THE RIMINI CENTRE FOR ECONOMIC ANALYSIS

WP 23_12

K.M. Abadir
Imperial College London, UK
The Rimini Centre for Economic Analysis (RCEA), Italy

W. Distaso
Imperial College London, UK

L. Giraitis
Queen Mary, University of London, UK

H.L. Koul
Michigan State University, USA

ASYMPTOTIC NORMALITY FOR WEIGHTED SUMS OF LINEAR PROCESSES

Asymptotic normality for weighted sums of linear processes [1]

K.M. 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 stationary 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 as long as the process of conditional variances of innovations is covariance stationary with lag k auto-covariances tending to zero, as k tends to infinity. 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(∞) models. They are also useful in deriving asymptotic normality of kernel type estimators of a nonparametric regression function when errors may have long memory.

1 Introduction

Numerous inference procedures in statistics and econometrics are based on weighted sums of dependent random variables. In this paper we focuss on deriving asymptotic distributions of the sums

$$(1.1) \qquad S_n = \sum_{j=1}^n X_j, \qquad W_n = \sum_{j=1}^n z_{nj} X_j$$

of a linear process $\{X_j\}$ weighted by an array of real known weights z_{nj}. In addition, the weak convergence property of the corresponding partial sum processes are also discussed. The linear process is assumed to be a moving average with martingale differences innovations, and may exhibit both short and long range dependence. It will be shown that $\{\mathrm{Var}(T_n)\}^{-1/2}(T_n - ET_n)$, with $T_n = S_n$ or W_n, converges weakly to a normal distribution under minimal assumptions. The proofs of these claims use central limit theorem for martingale differences.

The book of Ibragimov and Linnik (1971) contains a number of useful results on classical asymptotic theory of weekly dependent random variables. Davydov (1970) obtained weak convergence result for the partial sum process with i.i.d. innovations whereas Phillips and Solo (1992) developed CLT and invariance principles for sums of linear processes based on the

[1]Research in part supported by ESRC grant RES062230790 and USA NSF DMS Grant 0704130.
Keywords: Linear process, weighted sum, Lindeberg-Feller.

Beveridge-Nelson decomposition. Peligrad and Utev (2006) extended Ibragimov and Linnik (1971, Theorem 18.6.5) for linear processes with dependent innovations. Wu and Woodroofe (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 the asymptotic normality of sums, whereas in Section 3 we discuss the weak convergence of partial sum processes. Sections 4 and 5 contain examples, applications and simulation study. For any two sequences, $a_n\, b_n$, $a_n \sim b_n$ means $a_n/b_n \to 1$, as $n \to \infty$.

2 CLT for weighted sums

In this section we consider the asymptotic normality of the sums S_n and W_n of (1.1). If $\{X_j\}$ is a zero-mean Gaussian process, then $(\mathrm{Var}(S_n))^{-1/2}S_n \sim \mathcal{N}(0,1)$ and the question about the asymptotic distribution of S_n reduces to finding the asymptotics of the $\mathrm{Var}(S_n)$. Let $\{X_j\}$ be a linear process

$$X_j = \sum_{k=0}^{\infty} a_k \zeta_{j-k} = \sum_{k=-\infty}^{j} a_{j-k}\zeta_k, \quad j \in \mathbb{Z}, \tag{2.1}$$

with white noise innovations $\{\zeta_j\} \sim WN(0,\sigma_\zeta^2)$ and finite variance $EX_j^2 = \sigma_\zeta^2 \sum_{k=0}^{\infty} a_k^2 < \infty$.

The existing literature on asymptotic distributions of the sums of X_j's often assumes that the innovations ζ_j, $j \in \mathbb{Z}$ are i.i.d. random variables with zero mean and variance σ_ζ^2, $\{\zeta_j\} \sim IID(0,\sigma_\zeta^2)$. In some applications this assumption may be too restrictive. The present paper allows for a wide class of heteroscedastic martingale difference innovations.

Assumption 2.1 $\{\zeta_j\}$ *is a stationary sequence of martingale differences with respect to the natural filtration* $\mathcal{F}_{j-1} = \sigma(\zeta_{j-1}, \zeta_{j-2,}...)$. *The conditional variance* $V_j := E(\zeta_j^2|\mathcal{F}_{j-1})$, $j \in \mathbb{Z}$ *is a covariance stationary process such that* $\gamma_V(k) := \mathrm{Cov}(V_k, V_0) \to 0$, *as* $k \to \infty$.

We shall write $\{\zeta_j\} \sim MD(0,\sigma_\zeta^2)$, to indicate that $\{\zeta_j\}$ satisfies Assumption 2.1. Note that $E(\zeta_j|\mathcal{F}_{j-1}) = 0$, $EV_j = E\zeta_j^2 = \sigma_\zeta^2$, whereas the conditional variance V_t may be a either constant, e.g. if ζ_j's are i.i.d. random variables, or a stochastic process. In Section 4 it is shown that Assumption 2.1 is satisfied by ARCH type white noises.

Theorem 18.6.5 of Ibragimov and Linnik (1971) gives a CLT for S_n in case of i.i.d. innovations under general conditions. Theorem 2.1 below extends it to martingale difference innovations $\{\zeta_j\} \sim MD(0,\sigma_\zeta^2)$.

Theorem 2.1 *Suppose* $\{X_j\}$ *is a stationary linear process (2.1) with* $\{\zeta_j\} \sim MD(0,\sigma_\zeta^2)$. *Then,*

$$\sigma_n^2 := \mathrm{Var}(S_n) \to \infty, \tag{2.2}$$

implies

$$\sigma_n^{-1} S_n \to_D \mathcal{N}(0,1). \tag{2.3}$$

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

$$\sigma_n^{-1} S_n = \sum_{j=-\infty}^{n} c_{nj} \zeta_j, \qquad \sigma_n^{-2} \mathrm{Var}(S_n) = \sum_{j=-\infty}^{n} c_{nj}^2 = 1, \quad \forall\, n \geq 1. \tag{2.4}$$

Next we show

$$c_n := \sup_{j \leq n} |c_{nj}| = o(1), \tag{2.5}$$

which together with Lemma 2.1 below implies (2.3) and completes the proof of the theorem.

To prove (2.5), note that for any $s \in \mathbb{Z}$, $t = 1, 2, \dots,$

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

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

$$\begin{aligned} c_{n,s}^2 &= 2 \sum_{j=-\infty}^{s} c_{n,j-1}(c_{nj} - c_{n,j-1}) + \sum_{j=-\infty}^{s} (c_{nj} - c_{n,j-1})^2 \\ &\leq 2B_n + B_n^2, \end{aligned} \tag{2.6}$$

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

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

because $\sigma_n^2 \to \infty$ and $EX_1^2 = \sum_{j=0}^{\infty} a_j^2 < \infty$, proving (2.5).

Lemma 2.1 *Suppose* $S_n = \sum_{j=-\infty}^{n} d_{nj} \zeta_j$, $n \geq 1$, *where* $\{\zeta_j\} \sim MD(0,1)$ *and* $\mathrm{Var}(S_n) = \sum_{j=-\infty}^{n} d_{nj}^2 = 1$. *Then,*

$$d_n := \sup_{j \leq n} |d_{nj}| = o(1). \tag{2.7}$$

implies

$$S_n \to_D \mathcal{N}(0,1). \tag{2.8}$$

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

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

Write

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

Then $Es_{n,1}^2 = \sum_{j=-\infty}^{-M-1} d_{nj}^2 \le 1/\log(n) \to 0$ implies $s_{n,1} = o_p(1)$. To prove (2.8), it remains to show

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

By the CLT for martingale differences, see Hall and Heyde (1980, Corollary 3.1), to prove (2.10) it suffices to check that, as $n \to \infty$,

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

$$\sum_{j=-M}^{n} E[|d_{nj}\zeta_j|^2 1_{\{|d_{nj}\zeta_j| \ge \delta\}} | \mathcal{F}_{j-1}] \to_p 0, \tag{2.12}$$

for all $\delta > 0$. Let $q_n := \sum_{j=-M}^{n} d_{nj}^2 V_j$ denote the left hand side of (2.11), where $V_j = E[\zeta_j^2 | \mathcal{F}_{j-1}]$. We shall show that

$$Eq_n \to 1, \qquad \mathrm{Var}(q_n) \to 0, \quad n \to \infty,$$

which will imply (2.11). The first claim follows from $EV_j = E\zeta_j^2 = 1$, (2.9) and $\sum_{j=-\infty}^{n} d_{nj}^2 = 1$:

$$Eq_n = \sum_{j=-M}^{n} d_{nj}^2 = \sum_{j=-\infty}^{n} d_{nj}^2 + O(1/\log(n)) \to 1.$$

To show the second fact, set $K = K(n) = \lfloor 1/d_n \rfloor$. By (2.7), $K \to \infty$, $Kd_n^2 \le Cd_n \to 0$. Then, in view of Assumption 2.1,

$$\max_{s>K} |\gamma_V(s)| \to 0, \quad \max_{0\le s\le K} |\gamma_V(s)| \le EV_1^2 < \infty.$$

Thus,

$$\begin{aligned}
\mathrm{Var}(q_n) &= \sum_{j,k=-M}^{n} d_{nj}^2 d_{nk}^2 \gamma_V(j-k) = \sum_{j,k=-M:\,|j-k|>K}^{n} [\cdots] + \sum_{j,k=-M:\,|j-k|\le K}^{n} [\cdots] \\
&\le \max_{s>K} |\gamma_V(s)| \Big(\sum_{j=-\infty}^{n} d_{nj}^2 \Big)^2 + \max_{0\le s\le K} |\gamma_V(s)| d_n^2 \sum_{j=-M}^{n} d_{nj}^2 \sum_{k:|j-k|\le K} 1 \\
&\le \max_{s>K} |\gamma_V(s)| + (2K+1) d_n^2 \to 0,
\end{aligned}$$

which completes proof of (2.11).

Next, to prove (2.12), by the definition of K,

$$\max_{j\le n}|d_{nj}\zeta_j|^2 1_{\{|\zeta_j|\le K^{1/2}\}} \le d_n^2 K \to 0, \quad E\zeta_1^2 1_{\{|\zeta_1|>K^{1/2}\}} \to 0, \quad n\to\infty.$$

Therefore,

$$\begin{aligned} E\sum_{j=-M}^{n} E[|d_{nj}\zeta_j|^2 1_{\{|d_{nj}\zeta_j|\ge\delta\}}|\mathcal{F}_{j-1}] &= \sum_{j=-M}^{n} E[|d_{nj}\zeta_j|^2 1_{\{|d_{nj}\zeta_j|\ge\delta\}}] \\ &\le d_n^2 K \sum_{j=-M}^{n} E1_{\{|d_{nj}\zeta_j|\ge\delta\}} + E\zeta_1^2 1_{\{|\zeta_1|>K^{1/2}\}}\sum_{j=-M}^{n} d_{nj}^2 \\ &\le o(1)\delta^{-2}\sum_{j=-M}^{n} d_{nj}^2 + o(1) \to 0, \quad n\to\infty, \end{aligned}$$

which yields (2.12) and completes the proof of the lemma.

Remark 2.1 Note that the above lemma clearly covers the case of $\{\zeta_j\} \sim IID(0,1)$. Furthermore, the above theorem is a generalization of Theorem V.1.2.1 of Hájek and Šidák (1967) where $d_{nj}=0$, $j\le 0$ and $\{\zeta_j\}\sim IID(0,1)$.

Weighted sums. Theorem 2.1 can be generalized to weighted sums of linear processes. Let z_{nj}, $j,n\ge 1$ be an array of real numbers, and consider the weighted sums

$$W_n = \sum_{j=1}^{n} z_{nj}X_j. \tag{2.13}$$

The following proposition gives the limiting distribution of W_n. Now $\sigma_n^2 := \mathrm{Var}(W_n)$.

Theorem 2.2 *Suppose $\{X_j\}$ is a stationary linear process (2.1) with $\{\zeta_j\}\sim MD(0,\sigma_\zeta^2)$. Suppose the weights z_{nj} in W_n and a_j in (2.1) satisfy one of the following three conditions:*

(i) $\max_{1\le j\le n}|z_{nj}| = o(\sigma_n)$, *and* $\sum_{j=1}^{n} z_{nj}^2 \le C\sigma_n^2$,

(ii) $\max_{1\le j\le n}|z_{nj}| = o(\sigma_n)$, *and* $\sum_{j=0}^{\infty}|a_j| < \infty$,

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

Then, with $\sigma_n^2 := Var(W_n)$,

$$\sigma_n^{-1}W_n \to_D \mathcal{N}(0,1). \tag{2.14}$$

Proof. Similarly as in (2.4), set $a_k=0$, $k<0$, and let $c_{nj}=\sigma_n^{-1}\sum_{k=1}^{n} z_{nk}a_{k-j}$, $j\in\mathbb{Z}$. Then,

$$\sigma_n^{-1}W_n = \sum_{j=-\infty}^{n} c_{nj}\zeta_j. \tag{2.15}$$

Assume for simplicity $\sigma_\zeta^2 = 1$. Since $\mathrm{Var}(\sigma_n^{-1} W_n) = \sum_{j=-\infty}^{n} c_{nj}^2 = 1$, by Lemma 2.1, it suffices to verify relation (2.7) for the coefficients c_{nj}, i.e. to show that $\sup_{j\in\mathbb{Z}} |c_{nj}| = o(1)$, or (2.5).

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

$$\begin{aligned} |c_{nj}| &\le \sigma_n^{-1} \sum_{k=1}^{n} |z_{nk} a_{k-j}| 1_{\{|k-j|\ge K_n\}} \\ &\quad + \sigma_n^{-1} \sum_{k=1}^{n} |z_{nk} a_{k-j} 1_{\{|k-j|<K_n\}} := q_{n,1j} + q_{n,2j}, \quad \text{say.} \end{aligned}$$

By the Cauchy-Schwarz inequality,

$$\begin{aligned} q_{n,1j} &\le \frac{(\sum_{k=1}^{n} z_{nk}^2)^{1/2}}{\sigma_n} (\sum_{k=1}^{n} a_{k-j}^2 1_{\{|k-j|\ge K_n\}})^{1/2} \le C \sum_{i\ge K_n} a_i^2, \\ q_{n,2j} &\le \frac{\max_{1\le k\le n} |z_{nk}|}{\sigma_n} \sum_{k=1}^{n} |a_{k-j}| 1_{\{|k-j|<K_n\}} \\ &\le K_n^{-1} (2K_n)^{1/2} \Big(\sum_{i=0}^{\infty} a_i^2 \Big)^{1/2} \le C K_n^{-1/2}, \qquad \forall\, j \le n. \end{aligned}$$

Hence, assumption (i) and $\sum_{i\in\mathbb{Z}} a_i^2 < \infty$ yield

$$\max_{j\le n} |c_{nj}| \le C \big(\sum_{i\ge K_n} a_i^2 + K_n^{-1/2} \big) \to 0,$$

thereby proving (2.5).

In Case (ii), (2.5) follows because

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

To verify (2.5) in Case (iii), we use the bound (2.6) which does not depend on any particular form of c_{nj}. It suffices to show that $B_n \to 0$. Define for simplicity $z_{n0} = z_{n,n+1} = 0$. Then one can write

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

Hence,

$$\begin{aligned} B_n^2 &= \sum_{j\in\mathbb{Z}} (c_{nj} - c_{n,j-1})^2 \\ &= \sigma_n^{-2} \sum_{k,s=1}^{n+1} (z_{nk} - z_{n,k-1})(z_{ns} - z_{n,s-1}) \sum_{j\in\mathbb{Z}} a_{k-j} a_{s-j} \\ &\le \sigma_n^{-2} \sum_{k,s=1}^{n+1} |z_{nk} - z_{n,k-1}| |z_{ns} - z_{n,s-1}| \sum_{j=0}^{\infty} a_j^2 \end{aligned}$$

$$\leq \; C\sigma_n^{-2}\Big(|z_{n1}+|z_{nn}|+\sum_{k=2}^{n}|z_{nk}-z_{n,k-1}|\Big)^2 \to 0,$$

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

Verification of conditions (i)-(iii) of Theorem 2.2 requires further analysis of the variance $\sigma_n^2 \equiv \mathrm{Var}(W_n) = \sum_{j,k=1}^n z_{nj}\gamma_X(j-k)z_{nk}$. The next proposition provides sufficient conditions for (i) and (iii) to hold in terms of the spectral density f of $\{X_j\}$ and weights z_{nk}. Part a) requires f to be only continuous at 0. This includes spectral densities that may have infinite peaks or may be vanishing on an interval not including the origin, e.g. spectral densities of seasonal GARMA long memory models. Condition (2.16) on weights is mild and satisfied in most of the applications. Part b) allows the spectral density to be unbounded at 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.

Proposition 2.1 *Assume $\{X_j\}$, z_{nj} and W_n are as in Theorem 2.2, with f denoting the spectral density of $\{X_j\}$.*

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

$$|z_{n1}|+|z_{nn}|+\sum_{j=2}^{n}|z_{nj}-z_{n,j-1}| = o\Big(\big(\sum_{j=1}^{n} z_{nj}^2\big)^{1/2}\Big). \tag{2.16}$$

Then, (iii) *of Theorem 2.2 is satisfied, and hence, (2.14) holds. Moreover,*

$$\mathrm{Var}(W_n) \sim 2\pi f(0)\sum_{j=1}^{n} z_{nj}^2. \tag{2.17}$$

(b) *If $f(u)\geq c>0$, $|u|\leq u_0$, for some $c>0$ and $u_0>0$, and (2.16) is valid, then* (iii) *of Theorem 2.2 is satisfied and (2.14) holds.*

(c) *Suppose there exists a $c>0$ such that $f(u)\geq c>0$, $u\in\Pi:=[-\pi,\pi]$, and*

$$\max_{1\leq j\leq n}|z_{nj}| = o\Big(\big(\sum_{j=1}^{n} z_{nj}^2\big)^{1/2}\Big). \tag{2.18}$$

Then, (i) *of Theorem 2.2 is satisfied, and hence, (2.14) holds.*

Proof. (a) Let $G(u) := \sum_{j=1}^n e^{-\mathbf{i}ju}z_{nj}$, $u\in\Pi$. Since $\gamma_X(k)=\int_\Pi e^{\mathbf{i}ku}f(u)\,du$, one can write

$$\sigma_n^2 = \sum_{j,k=1}^{n} z_{nj}\gamma_X(j-k)z_{nk} = \int_\Pi f(u)|G(u)|^2du. \tag{2.19}$$

Observe that

$$\begin{aligned}(2.20)\qquad \int_\Pi |G(u)|^2 du &= \int_\Pi \sum_{j,k=1}^n e^{\mathbf{i}(j-k)u} z_{nj} z_{nk} du = 2\pi \sum_{j=1}^n z_{nj}^2, \\ \sigma_n^2 &= 2\pi f(0) \sum_{j=1}^n z_{nj}^2 + i_n, \quad i_n := \sigma_n^2 - 2\pi f(0) \sum_{j=1}^n z_{nj}^2.\end{aligned}$$

We shall show that

$$(2.21)\qquad i_n = \big|\sigma_n^2 - 2\pi f(0) \sum_{j=1}^n z_{nj}^2\big| = o(\sum_{j=1}^n z_{nj}^2),$$

which proves (2.17) and together with (2.16) implies condition (iii) Theorem 2.2 and thus (2.14).

Let $\epsilon > 0$. Choose $\delta > 0$, such that $\sup_{0\le u\le\delta} |f(u) - f(0)| \le \epsilon/2\pi$. Then by (2.19) and (2.20),

$$\begin{aligned} i_n &\le \int_{|u|\le\pi} |f(u) - f(0)||G(u)|^2 du \\ &= \int_{|u|\le\delta} [\cdots]du + \int_{\delta<|u|\le\pi} [\cdots]du := s_{n,1} + s_{n,2},\end{aligned}$$

where

$$s_{n,1} \le \frac{\epsilon}{2\pi} \int_{|u|\le\delta} |G(u)|^2 du \le \frac{\epsilon}{2\pi} \int_\Pi |G(u)|^2 du = \epsilon \sum_{j=1}^n z_{nj}^2.$$

To prove (2.21), it suffices to check that for any $\delta > 0$,

$$(2.22)\qquad s_{n,2} = o\big(\sum_{j=1}^n z_{nj}^2\big).$$

Using summation by parts, write

$$G(u) = \sum_{j=1}^{n-1} (\sum_{l=1}^j e^{-\mathbf{i}lu})(z_{nj} - z_{n,j+1}) + z_{nn} \sum_{l=1}^n e^{-\mathbf{i}lu}.$$

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

$$\begin{aligned} |\sum_{l=1}^j e^{\mathbf{i}lu}| &= |\frac{\sin(ju/2)}{\sin(u/2)}| \le |\sin(nu/2)|^{-1} \\ &\le \pi u^{-1} \le \pi\delta^{-1}, \qquad \delta \le u \le \pi.\end{aligned}$$

Therefore,

$$|G(u)| \le \pi\delta^{-1}\big(\sum_{j=1}^{n-1}|z_{nj}-z_{n,j+1}|+|z_{nn}|\big) = o\big(\sum_{j=1}^{n} z_{nj}^2\big),$$

by (2.16). Hence,

$$s_{n,2} = o\big(\sum_{j=1}^{n} z_{nj}^2\big)\int_{\delta<|u|\le\pi}(f(u)+f(0))du = o\big(\sum_{j=1}^{n} z_{nj}^2\big),$$

which implies (2.22).

(b) Define $\tilde f(u) = cI(|u| \le u_0)$, $u \in \Pi$. Notice that $f \ge \tilde f$, and $\tilde f(u) \to \tilde f(0) = c > 0$, $u \to 0$. Thus, by (a),

$$\sigma_n^2 = \int_\Pi f(u)|G(u)|^2du \ge \int_\Pi \tilde f(u)|G(u)|^2du \sim 2\pi\tilde f(0)\sum_{j=1}^{n} z_{nj}^2.$$

This together with (2.16) verifies (iii) of Theorem 2.2 and implies (2.14).

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

$$\sigma_n^2 \ge c\int_\Pi |G(u)|^2du = 2\pi\, c\sum_{j=1}^{n} z_{nj}^2, \tag{2.23}$$

which together with (2.18) yields (i) of Theorem 2.2 and therefore proves (2.14) and completes the proof of the proposition.

The following proposition, which is valid for any stationary process $\{X_j\}$ not necessarily linear, provides an upper bound for $\mathrm{Var}(W_n)$ and analyzes its asymptotic behavior. These results are found useful when analyzing the variance of kernel-type estimators of nonparametric regression functions.

Proposition 2.2 *Let $\{X_j\}$ be a covariance-stationary process with zero mean, finite variance and covariance function γ such that*

$$\sum_{k\in\mathbb{Z}}|\gamma(k)| < \infty. \tag{2.24}$$

Let W_n be as in (2.13) with z_{nj} satisfying conditions (2.16) and (2.18). Then,

$$\begin{aligned}\big(\sum_{j=1}^{n} z_{nj}^2\big)^{-1}\mathrm{Var}(W_n) &\le \sum_{k\in\mathbb{Z}}|\gamma(k)|,\\ \big(\sum_{j=1}^{n} z_{nj}^2\big)^{-1}\mathrm{Var}(W_n) &\to \sigma^2 := \sum_{k\in\mathbb{Z}}\gamma(k).\end{aligned} \tag{2.25}$$

Proof. Under (2.24),

$$f(u) = (2\pi)^{-1}\sum_{k\in\mathbb{Z}} e^{-\mathbf{i}ku}\gamma(k) \le (2\pi)^{-1}\sum_{k\in\mathbb{Z}}|\gamma(k)| < \infty, \quad u\in\Pi,$$
$$f(u) \to f(0) = (2\pi)^{-1}\sum_{k\in\mathbb{Z}}\gamma(k) = (2\pi)^{-1}\sigma^2, \quad u\to 0.$$

Thus, by (2.19) and (2.20),

$$\mathrm{Var}(W_n) = \int_\Pi f(u)|G(u)|^2du \le \sup_{u\in\Pi} f(u)\int_\Pi |G(u)|^2du = \sum_{k\in\mathbb{Z}}|\gamma(k)|\sum_{j=1}^n z_{nj}^2,$$

which proves the first bound of (2.25). The proof of the second bound is the same as that of (2.17). This completes the proof of the proposition.

The following result is a multivariate generalization of Theorem 2.2.

Theorem 2.3 *Let $z_{n,j}^{(i)}$, $i = 1,\cdots,k$ be now k arrays of real weights, and $\{X_j\}$ same as in Theorem 2.2. Assume that each sum $W_n^{(i)} := \sum_{j=1}^n z_{n,j}^{(i)}X_j$ and $(\sigma_n^{(i)})^2 := \mathrm{Var}(W_n^{(i)})$, $i=1,\dots,k$, satisfies one of the conditions of (i) -(iii) of Theorem 2.2, and for some (positive definite) matrix Σ,*

(2.26)
$$\Big(\mathrm{Cov}(W_n^{(i)}/\sigma_n^{(i)},\ W_n^{(j)}/\sigma_n^{(j)})\Big)_{i,j=1,\dots,k} \to \Sigma.$$

Then,

(2.27)
$$\Big(W_n^{(1)}/\sigma_n^{(1)},\dots,W_n^{(k)}/\sigma_n^{(k)}\Big) \to_D \mathcal{N}_k(0,\Sigma).$$

Proof. Similarly as in (2.15), write

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

To prove (2.27), in view of the Cramér-Wold device, it suffices to show that for every $a = (a_1,\dots,a_k)\in\mathbb{R}^k$,

(2.28)
$$S_n := a_1S_n^{(1)} + \cdots + a_kS_n^{(k)} \to_D N(0, a'\Sigma a).$$

Write $S_n = \sum_{j=-\infty}^n c_{nj}\zeta_j$ where $c_{nj} = a_1c_{nj}^{(1)} + \cdots + a_kc_{nj}^{(k)}$. Condition (2.26) implies that

$$\mathrm{Var}(S_n) = E\zeta_1^2\sum_{j=-\infty}^n c_{nj}^2 \to a'\Sigma a.$$

The proof of Theorem 2.2 demonstrated that any one of the conditions (i), (ii) or (iii) assures that $\sup_{-\infty<j\le n}|c^{(i)}_{nj}| \to 0$, $i = 1, \cdots, k$, which yields $\sup_{-\infty<j\le n}|c_{nj}| \to 0$. Hence the coefficients c_{nj} of S_n satisfy the assumptions of Lemma 2.1 which implies (2.28) and completes the proof.

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 σ_n^2, since the remaining conditions on the weights z_{nj} are usually easy to verify. Assumption (ii) can be applied in the case of short and negative memory linear processes $\{X_j\}$ that satisfy $\sum_{j=0}^{\infty}|a_j| < \infty$, whereas condition (i) can be useful in case when $\{X_j\}$ has long or short memory.

As can be seen from the proofs of Theorems 2.2 and 2.3, their conclusions remain valid if the linear process $\{X_j\}$ in W_n is replaced by the triangular arrays of linear processes $\{X_{nj} = \sum_{k=0}^{\infty} a_{nk}\zeta_{j-k},\ j \in \mathbb{Z}\}$, $n = 1, 2, \cdots$, with innovations $\{\zeta_j\}$ not depending on n, and such that $\mathrm{Var}(X_{n1}) = \sigma_\zeta^2 \sum_{k=0}^{\infty} a_{nk}^2 < \infty$, $\forall\, n \ge 1$, and conditions (i)-(iii) of Theorem 2.2 are replaced by the following assumptions.

(i′) $\Big(\sum_{j=1}^{n} z_{nj}^2\Big)\Big(\sum_{k=0}^{\infty} a_{nk}^2\Big) = o(\sigma_n^2)$,

(ii′) $\Big(\max_{1\le j\le n}|z_{nj}|\Big)\sum_{k=0}^{\infty}|a_{nk}| = o(\sigma_n)$,

(iii′) $\Big(|z_{n1}| + |z_{nn}| + \sum_{j=2}^{n}|z_{nj} - z_{n,j-1}|\Big)\Big\{\sum_{k=0}^{\infty} a_{nk}^2\Big\}^{1/2} = o(\sigma_n)$.

3 Weak convergence of partial sum processes

A number of econometric applications require establishing convergence of finite dimensional distributions of the partial sums process $S_n(\tau) = \sum_{j=1}^{[n\tau]} X_j$, $\tau > 0$ to those of some limit process $S(\tau)$, $0 \le \tau \le 1$. We denote such convergence by $\to_{fdd}$.

According to Lamperti Theorem (1962), with $\sigma_n^2 := \mathrm{Var}(S_n(1))$, if the process $\{\sigma_n^{-1}S_n(\tau)\} \to_{fdd} \{S(\tau)\}$ has a f.d.d. limit, then for some $H \in (0, 1)$ and a positive slowly varying function L,

$$\sigma_n^2 = \mathrm{Var}(S_n) = n^{2H}L(n), \qquad S_n = \sum_{j=1}^{n} X_j. \tag{3.1}$$

In most applications

$$\sigma_n^2 = \mathrm{Var}(S_n) \sim s^2 n^{2H} \to \infty, \quad \text{for some } 0 < H < 1, \tag{3.2}$$

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

In case of a linear process $\{X_j\}$ of (2.1), with innovations $\{\zeta_j\} \sim MD(0, \sigma_\zeta^2)$, (3.2) is also a sufficient condition for this convergence.

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

$$r_H(s,t) := \frac{1}{2}\Big\{|s|^{2H} + |t|^{2H} - |s-t|^{2H}\Big\}, \quad 0 \le s,t \le 1. \tag{3.3}$$

Note that if $H = 1/2$, then $B_{1/2} = B$ is Brownian motion.

Proposition 3.1 *In addition to the conditions of Theorem 2.1 assume (3.2) holds. Then,*

$$\{\sigma_n^{-1} S_n(\tau)\}_{\tau>0} \to_{fdd} \{B_H(\tau)\}_{\tau>0}, \tag{3.4}$$

where B_H is the fBm with parameter H.

Proof. Let $T_n(\tau) := \sigma_n^{-1} S_n(\tau)$. Assumption (3.2) implies that for any $0 < \tau < 1$,

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

and hence,

$$\begin{aligned} \mathrm{Cov}(T_n(t), T_n(s)) &= (1/2)\{\mathrm{Var}(T_n(t)) + \mathrm{Var}(T_n(s)) - \mathrm{Var}\big(T_n(t) - T_n(s)\big)\} \\ &\to r_H(t,s), \qquad \forall\, 0 < s < t, \end{aligned} \tag{3.5}$$

where r_H is as in (3.3). 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 $k \ge 1$, $\mathcal{S}_n = a_1 S_n(\lfloor nt_1 \rfloor) + \cdots + a_k S_n(\lfloor nt_k \rfloor)$ satisfy

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

By (3.5), $\mathrm{Var}(\mathcal{S}_n) \to \infty$, and the rest of the proof repeats the lines of proof of Theorem 2.1.

Weak convergence. Here we shall discuss the weak convergence of a suitably standardized partial sum process $S_n(\tau)$, $0 \le \tau \le 1$. Observe that for each n, the re-normalized partial sum process $\sigma_n^{-1} S_n(\tau)$, $0 \le \tau \le 1$, is a step function in τ, belonging to the Skorokhod functional space $\mathcal{D}[0,1]$.

From Billingsley (1968), we recall that a sequence of stochastic processes $\{Z_n(\cdot)\}$, $n \ge 1$ in $\mathcal{D}[0,1]$ is said to converge weakly to a stochastic process $Z(\cdot) \in \mathcal{C}[0,1]$ 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.

Using the arguments in section 12 and Theorem 15.5, p. 127 of Billingsley (1968), one can show that a sufficient condition for tightness 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 $\forall\, 0 \le s < t \le 1$,

$$E|Z_n(t) - Z_n(s)|^\gamma \le C[F_n(t) - G_n(s)]^\beta, \qquad \forall\, n \ge 1, \tag{3.6}$$

where C may depend on γ, but not on s, t and n.

We are now ready to discuss the weak convergence of $S_n(\cdot)/\sigma_n$. 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 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 (3.2), with an $1/2 < H < 1$. Then, $\sigma_n^{-1}S_n(\cdot)$ is tight w.r.t. uniform norm. In addition, if $\sigma_n^{-1}S_n \to_{fdd} S$, where $S(u)$, $0 \le u \le 1$ is a stochastic process, then,*

$$\sigma_n^{-1}S_n(\cdot) \Rightarrow S(\cdot), \quad \text{in } \mathcal{D}[0,1] \text{ and the uniform metric.} \tag{3.7}$$

Proof. To check tightness, we shall verify (3.6) for the process $T_n(t) := \sigma_n^{-1}S_n(t)$. Let $F_n(t) := \lfloor nt \rfloor/n$, $F(t) := t$, $0 \le t \le 1$. Observe that $\sup_t |F_n(t) - F(t)| \to 0$ and F is continuous on $[0,1]$. By stationarity of increments,

$$\begin{aligned} E|T_n(t) - T_n(s)|^2 &= E\Big(\sigma_n^{-1}\sum_{j=1}^{\lfloor nt \rfloor - \lfloor ns \rfloor} X_j\Big)^2 \\ = &\ \Big(\frac{\lfloor nt \rfloor - \lfloor ns \rfloor}{n}\Big)^{2H}\frac{(1+o(1))}{(1+o(1))} \le C\Big(\frac{\lfloor nt \rfloor - \lfloor ns \rfloor}{n}\Big)^{2H} = C\,[F_n(t) - F_n(s)]^{2H}, \end{aligned} \tag{3.8}$$

for some $\delta > 0$. Since $\beta := 2H > 1$, this verifies (3.6) for the T_n process with $\gamma = 2$ and completes the proof.

The following lemma, where $S_n \equiv S_n(1)$, gives a useful bound for moments of the sums of a linear process. Its proof follows from the Rosenthall's inequality, see e.g., p.24, Hall and Heyde (1984).

Lemma 3.1 *Let $\{X_j\}$ be a linear process (2.1) with $\{\zeta_j\} \sim MD(0,\sigma_\zeta^2)$, and $E|\zeta_0|^p < \infty$, for some $p \ge 2$. Then*

$$E|S_n|^p \le c(ES_n^2)^{p/2}, \quad \forall\, n \ge 1, \tag{3.9}$$

where $c > 0$ depends only on p and $E|\zeta_0|^p$.

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

Proposition 3.3 *Assume that the linear process $\{X_j\}$ of (2.1) with $\{\zeta_j\} \sim MD(0,\sigma_\zeta^2)$ satisfies (3.2) with $0 < H < 1$. For $0 < H \le 1/2$, assume, in addition, that $E|\zeta_0|^p < \infty$, for some $p > 1/H$. Then,*

$$\sigma_n^{-1}S_n(\cdot) \Rightarrow B_H(\cdot), \quad \text{in } \mathcal{D}[0,1] \text{ and uniform metric,} \tag{3.10}$$

where B_H is a fBm.

Proof. Proposition 3.1 implies the finite dimensional convergence. For $H \le 1/2$, by (3.9) and (3.8),

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

This verifies (3.6) for the T_n process with $\gamma = p$, $\beta = Hp > 1$. For $H > 1/2$, (3.10) follows from Proposition 3.2. This completes the proof.

4 Applications

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

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

$$f(u) \sim c_f |u|^{-2d}, \quad u \to 0, \qquad |d| < 1/2, \quad c_f > 0. \tag{4.1}$$

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

$$\begin{aligned} \sum_{j\in\mathbb{Z}} |\gamma(j)| \ &< \ \infty, \quad \sigma^2 := \sum_{j\in\mathbb{Z}} \gamma(j) > 0, \quad \text{for } d = 0, \\ \gamma(j) \ &\sim \ c_\gamma |j|^{-1+2d}, \quad 0 < d < 1/2, \\ \gamma(j) \ &\sim \ c_\gamma |j|^{-1+2d}, \quad \sum_{j\in\mathbb{Z}} \gamma(j) = 0, \quad -1/2 < d < 0, \end{aligned} \tag{4.2}$$

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 $\{\zeta_j\} \sim MD(0, \sigma_\zeta^2)$. Assume that the spectral density f or autocovariance γ satisfies (4.1) or (4.2) with some $|d| < 1/2$. Then,*

$$\begin{aligned} \sigma_n^2 \ &\sim \ s^2 n^H, \qquad H = 1/2 + d, \\ n^{-1/2-d} S_n \ &\to_D \ \mathcal{N}(0, s^2), \end{aligned} \tag{4.3}$$

where

$$\begin{aligned} s^2 \ &= \ 2\pi f(0),\ d = 0; \quad s^2 = c_f \int_{\mathbb{R}} \frac{\sin^2(u/2)}{(u/2)^2} |u|^{-2d} du,\ 0 < |d| < 1/2, \quad \textit{under (4.1)}, \\ s^2 \ &= \ \sum_{k\in\mathbb{Z}} \gamma(k),\ d = 0; \quad s^2 = c_\gamma/(d(1+2d)),\ 0 < |d| < 1/2, \quad \textit{under (4.2)}. \end{aligned}$$

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

$$n^{-1/2-d} S_n(\cdot) \Rightarrow s B_{1/2+d}(\cdot), \tag{4.4}$$

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

Proof. The claim (4.3) about σ_n^2 under (4.1) or (4.2) is known. Together with Theorem 2.1 and Proposition 3.3, it proves (4.3) and (4.4).

Recall also the following known relationship between the weights a_k and $\gamma(j)$ when $\{X_j\}$ is a linear process (2.1) with $\{\zeta_j\} \sim MD(0, \sigma_\zeta^2)$. 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 $c_a \ne 0$, then $\gamma(k)$ satisfies (4.2) 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 $\sum_{k=0}^\infty a_k \ne 0$ then

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

Applications of Theorem 2.2. Let $\{X_j\}$ be a linear process with memory parameter $0 \le d < 1/2$. 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 β. Suppose

$$z_{nj} = g(j/n), \quad j = 1, \cdots, n, \tag{4.5}$$

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

$$\sum_{k=0}^\infty |\gamma(k)| < \infty, \qquad \sum_{k \in \mathbb{Z}} \gamma(k) > 0. \tag{4.6}$$

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

$$\gamma(k) \sim c_\gamma |k|^{-1+2d}, \quad k \to \infty. \tag{4.7}$$

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

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

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

Corollary 4.2 *Suppose the linear process* $\{X_j\}$ *of (2.1) with innovations* $\{\zeta_j\} \sim MD(0, \sigma_\zeta^2)$ *satisfies (4.6) or (4.7), and* z_{nj}*'s are as in (4.5). Then, with* $W_n := \sum_{j=1}^n z_{nj} X_j$,

$$n^{-1/2-d} W_n \to_D \mathcal{N}(0, v_d^2), \quad n^{1/2-d}(\hat\beta - \beta) \to_D \mathcal{N}(0, \tau_d^2), \quad 0 \le d < 1/2. \tag{4.8}$$

Proof. The second claim in (4.8) follows from the first claim and the fact that $\sum_{j=1}^n z_{nj}^2/n \to \int_0^1 g^2(u)du$, which is assured by the continuity of g. To prove the first claim in (4.8), we shall verify condition (i) of Theorem 2.2. Let $\sigma_n^2 := \mathrm{Var}(W_n) = \sum_{j,k=1}^n z_{nj} z_{nk} \gamma(j-k)$. We shall prove that

$$n^{-1-2d} \sigma_n^2 \to v_d^2, \quad \forall\, 0 \le d < 1/2. \tag{4.9}$$

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 $\max_{1\le k\le n} |z_{nk}| \le \sup_{0\le u\le 1} |g(u)| = o(\sigma_n)$, thereby verifying condition (i) for W_n.

Suppose $d = 0$. Then,

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

whereas for any $|i| \le K$,

$$n^{-1} \sum_{k=1}^n z_{n,k+i} z_{nk} \gamma(i) \to \int_0^1 g^2(u)du, \qquad n \to \infty.$$

Whence,

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

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

Next, consider the case $0 < d < 1/2$. Here, (4.5), (4.7), change of variables and the dominated convergence theorem yield

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

This completes the proof of (4.9) and corollary.

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 type process $Y_j := \sum_{s=1}^{j} X_s$, $j = 1, 2, \cdots$. Set $\psi(u) = \int_u^1 g(v)dv$, $0 \le u \le 1$. Also define,

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

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 innovations $\{\zeta_j\} \sim MD(0, \sigma_\zeta^2)$ satisfies (4.6) or (4.7) and z_{nj} are as in (4.5). Then, $n^{-3/2-d} V_n \to_D \mathcal{N}(0, \bar{v}_d^2)$.*

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 μ 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

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

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

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

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 and let

$$z_{nj} := K_{nx}^{-1} K\Big(\frac{x-nj}{nb}\Big). \tag{4.10}$$

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

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

We have

Corollary 4.4 *Suppose the linear process* $\{X_j\}$ *of (2.1) with innovations* $\{\zeta_j\} \sim MD(0,\sigma_\zeta^2)$ *satisfies (4.6) or (4.7). In addition, suppose* $b \to 0$, $nb \to \infty$, *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* $0 \le d < 1/2$,

$$(4.11) \qquad (nb)^{1-2d}\,Var(D_n(x)) \to \tau_{d,K}^2, \qquad n^{1/2-d}D_n(x) \to_D \mathcal{N}(0,\tau_{d,K}^2).$$

Proof. With $\sigma_n^2 := \mathrm{Var}(D_n(x))$ and z_{nj} as in (4.10),

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

A routine argument shows that continuity of K with $\|K\|_2^2 < \infty$ implies $(nb)^{1-2d}\sigma_n^2 \to \tau_{d,K}^2$, for all $0 \le d < 1/2$, and

$$(nb)\sum_{j=1}^{n} z_{nj}^2 \;\sim\; \frac{1}{nb}\sum_{j=1}^{n} K^2\big((x-(j/n))/b\big) \to \|K\|_2^2.$$

These facts together yield $\sigma_n = O((nb)^{-1/2+d})$, $\sigma_n^{-2}\sum_{j=1}^n z_{nj}^2 = O((nb)^{-2d}) = O(1)$ and $\max_{1\le j\le n}\sigma_n^{-1}|z_{nj}| = O((nb)^{-1/2-d}) = o(1)$, thereby verifying condition (i) of Theorem 2.2, and hence the corollary.

ARCH(∞) and stochastic volatility type white noises. A class of martingale differences satisfying Assumption 2.1 includes GARCH(p,q), ARCH(∞) processes and stochastic volatility models. ARCH(∞) process $\{\zeta_j,\ j \in \mathbb{Z}\}$ is defined as a solution to the equations

$$(4.12) \qquad \zeta_j = \sigma_j\varepsilon_j, \quad \sigma_j^2 = b_0 + \sum_{k=1}^{\infty} b_k\zeta_{j-k}^2, \quad j \in \mathbb{Z},$$

where $\{\varepsilon_j\} \sim IID(0,1)$ and $b_j \ge 0$, $j = 0,1,\dots$ are deterministic coefficients. Clearly, $\{\zeta_j\}$ has conditional mean zero and is heteroscedastic with the conditional variance σ_j^2, i.e.

$$E(\zeta_j|\zeta_k, k<j) = 0, \qquad \mathrm{Var}(\zeta_j|\zeta_k, k<j) = \sigma_j^2, \quad j \in \mathbb{Z}.$$

The class of ARCH (∞) processes includes the parametric ARCH and GARCH models of Engle (1982) and Bollerslev (1996). The general framework leading to model (4.12) was introduced by Robinson (1991). See also the review paper by Giraitis, Leipus and Surgailis (2007) for more on these processes. The volatility process σ_j^2 in these models is a stationary process with absolutely summable autocovariance and satisfies Assumption 2.1. This assumption is also satisfied by the linear ARCH model, introduced by Robinson (1991) that allows modelling long memory in $\{\sigma_j^2\}$, as expounded further in Giraitis *et al.* (2007).

By a stochastic volatility model one usually understands a stationary process $\zeta_j, j \in \mathbb{Z}$ of the form

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

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 σ-field generated by past information $\zeta_s, s \leq j-1$. If $\varepsilon_s, s \geq j$ are independent of $\mathcal{F}_{j-1}$, 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: for any $j > k$, $E[\zeta_j\zeta_k] = E\big[\zeta_k E[\zeta_j|\mathcal{F}_{j-1}]\big] = 0$.

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 $\{\eta_j\}$. Robinson (2001) showed that a wide class of these types of models with Gaussian $\{\eta_j\}$ allows for long memory in their volatility process. 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), Surgailis and Viano (2002). As a rule, the volatility process $V_j \equiv \sigma_j^2$ in these models is a stationary process with decaying autocovariance, and thus satisfies Assumption 2.1.

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) simulated 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 3.3 to the case of MD 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 the asymptotic results derived in the paper. We will consider three different experiments, all of them based on a sample size $n = 500$ and $S = 10,000$ replications.

ARFIMA-ARCH. We start from the case where X_j is generated as in (2.1), such that $X_j \sim ARFIMA(1, d, 0)$ and ζ_j are $ARCH(1)$ processes:

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

We set $\alpha_0 = 0.2$, $\alpha_1 = 0.8$, so that the ζ_j have unconditional variance equal to 1. We also set the AR parameter $r = 0.8$ and simulated the cases $d = 0.1, 0.2, 0.3, 0.4$. Given that the simulation results were qualitatively similar for all these choices of d, for the sake of brevity, we report the results for $d = 0.3$ only. We analyze the behavior 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 $Y_j = \sum_{s=1}^{j} X_s$, and the weights $z_{nj} = (j/n)^2 + \cos(j/n)$, $j = 1, 2, \cdots, n$. $\mathrm{Var}(W_n)$ and $\mathrm{Var}(V_n)$ are estimated from the vector of simulated sums. We then plot a kernel estimate of the densities of $\widehat{\mathrm{Var}}(W_n)^{-1/2} W_n$ and $\widehat{\mathrm{Var}}(V_n)^{-1/2} V_n$ and superimpose the standard normal density. The kernel estimate is obtained using a Gaussian kernel with bandwidth b chosen according the Silverman's (1986) rule, $b_n = (4/3)^{1/5} n^{-1/5}$. Results for $\{X_j\} \sim ARFIMA(1, 0.3, 0)$, $r = 0.8$ with $ARCH(1)$ innovations are given in Figure 1 and show a very close resemblance between the two curves. In Figures 1 and 2 below solid line graphs are kernel density estimates while dashed line graphs represent $\mathcal{N}(0,1)$ density.

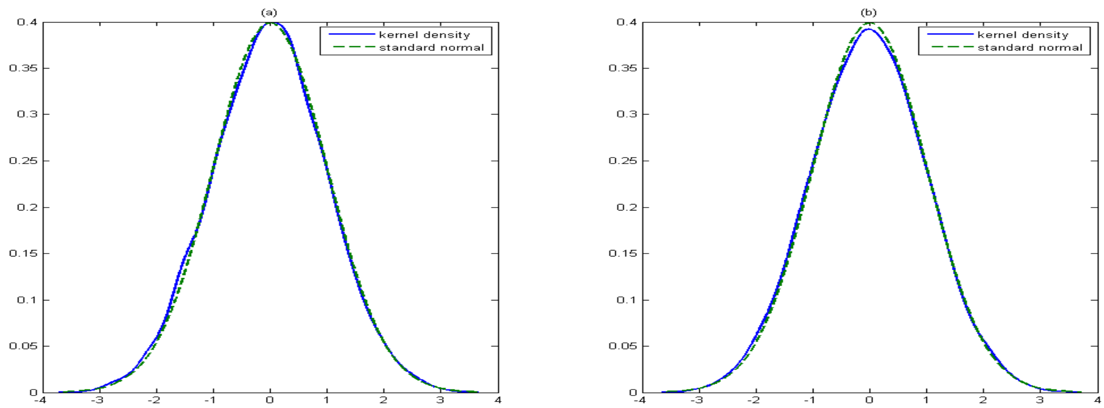

Figure 1: Panel (a): Kernel density of $\widehat{\mathrm{Var}}(W_n)^{-1/2} W_n$. Panel (b): Kernel density of $\widehat{\mathrm{Var}}(V_n)^{-1/2} V_n$.

AR-ARCH. Next we consider the case where X_j follows an AR(1) process

$$X_j = \rho X_{j-1} + \zeta_j, \quad j \in \mathbb{Z},$$

with $\rho = 0.6$ and ζ_j are $ARCH(1)$ errors generated using the same parameters as those in (5.1). Results for $\{X_j\} \sim AR(1)$, $\rho = 0.6$ with $ARCH(1)$ innovations are given in Figure 2. The kernel density of $\widehat{\mathrm{Var}}(W_n)^{-1/2} W_n$ seem to have less probability mass close to the mean than the $\mathcal{N}(0,1)$ density and a bit more on the right part of the distribution (skewness is, in fact, slightly positive).

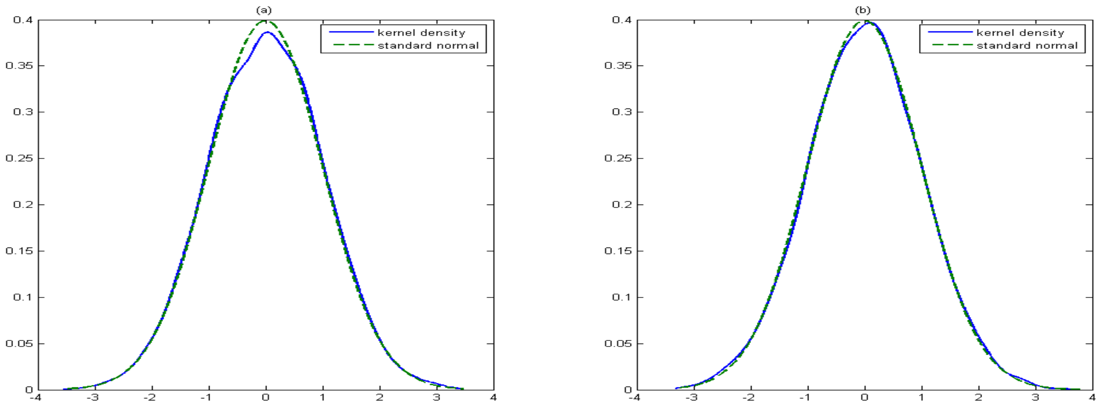

Figure 2: Panel (a): Kernel density of $\widehat{\text{Var}}(W_n)^{-1/2}W_n$. Panel (b): Kernel density of $\widehat{\text{Var}}(V_n)^{-1/2}V_n$.

Nonparametric regression. Finally, we consider the case of nonparametric estimation. We use the following regression model

$$Y_j = \mu(j/n) + X_j, \quad j = 1, \cdots, n,$$

where the errors X_j follow an $ARFIMA(1, d, 0) - ARCH(1)$ process and are generated as in (5.1). Here, $\mu(j/n) = (j/n)^2 + \cos(j/n)$ and $\mu(x)$ is estimated at $x = 1/4$ by a kernel type estimator

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

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 $J = j/n, j = 1, 2, \ldots, n$. In Figure 3, we plot the kernel estimate of the density of $\widehat{\text{Var}}(\hat{\mu}_n(x))^{-1/2}(\hat{\mu}_n(x) - \mu(x))$ for $\{X_j\} \sim ARFIMA(1, 0.3, 0)$, $r = 0.8$ with $ARCH(1)$. $\widehat{\text{Var}}(\hat{\mu}_n(x))$ is estimated from the vector of simulated regression estimators. Even in this case, resemblance of the two curves is quite clear.

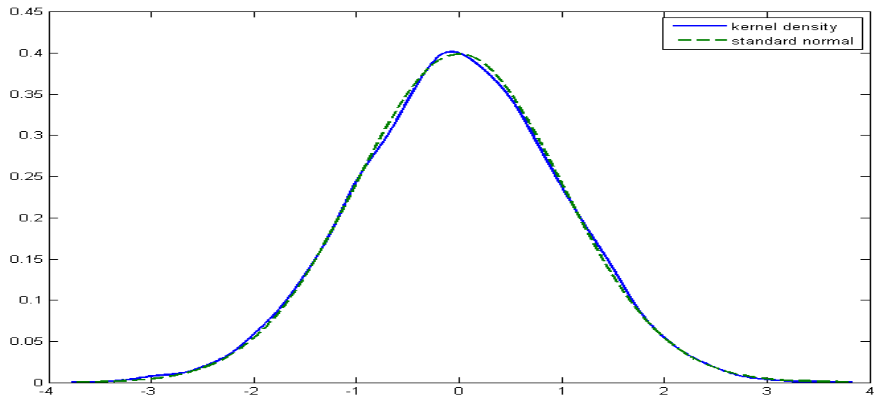

Figure 3: Kernel density of $\widehat{\mathrm{Var}}(\hat{\mu}_n(x))^{-1/2}(\hat{\mu}_n(x) - \mu(x))$ (solid line), $\mathcal{N}(0,1)$ density (dashed line).

Acknowledgements. We thank Donatas Surgalis for providing part (iii) of Theorem 2.2 and other useful comments.

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.

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.

Crato, N. and de Lima, P.J.F. (1994). Long-range dependence in the conditional variance of stock returns. *Economics Letters*, **45**, 281 – 285.

Davydov, Y.A. (1970). The invariance principle for stationary processes. *Theor. Probab. Appl.*, **15**, 487 – 498.

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, 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.

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.

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. *Journal of Statistical 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. (2006). Central limit theorem for stationary linear processes. *Ann. Probab.*, **34**, 1608 - 1622.

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.

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.

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.

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 Woodroofe, M. (2004). Martingale Approximations for Sums of Stationary Processes. *Ann. Probab.*, **32**, 1674 – 1690.