) \int_{\Pi} |G(u)|^2 du \\ &= \sum_{k \in \mathbb{Z}} |\gamma(k)| \sum_{j=1}^n z_{nj}^2, \end{aligned}$$

which proves the first bound of (2.28). The proof of the second bound is the same as that of (2.21). This completes the proof of proposition.  $\square$

The following result is a multivariate generalization of Theorem 2.2.

**Theorem 2.3** *Let  $z_{n,j}^{(i)}$ ,  $i = 1, \dots, k$  be  $k$  arrays of real weights, and  $\{X_j\}$  is a linear process (2.1) with m.d.s.  $\{\zeta_j\}$ . Assume that sums  $W_n^{(i)} := \sum_{j=1}^n z_{n,j}^{(i)} X_j$  and  $(\sigma_n^{(i)})^2 := \text{Var}(W_n^{(i)})$ ,  $i = 1, \dots, k$  satisfy one of conditions (a) or (b).*

(a)  $\{\zeta_j\}$  is stationary and ergodic, and each sum  $W_n^{(i)}$ ,  $i = 1, \dots, k$  satisfies condition (iii) of Theorem 2.2.

(b)  $\{\zeta_j\}$  satisfies Assumption 2.1 and each sum  $W_n^{(i)}$ ,  $i = 1, \dots, k$ , satisfies one of the conditions of (i) - (iii) of Theorem 2.2.

Let for some (positive definite) matrix  $\Sigma$ ,

$$(2.29) \quad \left( \text{Cov}(W_n^{(i)}/\sigma_n^{(i)}, W_n^{(j)}/\sigma_n^{(j)}) \right)_{i,j=1,\dots,k} \rightarrow \Sigma.$$

Then

$$(2.30) \quad \left( W_n^{(1)}/\sigma_n^{(1)}, \dots, W_n^{(k)}/\sigma_n^{(k)} \right) \rightarrow_D \mathcal{N}_k(0, \Sigma).$$

**Proof.** Similarly as in (2.19), write

$$S_n^{(i)} := W_n^{(i)}/\sigma_n^{(i)} = \sum_{j=-\infty}^n d_{nj}^{(i)} \zeta_j, \quad i = 1, \dots, k.$$

To prove (2.30), in view of the Cramér-Wold device, it suffices to show that for every  $a = (a_1, \dots, a_k) \in \mathbb{R}^k$ ,

$$(2.31) \quad S_n := a_1 S_n^{(1)} + \dots + a_k S_n^{(k)} \rightarrow_D N(0, a \Sigma a^T).$$

Write  $S_n = \sum_{j=-\infty}^n d_{nj} \zeta_j$  where  $d_{nj} = a_1 d_{nj}^{(1)} + \dots + a_k d_{nj}^{(k)}$ . Condition (2.29) implies that

$$\text{Var}(S_n) = E \zeta_1^2 \sum_{j=-\infty}^n d_{nj}^2 \rightarrow a \Sigma a^T.$$

Assume that m.d.s. satisfies Assumption 2.1. As seen in the proof of Theorem 2.2, any one of the conditions (i), (ii) or (iii) assures that  $\sup_{-\infty < j \leq n} |d_{nj}^{(i)}| \rightarrow 0$ ,  $i = 1, \dots, k$ , which yields  $\sup_{-\infty < j \leq n} |d_{nj}| \rightarrow 0$ . Hence coefficients  $d_{nj}$  of  $S_n$  satisfy assumptions of Lemma 2.1 which implies (2.31).

Assume that m.d.s. is stationary ergodic. In the proof of Theorem 2.2, it was shown that condition (iii) yields  $\sum_{-\infty < j \leq n} (d_{nj}^{(i)} - d_{n,j-1}^{(i)})^2 \rightarrow 0$ ,  $i = 1, \dots, k$ . Consequently,  $\sum_{-\infty < j \leq n} (d_{nj} - d_{n,j-1})^2 \rightarrow 0$ , and (2.31) follows by Lemma 2.1.  $\square$

### 3 Weak convergence of partial sum processes

A number of econometric applications require the weak convergence of a suitably standardized partial sums process  $S_n(\tau) = \sum_{j=1}^{[n\tau]} X_j$ ,  $\tau > 0$  to some limit process  $S(\tau)$ ,  $0 \leq \tau \leq 1$ . Observe that for each  $n$ ,  $S_n(\tau)$ ,  $0 \leq \tau \leq 1$ , is a step function in  $\tau$ , belonging to the Skorokhod functional space  $\mathcal{D}[0, 1]$ .

From Billingsley (1968), we recall that a sequence of stochastic processes  $\{Z_n(\cdot)\}$ ,  $n \geq 1$  in  $\mathcal{D}[0, 1]$  is said to converge weakly to a stochastic process  $Z(\cdot) \in \mathcal{C}[0, 1]$ , and we write  $Z_n \Rightarrow Z$ , if every finite dimensional distribution of  $\{Z_n(\cdot)\}$  converges to that of  $Z(\cdot)$  and if  $\{Z_n(\cdot)\}$  is tight with respect to the uniform metric, see also Pollard (1984). The uniform

topology is stronger than the Skorokhod  $J_1$ -topology, and the verification of tightness in the uniform metric is relatively easier.

Using the arguments in Section 12 and Theorem 15.5, p. 127 of Billingsley (1968), see also Pollard (1984, ch. V.1, Theorem 3), one can show that a sufficient condition for tightness of  $\{Z_n\}$  is the following: There exists a sequence of non-decreasing right continuous functions  $F_n$  on  $[0, 1]$  that are uniformly bounded and converge uniformly to a continuous function  $F$  such that for some  $\beta > 1$ ,  $\gamma > 0$ , and  $\forall 0 \leq s < t \leq 1$ ,

$$(3.1) \quad E|Z_n(t) - Z_n(s)|^\gamma \leq C[F_n(t) - F_n(s)]^\beta, \quad \forall n \geq 1,$$

where  $C$  may depend on  $\gamma$ , but not on  $s$ ,  $t$  and  $n$ .

Now, let  $S_n = S_n(1)$  and  $\sigma_n^2 := \text{Var}(S_n)$ . Our goal here is to establish the weak convergence of  $\{\sigma_n^{-1}S_n(\cdot)\}$ . For this purpose we need to establish that its finite dimensional distributions converge to those of the limit process, denoted by  $\rightarrow_{fdd}$ , and that the process is tight in uniform metric. We first focus on the finite dimensional convergence.

According to Lamperti Theorem (1962), if

$$(3.2) \quad \{\sigma_n^{-1}S_n(\tau)\} \rightarrow_{fdd} \{S(\tau)\},$$

then for some  $H \in (0, 1)$  and a positive slowly varying function  $L$ ,

$$(3.3) \quad \sigma_n^2 = \text{Var}(S_n) = n^{2H}L(n).$$

In most applications

$$(3.4) \quad \sigma_n^2 = \text{Var}(S_n) \sim s^2 n^{2H} \rightarrow \infty, \quad \text{for some } 0 < H < 1,$$

where  $0 < s^2 < \infty$  is the long-run variance of  $S_n$ .

In the case of a linear process  $\{X_j\}$  of (2.1), with m.d. stationary and ergodic innovations  $\{\zeta_j\}$ , (3.4) is also sufficient for (3.2).

The limits in this section will be described by the fractional Brownian motion (fBm),  $B_H(\tau)$ ,  $0 \leq \tau \leq 1$ , with parameter  $0 < H < 1$ , which is a Gaussian process with the mean  $EB_H(t) \equiv 0$ , and covariance function

$$(3.5) \quad r_H(s, t) := \frac{1}{2} \left\{ |s|^{2H} + |t|^{2H} - |s - t|^{2H} \right\}, \quad 0 \leq s, t \leq 1.$$

Note that if  $H = 1/2$ , then  $B_{1/2} = B$  is Brownian motion. The convergence of the finite dimensional distributions of  $\sigma_n^{-1}S_n(\tau)$ ,  $0 \leq \tau \leq 1$  is established in the following

**Proposition 3.1** *Suppose  $\{X_j\}$  is a linear process (2.1) with the m.d.s.  $\{\zeta_j\}$  being either stationary and ergodic, or satisfying Assumption 2.1, and that (3.4) holds.*

*Then,*

$$(3.6) \quad \{\sigma_n^{-1}S_n(\tau)\}_{\tau>0} \rightarrow_{fdd} \{B_H(\tau)\}_{\tau>0},$$

*where  $B_H$  is the fBm with parameter  $H$ .*

**Proof.** Let  $T_n(\tau) := \sigma_n^{-1}S_n(\tau)$ . Assumption (3.4) implies that for any  $0 < \tau < 1$ ,

$$\text{Var}(T_n(\tau)) = \frac{\text{Var}(S_n(\tau))}{\text{Var}(S_n)} = \frac{(n\tau)^{2H}(1+o(1))}{n^{2H}(1+o(1))} \rightarrow \tau^{2H},$$

and hence,

$$(3.7) \quad \begin{aligned} \text{Cov}(T_n(t), T_n(s)) &= (1/2)\{\text{Var}(T_n(t)) + \text{Var}(T_n(s)) - \text{Var}(T_n(t) - T_n(s))\} \\ &\rightarrow r_H(t, s), \quad \forall 0 < s < t, \end{aligned}$$

where  $r_H$  is as in (3.5). In view of the Cramér-Wold device, it suffices to show that  $\forall a_1, \dots, a_k \in \mathbb{R}, t_1, \dots, t_k > 0$  and  $k \geq 1$ ,  $\mathcal{S}_n = a_1 S_n([nt_1]) + \dots + a_k S_n([nt_k])$  satisfy

$$\sigma_n^{-1}\mathcal{S}_n \rightarrow_D \mathcal{S} := a_1 B_H(t_1) + \dots + a_k B_H(t_k).$$

By (3.7),  $\text{Var}(\mathcal{S}_n) \rightarrow \infty$ , and the rest of the proof repeats the lines of proof of Theorem 2.1.

The following proposition is a simplified version of the result obtained by Taqqu (1975). It shows that long memory of the summands  $\{X_j\}$  *a priori* guarantees the tightness of the normalized partial sum process. It does not require  $\{X_j\}$  to be a linear process.

**Proposition 3.2** *Let  $\{X_j\}$  be a second order stationary process satisfying (3.4), with an  $1/2 < H < 1$ . Then,  $\sigma_n^{-1}S_n(\cdot)$  is tight w.r.t. uniform metric. In addition, if  $\sigma_n^{-1}S_n \rightarrow_{fdd} S$ , where  $S(u), 0 \leq u \leq 1$  is a stochastic process, then,*

$$(3.8) \quad \sigma_n^{-1}S_n(\cdot) \Rightarrow S(\cdot), \quad \text{in } \mathcal{D}[0, 1] \text{ and uniform metric.}$$

**Proof.** To check tightness, we shall verify (3.1) for the process  $T_n(t) := \sigma_n^{-1}S_n(t)$ . Let  $F_n(t) := [nt]/n, F(t) := t, 0 \leq t \leq 1$ . Observe that  $\sup_t |F_n(t) - F(t)| \rightarrow 0$  and  $F$  is continuous on  $[0, 1]$ . By covariance stationarity of the increments,

$$(3.9) \quad \begin{aligned} E|T_n(t) - T_n(s)|^2 &= E\left(\sigma_n^{-1} \sum_{j=1}^{[nt]-[ns]} X_j\right)^2 \\ &= \left(\frac{[nt] - [ns]}{n}\right)^{2H} \frac{(1+o(1))}{(1+o(1))} \leq C \left(\frac{[nt] - [ns]}{n}\right)^{2H} = C [F_n(t) - F_n(s)]^{2H}, \end{aligned}$$

for some  $\delta > 0$ . Since  $\beta := 2H > 1$ , this verifies (3.1) for the  $T_n$  process with  $\gamma = 2$  and completes the proof.

The following lemma, where  $S_n = S_n(1)$ , gives a useful bound for moments of the sums of a linear process with m.d. innovations and is useful in proving the tightness of the process  $\{T_n(t)\}$  in the cases of short and negative memory processes  $\{X_j\}$ . It extends the well-known Burkholder-Rosenthal inequality for martingales and some other inequalities involving m.d.s. of Dharmadhikari, Fabian and Jogdeo (1968, Theorem) and Borovskikh and Korolyuk (1997, Chapter 3).



**Lemma 3.1** Let  $\{X_j\}$  be a linear process (2.1) with m.d. innovations  $\{\zeta_j\}$ , such that  $E\zeta_j^2 = \sigma_\zeta^2$  for all  $j$ , and  $\delta := \max_j E|\zeta_j|^p < \infty$ , for some  $p \geq 2$ . Then

$$(3.10) \quad E|S_n|^p \leq c(ES_n^2)^{p/2}, \quad \forall n \geq 1,$$

where  $c > 0$  depends only on  $p$  and  $\delta$ .

**Proof.** Setting  $d_{nj} = \sum_{k=\max(j,1)}^n a_{k-j}$ , write  $S_n = \sum_{j=-\infty}^n d_{nj}\zeta_j$ . Because  $E\zeta_j^2 = \sigma_\zeta^2$ ,  $ES_n^2 = \sigma_\zeta^2 \sum_{j=-\infty}^n d_{nj}^2 < \infty$  and by Lemma 3.2 below,

$$E|S_n|^p = E \left| \sum_{j=-\infty}^n d_{nj}\zeta_j \right|^p \leq C \max_j E|\zeta_j|^p \left( \sum_{j=-\infty}^n d_{nj}^2 \right)^{p/2} = c(ES_n^2)^{p/2},$$

where  $c = C \max_j E|\zeta_j|^p / \sigma_\zeta^p$  does not depend on  $n$ . □

**Lemma 3.2** Let  $p \geq 2$  and  $\{Y_j, \mathcal{F}_j, 1 \leq j \leq n\}$  be a m.d.s. with  $\max_j E|Y_j|^p < \infty$ . Then, for every  $n \geq 1$ ,

$$(3.11) \quad E \left| \sum_{j=1}^n Y_j \right|^p \leq C_p \left( \sum_{j=1}^n (E|Y_j|^p)^{2/p} \right)^{p/2}, \quad p > 2,$$

with a constant  $C_p > 0$  depending only on  $p$ .

The inequality (3.11) remains valid also for  $n = \infty$ .

**Proof.** Let  $n < \infty$ . For  $p \geq 2$ , by the Burkholder-Rosenthal's inequality (see Hall and Heyde (1980), p. 24),

$$(3.12) \quad E \left| \sum_{j=1}^n Y_j \right|^p \leq C_p \left[ \sum_{j=1}^n E|Y_j|^p + E \left( \sum_{j=1}^n E[Y_j^2 | \mathcal{F}_{j-1}] \right)^{p/2} \right].$$

Recall the fact that for any real numbers  $a, b$ , and for  $0 < \alpha \leq 1$ ,  $|a + b|^\alpha \leq |a|^\alpha + |b|^\alpha$ . Apply this fact with  $\alpha = 2/p \leq 1$  to obtain

$$(3.13) \quad \left( \sum_{j=1}^n E|Y_j|^p \right)^{2/p} \leq \sum_{j=1}^n (E|Y_j|^p)^{2/p}.$$

To bound the second term on the r.h.s. of (3.12), first use the Cauchy-Schwarz inequality for the conditional expectation with  $p/2 \geq 1$ , to obtain  $E[Y_j^2 | \mathcal{F}_{j-1}] \leq (E[|Y_j|^p | \mathcal{F}_{j-1}])^{2/p}$ ,  $1 \leq j \leq n$ . Next, by Minkowski inequality for any r.v.'s  $X$  and  $Y$ ,

$$(E|X + Y|^r)^{1/r} \leq (E|X|^r)^{1/r} + (E|Y|^r)^{1/r}, \quad r \geq 1.$$

These facts in turn imply

$$\begin{aligned} E\left(\sum_{j=1}^n E[Y_j^2|\mathcal{F}_{j-1}]\right)^{p/2} &\leq E\left(\sum_{j=1}^n (E[|Y_j|^p|\mathcal{F}_{j-1}])^{2/p}\right)^{p/2} \\ &\leq \left(\sum_{j=1}^n (E|Y_j|^p)^{2/p}\right)^{p/2}, \end{aligned}$$

which together with (3.13) and (3.12) proves (3.11).

Consider now the case of an infinite sum. Let  $p \geq 2$ . We claim

$$(3.14) \quad E\left|\sum_{j=1}^{\infty} Y_j\right|^p \leq C_p \left(\sum_{j=1}^{\infty} (E|Y_j|^p)^{2/p}\right)^{p/2}.$$

This inequality is trivially true if the r.h.s. of (3.14) is infinite. Assume it is finite. By (3.11), for any  $1 < m \leq n < \infty$ ,

$$E\left|\sum_{j=m}^n Y_j\right|^p \leq C_p \left(\sum_{j=m}^n (E|Y_j|^p)^{2/p}\right)^{p/2} \rightarrow 0, \quad m, n \rightarrow \infty.$$

Hence, by the Cauchy convergence criterion,  $\sum_{j=1}^{\infty} Y_j$  converges in  $L_p$  norm, and

$$E\left|\sum_{j=n+1}^{\infty} Y_j\right|^p \rightarrow 0.$$

Next, by the Minkowski inequality and (3.11),

$$\begin{aligned} E^{1/p}\left|\sum_{j=1}^{\infty} Y_j\right|^p &\leq E^{1/p}\left|\sum_{j=1}^n Y_j\right|^p + E^{1/p}\left|\sum_{j=n+1}^{\infty} Y_j\right|^p \\ &\leq C_p \left(\sum_{j=1}^n (E|Y_j|^p)^{2/p}\right)^{p/2} + E^{1/p}\left|\sum_{j=n+1}^{\infty} Y_j\right|^p. \end{aligned}$$

Claim (3.14) now is proved upon taking limit as  $n \rightarrow \infty$  in this bound.  $\square$

We are now ready to state and prove the following weak convergence result.

**Theorem 3.1** *Assume that the m.d.s.  $\{\zeta_j\}$  in the linear process  $\{X_j\}$  is either stationary and ergodic, or satisfies Assumption 2.1. Suppose that  $S_n$  satisfies (3.4) with  $0 < H < 1$ . For  $0 < H \leq 1/2$ , assume, in addition, that  $\max_j E|\zeta_j|^p < \infty$ , for some  $p > 1/H$ . Then,*

$$(3.15) \quad \sigma_n^{-1} S_n(\cdot) \Rightarrow B_H(\cdot), \quad \text{in } \mathcal{D}[0, 1] \text{ and uniform metric,}$$

where  $B_H$  is a fBm.

**Proof.** Proposition 3.1 implies the finite dimensional convergence. To prove tightness we shall use (3.10). For  $H \leq 1/2$ , by (3.10) and (3.9),

$$E|T_n(t) - T_n(s)|^p \leq c[E(T_n(t) - T_n(s))^2]^{p/2} \leq c|F_n(t) - F_n(s)|^{Hp}.$$

This verifies (3.1) for the  $T_n$  process with  $\gamma = p$ ,  $\beta = Hp > 1$ . For  $H > 1/2$ , (3.15) follows from Proposition 3.2. This completes the proof.  $\square$

**Remark 3.1** A functional CLT for partial sums process of a short memory linear process  $\{X_j\}$  of (2.1) with m.d. innovations and  $a_j$ 's such that  $\sum_{j=0}^{\infty} j|a_j| < \infty$ , was established in Phillips and Solo (1992), and for more general stationary processes in Peligrad and Utev (2005).

## 4 Applications

In this section we shall illustrate the usefulness of the above results in some important models. In Subsection 4.1 these results are shown to be applicable to ARCH and stochastic volatility processes, and the sample ACF of a m.d.s. Subsection 4.2 discusses applications to parametric and non-parametric regression models.

### 4.1 Conditionally heteroscedastic processes

**ARCH process.** An important example of a stationary and ergodic m.d.s. is an ARCH( $\infty$ ) process, where

$$(4.1) \quad \zeta_j \equiv r_j = \varepsilon_j \sigma_j, \quad \sigma_j^2 = b_0 + \sum_{k=1}^{\infty} b_k r_{j-k}^2, \quad j \in \mathbb{Z}, \quad \{\varepsilon_j\} \sim \text{IID}(0, 1),$$

$$b_0 > 0, \quad b_k \geq 0, \quad k \geq 1, \quad \sum_{k=1}^{\infty} b_k < 1.$$

This process was introduced by Robinson (1991). It includes the parametric ARCH and GARCH models of Engle (1982) and Bollerslev (1996). Obviously,  $\{r_j\}$  is a m.d.s.  $E[r_j | \mathcal{F}_{j-1}] = 0$ ,  $V_j = E[r_j^2 | \mathcal{F}_{j-1}] = \sigma_j^2$ , and the equations in (4.1) have a unique second order stationary solution; see e.g., Giraitis, Leipus and Kokoszka (2000). Since  $r_j = \phi(\varepsilon_j, \varepsilon_{j-1}, \varepsilon_{j-2}, \dots)$ , where  $\phi$  is a measurable function, by Theorem 3.5.8 of Stout (1974),  $r_j$  is stationary and ergodic.

**Squared ARCH process.** A centered squared ARCH process

$$X_j = r_j^2 - E r_j^2,$$

with  $r_j$  as in (4.1) and satisfying  $E^{1/2}(\varepsilon_0^4) \sum_{k=1}^{\infty} b_k < 1$ , has a unique fourth order stationary solution, see, e.g., Giraitis, Leipus and Surgailis (2007), and is covered by the set up in (2.1). Indeed, by (4.1),

$$(4.2) \quad X_j = \sum_{k=1}^{\infty} b_k X_{j-k} + \eta_j, \quad \eta_j = (\varepsilon_j^2 - 1)\sigma_j^2,$$

where  $\{\eta_j\}$  is a m.d.s. with  $E[\eta_j|\mathcal{F}_{j-1}] = 0$ ,  $V_j = E[\eta_j^2|\mathcal{F}_{j-1}] = \text{Var}(\varepsilon_0^2)\sigma_j^4$  and  $E\eta_0^2 < \infty$ . Again, since  $\eta_j = \phi(\varepsilon_j, \varepsilon_{j-1}, \varepsilon_{j-2}, \dots)$ , where  $\phi$  is a measurable function,  $\{\eta_j\}$  is a stationary and ergodic process.

By Proposition 4.1, under some additional conditions,  $\{r_j\}$  and  $\{\eta_j\}$  satisfy Assumption 2.1. Hence, the results for the sums  $S_n$  and weighted sums  $W_n$  of this paper are applicable also to a centered squared ARCH process (4.2), and to a linear process (2.1) with ARCH innovations  $\zeta_j = r_j$  of (4.1).

**Proposition 4.1** *Let  $r_j$  be as in (4.1), and  $\eta_j$  as in (4.2). Then,*

- (i)  $\{r_j\}$  and  $\{\eta_j\}$  are stationary and ergodic m.d.s.
- (ii)  $r_j$  satisfies Assumption 2.1, if  $E^{1/2}(\varepsilon_0^4) \sum_{k=1}^{\infty} b_k < 1$ .
- (iii)  $\eta_j$  satisfies Assumption 2.1, if  $E^{1/4}(\varepsilon_0^8) \sum_{k=1}^{\infty} b_k < 1$ .

**Proof.** Claim (i) was proved above, just before the statement of this proposition.

(ii) First we verify Assumption 2.1(a) for  $V_j = \sigma_j^2 = E[\zeta_j^2|\mathcal{F}_{j-1}]$ . Recursion (4.1) yields Volterra expansion of  $\sigma_j^2$ , see, e.g., Giraitis, Leipus and Kokoszka (2000):

$$(4.3) \quad \sigma_j^2 \equiv b_0 \left(1 + \sum_{k=1}^{\infty} h_{k,j}\right), \quad h_{k,j} := \sum_{s_1, \dots, s_k=1}^{\infty} b_{s_1} \dots b_{s_k} \varepsilon_{j-s_1}^2 \dots \varepsilon_{j-s_1-\dots-s_k}^2, \quad j \in \mathbb{Z}.$$

To show  $E \max_j \sigma_j^4 < \infty$ , let  $B = \sum_{s=1}^{\infty} b_s$ . Since  $b_j$ 's are non-negative and  $\varepsilon_j$ 's are i.i.d. r.v.'s, by the Cauchy-Schwarz inequality (C-S)

$$(4.4) \quad \begin{aligned} E h_{k,j}^2 &\leq \sum_{s_1, \dots, s_k=1}^{\infty} \sum_{t_1, \dots, t_k=1}^{\infty} b_{s_1} \dots b_{s_k} b_{t_1} \dots b_{t_k} \\ &\quad \times \{E[(\varepsilon_{j-s_1}^2 \dots \varepsilon_{j-s_1-\dots-s_k}^2)^2]\}^{1/2} \{E[(\varepsilon_{j-t_1}^2 \dots \varepsilon_{j-t_1-\dots-t_k}^2)^2]\}^{1/2} \\ &= \left( \sum_{s_1, \dots, s_k=1}^{\infty} b_{s_1} \dots b_{s_k} \right)^2 (E\varepsilon_0^4)^k = (B^2 E\varepsilon_0^4)^k. \end{aligned}$$

Use this bound, C-S, and  $BE^{1/2}\varepsilon_0^4 < 1$  of assumption (ii), to obtain

$$(4.5) \quad \begin{aligned} E[h_{k,j} h_{p,j}] &\leq (E[h_{k,j}^2] E[h_{p,j}^2])^{1/2} \leq (BE^{1/2}[\varepsilon_0^4])^{k+p}, \\ E\sigma_j^4 &\leq 2b_0^2 \left(1 + E\left(\sum_{k=1}^{\infty} h_{k,j}\right)^2\right) \leq 2b_0^2 \left(1 + \left\{\sum_{k=1}^{\infty} (BE^{1/2}[\varepsilon_0^4])^k\right\}^2\right) < \infty, \end{aligned}$$

which verifies Assumption 2.1(a).

To verify Assumption 2.1(b) we approximate  $\sigma_j^2$  by  $m$ -dependent r.v.'s as follows. For  $m \geq 1$ , define

$$(4.6) \quad \sigma_{mj}^2 = b_0 \left(1 + \sum_{k=1}^{\infty} h_{m,k,j}\right), \quad h_{m,k,j} := \sum_{s_1, \dots, s_k=1}^m b_{s_1} \dots b_{s_k} \varepsilon_{j-s_1}^2 \dots \varepsilon_{j-s_1-\dots-s_k}^2.$$

Then  $\sigma_{mj}^2$ ,  $j \in \mathbb{Z}$  is a stationary  $m$ -dependent process,  $E\sigma_{mj}^4 \leq E\sigma_j^4 < \infty$ , and for  $|t-j| > m$ ,  $\text{Cov}(\sigma_{mt}^2, \sigma_{mj}^2) = 0$ . We will show that for  $|t-j| > m$ ,

$$(4.7) \quad \max_j E(\sigma_j^2 - \sigma_{mj}^2)^2 \leq \delta_m \rightarrow 0, \quad m \rightarrow \infty.$$

The latter with the bound  $|\text{Cov}(X, Y)| \leq (EX^2EY^2)^{1/2}$  and (4.5) implies

$$(4.8) \quad |\text{Cov}(\sigma_t^2, \sigma_j^2)| = \left| \text{Cov}(\sigma_{mt}^2, \sigma_{mj}^2) + \text{Cov}(\sigma_t^2 - \sigma_{mt}^2, \sigma_{mj}^2) + \text{Cov}(\sigma_{mt}^2, \sigma_j^2 - \sigma_{mj}^2) + \text{Cov}(\sigma_t^2 - \sigma_{mt}^2, \sigma_j^2 - \sigma_{mj}^2) \right| \\ \leq C(\delta_m^{1/2} + \delta_m) \rightarrow 0, \quad m \rightarrow \infty,$$

which proves Assumption 2.1(b) with  $K = m$ . To verify (4.7), notice that  $b_j \geq 0$  implies  $h_{k,j} \geq h_{m,k,j}$ . Hence, with  $B_m := \sum_{s=1}^m b_s$  and  $\bar{b}_s = b_s I(1 \leq s \leq m)$ , as in (4.4),

$$(4.9) \quad E(h_{k,j} - h_{m,k,j})^2 = \left( \sum_{s_1, \dots, s_k=1}^{\infty} (b_{s_1} \cdots b_{s_k} - \bar{b}_{s_1} \cdots \bar{b}_{s_k}) \varepsilon_{j-s_1}^2 \cdots \varepsilon_{j-s_1-\dots-s_k}^2 \right)^2 \\ \leq \left( \sum_{s_1, \dots, s_k=1}^{\infty} \{b_{s_1} \cdots b_{s_k} - \bar{b}_{s_1} \cdots \bar{b}_{s_k}\} \right)^2 (E\varepsilon_0^4)^k \\ = (B^k - B_m^k)^2 (E\varepsilon_0^4)^k.$$

Next, since  $0 < B_m \leq B$ , by the mean value theorem,  $B^k - B_m^k \leq kB^{k-1}(B - B_m) \leq kB^k \bar{\delta}_m$ , where  $\bar{\delta}_m := (B - B_m)/B \rightarrow 0$ , as  $m \rightarrow \infty$ . Hence, as in (4.5),

$$(4.10) \quad E(\sigma_j^2 - \sigma_{mj}^2)^2 = b_0^2 E \left( \sum_{k=1}^{\infty} \{h_{k,j} - h_{m,k,j}\} \right)^2 \\ \leq b_0^2 \left( \sum_{k=1}^{\infty} k (BE^{1/2}[\varepsilon_0^4])^k \right)^2 \times \bar{\delta}_m^2 = C\bar{\delta}_m^2 \rightarrow 0, \quad m \rightarrow \infty,$$

which proves (4.7).

Assumption 2.1(c) follows by Markov inequality noting that  $Er_j^4 = E\varepsilon_1^4 E\sigma_1^4 < \infty$ .

(iii) Here  $V_j = \sigma_j^4 \text{Var}(\varepsilon_0^2) = E[\eta_j^2 | \mathcal{F}_{j-1}]$ . Thus verifying Assumption 2.1(a) is equivalent to proving  $\max_j E\sigma_j^8 < \infty$ . Towards this goal, first note that by the Hölder inequality and by arguing as for (4.4),  $Eh_{k,j}^4 \leq B^{4k} (E\varepsilon_0^8)^k$ . Together with assumption  $BE^{1/4}[\varepsilon_0^8] < 1$ , this bound yields

$$(4.11) \quad E[h_{k_1,j} \cdots h_{k_4,j}] \leq (E[h_{k_1,j}^4] \cdots E[h_{k_4,j}^4])^{1/4} \leq (BE^{1/4}[\varepsilon_0^8])^{k_1+\dots+k_4}, \\ E\sigma_j^8 \leq 4b_0^4 (1 + E(\sum_{k=1}^{\infty} h_{k,j})^4) \leq 4b_0^4 (1 + \{\sum_{k=1}^{\infty} (BE^{1/4}[\varepsilon_0^8])^k\}^4) < \infty.$$

To verify Assumption 2.1(b), we again use approximating  $m$ -dependent variables  $\sigma_{mj}^2$ , (4.6). We show that for  $|t - j| > m$ ,

$$(4.12) \quad \max_j E(\sigma_j^4 - \sigma_{mj}^4)^2 \leq \delta_m \rightarrow 0, \quad m \rightarrow \infty,$$

which together with (4.11), as in (4.8), implies that

$$\begin{aligned} |\text{Cov}(\sigma_t^4, \sigma_j^4)| &= \left| \text{Cov}(\sigma_{mt}^4, \sigma_{mj}^4) + \text{Cov}(\sigma_t^4 - \sigma_{mt}^4, \sigma_{mj}^4) + \text{Cov}(\sigma_{mt}^4, \sigma_j^4 - \sigma_{mj}^4) \right. \\ &\quad \left. + \text{Cov}(\sigma_t^4 - \sigma_{mt}^4, \sigma_j^4 - \sigma_{mj}^4) \right| \leq C(\delta_m^{1/2} + \delta_m) \rightarrow 0, \quad m \rightarrow \infty. \end{aligned}$$

Finally, to prove (4.12), use Hölder inequality, similarly as in (4.9) and (4.10), to obtain

$$\begin{aligned} E(h_{k,j} - h_{m,k,j})^4 &\leq \left( \sum_{s_1, \dots, s_k=1}^{\infty} b_{s_1} \cdots b_{s_k} - \sum_{s_1, \dots, s_k=1}^m b_{s_1} \cdots b_{s_k} \right)^4 (E\varepsilon_0^8)^k \\ &\leq (B^k - B_m^k)^4 (E\varepsilon_0^8)^k \leq (E\varepsilon_0^8)^k k^4 B^{4k} \bar{\delta}_m^4, \\ E(\sigma_j^2 - \sigma_{mj}^2)^4 &\leq b_0^4 E \left( \sum_{k=1}^{\infty} \{h_{k,j} - h_{m,k,j}\} \right)^4 \\ &\leq b_0^4 \left( \sum_{k=1}^{\infty} k (BE^{1/4}[\varepsilon_0^8])^k \right)^4 \times \bar{\delta}_m^4 = C\bar{\delta}_m^4 \rightarrow 0, \quad m \rightarrow \infty, \end{aligned}$$

because  $BE^{1/4}[\varepsilon_0^8] < 1$  by assumption (iii), where  $\bar{\delta}_m = (B - B_m)/B \rightarrow 0$ . This together with (4.11) and  $\sigma_{mj}^2 \leq \sigma_j^2$  implies

$$\begin{aligned} E(\sigma_j^4 - \sigma_{mj}^4)^2 &\leq E[(\sigma_j^2 - \sigma_{mj}^2)^2 (\sigma_j^2 + \sigma_{mj}^2)^2] \\ &\leq E[(\sigma_j^2 - \sigma_{mj}^2)^2 4\sigma_j^4] \leq (E[(\sigma_j^2 - \sigma_{mj}^2)^4])^{1/2} 4(E\sigma_j^8)^{1/2} \leq C\bar{\delta}_m^2 \rightarrow 0 \quad m \rightarrow \infty, \end{aligned}$$

which proves (4.12) and completes verification of Assumption 2.1(b).

Assumption 2.1(c) follows by Markov inequality noting that  $E\eta_j^4 \leq CE\varepsilon_1^8 E\sigma_1^8 < \infty$ , which completes the proof of the proposition.  $\square$

**Stochastic volatility m.d. processes.** By a stochastic volatility model one usually understands a stationary process  $\zeta_j, j \in \mathbb{Z}$  of the form

$$\zeta_j = \varepsilon_j \sigma_j, \quad \varepsilon_j \sim \text{IID}(0, 1), \quad j \in \mathbb{Z},$$

where the (volatility) process  $\sigma_j > 0$  is a function of the past information up to time  $j - 1$ . Let  $\mathcal{F}_{j-1}$  be the  $\sigma$ -field generated by past innovations  $\varepsilon_s, s \leq j - 1$ , and  $E\sigma_j^2 < \infty$ . Then  $E[\zeta_j | \mathcal{F}_{j-1}] = 0$ ,  $\sigma_j^2 = \text{Var}(\zeta_j | \mathcal{F}_{j-1})$ , and  $\{\zeta_j\}$  is a white noise process, which by Theorem 3.5.8 of Stout (1974) is also stationary and ergodic.

It is often assumed that the volatility process  $\sigma_j = h(\eta_j), j \in \mathbb{Z}$  is a nonlinear function of a stationary Gaussian (or linear) process  $\{\eta_j\}$ , see e.g. Robinson (2001). The choice of

$h(\eta_j) = \exp(\eta_j)$  includes the Exponential Generalized ARCH (EGARCH) model, proposed by Nelson (1991). A related class of stochastic volatility models with long memory in  $\{\sigma_j\}$  was introduced and studied in Breidt, Crato and de Lima (1998), Harvey (1998), and Surgailis and Viano (2002). As a rule, the volatility process  $V_j \equiv \sigma_j^2$  in these models is a stationary process with autocovariances decaying to zero, and thus satisfies Assumption 2.1.

**Sample ACF.** Let  $\{\zeta_j\}$  be a stationary and ergodic m.d.s. such that  $\sigma_\zeta^2 = E\zeta_0^2 < \infty$ . Consider the sample autocorrelation  $\hat{\rho}_k = \hat{\gamma}_k/\hat{\gamma}_0$ ,  $k \geq 1$  where  $\hat{\gamma}_k = n^{-1} \sum_{j=k+1}^n (\zeta_j - \bar{\zeta})(\zeta_{j-k} - \bar{\zeta})$  and  $\bar{\zeta} = n^{-1} \sum_{j=1}^n \zeta_j$ . It is well known that in the case of i.i.d. random variables,  $\sqrt{n}\hat{\rho}_k \rightarrow_D \mathcal{N}(0, 1)$  and  $(\sqrt{n}\hat{\rho}_1, \dots, \sqrt{n}\hat{\rho}_k) \rightarrow_D \mathcal{N}(0, I)$  where the limit is a vector of  $k$  independent standard normal variables. These results are widely used for testing for the absence of correlation, see e.g. Brockwell and Davis (1991). The next proposition shows that, in general, this asymptotic normality result is true only for i.i.d. sequences.

Let  $\Omega$  be  $m \times m$  matrix with  $(j, k)$ th element  $\omega_{j,k} = E[\zeta_1^2 \zeta_{1-j} \zeta_{1-k}]/\sigma_\zeta^4$ .

**Proposition 4.2** *Let  $\{\zeta_j\}$  be a stationary and ergodic m.d.s. with  $\sigma_\zeta^2 := E\zeta_0^2 < \infty$  such that  $E[\zeta_1^2 \zeta_{1-k}^2] < \infty$ , for  $k = 1, \dots, m$ , where  $m$  is a given positive integer. Then*

$$(4.13) \quad (\sqrt{n}\hat{\rho}_1, \dots, \sqrt{n}\hat{\rho}_m) \rightarrow_D \mathcal{N}_m(0, \Omega).$$

In particular, (4.13) holds in case of ARCH process  $\zeta_j = r_j$  of (4.1), such that  $E^{1/2}(\varepsilon_0^4) \sum_{k=1}^{\infty} b_k < 1$ . If in addition,  $E\varepsilon_j^3 = 0$ , then  $\Omega$  is diagonal with  $\omega_{k,k} \geq 1$ ,  $k = 1, \dots, m$ .

**Proof.** Let  $1 \leq k \leq m$ . Under the given assumptions, by Stout (1974, Theorem 3.5.8),  $\xi_j = \zeta_j \zeta_{j-k}$  is a stationary and ergodic m.d.s., and  $E\xi_0^2 < \infty$ . Thus, by Theorem 2.1,  $\bar{\zeta} = O_p(n^{-1/2})$ ,  $n^{-1/2} \sum_{j=k+1}^n \zeta_j \zeta_{j-k} \rightarrow_D \mathcal{N}(0, E\xi_0^2)$ , while by ergodicity of  $\zeta_j^2$ ,  $\hat{\gamma}_0 \rightarrow_p E\zeta_0^2$ . Hence  $\sqrt{n}\hat{\rho}_k = (\sigma_\zeta^2)^{-1} n^{-1/2} \sum_{j=k+1}^n \zeta_j \zeta_{j-k} + O(n^{-1/2}) \rightarrow \mathcal{N}(0, \omega_{k,k})$ .

To prove (4.13), let  $\xi_j^{(k)} = \zeta_j \zeta_{j-k}$ ,  $k = 1, \dots, m$ . Then for any real numbers  $c_1, \dots, c_m$ ,  $\eta_j := c_1 \xi_j^{(1)} + \dots + c_m \xi_j^{(m)}$  is a stationary ergodic m.d.s., with  $E\eta_j^2 = \sigma_\eta^2 := c_1^2 \omega_{1,1} + \dots + c_m^2 \omega_{m,m}$ . Let  $S_n := \sum_{j=m+1}^n \eta_j$ . Then  $ES_n^2 = \sigma_\eta^2 n \rightarrow \infty$ , and by Theorem 2.1,  $n^{-1/2} S_n \rightarrow_D \mathcal{N}(0, \sigma_\eta^2)$ , which by the Cramer-Wold device implies the claim of (4.13).

In case of ARCH process  $r_j$ , (4.1) and (4.3) imply  $\omega_{k,k} \geq 1$ , and show that  $\Omega$  is diagonal if, in addition,  $E\varepsilon_j^3 = 0$ .  $\square$

**Remark 4.1** Proposition 4.2 shows that for the ARCH m.d.s. (4.1), the limit limit variance of  $\sqrt{n}\hat{\rho}_k$ , as the rule, is greater than one, and so the 95% confidence band for  $\rho_k = 0$  is wider than in i.i.d. case. Moreover, the limit matrix  $\Omega$  may be non-diagonal, unless the distribution of  $\varepsilon_j$  in (4.1) is symmetric.

## 4.2 Regression models

In this section we discuss application of the above results for obtaining asymptotic normality of least squares (LS) estimators in parametric regression models and kernel type estimators in non-parametric regression models.

But first we verify the above conditions for asymptotic normality  $S_n$  and weak convergence of the partial sum process  $\{S_n(\tau), 0 \leq \tau \leq 1\}$  of a linear process  $\{X_j\}$  of (2.1) in some typical cases.

Let  $\gamma(j), j = 0, 1, 2, \dots$  denote the autocovariance function of  $\{X_j\}$  and  $f$  its spectral density. Note that  $\gamma(j) := \text{Cov}(X_j, X_0) = \sigma_\zeta^2 \sum_{k=0}^{\infty} a_k a_{k+j}, j = 0, 1, 2, \dots$ . Consider the following assumption in terms of  $f$ .

$$(4.14) \quad f(u) \sim c|u|^{-2d}, \quad u \rightarrow 0, \quad |d| < 1/2, \quad c > 0.$$

In terms of  $\gamma(j)$ , consider the following condition:

$$(4.15) \quad \begin{aligned} \sum_{j \in \mathbb{Z}} |\gamma(j)| &< \infty, \quad \sigma^2 := \sum_{j \in \mathbb{Z}} \gamma(j) > 0, \quad \text{for } d = 0, \\ \gamma(j) &\sim c_\gamma |j|^{-1+2d}, \quad 0 < d < 1/2, \\ \gamma(j) &\sim c_\gamma |j|^{-1+2d}, \quad \sum_{j \in \mathbb{Z}} \gamma(j) = 0, \quad -1/2 < d < 0, \end{aligned}$$

where  $c_\gamma \neq 0$ . The cases  $d = 0, 0 < d < 1/2$  and  $-1/2 < d < 0$  define short, long and negative memory of the process  $\{X_j\}$ .

**Corollary 4.1** *Let  $\{X_j\}$  be a linear process (2.1) with stationary and ergodic m.d. innovations  $\{\zeta_j\}$  satisfying (4.14) or (4.15) with some  $|d| < 1/2$ . Then,*

$$(4.16) \quad \sigma_n^2 \sim s^2 n^{1/2+d}, \quad n^{-1/2-d} S_n \rightarrow_D \mathcal{N}(0, s^2),$$

where

$$\begin{aligned} s^2 &= 2\pi f(0), \quad d = 0; \quad s^2 = c_f \int_{\mathbb{R}} \frac{\sin^2(u/2)}{(u/2)^2} |u|^{-2d} du, \quad 0 < |d| < 1/2, \quad \text{under (4.14)}, \\ s^2 &= \sum_{k \in \mathbb{Z}} \gamma(k), \quad d = 0; \quad s^2 = c_\gamma / (d(1+2d)), \quad 0 < |d| < 1/2, \quad \text{under (4.15)}. \end{aligned}$$

In addition, if for  $-1/2 < d \leq 0$ ,  $E|\zeta_0|^p < \infty$ , for some  $p > 1/(1/2 + d)$ , then

$$(4.17) \quad n^{-1/2-d} S_n(\cdot) \Rightarrow sB_{1/2+d}(\cdot),$$

in  $\mathcal{D}[0, 1]$  and uniform metric.



**Proof.** The claim about  $\sigma_n^2$  in (4.16) under (4.14) or (4.15) is known. Together with Theorem 2.1 and Proposition 3.1, it proves the second claim in (4.16) and (4.17).  $\square$

Recall also the following known relationship between the weights  $a_k$  and  $\gamma(j)$  when  $\{X_j\}$  is a linear process (2.1) with stationary m.d. innovations. If

$$\begin{aligned} a_k &\sim c_a |k|^{-1+d}, \quad 0 < d < 1/2, \\ &= c_a |k|^{-1+d} (1 + O(k^{-1})), \quad -1/2 < d < 0, \end{aligned}$$

where  $c_a \neq 0$ , then  $\gamma(k)$  satisfies (4.15) with  $c_\gamma = \sigma_\zeta^2 c_a^2 B(d, 1 - 2d)$ , where  $B(\cdot, \cdot)$  is the beta function.

If  $\sum_{k=0}^{\infty} |a_k| < \infty$  and  $\sum_{k=0}^{\infty} a_k \neq 0$  then

$$\sum_{k=0}^{\infty} |\gamma(k)| < \infty, \quad \sum_{k=0}^{\infty} \gamma(k) = \sigma_\zeta^2 \left( \sum_{k=0}^{\infty} a_k \right)^2 > 0.$$

**LS estimation.** Let  $\{X_j\}$  be a linear process with memory parameter  $0 \leq d < 1/2$ . Consider the simple parametric regression model where for some  $\beta \in \mathbb{R}$ ,  $Y_j = z_{nj}\beta + X_j$ . A problem of interest is to obtain asymptotic distribution of the least square estimator  $\hat{\beta} = \sum_{j=1}^n z_{nj} Y_j / \sum_{j=1}^n z_{nj}^2$  of  $\beta$ . Suppose

$$(4.18) \quad z_{nj} = g(j/n), \quad j = 1, \dots, n,$$

where  $g$  is a continuous real valued function on  $[0, 1]$ . Moreover, in the short memory case where  $d = 0$ , assume that the covariance function  $\gamma$  of  $\{X_j\}$  satisfies

$$(4.19) \quad \sum_{k=0}^{\infty} |\gamma(k)| < \infty, \quad \sum_{k \in \mathbb{Z}} \gamma(k) > 0.$$

In the long memory case  $0 < d < 1/2$ , assume

$$(4.20) \quad \gamma(k) \sim c_\gamma |k|^{-1+2d}, \quad k \rightarrow \infty.$$

Now, note that  $\hat{\beta} - \beta = \sum_{j=1}^n z_{nj} X_j / \sum_{j=1}^n z_{nj}^2 =: W_n / \sum_{j=1}^n z_{nj}^2$ . Define

$$\begin{aligned} v_D^2 &:= \int_0^1 g^2(u) du \left( \sum_{k \in \mathbb{Z}} \gamma(k) \right), & d = 0, \\ &:= c_\gamma \int_0^1 \int_0^1 g(u) g(v) |u - v|^{-1+2d} dudv, & 0 < d < 1/2, \\ \tau_d^2 &:= v_d^2 / \left( \int_0^1 g^2(u) du \right)^2. \end{aligned}$$

The following corollary gives limiting distribution of  $\hat{\beta}$ .

**Corollary 4.2** Suppose the linear process  $\{X_j\}$  of (2.1) with stationary and ergodic m.d. innovations  $\{\zeta_j\}$  satisfies (4.19) or (4.20), and  $z_{nj}$ 's are as in (4.18). Then, with  $W_n := \sum_{j=1}^n z_{nj} X_j$ ,

$$(4.21) \quad n^{-1/2-d} W_n \rightarrow_D \mathcal{N}(0, v_d^2), \quad n^{1/2-d}(\hat{\beta} - \beta) \rightarrow_D \mathcal{N}(0, \tau_d^2), \quad 0 \leq d < 1/2.$$

**Proof.** The second claim in (4.21) follows from the first claim and the fact that  $\sum_{j=1}^n z_{nj}^2/n \rightarrow \int_0^1 g^2(u)du$ , which is assured by the continuity of  $g$ . To prove the first claim in (4.21), we shall verify condition (i) of Theorem 2.2. Let  $\sigma_n^2 := \text{Var}(W_n) = \sum_{j,k=1}^n z_{nj} z_{nk} \gamma(j-k)$ . We shall prove that

$$(4.22) \quad n^{-1-2d} \sigma_n^2 \rightarrow v_d^2, \quad \forall 0 \leq d < 1/2.$$

Then  $\sigma_n^2 \sim v_d^2 n^{1+2d}$ , which implies  $\sum_{j=1}^n z_{nj}^2 = O(\sigma_n^2)$  and  $\max_{1 \leq k \leq n} |z_{nk}| \leq \sup_{0 \leq u \leq 1} |g(u)| = o(\sigma_n)$ , thereby verifying condition (i) of Theorem 2.2 for  $W_n$ .

We now prove (4.22). Suppose  $d = 0$ . Then,

$$n^{-1} \sum_{j,k=1: |j-k| > K}^n |z_{nj} z_{nk} \gamma(j-k)| \leq \sup_{0 \leq u \leq 1} |g(u)|^2 \sum_{|s| > K}^n |\gamma(s)| \rightarrow 0, \quad K \rightarrow \infty,$$

whereas for any  $|i| \leq K$ ,  $n^{-1} \sum_{k=1}^n z_{n,k+i} z_{nk} \gamma(i) \rightarrow \int_0^1 g^2(u)du$ . Whence,

$$\lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \sum_{j,k=1: |j-k| \leq K}^n z_{nj} z_{nk} \gamma(j-k) = \int_0^1 g^2(u)du \lim_{K \rightarrow \infty} \sum_{|i| \leq K} \gamma(i) = v_0^2,$$

which proves  $\sigma_n^2 \rightarrow v_0^2$ .

Next, consider the case  $0 < d < 1/2$ . Here, (4.18), (4.20), change of variables and the dominated convergence theorem yield

$$\begin{aligned} n^{-1-2d} \sigma_n^2 &= n^{-1-2d} \sum_{k,j=1}^n z_{nj} z_{n,k} \gamma_X(j-k) \\ &= n^{-1-2d} c_\gamma \sum_{k,j=1: k \neq j}^n g\left(\frac{j}{n}\right) g\left(\frac{k}{n}\right) |j-k|^{-1+2d} + o(1) \\ &\rightarrow c_\gamma \int_0^1 \int_0^1 g(u) g(v) |u-v|^{-1+2d} dudv = v_d^2. \end{aligned}$$

This completes the proof of (4.22) and the corollary.  $\square$

**A non-stationary process.** As another application of Theorem 2.2, consider the weighted sum  $V_n := \sum_{j=1}^n z_{nj} Y_j$  of a non-stationary unit root process  $Y_j := \sum_{s=1}^j X_s$ ,  $j = 1, 2, \dots$ . Set  $\psi(u) = \int_u^1 g(v)dv$ ,  $0 \leq u \leq 1$ . Also define,

$$\begin{aligned} \bar{v}_d^2 &:= \int_0^1 \psi^2(u) du \left( \sum_{k \in \mathbb{Z}} \gamma(k) \right), & d = 0, \\ &:= c_\gamma \int_0^1 \int_0^1 \psi(u) \psi(v) |u-v|^{-1+2d} dudv, & 0 < d < 1/2. \end{aligned}$$

In view of Remark 2.4, the proof of the following corollary is similar to that of Corollary 4.2.

**Corollary 4.3** *Suppose the linear process  $\{X_j\}$  of (2.1) with stationary and ergodic m.d.s.  $\{\zeta_j\}$  satisfies (4.19) or (4.20) and  $z_{nj}$  are as in (4.18). Then,  $n^{-3/2-d}V_n \rightarrow_D \mathcal{N}(0, \bar{v}_d^2)$ .*

**Nonparametric regression.** We shall now show the usefulness of Theorem 2.2 in deriving limiting distribution of a kernel type estimator of the regression function  $\mu$  in the nonparametric regression model  $Y_j = \mu(j/n) + X_j$ , when errors  $X_j$  may have long memory. Let  $K$  be a density kernel on  $\mathbb{R}$  with  $\|K\|_2^2 := \int_{\mathbb{R}} K^2(v)dv < \infty$ , and  $b \equiv b_n$  be a sequence of window widths. A kernel type estimator of  $\mu(x)$  is given by

$$\hat{\mu}_n(x) := K_{nx}^{-1} \sum_{j=1}^n K\left(\frac{nx-j}{nb}\right)Y_j, \quad K_{nx} := \sum_{j=1}^n K\left(\frac{nx-j}{nb}\right).$$

Note, that as  $n \rightarrow \infty$ ,  $K_{nx} \sim nb \int_{\mathbb{R}} K(u)du = nb$ , for all  $0 < x < 1$ . Let

$$\bar{\mu}_n(x) := K_{nx}^{-1} \sum_{j=1}^n K\left(\frac{nx-j}{nb}\right)\mu\left(\frac{j}{n}\right), \quad D_n(x) := \hat{\mu}_n(x) - \bar{\mu}_n(x) = K_{nx}^{-1} \sum_{j=1}^n K\left(\frac{nx-j}{nb}\right)X_j.$$

Then,  $\hat{\mu}_n(x) - \mu(x) = D_n(x) + \bar{\mu}_n(x) - \mu(x)$ . Typically the bias term  $\bar{\mu}_n(x) - \mu(x)$  is negligible compared to  $D_n(x)$  and asymptotic distribution of  $\hat{\mu}_n(x) - \mu(x)$  is determined by that of  $D_n(x)$ . Fix an  $x \in (0, 1)$  and let

$$(4.23) \quad z_{nj} := K_{nx}^{-1} K\left(\frac{x-nj}{nb}\right).$$

Then, clearly  $D_n(x)$  is like a  $W_n$ . Define

$$\begin{aligned} \tau_{d,K}^2 &:= \|K\|_2^2 \sum_{k \in \mathbb{Z}} \gamma(k), & d = 0, \\ &:= c_\gamma \int_0^1 \int_0^1 K(u)K(v)|u-v|^{-1+2d} dudv, & 0 < d < 1/2. \end{aligned}$$

We have

**Corollary 4.4** *Suppose the linear process  $\{X_j\}$  of (2.1) with stationary and ergodic m.d.s.  $\{\zeta_j\}$  satisfies (4.19) or (4.20). In addition, suppose  $b \rightarrow 0$ ,  $nb \rightarrow \infty$ , and  $K$  is a continuous density on  $\mathbb{R}$  with  $\int_{\mathbb{R}} K^2(v)dv < \infty$ .*

*Then, for every  $0 < x < 1$ , and  $0 \leq d < 1/2$ ,*

$$(4.24) \quad (nb)^{1-2d} \text{Var}(D_n(x)) \rightarrow \tau_{d,K}^2, \quad n^{1/2-d} D_n(x) \rightarrow_D \mathcal{N}(0, \tau_{d,K}^2).$$

**Proof.** With  $\sigma_n^2 := \text{Var}(D_n(x))$  and  $z_{nj}$  as in (4.23),

$$(nb)^{1-2d} \sigma_n^2 \sim (nb)^{-1-2d} \sum_{i=1}^n \sum_{j=1}^n K\left(\frac{nx-i}{nb}\right) K\left(\frac{nx-j}{nb}\right) \gamma(i-j)$$

A routine argument shows that continuity of  $K$  with  $\|K\|_2^2 < \infty$  implies  $(nb)^{1-2d}\sigma_n^2 \rightarrow \tau_{d,K}^2$ ,  $\forall 0 \leq d < 1/2$ , and  $(nb) \sum_{j=1}^n z_{nj}^2 \sim (nb)^{-1} \sum_{j=1}^n K^2((x - (j/n))/b) \rightarrow \|K\|_2^2$ . These facts together yield  $\sigma_n \sim \tau_{d,K}(nb)^{-2d}$ ,  $\sigma_n^{-2} \sum_{j=1}^n z_{nj}^2 = O((nb)^{-2d}) = O(1)$  and  $\max_{1 \leq j \leq n} \sigma_n^{-1} |z_{nj}| = O((nb)^{-1/2-d}) = o(1)$ , verify condition (i) of Theorem 2.2, and hence the corollary.  $\square$

**Dickey-Fuller distributions and their fractional versions.** The results of Section 3 also imply a number of existing findings that are widely used in the econometric literature. Dickey and Fuller (1979, 1981) derived the distributions of the normalized autoregressive coefficient and the t-ratio, when the generating process has a unit root. Phillips (1987) described their limits in terms of functionals of Brownian motion, while Abadir (1993, 1995) obtained the explicit expressions for their density and distribution functions. More recently, Dolado, Gonzalo, and Mayoral (2002) generalized the Dickey-Fuller tests to allow for fractional roots in the null hypothesis to be tested. Their limiting distribution results on pp.1969–1070 can be extended by means of our Proposition 3.1 to the case of m.d. innovations instead of just i.i.d. ones, except that they use a different type of fractional Brownian motion; see Marinucci and Robinson (1999).

## 5 Simulations

In this Section we examine the small sample performance of some of the above asymptotic results. We will consider three different experiments, all of them based on a sample size  $n = 500$  and 10,000 replications.

*ARFIMA-ARCH.* We start from the case where  $X_j$  in (2.1) is ARFIMA(1,  $d$ , 0) model with the AR parameter  $r = 0.8$  and m.d. innovations  $\zeta_j$  generated as ARCH(1) process:

$$(5.1) \quad \begin{aligned} \zeta_j &= \varepsilon_j \sigma_j, \quad \varepsilon_j \sim \mathcal{N}(0, 1), \\ \sigma_j^2 &= \alpha_0 + \alpha_1 \zeta_{j-1}^2, \quad j = 1, \dots, n. \end{aligned}$$

We take  $\alpha_0 = 0.2$ ,  $\alpha_1 = 0.8$ , so that the  $E\zeta_j^2 = 1$ . We simulated the cases  $d = 0.1, 0.2, 0.3, 0.4$ , which gave qualitatively similar results. For the sake of brevity, we report the results for  $d = 0.3$  only. We analyze the asymptotic normality of the suitably standardized sums

$$W_n = \sum_{j=1}^n z_{nj} X_j, \quad V_n = \sum_{j=1}^n z_{nj} Y_j,$$

where  $Y_j = \sum_{s=1}^j X_s$ , with the weights  $z_{nj} = (j/n)^2 + \cos(j/n)$ ,  $j = 1, 2, \dots, n$ .  $\text{Var}(W_n)$  and  $\text{Var}(V_n)$  are estimated by the Monte Carlo variances. We then plot a kernel estimate of the densities of  $\widehat{\text{Var}}(W_n)^{-1/2} W_n$  and  $\widehat{\text{Var}}(V_n)^{-1/2} V_n$  using a Gaussian kernel with bandwidth  $b$  chosen according the Silverman's (1986) rule,  $b_n = (4/3)^{1/5} n^{-1/5}$ , and superimpose the standard normal density. The results for ARFIMA(1, 0.3, 0),  $r = 0.8$  process  $\{X_j\}$  with

ARCH(1) innovations in Figure 1 show a very close resemblance between the solid line representing the kernel density estimate and the dashed line representing  $\mathcal{N}(0, 1)$  density.

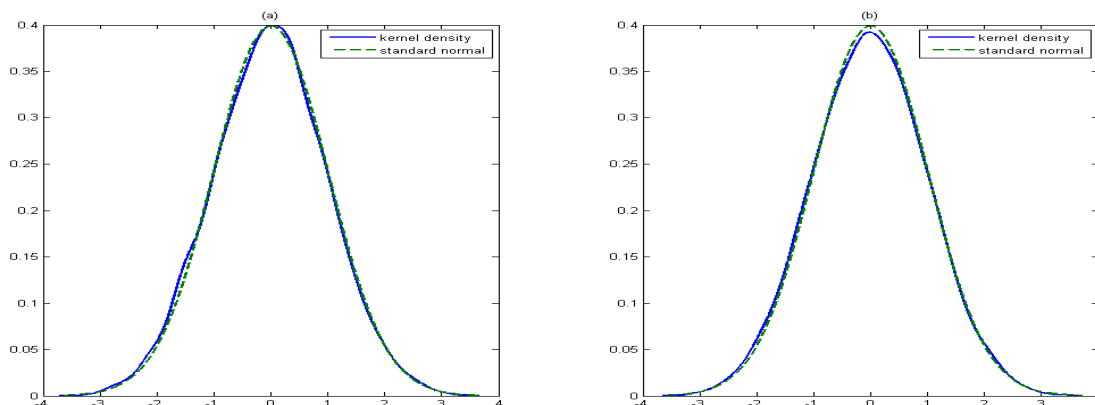


Figure 1: Panel (a): Kernel density of  $\widehat{\text{Var}}(W_n)^{-1/2}W_n$ . Panel (b): Kernel density of  $\widehat{\text{Var}}(V_n)^{-1/2}V_n$ .

*AR-ARCH.* Next we analyze asymptotic normality of  $W_n$  and  $V_n$  in the case of AR(1) process  $X_j = \rho X_{j-1} + \zeta_j$ ,  $j \in \mathbb{Z}$ , with  $\rho = 0.6$  and ARCH(1) errors  $\zeta_j$  as in (5.1). The results in Figure 2 show that the kernel density estimate of  $\widehat{\text{Var}}(W_n)^{-1/2}W_n$  seems to have less probability mass close to the mean than the  $\mathcal{N}(0, 1)$  density and is slightly positively skewed.

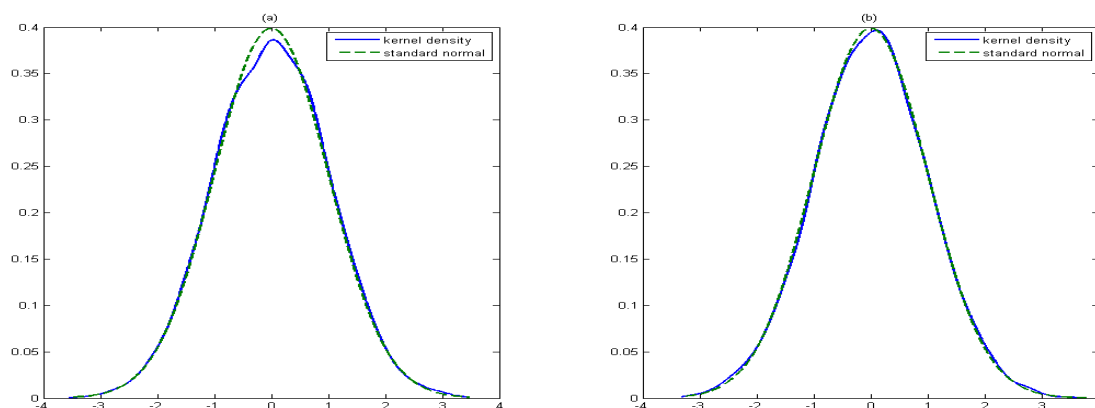


Figure 2: Panel (a): Kernel density estimate of  $\widehat{\text{Var}}(W_n)^{-1/2}W_n$ . Panel (b): Kernel density estimate of  $\widehat{\text{Var}}(V_n)^{-1/2}V_n$ .

*Nonparametric regression.* To illustrate the fit of normal approximation in nonparametric estimation we use the following regression model  $Y_j = \mu(j/n) + X_j$ ,  $j = 1, \dots, n$ , where  $X_j$ 's follow an ARFIMA(1, 0.3, 0)-ARCH(1) process with AR parameter  $r = 0.8$  generated as in (5.1). We set  $\mu(j/n) = (j/n)^2 + \cos(j/n)$  and  $\mu(x)$  is estimated at  $x = 1/4$  by a kernel type estimator

$$\hat{\mu}_n(x) = \frac{\sum_{j=1}^n K\left(\frac{nx-j}{nb}\right)Y_j}{\sum_{j=1}^n K\left(\frac{nx-j}{nb}\right)},$$

using a Gaussian kernel and setting  $b = n^{-1/4}(4/3)^{1/5}SD(J)$ , where  $SD(J)$  is the standard deviation of the regressor  $J = j/n$ ,  $j = 1, 2, \dots, n$ .  $\text{Var}(\hat{\mu}_n(x))$  is estimated by the Monte Carlo variance. Figure 3 shows close fit of the estimated density of  $\widehat{\text{Var}}(\hat{\mu}_n(x))^{-1/2}(\hat{\mu}_n(x) - \mu(x))$  and the standard normal density.

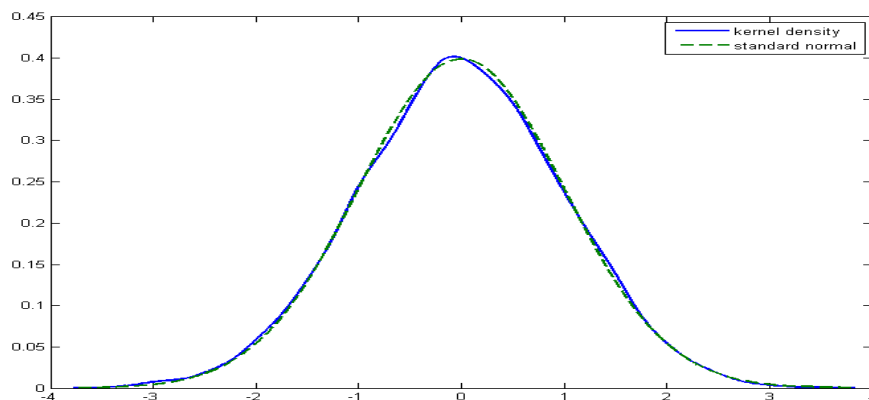


Figure 3: Kernel density of  $\widehat{\text{Var}}(\hat{\mu}_n(x))^{-1/2}(\hat{\mu}_n(x) - \mu(x))$  (solid line),  $\mathcal{N}(0, 1)$  density (dashed line).

**Acknowledgements.** The authors would like to thank Donatas Surgailis for providing part (iii) of Theorem 2.2, Lemma 3.1, and other useful comments, and to the Editor and the three reviewers for their constructive suggestions.

## References

- Abadir, K.M. (1993). The limiting distribution of the autocorrelation coefficient under a unit root. *Annals of Statistics*, **21**, 1058 – 1070.
- Abadir, K.M. (1995). The limiting distribution of the  $t$  ratio under a unit root. *Econometric Theory*, **11**, 775–793.
- Anderson, T.W. (1959). On Asymptotic Distributions of Estimates of Parameters of Stochastic Difference Equations. *Ann. Math. Statist.*, **30**, 676–687.
- Billingsley, P. (1968). *Convergence of Probability Measures*. John Wiley & Sons, Inc., New York.
- Bollerslev, T. (1988). On the correlation structure for the generalized autoregressive conditional heteroskedastic process, *J. Time Ser. Anal.*, **9**, 121–131.
- Borovskikh, Yu. V., and Korolyuk, V. S. (1997). *Martingale Approximation*, VSP, Utrecht.
- Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods (2nd ed.)*. Springer Series in Statistics. Springer-Verlag, New York.
- Crato, N. and de Lima, P.J.F. (1994). Long-range dependence in the conditional variance of stock returns. *Econ. Lett.*, **45**, 281–285.
- Davydov, Y.A. (1970). The invariance principle for stationary processes. *Theor. Probab. Appl.*, **15**, 487–498.
- Dharmadhikari, S. W., Fabian, V. and Jogdeo, K. (1968). Bounds on the moments of martingales *Ann. Math. Statist.*, **39**, 1719–1723.
- Dickey, D.A., and Fuller, W.A. (1979). Distribution of estimators of autoregressive time series with a unit root. *Journal of the American Statistical Association*, **74**, 427 – 431.
- Dickey, D.A. and Fuller, W.A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, **49**, 1057–1072.
- Dolado, J.J., Gonzalo, J. and Mayoral, L. (2002). A fractional Dickey-Fuller test for unit roots. *Econometrica*, **70**, 1963 – 2006.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**, 987–1008.
- Giraitis, L., Kokoszka, P., Leipus, R. (2000). Stationary ARCH models: dependence structure and central limit theorems. *Econometric Theory*, **16**, 3–22.

- Giraitis, R. Leipus and D. Surgailis. (2007). Recent advances in ARCH modelling. In: G. Teyssiere and A.P. Kirman, Editors, *Long Memory in Economics*, Springer, 3-38.
- Gordin, M. I. (1969). The central limit theorem for stationary processes. *Soviet Math. Dokl.*, **10**, 1174–1176.
- Granger, C.W.J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *J. Time Series Anal.*, **1**, 15–29.
- Gray, H.L., Zhang, N.-F., and Woodward, W.A. (1989). On generalized fractional processes. *J. Time Series Anal.*, **10**, 233–257.
- Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests*. Academic Press, New York, N.Y.
- Hall, P. and Heyde, C.C. (1980). *Martingale limit theory and applications*. Academic Press, New York.
- Harvey, A.C. (1998). Long memory in stochastic volatility. In: Knight, J. and Satchell, S., (eds.), *Forecasting Volatility in the Financial Markets*, 307–320, Butterworth & Heineman, Oxford.
- Hosking, J.R.M. (1981). Fractional differencing. *Biometrika*, **68**, 165–176.
- Ibragimov, I.A. and Linnik, Yu.V. (1971). *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen.
- Lamperti, J.W. (1962). Semi-stable stochastic processes. *Trans. Amer. Math. Soc.*, **104**, 62-78.
- Marinucci, D. and Robinson, P.M. (1999). Alternative forms of fractional Brownian motion. *J. Statist. Planning and Inference*, **80**, 111 – 122.
- Merlevède, F., Peligrad, M. and Utev, S. (2006). Recent advances in invariance principles for stationary sequences. *Probability Surveys*, **3**, 1-36.
- Nelson, D.B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, **59**, 347–370.
- Peligrad, M. and Utev, S. (1997). Central limit theorem for stationary linear processes. *Ann. Probab.*, **25**, 443–456.
- Peligrad, M. and Utev, S. (2005). A new maximal inequality and invariance principle for stationary sequences. *Ann. Probab.*, **33**, 798-815.
- Peligrad, M. and Utev, S. (2006). Central limit theorem for stationary linear processes. *Ann. Probab.*, **34**, 1608–1622.
- Phillips, P.C.B. (1987). Time series regression with a unit root. *Econometrica*, **55**, 277 – 301.
- Phillips, P.C.B. and Solo, V. (1992). Asymptotics for linear processes. *Ann. Statist.*, **20**, 971–1001.



- Phillips, P.C.B. and Magdalinos, T. (2007). Limit theory for moderate deviations from a unit root. *Journal of Econometrics*, **136**, 115-130.
- Pollard, D. (1984). *Convergence of Stochastic Processes*. Springer Series in Statistics. Springer-Verlag, New York.
- Robinson, P. M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *J. of Econometrics*, **47**, 67-84.
- Robinson, P.M. (2001). The memory of stochastic volatility models. *J. of Econometrics*, **101**, 195-218.
- Silverman, B.W. (1986). *Density Estimation*. Chapman & Hall, London.
- Surgailis, D. and Viano, M.-C. (2002). Long memory properties and covariance structure of the *EGARCH* model. *ESAIM: Probability and Statistics*, **6**, 311-329.
- Stout, W. (1974). *Almost Sure Convergence*. Academic Press, New York.
- Taqqu, M.S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahr. verw. Gebiete*, **31**, 287-302.
- Wu, W.B. and Woodroffe, M. (2004). Martingale Approximations for Sums of Stationary Processes. *Ann. Probab.*, **32**, 1674-1690.