ASYMPTOTIC THEORY FOR LOCAL TIME DENSITY ESTIMATION AND NONPARAMETRIC COINTEGRATING REGRESSION

By

Qiying Wang and Peter C. B. Phillips

December 2006

COWLES FOUNDATION DISCUSSION PAPER NO. 1594

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Asymptotic theory for local time density estimation and nonparametric cointegrating regression *

Qiying Wang
School of Mathematics and Statistics
The University of Sydney

Peter C. B. Phillips
Cowles Foundation, Yale University
University of Auckland & University of York

October 30, 2006

Abstract

We provide a new asymptotic theory for local time density estimation for a general class of functionals of integrated time series. This result provides a convenient basis for developing an asymptotic theory for nonparametric cointegrating regression and autoregression. Our treatment directly involves the density function of the processes under consideration and avoids Fourier integral representations and Markov process theory which have been used in earlier research on this type of problem. The approach provides results of wide applicability to important practical cases and involves rather simple derivations that should make the limit theory more accessible and useable in econometric applications. Our main result is applied to offer an alternative development of the asymptotic theory for non-parametric estimation of a non-linear cointegrating regression involving non-stationary time series. In place of the framework of null recurrent Markov chains as developed in recent work of Karlsen, Myklebust and Tjostheim (2007), the direct local time density argument used here more closely resembles conventional nonparametric arguments, making the conditions simpler and more easily verified.

Key words and phrases: Brownian Local time, Cointegration, Integrated process, Local time density estimation, Nonlinear functionals, Nonparametric regression, Unit root.
JEL Classification: C14, C22.

1 Introduction

Since the introduction of unit root and cointegration analysis in time series econometrics, linear models have dominated empirical work in the application of these methods. This

*Wang acknowledges partial research support from Australian Research Council. Phillips acknowledges partial research support from a Kelly Fellowship and the NSF under Grant No. SES 04-142254.

emphasis on linearity is convenient for practical implementation and accords well with the linear framework of partial summation in which the integrated process and cointegration concepts have been developed. Nonetheless, it is restrictive, especially in view of the attention given elsewhere in modern econometric methodology to nonlinear and nonparametric estimation, and the fact that economic theory often suggests nonlinear responses without being specific regarding functional form. In such situations, nonparametric function estimation offers an alternative that is appealing in applied work.

For stationary time series data, the theory of nonparametric function estimation and inference is well developed and the methods are widely used in practice. By contrast, density function estimation and nonparametric regression involving stochastically nonstationary time series are presently rather undeveloped. An early contribution to the study of nonparametric autoregression in the context of a random walk was undertaken in Phillips and Park (1998). Their results showed that, in contrast to parametric autoregressions, nonstationarity slows down the rate of convergence in nonparametric estimation because of the signal reduction (in the local behavior) that results from the random wandering characteristic of processes such as a random walk.

Some related analytic tools on the local time density and hazard functions of the limiting Brownian motion of a standardized integrated process were developed and applied in Phillips (1998/2005, 2001) and have recently been used in Park (2006) to study stochastic dominance relations for nonstationary time series. Nonlinear transformations of integrated time series and an asymptotic theory of inference for nonlinear regression were developed in Park and Phillips (1999, 2001). de Jong (2002), Pötscher (2004), and Berkes and Horváth (2006) extended this limit theory for nonlinear transformations to cover a wider class of functionals. Bandi and Phillips (2003) developed an asymptotic theory of function estimation and inference in possibly nonstationary diffusions. Tests for nonlinearity in cointegrating relations have been developed by Hong and Phillips (2005) and Kasparis (2005). Karlsen and Thostheim (2001) and Guerre (2004) studied nonparametric estimation for certain nonstationary processes in the framework of recurrent Markov chains. This work has been overviewed in relation to the approach of Phillips and Park (1998) by Bandi (2004). Most recently, Karlsen, Myklebust and Tjostheim (2007, hereafter KMT) developed an asymptotic theory for nonparametric estimation of a time series regression equation involving stochastically nonstationary time series. KMT specifically address the function estimation problem for a possibly nonlinear cointegrating relation, providing an asymptotic theory of estimation and inference for nonparametric forms of cointegration.

The present paper has a similar goal to KMT but offers an alternative approach to the asymptotic theory that we hope is simpler and more accessible. While KMT use the framework of null recurrent Markov chains, we use a direct local time density argument that makes the approach more closely related to conventional nonparametric arguments. The starting point in our development is to show the weak convergence of a general class of funtionals to the local time density of a certain limiting stochastic process. The functional class is designed to include the type of kernel averages that appear in standard kernel density estimation, thereby making the results applicable to nonparametric density estimation and regression with nonstationary time series.

To begin, consider a triangular array $x_{k,n}, 1 \leq k \leq n, n \geq 1$ constructed from some underlying time series and assume that there is a continuous limiting Gaussian process $G(t), 0 \leq t \leq 1$, such that

$$x_{[nt],n} \Rightarrow G(t), \quad \text{on} \;\; D[0,1],$$

where $[a]$ denotes the integer part of a and $\Rightarrow$ denotes weak convergence. The functional of interest S_n of $x_{k,n}$ is defined by the sample average

$$S_n = \frac{c_n}{n} \sum_{k=1}^{n} g(c_n \, x_{k,n}),$$

where c_n is a certain sequence of positive constants and g is a real function on R. Such functionals commonly arise in non-linear regression with integrated time series [Park and Phillips (1999, 2001)] and non-parametric estimation in relation to nonlinear cointegration models [Phillips and Park (1998), Karlsen and Tjostheim (2001), and KMT]. The limit behavior of S_n in the situation that $c_n \to \infty$ and $n/c_n \to \infty$ is particularly interesting and important for practical applications as it provides a setting that accommodates a sufficiently wide range of bandwidth choices to be relevant for non-parametric kernel estimation, as discussed later.

Accordingly, the present paper derives by direct calculation the limit distribution of S_n when $c_n \to \infty$ and $n/c_n \to \infty$, showing that under very general conditions on the function g and the process $x_{k,n}$

$$S_n \to_D \int_{-\infty}^{\infty} g(x)dx \, L(1,0), \tag{1.1}$$

where $L(t,s)$ is the local time of the process $G(t)$ at the spatial point s. When the function g is a kernel density, the limit (1.1) is simply the local time of G at the origin, and this limit may be recentred at an arbitrary spatial point s. These results relate to those of Jeganathan (2004), who investigated the asymptotic form of similar functionals when $x_{k,n}$ is the partial sum of a linear process. For the particular situation where $c_n x_{k,n}$ is a partial sum of iid random variables, some other related results can be found in the work of Borodin and Ibragimov (1995), Akonom (1993) and Phillips and Park (1998).

As in Jeganathan (2004), the approach in this paper involves approximating the difference

$$\frac{c_n}{n}\sum_{k=1}^{n} g\big(c_n\, x_{k,n}\big) - \frac{c_n}{n}\sum_{k=1}^{n}\int_{-\infty}^{\infty} g\big[c_n\,(x_{k,n}+z\epsilon)\big]\phi(z)dz,$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}}\exp\big\{-\frac{x^2}{2}\big\}$. However, unlike Jeganathan (2004) who used a traditional Fourier transformation like that of Borodin and Ibragimov (1995) for dealing with this kind of problem, our treatment directly involves the density function of $x_{k,n}$. In this respect our work is related to the approach used in Pötscher (2004) and Berkes and Horváth (2006). The application of this idea gives the results wide applicability to important practical cases where $x_{k,n}$ is an integrated time series and the limit process is Gaussian, and it also makes for rather simple and neat derivations.

We mention that the limit distribution of S_n in the situation that $c_n = 1$ is very different from that when $c_n \to \infty$ and $n/c_n \to \infty$. When $c_n = 1$, in a series of papers of increasing generality on the conditions for $x_{k,n}$, $g(x)$ and $G(t)$, Park and Phillips (1999), de Jong (2002), Pötscher (2004), De Jong and Wang (2005), and Berkes and Horváth (2006) proved that

$$\frac{1}{n}\sum_{k=1}^{n} g(x_{k,n}) \to_D \int_0^1 g(G(t))dt. \tag{1.2}$$

The limit distribution of S_n in this case is an integral of $G(t)$ and the result may be interpreted as an application of weak convergence in conjunction with a version of the continuous mapping theorem. When $c_n \to \infty$, not only is the limit result different, but the rate of convergence is affected and the result no longer has a form associated with a continuous map.

Some heuristic arguments help to reveal the nature of these differences. Note first that by virtue of the occupation times formula (see (2.1) below) we may write

$$\int_0^1 g(G(t))dt = \int_{-\infty}^{\infty} g(s)L_G(1,s)ds, \tag{1.3}$$

where $L_G(1,s)$ is the local time at s of the limit process G over the time interval $[0,1]$, as discussed in Section 2 below. Next, rewrite the average S_n so that it is indexed by twin sequences c_m and n defining $S_{m,n} = \frac{c_m}{n} \sum_{k=1}^n g(c_m\, x_{k,n})$ and noting that $S_{m,n} = S_n$ when $m = n$. If we hold c_m fixed as $n \to \infty$, then from (1.2) - (1.3) we have

$$\begin{aligned} S_{m,n} \quad &\to_D \quad c_m \int_0^1 g(c_m G(t))dt = c_m \int_{-\infty}^{\infty} g(c_m s) L_G(1, s) ds \\ &= \quad \int_{-\infty}^{\infty} g(r) L_G(1, \frac{r}{c_m}) dr := S_{m,\infty}. \end{aligned}$$

Clearly, $S_{m,\infty} \to_D \int_{-\infty}^{\infty} g(r) dr L_G(1, 0)$ as $m \to \infty$, so that (1.1) may be regarded as a limiting version of (1.2). The goal is to turn this sequential argument as $n \to \infty$, followed by $m \to \infty$, into a joint limit argument so that c_n may play an active role as a bandwidth parameter in density estimation and kernel regression.

The paper is organized as follows. The next section presents our main results. Theorem 2.1 provides a general framework for the limit theory, and its applications to integrated time series and Gaussian limit processes are given in the following Corollaries. Section 3 further investigates applications of Theorem 2.1, which include nonlinear nonparametric cointegrating regressions and the nonparametric estimation of a unit root autoregression. These applications provide a basis for practical nonparametric work with nonstationary series. Section 4 concludes by discussing these results and some possible extensions. Section 5 gives proofs of the main results and corollaries. Throughout the paper we use conventional notation, so that $\to_D$ stands for the convergence in distribution and $\to_P$ for the convergence in Probability. $A, A_1, ...$ denote constants which may be different at each appearance.

2 Main results

We start by recalling the definition of local time. The process $\{L_\zeta(t, s), t \geq 0, s \in R\}$ is said to be the local time of a measurable process $\{\zeta(t), t \geq 0\}$ if, for any locally integrable function $T(x)$,

$$\int_0^t T[\zeta(s)] ds = \int_{-\infty}^{\infty} T(s) L_\zeta(t, s) ds, \quad \text{all } t \in R, \tag{2.1}$$

with probability one. Equation from (2.1) is known as the occupation times formula. Roughly speaking, $L_\zeta(t, s)$ is a spatial density that records the relative sojourn time of the process $\zeta(t)$ at the spatial point s over the time interval $[0, t]$. For further discussion and the properties of local time, we refer to Geman and Horowitz (1980) and Revuz and Yor (1999) and to Phillips (2001) for economic applications. We also define a fractional

Brownian motion with $0 < \beta < 1$ on $D[0,1]$ as follows:

$$W_\beta(t) = \frac{1}{A(\beta)} \int_{-\infty}^0 \Big[(t-s)^{\beta-1/2} - (-s)^{\beta-1/2}\Big] dW(s) + \int_0^t (t-s)^{\beta-1/2} dW(s),$$

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

$$A(\beta) = \Big(\frac{1}{2\beta} + \int_0^\infty \Big[(1+s)^{\beta-1/2} - s^{\beta-1/2}\Big]^2 ds\Big)^{1/2}.$$

Note that $W_{1/2}(t)$ is a standard Brownian motion and $W_\beta(t)$ has a continuous local time $L_{W_\beta}(t,s)$ with regard to (t,s) in $[0,\infty) \times R$. See German and Horowitz (1980), Theorem 22.1, for example.

As in Section 1, let $x_{k,n}, 0 \le k \le n, n \ge 1$ (define $x_{0,n} \equiv 0$) be a random triangular array and $g(x)$ be a real measurable function on R. We make the following assumptions.

Assumption 2.1. *$|g(x)|$ and $g^2(x)$ are Lebesque integrable functions on R with $\tau \equiv \int g(x)dx \ne 0$.*

Assumption 2.2. *There exists a stochastic process $G(t)$ having a continuous local time $L_G(t,s)$ such that $x_{[nt],n} \Rightarrow G(t)$, on $D[0,1]$, where weak convergence is understood w.r.t the Skorohod topology on the space $D[0,1]$.*

Assumption 2.2*. *On a suitable probability space, there exists a stochastic process $G(t)$ having a continuous local time $L_G(t,s)$ such that $\sup_{0\le t\le 1} \big|x_{[nt],n} - G(t)\big| = o_P(1)$.*

In the Assumption 2.3 below we shall make use of the notation: $\Omega_n(\eta) = \big\{(l,k) : \eta\, n \le k \le (1-\eta)\, n,\ k + \eta\, n \le l \le n\big\}$, where $0 < \eta < 1$.

Assumption 2.3. *For all $0 \le k < l \le n, n \ge 1$, there exist a sequence of constants $d_{l,k,n}$ and a sequence of σ-fields $\mathcal{F}_{k,n}$ (define $\mathcal{F}_{0,n} = \sigma\{\phi, \Omega\}$, the trivial σ-field) such that,*

(i) for some $m_0 > 0$ and $C > 0$, $\inf_{(l,k)\in\Omega_n(\eta)} d_{l,k,n} \ge \eta^{m_0}/C$ as $n \to \infty$,

$$\lim_{\eta\to 0} \lim_{n\to\infty} \frac{1}{n} \sum_{k=(1-\eta)\, n}^{n} (d_{k,0,n})^{-1} \to 0, \tag{2.2}$$

$$\lim_{\eta\to 0} \lim_{n\to\infty} \frac{1}{n} \max_{0\le k\le (1-\eta)\, n} \sum_{l=k+1}^{k+\eta\, n} (d_{l,k,n})^{-1} \to 0, \tag{2.3}$$

$$\limsup_{n\to\infty} \frac{1}{n} \max_{0\le k\le n-1} \sum_{l=k+1}^{n} \big(d_{l,k,n}\big)^{-1} < \infty; \tag{2.4}$$

(ii) *$x_{k,n}$ are adapted to $\mathcal{F}_{k,n}$ and, conditional on $\mathcal{F}_{k,n}$, $(x_{l,n} - x_{k,n})/d_{l,k,n}$ has a density $h_{l,k,n}(x)$ satisfying that $h_{l,k,n}(x)$ is uniformly bounded by a constant K and*

$$\sup_{(l,k)\in\Omega_n[\delta^{1/(2m_0)}]} \sup_{|u|\le\delta} \big|h_{l,k,n}(u) - h_{l,k,n}(0)\big| = o_P(1), \tag{2.5}$$

when $n \to \infty$ first and then $\delta \to 0$.

We remark that Assumptions 2.1 and 2.2 are quite weak and likely very close to necessary conditions for this kind of problem. Assumptions 2.1 excludes the zero energy case $\int g(x)dx = 0$, where the limit theory is different and a different convergence rate applies. Assumption 2.2* is a stronger version of Assumption 2.2. In certain situations Assumptions 2.2 and 2.2* are equivalent (for example, in the situation that $x_{i,n} = \sum_{j=1}^{i} \epsilon_j/\sqrt{n}$, where ϵ_j are iid random variable with $E\epsilon_1 = 0$ and $E\epsilon_1^2 = 1$). If Assumption 2.2 holds and $G(t)$ is a continuous Gaussian process, it follows from the so-called Skorohod-Dudley-Wichura representation theorem (e.g., Shorack and Wellner, 1986, p. 49, Remark 2) that $x_{i,n}$ may be replaced by an equivalent process $x_{i,n}^*$ (i.e., $x_{i,n}^* =_d x_{i,n}, 1 \le i \le n, n \ge 1$, where $=_d$ denotes equivalence in distribution) for which $x_{i,n}^*$ satisfies Assumption 2.2*. This is sufficient for many applications if we are only interested in weak convergence. As for Assumption 2.3, we may choose $\mathcal{F}_{k,n} = \sigma(x_{1,n}, ..., x_{k,n})$, the natural σ-fields, and the $d_{l,k,n}$ being a numerical sequence such that, conditional on $\mathcal{F}_{k,n}$, $(x_{l,n} - x_{k,n})/d_{l,k,n}$ has a limit distribution as $l-k \to \infty$ in many applications. For instance, if $x_{i,n} = \sum_{j=1}^{i} \epsilon_j/\sqrt{n}$, where ϵ_j are iid random variable with $E\epsilon_1 = 0$ and $E\epsilon_1^2 = 1$, we may choose $\mathcal{F}_{k,n} = \sigma(\epsilon_1, ..., \epsilon_k)$ and $d_{l,k,n} = \sqrt{n}/\sqrt{l-k}$. More examples are given in Corollaries 2.1 and 2.2 below.

We now state our main result.

THEOREM 2.1. *Suppose Assumptions 2.1-2.3 hold . Then, for any $c_n \to \infty$, $c_n/n \to 0$ and $r \in [0, 1]$,*

$$\frac{c_n}{n} \sum_{k=1}^{[nr]} g\big(c_n\, x_{k,n}\big) \to_D \tau\, L_G(r, 0). \tag{2.6}$$

If Assumption 2.2 is replaced by Assumption 2.2, then, for any $c_n \to \infty$ and $c_n/n \to 0$,*

$$\sup_{0\le r\le 1} \Big| \frac{c_n}{n} \sum_{k=1}^{[nr]} g\big(c_n\, x_{k,n}\big) - \tau\, L_G(r, 0) \Big| \to_P 0, \tag{2.7}$$

under the same probability space defined as in Assumption 2.2.*

REMARK 2.1. Many examples occur in applications where limit results at spatial points other than the origin are relevant. Phillips (2001) gave examples of hazard rate analyses for inflation series and Hu and Phillips (2004) analyzed Federal funds rate market intervention policy on interest rates. To suit such applications, versions of results (2.6) and

(2.7) still hold if the $x_{i,n}$ is replaced by $y_{i,n} = x_{i,n} + x\, c_n'$ where $c_n' \to 0$ or $c_n' = 1$, and respectively $L_G(r, 0)$ is replaced by

$$L_G^*(r) = \begin{cases} L_G(r, 0), & \text{if } c_n' \to 0, \\ L_G(r, -x), & \text{if } c_n' = 1. \end{cases}$$

Indeed, if $x_{i,n}$ satisfies Assumptions 2.2 (similarly for Assumption 2.2*), then for any given $x \in R$

$$y_{[nt],n} \Rightarrow \begin{cases} G(t), & \text{if } c_n' \to 0, \\ G(t) + x, & \text{if } c_n' = 1; \end{cases}$$

If $x_{i,n}$ satisfies Assumptions 2.3 then $y_{i,n}$ also satisfies Assumption 2.3. The claim follows directly from Theorem 2.1 and the fact that $G(t) + x$ has local time $L_G(t, s - x)$.

In the following we consider the applications of Theorem 2.1 to Gaussian processes and general linear processes. Further applications will be investigated in Section 3 where we consider the non-parametric estimate in a non-linear cointegration regression model.

COROLLARY 2.1. *Suppose Assumption 2.1 holds. Let $\{\xi_j, j \geq 1\}$ be a stationary sequence of Gaussian random variables with $E\xi_1 = 0$ and the co-variance $\gamma(j-i) = E\xi_i\xi_j$ satisfying the following condition, for some $0 < \alpha < 2$ and $\lambda < 1$,*

$$d_n^2 \equiv \sum_{1 \leq i,j \leq n} \gamma(j-i) \;\sim\; n^\alpha\, h(n) \quad \textit{and} \quad |\tilde{\gamma}_{l,k}| \leq \lambda\, d_k\, d_{l-k}, \tag{2.8}$$

as $\min\{k, l-k\} \to \infty$, where $h(n)$ is a slowly varying function at ∞ and

$$\tilde{\gamma}_{l,k} = \sum_{i=1}^{k} \sum_{j=k+1}^{l} \gamma(j-i).$$

Let $S_i = \sum_{j=1}^{i} \xi_j$ and $x_{i,n} = S_i/d_n$, for $1 \leq i \leq n, n \geq 1$. Then, for any $c_n \to \infty$, $c_n/n \to 0$ and $r \in [0, 1]$,

$$\frac{c_n}{n} \sum_{k=1}^{[nr]} g\big(c_n\, x_{k,n}\big) \to_D \tau\, L_{W_{\alpha/2}}(r, 0). \tag{2.9}$$

REMARK 2.2. Note that $d_n^2 = ES_n^2$ and $\tilde{\gamma}_{l,k} = cov(S_k, S_l - S_k)$. Condition (2.8) is quite weak. For instance, if one of the following conditions is satisfied, then (2.8) holds:

(i) $\gamma(j) = E(\xi_1\xi_{1+j}) \geq 0$ for all $j \geq 0$ and $\sum_{j=0}^{\infty} \gamma(j) < \infty$;

(ii) $\gamma(k) = E(\xi_1\xi_{1+k}) \sim C\, k^{-\mu}$ with some $0 < \mu < 1$ and $C > 0$;

(iii) $\gamma(k) = E(\xi_1\xi_{1+k}) \sim -C\, k^{-\mu}$ with some $1 < \mu < 2$, $C > 0$ and $\gamma(0) + 2\sum_{k=1}^{\infty} \gamma(k) = 0$.

Indeed, in situation (i), it is readily seen that $d_n^2 \sim C\, n$ with some constant $C > 0$ and as $\min\{k, l-k\} \to \infty$,

$$\tilde{\gamma}_{l,k} = \sum_{i=0}^{k-1} \sum_{j=1}^{l-k} \gamma(j+i) = o(1)\, \min\{k, l-k\} \le \frac{1}{2}\, d_k^{1/2}\, d_{l-k}^{1/2}.$$

In both situations (ii) and (iii), it follows from Taqqu (1975, Lemma 5.1) [also see Example 2.3 of Berkes and Horváth (2006)] that $d_n^2 = ES_n^2 \sim K\, n^{2-\mu}$, where K is constant depending only on μ and C. This yields the first part of (2.8). On the other hand, it can be easily seen that, as $\min\{k, l-k\} \to \infty$,

$$\begin{aligned} |\tilde{\gamma}_{l,k}| &= \frac{1}{2}\, \big| ES_l^2 - E(S_l - S_k)^2 - ES_k^2 \big| \\ &\sim \frac{1}{2}\, K\, \big| l^\alpha - (l-k)^\alpha - k^\alpha \big| \;\le\; \frac{1}{2}(1+\varsigma)\, \max\{1, 2-\mu\}\, d_k\, d_{l-k}, \end{aligned}$$

for arbitrary $\varsigma > 0$, where we have used the fact that

$$|(x+y)^\alpha - x^\alpha - y^\alpha| \le \max\{1, \alpha\} x^{\alpha/2} y^{\alpha/2}, \quad x, y \ge 0,\ 0 < \alpha < 2.$$

Recall $0 < \mu < 2$. By letting $\varsigma = \varsigma_0$ sufficient small, we prove the second part of (2.8) with $\lambda = \frac{1}{2}(1+\varsigma_0)\, \max\{1, 2-\mu\} < 1$.

COROLLARY 2.2. *Let Assumption 2.1 hold. Let $\{\xi_j, j \ge 1\}$ be a sequence of linear processes defined by*

$$\xi_j = \sum_{k=0}^{\infty} \psi_k\, \epsilon_{j-k},$$

where $\{\epsilon_j, -\infty < j < \infty\}$ is a sequence of iid random variables with $E\epsilon_0 = 0$, $E\epsilon_0^2 = 1$ and characteristic function $\varphi(t)$ of ϵ_0 satisfying $\int_{-\infty}^{\infty} |\varphi(t)| dt < \infty$. Let $S_i = \sum_{j=1}^{i} \xi_j$ and $x_{i,n} = S_i/d_n$, for $1 \le i \le n, n \ge 1$, where $d_n^2 = ES_n^2$.

(i) If $\psi_k \sim k^{-\mu}\, h(k)$, where $1/2 < \mu < 1$ and $h(k)$ is a function slowly varying at ∞, then $d_n^2 \sim c_\mu\, n^{3-2\mu}\, h^2(n)$ with $c_\mu = \frac{1}{(1-\mu)(3-2\mu)} \int_0^\infty x^{-\mu}(x+1)^{-\mu} dx$ and, for any $c_n \to \infty$, $c_n/n \to 0$ and $r \in [0,1]$,

$$\frac{c_n}{n} \sum_{k=1}^{[nr]} g\big(c_n\, x_{k,n}\big) \to_D \tau\, L_{W_{3/2-\mu}}(r, 0). \tag{2.10}$$

(ii) If $\sum_{k=0}^{\infty} |\psi_k| < \infty$ and $\psi \equiv \sum_{k=0}^{\infty} \psi_k \ne 0$, then $d_n^2 \sim \psi^2\, n$ and, for any $c_n \to \infty$, $c_n/n \to 0$ and $r \in [0,1]$,

$$\frac{c_n}{n} \sum_{k=1}^{[nr]} g\big(c_n\, x_{k,n}\big) \to_D \tau\, L_W(r, 0). \tag{2.11}$$

REMARK 2.3. Corollary 2.2 (i) provides a result similar to Theorem 3 of Jeganathan (2004) who considered the more general situation where ϵ_0 is in the domain of attraction of the stable law. It is possible to restate our corollary in the same setting. However, this is not essential for our purpose in the present paper and we therefore omit the details. Corollary 2.2 (ii) essentially improves and extends similar results obtained in Akonom (1993), Park and Phillips (1999) and others.

3 Nonparametric cointegrating regression

Consider a non-linear cointegrating regression model:

$$y_t = f(x_t) + u_t, \quad t = 1, 2, ..., n, \tag{3.1}$$

where $x_0 = 0$ and

$$x_t = x_{t-1} + \epsilon_t, \quad t = 1, 2, ..., n.$$

Let $K(x)$ be a non-negative real function and write $K_h(s) = \frac{1}{h}K(s/h)$ where $h \equiv h_n \to 0$. The conventional kernel estimate of $f(x)$ in model (3.1) is given by

$$\hat{f}(x) \;=\; \frac{\sum_{t=1}^n y_t K_h(x_t - x)}{\sum_{t=1}^n K_h(x_t - x)}. \tag{3.2}$$

The limit behavior of $\hat{f}(x)$ has currently been investigated in KMT in the situation where x_t is a recurrent Markov chain. [Also, see Phillips and Park (1998), Karlsen and Thostheim (2001), Guerre (2004) and Bandi (2004) for related work on non-linear, non-stationary autoregressions]. The main theorem in KMT (Theorem 3.1 of KMT) relies heavily on the asymptotic theory developed in Karlsen and Thostheim (2001) involving the conditions on the invariant measure associated with a recurrent Markov chain. These conditions are difficult to check and less accessible.

This section provides a different and simpler approach to nonparametric cointegration. In particular, we reconsider the limit behavior of $\hat{f}(x)$ by making direct use of Theorem 2.1 in developing the asymptotics. This approach gives an alternative route to the asymptotic theory that is more closely associated with traditional nonparametric asymptotics, and the conditions required for this development are simpler and more accessible.

Our first theorem assumes that the ϵ_t are independent of u_t. We relax this independence condition in the second theorem. Throughout the section we make use of the following assumptions.

Assumption 3.1. *The kernel K satisfies that $\int_{-\infty}^{\infty} K(s)ds = 1$ and $\sup_s K(s) < \infty$.*

Assumption 3.2. *For given x, there exists a real function $f_1(s,x)$ such that, when h sufficiently small, $|f(hy+x)-f(x)| \le h\, f_1(y,x)$ for all $y \in R$ and $\int_{-\infty}^{\infty} K(s)\, f_1(s,x)ds < \infty$.*

Assumption 3.3. *$(u_t, \mathcal{F}_t, 1 \le t \le n)$ is a martingale difference with $E(u_t^2|F_{t-1}) \to_{a.s.} \sigma^2 > 0$ as $t \to \infty$ and $\sup_{1\le t\le n} E(|u_t|^q|F_{t-1}) < \infty$ a.s. for some $q > 2$.*

Assumption 3.4. *There exists $0 < d_n \to \infty$ and $d_n = o(n)$ such that $x_{i,n} = x_i/d_n, 1 \le i \le n, n \ge 1$, satisfies Assumption 2.3.*

Our first result is as follows.

THEOREM 3.1. *Suppose Assumptions 3.1-3.4 hold. Suppose that $(\epsilon_t)_1^n$ is independent of $(u_t)_1^n$ and, on a suitable probability space, there exists a stochastic process $G(t)$ having a continuous local time $L_G(t,s)$ such that*

$$\sup_{0\le t\le 1} \left| x_{[nt],n} - G(t) \right| = o_P(1), \tag{3.3}$$

where d_n and $x_{i,n} = x_i/d_n$ are defined as in Assumption 3.4. Then, for any h satisfying $nh/d_n \to \infty$ and $nh^3/d_n \to 0$,

$$\Big(h \sum_{t=1}^{n} K_h(x_t - x)\Big)^{1/2} (\hat{f}(x) - f(x)) \;\to_D\; N(0, \sigma_1^2), \tag{3.4}$$

where $\sigma_1^2 = \sigma^2 \int_{-\infty}^{\infty} K^2(s)ds$.

REMARK 3.1. The conditions in Assumptions 3.1 and 3.2 are quite weak and simply verified for various kernels $K(x)$ and regression functions $f(x)$. For instance, if $K(x)$ is a standard normal kernel or has a compact support as in KMT, a wide range of regression functions $f(x)$ are included. Thus, commonly occuring functions like $f(x) = |x|^\alpha$ and $f(x) = 1/(1 + |x|^\alpha)$ for some $\alpha > 0$ satisfy the Assumption 3.2. Assumption 3.3 is a standard condition for the error processes. As in the proofs of Corollaries 2.1 and 2.2, if ϵ_t are iid random variable with $E\epsilon_1 = 0$, $E\epsilon_1^2 = 1$ and characteristic function $\varphi(t)$ satisfying $\int_{-\infty}^{\infty} |\varphi(t)|dt < \infty$, then ϵ_t (respectively x_t) satisfies (3.3) and Assumption 3.4. Furthermore, if the $x_{i,n} = x_i/d_n$ defined in Assumption 3.4 satisfies Assumption 2.2 with the $G(t)$ being a continuous Gaussian process, then ϵ_t (respectively x_t) satisfies (3.3) and Assumption 3.4. This fact follows from the Skorohod-Dudley-Wichura representation theorem, as observed earlier. Since fractional Brownian motion $W_\beta(t)$ is a continuous Gaussian process, the result (3.4) holds true for the ϵ_t being equal to the process ξ_t defined in Corollaries 2.1 and 2.2.

REMARK 3.2. It is interesting to notice that the bandwidth h needs to satisfy certain rate conditions to ensure the stated asymptotic normality applies. For instance, in the most common situation where $d_n = \sqrt{n}$ (e.g., when the ϵ_t are iid random variables), we require $nh^2 \to \infty$ and $nh^6 \to 0$. This can be explained as follows. In stationary non-parametric models the convergence rate of a kernel regression estimate is $\sqrt{nh}$ requiring that $nh \to \infty$. Undersmoothing in such regressions to avoid bias typically requires that $h = o(n^{-1/5})$. In the nonstationary case, the amount of time spent by the process around any particular spatial point is of order $\sqrt{n}$ rather than n, so that the corresponding rate in such regressions is now $\sqrt{\sqrt{n}h}$, which requires that $nh^2 \to \infty$. Undersmoothing to remove asymptotic bias in this situation typically requires a rate smaller than that in the stationary case. Here we find that the rate $h = o(n^{-1/6})$ is sufficient for undersmoothing.

It is possible to improve the range for the bandwidth h by adding a bias term in (3.4). Since it is not essential for the purpose of this paper and since applications will typically involve some undersmoothing for bias removal, we leave developments in this direction for later work. Also, it is clear from the proof of Theorem 3.1 that $\hat{f}(x) \to_P f(x)$ for any h satisfying $h \to 0$ and $d_n/(nh) \to 0$.

Our next theorem considers the effect of some relaxation of the restriction on the independence between ϵ_t and u_t. To do so, denote the stochastic processes U_n and V_n on $D[0,1]$ by

$$U_n(r) = x_{[nr],n} \qquad \text{and} \qquad V_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} u_t,$$

where d_n and $x_{i,n} = x_i/d_n$ are defined as in Assumption 3.4.

THEOREM 3.2. *Suppose Assumptions 3.1-3.4 hold. Suppose that, for each $n \geq 2$, $x_{i,n}$ is adapted to $\mathcal{F}_{i-1}$, $2 \leq i \leq n$, and $(U_n, V_n) \Rightarrow_D (U, V)$ on $D[0,1]^2$ as $n \to \infty$, where (U, V) is a standard vector Brownian motion. Then (3.4) still holds true for any h satisfying $nh/d_n \to \infty$ and $nh^3/d_n \to 0$.*

REMARK 3.3. Theorem 3.2 can be used to construct a non-parametric kernel estimate of $m(x)$ in the unit-root autoregressive model

$$y_t = m(y_{t-1}) + u_t, \quad m(y_{t-1}) = \alpha\, y_{t-1}, \ a.s..$$

with $\alpha = 1$ and $y_0 = 0$. To illustrate, let u_t be a sequence of iid random variables with $Eu_0 = 1$, $Eu_0^2 = 1$, $E|u_0|^q < \infty$ for some $q > 2$ and the characteristic function $\varphi(t)$ of

u_0 satisfying $\int_{-\infty}^{\infty} |\varphi(t)|dt < \infty$. As in (3.2), the conventional kernel estimate of $m(x)$ is given as follows:

$$\hat{m}(x) \quad = \quad \frac{\sum_{t=1}^{n} y_t K_h(y_{t-1} - x)}{\sum_{t=1}^{n} K_h(y_{t-1} - x)}.$$

In this case, $x_{i,n} = y_{i-1} = \sum_{t=1}^{i-1} u_t$ and the stochastic processes U_n and V_n on $D[0, 1]$ are defined by

$$U_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]-1} u_t \qquad \text{and} \qquad V_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} u_t,$$

By letting $\mathcal{F}_i = \sigma\{u_1, u_2, ..., u_i\}$, it is easy to check that $x_{i,n}$ are $\mathcal{F}_{i-1}$ measurable and $(U_n, V_n) \Rightarrow_D (W, W)$ on $D[0, 1]^2$ since $V_n(r) \Rightarrow_D W(r)$ on $D[0, 1]$ and $\sup_r |U_n(r) - V_n(r)| \leq \sup_r |u_{[nr]}|/\sqrt{n} \rightarrow_P 0$. It therefore follows from Theorem 3.2 that

$$\Big(h \sum_{t=1}^{n} K_h(y_{t-1} - x)\Big)^{1/2} (\hat{m}(x) - m(x)) \ \rightarrow_D \ N(0, \sigma_1^2), \tag{3.5}$$

where $\sigma_1^2 = \int_{-\infty}^{\infty} K^2(s)dt$. Result (3.5) provides a simple demonstration that kernel autogression in the case of a unit root is asymptotically normal upon standardization in the usual way. However, the implied convergence rate is slower than that in stationary nonparametric autoregession and much slower than parametric rate in the unit root case, as found in Phillips and Park (1998) and Guerre (2004).

4 Conclusion

The main advantage of the approach adopted here is its simplicity. Just as sample averages of a kernel function of a strictly stationary time series inform us about the probability density of the time series at some locality, the same sample averages of an integrated process provide local spatial density information about the trajectories of the process. The fact that the rates of convergence differ between the two cases simply reflects the fact that integrated time series wander over the entire sample space and spend only $O(\sqrt{n})$ of the sample time in the vicinity of particular points like the origin. The proofs of the results given here on local time density estimation and nonparametric cointegrating regression take advantage of these characteristics and, in other respects, more closely relate to conventional nonparametric arguments.

The nonparametric formulation of cointegrating relations seems important in many different empirical applications, especially in view of the fact that economic variables are frequently considered to be driven by fundamentals which have random wandering characteristics. Nonparametric treatment of such relations is appealing because the nature of the functional dependence on fundamentals is seldom specified. The limit distribution theory of KMT and the present paper on the kernel estimation of such relations provides a foundation for empirical work in this context. Further work seems desirable on many different econometric aspects of this central problem, such as dealing with endogeneous regressor issues.

5 Proof of Theorems

This section provides proofs of the main results. The proof of Theorem 2.1 is simple and uses conventional arguments in the main.

Proof of Theorem 2.1. Write

$$L_n^{(r)} = \frac{c_n}{n}\sum_{k=1}^{[nr]} g\big(c_n\, x_{k,n}\big), \quad L_{n,\epsilon}^{(r)} = \frac{c_n}{n}\sum_{k=1}^{[nr]}\int_{-\infty}^{\infty} g\big[c_n\,(x_{k,n}+z\epsilon)\big]\phi(z)dz,$$

where $\phi(x)=\phi_1(x)$ with $\phi_\epsilon(x)=\frac{1}{\epsilon\sqrt{2\pi}}\exp\big\{-\frac{x^2}{2\epsilon^2}\big\}$. By a similar argument to the proof of Lemma 7 of Jeganathan (2004), we have that, for any $\epsilon>0$,

$$L_{n,\epsilon}^{(r)} - \frac{\tau}{n}\sum_{k=1}^{[nr]}\phi_\epsilon(x_{k,n}) = o_P(1), \tag{5.1}$$

uniformly in $r\in[0,1]$. Now Theorem 2.1 will follow if we prove that, uniformly in $r\in[0,1]$,

$$\lim_{\epsilon\to 0}\lim_{n\to\infty} E|L_n^{(r)} - L_{n,\epsilon}^{(r)}| = 0. \tag{5.2}$$

Indeed it follows from the continuous mapping theorem that, for $\forall\epsilon>0$,

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{[nr]}\phi_\epsilon(x_{k,n}) \quad &= \quad \int_0^r \phi_\epsilon(x_{[nt],n})dt - \frac{1}{n}\phi_\epsilon(0) + \frac{1}{n}\phi_\epsilon(x_{n,[nr]}) \\ &\to_D \quad \int_0^r \phi_\epsilon(G(t))dt. \end{aligned}\tag{5.3}$$

Furthermore, by recalling $L(t,s)$ is a continuous local time process satisfying (2.1),

$$\int_0^r \phi_\epsilon(G(t))dt = \int_{-\infty}^{\infty}\phi(x)L(r,\epsilon\, x)dx = L(r,0)+o_{a.s.}(1), \tag{5.4}$$

as $\epsilon \to 0$. By (5.1)-(5.4), we obtain (2.6). The proof of (2.7) is the same except that we replace (5.3) by

$$\begin{aligned}\Big|\frac{1}{n}\sum_{k=1}^{[nr]}\phi_\epsilon(x_{k,n}) - \int_0^r \phi_\epsilon(G(t))dt\Big| &\le \int_0^r \big|\phi_\epsilon(x_{[nt],n}) - \phi_\epsilon(G(t))\big|dt + \frac{2}{n} \\ &\le A(\epsilon)\sup_{0\le t\le 1}|x_{[nt],n} - G(t)| + 2/n \ \to_P \ 0,\end{aligned}$$

as $n \to \infty$.

We next prove (5.2). Write $Y_{k,n}(z) = g[c_n x_{k,n}] - g[c_n(x_{k,n}+z\epsilon)]$. Since $\int_{-\infty}^{\infty}\phi(x)dx = 1$, it is readily seen that

$$E|L_n^{(r)} - L_{n,\epsilon}^{(r)}| \ \le \ \int_{-\infty}^{\infty}\frac{c_n}{n}E\Big|\sum_{k=1}^{[nr]}Y_{k,n}(z)\Big|\phi(z)dz. \tag{5.5}$$

Recall that $x_{k,n}/d_{k,0,n}$ has a density $h_{k,0,n}(x)$ which is bounded by a constant K for all x, $1 \le k \le n$ and $n \ge 1$. For all $z \in R$ and $1 \le k \le n$, we have

$$\begin{aligned}c_n E\big|Y_{k,n}(z)\big| &= c_n\int_{-\infty}^{\infty}\Big|g\big[c_n\,(d_{k,0,n}\,x + z\epsilon)\big] - g\big(c_n\,d_{k,0,n}\,x\big)\Big|h_{k,0,n}(x)dx \\ &\le \frac{A}{d_{k,0,n}}\int_{-\infty}^{\infty}\big|g(x+c_n z\epsilon) - g(x)\big|dx \le 2A\int_{-\infty}^{\infty}\big|g(x)\big|dx/d_{k,0,n}. \quad (5.6)\end{aligned}$$

Hence, for each $z \in R$, $\frac{c_n}{n}E\Big|\sum_{k=1}^{[nr]}Y_{k,n}(z)\Big| \le A_1\frac{1}{n}\sum_{k=1}^{n}1/d_{k,0,n} < \infty$, by (2.4). This, together with (5.5) and the dominated convergence theorem, implies that, to prove (5.2), it suffices to show that, for each fixed z, uniformly in $r \in [0,1]$,

$$\Lambda_n(\epsilon) \ \equiv \ \frac{c_n^2}{n^2}E\Big[\sum_{k=1}^{[nr]}Y_{k,n}(z)\Big]^2 \to 0, \tag{5.7}$$

when $n \to \infty$ first and then $\epsilon \to 0$. We may rewrite Λ_n as

$$\Lambda_n(\epsilon) = \frac{c_n^2}{n^2}\sum_{k=1}^{[nr]}E\,Y_{k,n}^2(z) + \frac{2\,c_n^2}{n^2}\sum_{k=1}^{[nr]}\sum_{l=k+1}^{[nr]}E\,Y_{k,n}(z)\,Y_{l,n}(z) = \Lambda_{1n}(\epsilon) + \Lambda_{2n}(\epsilon), \quad \text{say.}$$

Since $g^2(x)$ is integrable, by a similar argument as in the proof of (5.6), we have

$$\Lambda_{1n}(\epsilon) \ \le \ \frac{Ac_n^2}{n^2}\sum 1/d_{k,0,n} \le A_1 c_n/n \to 0.$$

We next prove $\lim_{\epsilon\to 0}\lim_{n\to\infty}\Lambda_{2n}(\epsilon) \to 0$, and then (5.7) follows accordingly. Write $\Omega_n = \Omega_n(\epsilon^{1/(2m_0)})$. Recall that $x_{k,n}$ are adapted to $\mathcal{F}_{k,n}$ and conditional on $\mathcal{F}_{k,n}$, $(x_{l,n} -$

$x_{k,n})/d_{l,k,n}$ has a density $h_{l,k,n}(x)$ which is bounded by a constant K. We obtain that

$$\begin{aligned}
c_n d_{l,k,n} \left|E(Y_{l,n} \mid \mathcal{F}_{k,n})\right| &= c_n d_{l,k,n} \Big| \int_{-\infty}^{\infty} \Big(g\big[c_n x_{k,n} + c_n d_{l,k,n}\, y\big] \\
&\qquad - g\big[c_n(x_{k,n} + z\epsilon) + c_n d_{l,k,n}\, y\big]\Big) h_{l,k,n}(y)dy \Big| \\
&\leq \int_{-\infty}^{\infty} |g(y)| \big|V(y, c_n\, x_{k,n})\big| dy \\
&\leq \begin{cases} A, & \text{if } (l,k) \notin \Omega_n \\ A \int_{|y| \geq \sqrt{c_n}} |g(y)|dy + \int_{|y|\leq\sqrt{c_n}} |g(y)|\, |V(y, c_n\, x_{k,n})|dy, & \text{if } (l,k) \in \Omega_n, \end{cases}
\end{aligned}$$

where $V(y,t) = h_{l,k,n}\Big(\frac{y-t}{c_n d_{l,k,n}}\Big) - h_{l,k,n}\Big(\frac{y-t-c_n\, z\, \epsilon}{c_n d_{l,k,n}}\Big)$. Furthermore, as in the proof of (5.6), whenever $|y| \leq \sqrt{c_n}$, n large enough and $(l,k) \in \Omega_n$,

$$\begin{aligned}
& E\Big[|Y_{k,n}(z)|\, |V(y, c_n\, x_{k,n})|\Big] \\
&= \int_{-\infty}^{\infty} \Big| g\big[c_n\, (d_{k,0,n}\, x + z\epsilon)\big] - g\big(c_n\, d_{k,0,n}\, x\big) \Big| |V(y, c_n\, d_{k,0,n} x)| h_{k,0,n}(x) dx \\
&\leq \frac{A}{c_n\, d_{k,0,n}} \int_{-\infty}^{\infty} \big|g(x + c_n z\epsilon) - g(x)\,\big|\, |V(y,x)|\, dx \\
&\leq \frac{A}{c_n\, d_{k,0,n}} \int_{-\infty}^{\infty} \big|g(x)\,\big|\, \big[|V(y,x)| + |V(y, x - c_n z\epsilon)|\big]\, dx \\
&\leq \frac{A}{c_n\, d_{k,0,n}} \Big[\int_{|x|\geq\sqrt{c_n}} |g(x)|dx + \sup_{|u|\leq Cz\epsilon^{1/2}} |h_{l,k,n}(u) - h_{l,k,n}(0)|\Big],
\end{aligned}$$

where we have used the facts that $\inf_{(l,k)\in\Omega_n} d_{l,k,n} \geq \epsilon^{1/2}/C$, $c_n \to \infty$ and $V(y,t)$ is bounded. In view of these facts, together with (5.6), we obtain that, if $(l,k) \notin \Omega_n$,

$$\begin{aligned}
\Big|E\Big[Y_{k,n}(z)\, Y_{l,n}(z)\Big]\Big| &= \Big|E\Big[Y_{k,n}(z)\, E\big(Y_{l,n}(z) \mid \mathcal{F}_{k,n}\big)\Big]\Big| \\
&\leq A\, (c_n d_{l,k,n})^{-1}\, E\, |Y_{k,n}(z)| \leq A_1\, (c_n^2\, d_{l,k,n}\, d_{k,0,n})^{-1}, \qquad (5.8)
\end{aligned}$$

and if $(l,k) \in \Omega_n$,

$$\begin{aligned}
& \Big|E\Big[Y_{k,n}(z)\, Y_{l,n}(z)\Big]\Big| \\
&\leq A\, (c_n\, d_{l,k,n})^{-1}\, E\, |Y_{k,n}(z)| \int_{|y|\geq\sqrt{c_n}} |g(y)|dy \\
&\qquad + A\, (c_n\, d_{l,k,n})^{-1} \int_{|y|\leq\sqrt{c_n}} |g(y)|\, E\Big[|Y_{k,n}(z)|\, |V(y, c_n\, x_{k,n})|\Big] dy \\
&\leq A\, (c_n^2\, d_{l,k,n}\, d_{k,0,n})^{-1} \Big(\int_{|y|\geq\sqrt{c_n}} |g(y)|dy + \sup_{|u|\leq Cz\epsilon^{1/2}} |h_{l,k,n}(u) - h_{l,k,n}(0)|\Big). \qquad (5.9)
\end{aligned}$$

It follows from (5.8)-(5.9) and (2.2)-(2.5) that, with $\eta = \epsilon^{1/2}/C$ below,

$$\begin{aligned}
|\Lambda_{2n}(\epsilon)| &\leq \frac{2\,c_n^2}{n^2}\Big(\sum_{l>k,(l,k)\notin\Omega_n} + \sum_{(l,k)\in\Omega_n}\Big)\Big|E\left[Y_{k,n}(z)\,Y_{l,n}(z)\right]\Big| \\
&\leq \frac{A}{n^2}\sum_{k=(1-\eta)n}^{n}(d_{k,0,n})^{-1}\max_{1\leq k\leq n-1}\sum_{l=k+1}^{n}(d_{l,k,n})^{-1} \\
&\quad + \frac{A}{n^2}\max_{0\leq k\leq n-1}\sum_{l=k+1}^{n}(d_{l,k,n})^{-1}\max_{0\leq k\leq (1-\eta)n}\sum_{l=k+1}^{k+\eta n}(d_{l,k,n})^{-1} \\
&\quad + \frac{A}{n^2}\sum_{k=1}^{n}(d_{k,0,n})^{-1}\max_{1\leq k\leq n-1}\sum_{l=k+1}^{n}(d_{l,k,n})^{-1} \\
&\qquad \times\Big[\int_{|y|\geq c_n^{1/2}}|g(y)|dy + \sup_{(l,k)\in\Omega_n}\sup_{|u|\leq Cz\epsilon}|h_{l,k,n}(u)-h_{l,k,n}(0)|\Big] \\
&\to 0,
\end{aligned}$$

as $n\to\infty$ first and then $\epsilon\to 0$, as required. The proof of Theorem 2.1 is now complete.

Proof of Corollary 2.1. Note that $d_n^2 = ES_n^2$. It follows from Lemma 5.1 in Taqqu (1975) that $x_{[nt],n} \Rightarrow W_{\alpha/2}(t), 0\leq t\leq 1$, on $D[0,1]$, where $W_\beta(t)$ is a fractional Brownian motion having a continuous local time $L_{W_\beta}(t,s)$ with regard to both coordinates (t,s) in $[0,\infty)\times R$. Therefore $x_{i,n}$ satisfies Assumption 2.2. We next show that $x_{i,n}$ also satisfies Assumption 2.3 and then (2.9) follows form Theorem 2.1 accordingly.

In order to check Assumption 2.3, let $\mathcal{F}_{t,n} = \sigma\{\xi_1,\xi_2,...,\xi_t\}$ and n_0 be so large that $|\tilde\gamma_{l,k}| \leq \lambda\, d_k\, d_{l-k}$ for all $\min\{k, l-k\}\geq n_0$. The choice of n_0 is possible because of the second part of condition (2.8). For any $0\leq k<l\leq n$, let

$$d_{l,k,n} = \begin{cases} d^*_{l,k}/d_n, & \text{if } \min\{k,l-k\}\geq n_0, \\ d_l/d_n, & \text{otherwise,}\end{cases}$$

where $d^*_{l,k} = \left[d_{l-k}^2 - \tilde\gamma_{l,k}^2/d_k^2\right]^{1/2}$. Recall $d_n^2 \sim n^\alpha h(n)$ and note $d_{l,k,n}^{-1} \leq d_n/d_l + (1-\lambda^2)^{-1/2}\,d_n/d_{l-k}$. It is readily seen that, as $n\to\infty$,

$$\inf_{(l,k)\in\Omega_n(\eta)} d_{l,k,n} \geq (1-\lambda^2)^{1/2}\inf_{(l,k)\in\Omega_n(\eta)} d_{l-k}/d_n \geq C\,(1-\lambda^2)^{1/2}\,\eta^{\alpha/2},$$

and $d_{l,k,n}$ satisfy (2.2)–(2.4). On the other hand, by noting that $(S_k, S_l-S_k)\sim N(0,\sum)$, where $\sum = \begin{pmatrix} d_k^2 & \tilde\gamma_{l,k} \\ \tilde\gamma_{l,k} & d_{l-k}^2\end{pmatrix}$. The conditional distribution of $S_l - S_k$ given S_k is $N(\tilde\gamma_{l,k}S_k/d_k^2,\ d_{l,k}^{*2})$. This implies that, conditional on $\mathcal{F}_{t,n}$,

$$(x_{l,n}-x_{k,n})/d_{l,k,n} = (S_l - S_k)/d^*_{l,k} \ \sim\ N\big(\tilde\gamma_{l,k}\,S_k/(d_k^2\,d^*_{l,k}),\ \ 1\big),$$

for $\min\{k, l-k\} \geq n_0$, and

$$(x_{l,n} - x_{k,n})/d_{l,k,n} = (S_l - S_k)/d_l \;\sim\; N\big(\tilde{\gamma}_{l,k}\, S_k/(d_k^2\, d_l),\;\; d_{l,k}^{*2}/d_l^2\big),$$

in other cases. Therefore $(x_{l,n} - x_{k,n})/d_{l,k,n}$ has a bounded density $h_{j,k,n}(x)$. The $h_{j,k,n}(x)$ satisfy (2.5) since, whenever $\min\{k, l-k\} \geq n_0$,

$$\sup_x \big|h_{j,k,n}(x) - h_{j,k,n}(x+u)\big| \leq \frac{1}{\sqrt{2\pi}} \sup_x |e^{-(u+x)^2/2} - e^{-x^2/2}| \leq A\,|u|.$$

This proves that the Assumption 2.3 holds true for x_{in}, and also completes the proof of Corollary 2.1.

Proof of Corollary 2.2. We first prove (2.10). We need some preliminaries. Write $\tilde{\psi}_i = \sum_{j=0}^{i} \psi_j$, $\tilde{S}_n = \sum_{i=0}^{n} \tilde{\psi}_i \epsilon_i$ and $\Lambda_n^2 = \sum_{i=0}^{n} (\tilde{\psi}_i)^2$. Also let $f_n(t) = Ee^{it\tilde{S}_n/\Lambda_n}$. Recalling the definitions of ψ_j, simple calculations show that $\tilde{\psi}_i \sim \frac{1}{1-\mu} i^{1-\mu} h(i)$ and $\Lambda_n^2 \sim \frac{1}{(1-\mu)^2(3-2\mu)} n^{3-2\mu} h^2(n)$. This, together with the facts that $E\epsilon_0 = 0, E\epsilon_0^2 = 1$ and $E\tilde{S}_n^2 = \Lambda_n^2$, implies that $\tilde{S}_n/\Lambda_n \to_D N(0,1)$. Furthermore we may prove the following:

(a) for each $n \geq 1$, if not all $\tilde{\psi}_i = 0, 0 \leq i \leq n$, then $\tilde{S}_n/\Lambda_n$ has a density $h_n(x)$ which is uniformly bounded by a constant K;

(b) as $n \to \infty$, the density function $h_n(x)$ satisfies that

$$\sup_x |h_n(x) - n(x)| \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} |f_n(t) - e^{-t^2/2}| dt \to 0, \tag{5.10}$$

where $n(x) = e^{-x^2/2}/\sqrt{2\pi}$ is the density of a standard normal.

In fact, it follows from $\int_{-\infty}^{\infty} |Ee^{it\epsilon_0}| dt < \infty$ and the independence of ϵ_i that, whenever $n \geq 1$ and $\tilde{\psi}_{i_0} \neq 0$ for some $i_0 \leq n$,

$$\int_{-\infty}^{\infty} |f_n(t)| dt \leq \int_{-\infty}^{\infty} |Ee^{it\tilde{\psi}_{i_0}\epsilon_{i_0}/\Lambda_n}| dt = \int_{-\infty}^{\infty} |Ee^{it\epsilon_0}| dt < \infty.$$

This yields the result (a) (see, e.g., Lukács, 1970, Theorem 3.2.2).

The left inequality of (5.10) is obvious. In order to prove the convergence in (5.10), we split the integral into three parts as

$$I_{1n} \equiv \int_{|t| \leq A} |f_n(t) - e^{-t^2/2}| dt, \quad I_{2n} \equiv \int_{A < |t| \leq \delta\sqrt{n}} |f_n(t) - e^{-t^2/2}| dt,$$

and $I_{3n} \equiv \int_{|t| > \delta\sqrt{n}} |f_n(t) - e^{-t^2/2}| dt$. It is clear that $I_{1n} \to 0$ for each $A > 0$ since $\tilde{S}_n/\Lambda_n \to_D N(0,1)$. To prove $I_{2n} + I_{3n} \to 0$ for some $A, \delta > 0$, we need the following facts:

(i) for n sufficiently large, there exist $0 < c_1 < c_2 < \infty$ such that $c_1\sqrt{n} < \tilde{\psi}_i/\Lambda_n \leq c_2\sqrt{n}$ for $n/2 \leq i \leq n$;

(ii) for some $\delta_0 > 0$, there exist an $0 < \eta < 1$ such that

$$|\varphi(t)| = |Ee^{it\epsilon_0}| \ \leq \ \begin{cases} e^{-t^2/4}, & \text{for } |t| \leq \delta_0, \\ \eta, & \text{for } |t| \geq \delta_0. \end{cases}$$

Fact (i) follows immediately from the estimates of $\tilde{\psi}_i$ and Λ_n. Recalling $E\epsilon_0 = 0, E\epsilon_0^2 = 1$ and $\int_{-\infty}^{\infty} |\varphi(t)|dt < \infty$, fact (ii) follows from (5.6) and the proof of Theorem 5.2 in Chapter 8 of Feller (1971, page 489). In view of (i) and (ii), for $\forall \epsilon > 0$, by choosing $\delta = \delta_0/c_2$ and A sufficiently large such that $\int_{|t|\geq A} e^{-t^2/8}dt < \epsilon/2$, we have

$$I_{2n} \ \leq \ \int_{A<|t|\leq\delta\sqrt{n}} \Big(| \prod_{j=[n/2]+1}^{n} \Big|Ee^{it\tilde{\psi}_j\epsilon_j/\Lambda_n}\Big| + e^{-t^2/2}\Big)dt \leq 2\int_{|t|\geq A} e^{-t^2/8}dt < \epsilon,$$

for n sufficiently large, and similarly

$$\begin{aligned} I_{3n} \ &\leq \ \int_{|t|\geq\delta\sqrt{n}} \Big(| \prod_{j=[n/2]+1}^{n} \Big|Ee^{it\tilde{\psi}_j\epsilon_j/\Lambda_n}\Big| + e^{-t^2/2}\Big)dt \\ &\leq \ \eta^{n/2-1} \int_{|t|\geq\delta/c_1} |Ee^{it\epsilon_0/\Lambda_n}|dt + \int_{|t|\geq\delta\sqrt{n}} e^{-t^2/2}dt \to 0, \end{aligned}$$

as $n \to \infty$. So we have proved the convergence of (5.10), and this completes the proof of result (b).

We are now ready to prove (2.10). The fact that $d_n^2 = ES_n^2 \sim c_\mu n^{3-2\mu} h^2(n)$ with $c_\mu = \frac{1}{(1-\mu)(3-2\mu)} \int_0^\infty x^{-\mu}(x+1)^{-\mu}dx$ can be found in the Proposition 2.1 of Wang, Ling and Gulati (2003a). Now it follows from Gorodetskii (1977) [also see Wang, Ling and Gulati (2003b)] that $x_{[nt],n} \Rightarrow W_\beta(t), 0 \leq t \leq 1$, on $D[0,1]$, where $\beta = (3-2\mu)/2$ and $W_\beta(t)$ is a fractional Brownian motion having a continuous local time $L_{W_\beta}(t,s)$ with regard to (t,s) in $[0,\infty) \times R$. This proves that $x_{i,n}$ satisfies the Assumption 2.2.

We next show that $x_{i,n}$ also satisfies the Assumption 2.3 and then (2.10) follows from Theorem 2.1 accordingly. In order to check the Assumption 2.3, let $\mathcal{F}_{t,n} = \sigma\{..., \epsilon_{t-1}, \epsilon_t\}$ and $d_{l,k,n} = \Lambda_{l-k}/d_n$. Recall $\Lambda_n^2/d_n^2 \sim \frac{1}{1-\mu}\int_0^1 x^{-\mu}(x+1)^{-\mu}dx$. It is readily seen that, as $n \to \infty$,

$$\inf_{(l,k)\in\Omega_n(\eta)} d_{l,k,n} = \inf_{(l,k)\in\Omega_n(\eta)} \Lambda_{l-k}/d_n \geq C\ \eta^{(3-2\mu)/2},$$

for some constant $C > 0$ and $d_{l,k,n}$ satisfy (2.2)–(2.4). On the other hand, by noting that

$$\begin{aligned} S_l &= \sum_{j=1}^{l}\sum_{i=-\infty}^{j} \epsilon_i \psi_{j-i} \\ &= \sum_{j=1}^{k}\sum_{i=-\infty}^{j} \epsilon_i \psi_{j-i} + \sum_{j=k+1}^{l}\sum_{i=-\infty}^{j} \epsilon_i \psi_{j-i} \\ &= S_k + \sum_{j=k+1}^{l}\sum_{i=-\infty}^{k} \epsilon_i \psi_{j-i} + \sum_{j=k+1}^{l}\sum_{i=k+1}^{j} \epsilon_i \psi_{j-i} \\ &:= S_k + S_{1l} + S_{2l}, \end{aligned}$$

it follows from the independence of ϵ_i, results (a) and (b) above, and the fact $S_{2l} =_d \tilde{S}_{l-k}$ (where $=_d$ denotes equivalence in distribution) that, conditional on $\mathcal{F}_{k,n}$, $(x_{l,n} - x_{k,n})/d_{l,k,n} = (S_{1l}+S_{2l})/\Lambda_{l-k}$ has a density $h_{l-k}(x - S_{1l}/\Lambda_{l-k})$ which is uniformly bounded by a constant K for all $n \geq 1$ and

$$\begin{aligned} & \sup_{(l,k)\in\Omega_n[\delta^{1/\alpha}]} \sup_{|u|\leq\delta} \big|h_{l-k}(u - S_{1l}/\Lambda_{l-k}) - h_{l-k}(-S_{1l}/\Lambda_{l-k})\big| \\ \leq\ & 2 \sup_{(l,k)\in\Omega_n[\delta^{1/\alpha}]} \sup_{x} |h_{l-k}(x) - \frac{1}{\sqrt{2\pi}} e^{-x^2/2}| + \frac{1}{\sqrt{2\pi}} \sup_{|u|\leq\delta} \sup_{x} |e^{-(x+u)^2/2} - e^{-x^2/2}| \\ \to\ & 0, \end{aligned}$$

as $n \to \infty$ first and then $\delta \to 0$, because of (5.10). This proves that the Assumption 2.3 holds true for x_{in}, and also completes the proof of (2.10).

By noting that $\Lambda_n^2 \sim d_n^2 \sim \psi^2\, n$ if $\sum_{k=0}^{\infty} |\psi_k| < \infty$ and $\psi \equiv \sum_{k=0}^{\infty} \psi_k \neq 0$, the proof of (2.11) is similar to that of (2.10) except that the weak convergence in Gorodetskii (1977) is replaced by Hannan (1979). We omit the details. The proof of Corollary 2.2 is now complete.

Proof of Theorem 3.1. Without loss of generality we assume that ϵ_t and u_t, $1 \leq t \leq n$ are defined on the same probability space $\{\Omega, \mathcal{F}, P\}$. If it were not so, it can be easily arranged since the result to be proved in (3.4) involves only weak convergence. In order to prove (3.4), we split $\hat{f}(x) - f(x)$ as

$$\hat{f}(x) - f(x) = \frac{\sum_{t=1}^{n} u_t K_h(x_t - x)}{\sum_{t=1}^{n} K_h(x_t - x)} + \frac{\sum_{t=1}^{n} \big[f(x_t) - f(x)\big] K_h(x_t - x)}{\sum_{t=1}^{n} K_h(x_t - x)}.$$

It is readily seen that

$$\Big(h \sum_{t=1}^{n} K_h(x_t - x)\Big)^{1/2} (\hat{f}(x) - f(x)) = \sum_{t=1}^{n} u_t\, Z_{nt} + \Theta_{1n}/\Theta_{2n}, \tag{5.11}$$

where $Z_{nt} = \big(\frac{d_n}{nh}\big)^{1/2} K\big(\frac{d_n}{h}\, x_{t,n} - \frac{x}{h}\big)/\Theta_{2n}$ with $\Theta_{2n}^2 = \frac{d_n}{nh}\sum_{t=1}^n K\big(\frac{d_n}{h}\, x_{t,n} - \frac{x}{h}\big)$ and

$$\Theta_{1n} = \big(\frac{d_n}{nh}\big)^{1/2} \sum_{t=1}^n \big[f(d_n\, x_{t,n}) - f(x)\big]\, K\big(\frac{d_n}{h}\, x_{t,n} - \frac{x}{h}\big).$$

Recall that $x_{i,n}$ satisfies Assumptions 2.2* and 2.3, and for any $\lambda \geq 1$, $g(s) = K^{\lambda}(s)$ satisfies Assumption 2.1 due to the Assumption 3.1. It follows from Theorem 2.1 and Remark 2.1 that, for any $\lambda \geq 1$ and h satisfying $nh^3/d_n \to 0$ (hence $h \to 0$) and $nh/d_n \to \infty$,

$$\frac{d_n}{nh}\sum_{t=1}^n K^{\lambda}\big(\frac{d_n}{h}\, x_{t,n} - \frac{x}{h}\big) \to_P L_G\,(1,0)\int_{-\infty}^{\infty} K^{\lambda}(s)ds. \tag{5.12}$$

Now, to prove (3.4), it suffices to show that, for any h satisfying $nh^3/d_n \to 0$ and $nh/d_n \to \infty$,

$$E|\Theta_{1n}| \to 0, \tag{5.13}$$

and

$$V_n \equiv \frac{1}{\Lambda_n}\sum_{t=1}^n u_t\, Z_{nt} \;\to_D\; N(0,\sigma^2), \tag{5.14}$$

where $\Lambda_n^2 = \Theta_{2n}^{-2}\,\frac{d_n}{nh}\,\sum_{t=1}^n K^2\big(\frac{d_n}{h}\, x_{t,n} - \frac{x}{h}\big)$. Indeed it follows from (5.12) with $\lambda = 1$ and $\lambda = 2$ respectively that $\Theta_{2n}^2 \to_P L_G(1,0)$ and $\Lambda_n^2 \to_P \int_{-\infty}^{\infty} K^2(s)ds$. These facts, together with (5.13) and (5.14), yield that $\Theta_{1n}/\Theta_{2n} \to 0$ and $\sum_{t=1}^n u_t\, Z_{nt} \to N(0,\sigma_1^2)$, and hence we have (3.4) due to (5.11).

We next prove (5.13) and (5.14), starting with (5.13). In fact, recalling that $x_{tn}/d_{t,0,n}$ has a density $h_{t,0,n}(x)$ [in the notation of Assumption 2.3 (ii) due to $x_{i,n}$ satisfying Assumption 2.3] which is uniformly bounded by a constant K, we have

$$\begin{aligned}
E|\Theta_{1n}| &\leq \big(\frac{d_n}{nh}\big)^{1/2}\sum_{t=1}^n E\Big\{\big|f(d_n\, x_{t,n}) - f(x)\big|\, K\big(\frac{d_n}{h}\, x_{t,n} - \frac{x}{h}\big)\Big\} \\
&= \big(\frac{d_n}{nh}\big)^{1/2}\sum_{t=1}^n \int_{-\infty}^{\infty}\Big\{\big|f(d_n\, d_{t,0,n}\, y) - f(x)\big|\, K\big(\frac{d_n\, d_{t,0,n}}{h}\, y - \frac{x}{h}\big)\Big\} h_{t,0,n}(y)\, dy \\
&\leq \big(\frac{d_n}{nh}\big)^{1/2}\frac{h}{d_n}\sum_{t=1}^n (d_{t,0,n})^{-1}\int_{-\infty}^{\infty}\Big\{\big|f(h\, y + x) - f(x)\big|\, K\big(y\big)\Big\}\, dy \\
&\leq h\,\big(\frac{nh}{d_n}\big)^{1/2}\frac{1}{n}\sum_{t=1}^n (d_{t,0,n})^{-1}\int_{-\infty}^{\infty} K\big(s\big)\, f_1(s)\, ds \\
&\to 0,
\end{aligned}$$

since $nh^3/d_n \to 0$ and the fact that $d_{t,0,n}$ satisfies (2.4). This yields (5.13). As for (5.14), by noting that, given $\epsilon_1, \epsilon_2, ..., \epsilon_t$, $(Z_{nt}\, u_t, t = 1, 2, ..., n)$ is a martingale difference since ϵ_t is independent of u_t, it follows from Theorem 3.9 [(3.75) there] in Hall and Heyde (1980) with $\delta = q/2 - 1$ that

$$\sup_x \big| P(V_n \le x\sigma \mid \epsilon_1, \epsilon_2, ..., \epsilon_n) - \Phi(x) \big| \le A(\delta)\, \mathcal{L}_n^{1/(1+q)}, \quad a.s., \tag{5.15}$$

where $A(\delta)$ is a constant depending only on δ and

$$\mathcal{L}_n = \frac{1}{\sigma^q \Lambda_n^q} \sum_{k=1}^n |Z_{nk}|^q E|u_k|^q + E\Big| \frac{1}{\sigma^2 \Lambda_n^2} \sum_{k=1}^n Z_{nk}^2 \big[E(u_k^2 | \mathcal{F}_{k-1}) - \sigma^2\big] \Big|^{q/2}.$$

Recall Assumption 3.3, $K(x)$ is uniformly bounded and $\Lambda_n^2 = \sum_{t=1}^n Z_{tn}^2$. Routine calculations show that

$$\mathcal{L}_n \le \frac{A}{\sigma^q \Lambda_n^{q-2}} \big(\frac{d_n}{nh}\big)^{(q-2)/2} + o_P(1) = o_P(1),$$

since $\Lambda_n^2 \to_P \int_{-\infty}^{\infty} K^2(s) ds$, $q > 2$ and $nh/d_n \to 0$. Therefore, we obtain

$$\sup_x \big| P(V_n \le x\sigma) - \Phi(x) \big| \le E\Big[\sup_x \big| P(V_n \le x\sigma \mid x_1, x_2, ..., x_n) - \Phi(x) \big| \Big] \to 0.$$

This proves (5.14) and also completes the proof of Theorem 3.1.

Proof of Theorem 3.2. The idea of this theorem is similar to Park and Phillips (2001). First notice that, under the assumption $(U_n, V_n) \Rightarrow_D (U, V)$, it follows from the so-called Skorohod-Dudley-Wichura representation theorem that there is a common probability space $(\Omega, \mathcal{F}, P)$ supporting (U_n^0, V_n^0) and (U, V) such that

$$(U_n, V_n) =_d (U_n^0, V_n^0) \quad \text{and} \quad (U_n^0, V_n^0) \to_{a.s.} (U, V) \tag{5.16}$$

in $D[0,1]^2$ with the uniform topology. Moreover, as in the proof of Lemma 2.1 in Park and Phillips (2001), (U_n^0, V_n^0) can be chosen such that, for each $n \ge 1$

$$U_n^0(k/n) =_d U_n(k/n) \quad \text{and} \quad V_n^0(k/n) =_d V(\tau_{nk}/n), \qquad k = 1, 2, ..., n, \tag{5.17}$$

where $\tau_{nt}, 0 \le t \le 1$, are stopping times in $(\Omega, \mathcal{F}, P)$ with $\tau_{n0} = 0$ satisfying

$$\sup_{0 \le t \le 1} \Big| \frac{\tau_{nt} - t}{n^\delta} \Big| \to_{a.s} 0 \tag{5.18}$$

as $n \to \infty$ for any $\delta > \max(1/2, 2/q)$. These facts, together with (5.11), yield that, under the extended probability space,

$$\Big(h \sum_{t=1}^n K_h(x_t - x)\Big)^{1/2} (\hat{f}(x) - f(x))$$

$$=_d \ \frac{1}{\Theta_{2n}^*} \sum_{t=1}^n Y_{nt} \big[V^0(\frac{\tau_{n,t}}{n}) - V^0(\frac{\tau_{n,t-1}}{n})\big] \ + \ \frac{\Theta_{1n}^*}{\Theta_{2n}^*}, \tag{5.19}$$

where $Y_{nt} = \big(\frac{d_n}{h}\big)^{1/2} K\big[\frac{d_n}{h} U_n^0(\frac{t}{n}) - \frac{x}{h}\big]$, $\Theta_{2n}^{*2} = \frac{d_n}{nh} \sum_{t=1}^n K\big[\frac{d_n}{h} U_n^0(\frac{t}{n}) - \frac{x}{h}\big]$ and

$$\Theta_{1n}^* = \big(\frac{d_n}{nh}\big)^{1/2} \sum_{t=1}^n \big[f(d_n U_n^0(\frac{t}{n})) - f(x)\big] K\Big[\frac{d_n}{h} U_n^0(\frac{t}{n}) - \frac{x}{h}\Big].$$

Since (5.16) implies that Assumption 2.2* holds true for $U_n^0(t/n)$ with $G(t)$ being a Brownian motion [that is, $G(t) = U(t)$], it follows from a similar argument to the proofs of (5.12) and (5.13) that, for any $\lambda \geq 1$,

$$\frac{d_n}{nh} \sum_{t=1}^{[nr]} K^\lambda\big[\frac{d_n}{h} U_n^0(\frac{t}{n}) - \frac{x}{h}\big] \to_P L_U(r, 0) \int_{-\infty}^{\infty} K^\lambda(s) ds, \tag{5.20}$$

uniformly in $r \in [0, 1]$ and $E|\Theta_{1n}^*| \to 0$. We mention that (5.20) also implies, for any $\lambda \geq 1$, uniformly in $r \in [0, 1]$,

$$\begin{aligned} \frac{d_n}{h} \int_0^r K^\lambda\big[\frac{d_n}{h} U_n^0(s) - \frac{x}{h}\big] ds \quad &= \quad \frac{d_n}{nh} \sum_{t=1}^{[nr]} K^\lambda\big[\frac{d_n}{h} U_n^0(\frac{t}{n}) - \frac{x}{h}\big] \\ &\quad + \frac{d_n}{nh} K^\lambda\big(-\frac{x}{h}\big) + K^\lambda\big(\frac{d_n}{h} U_n^0(\frac{[nr]}{n}) - \frac{x}{h}\big)(nr - [nr]) \\ &\to_P \quad L_U(r, 0) \int_{-\infty}^{\infty} K^\lambda(s) ds, \end{aligned} \tag{5.21}$$

since $K(x)$ is uniformly bounded and $\frac{d_n}{nh} \to 0$.

By virtue of (5.20) and $E|\Theta_{1n}^*| \to 0$, we have $\Theta_{2n}^* \to_P L_U(1, 0)$ and $\Theta_{1n}^*/\Theta_{2n}^* \to_P 0$. These facts, together with (5.19) and an argument similar to that in the proof of Theorem 3.3 in Hall and Heyde (1980) imply that (3.4) will follow if we prove

$$\sum_{t=1}^n Y_{nt}\big[V^0(\frac{\tau_{n,t}}{n}) - V^0(\frac{\tau_{n,t-1}}{n})\big] \to_D \eta\, N, \tag{5.22}$$

where $\eta^2 = L_U(1, 0) \int_{-\infty}^{\infty} K^2(s) ds$ and N is a standard normal variable independent of η.

In order to prove (5.22), write

$$M_n(r) \quad = \quad \sum_{t=1}^{j-1} Y_{nt}\big[V^0(\frac{\tau_{n,t}}{n}) - V^0(\frac{\tau_{n,t-1}}{n})\big] + Y_{n,j-1}\big[V^0(\frac{r}{n}) - V^0(\frac{\tau_{n,j-1}}{n})\big], \tag{5.23}$$

for $\tau_{n,j-1}/n < r \leq \tau_{n,j}/n, j = 1, 2, ..., k$. It is readily seen that M_n is a continuous martingale with the quadratic variation process $[M_n]$ given by

$$\begin{aligned} [M_n]_r \quad &= \quad \sum_{k=1}^{j-1} Y_{nt}^2 \Big(\frac{\tau_{n,k}}{n} - \frac{\tau_{n,k-1}}{n}\Big) + Y_{n,j-1}^2 \Big(r - \frac{\tau_{n,j-1}}{n}\Big) \\ &= \quad \frac{d_n}{h} \int_0^r K^2\big[\frac{d_n}{h} U_n^0(s) - \frac{x}{h}\big] ds \,\big[1 + o_P(1)\big] \\ &\to_P \quad L_U(r, 0) \int_{-\infty}^{\infty} K^2(s) ds, \end{aligned} \tag{5.24}$$

uniformly in $r \in [0,1]$, in view of (5.18) and (5.21) with $\lambda = 2$. For the covariance process $[M_n, U]$ of M_n and U, we also have

$$\begin{aligned}
[M_n, U]_r &= \sum_{k=1}^{j-1} Y_{nk} \Big(\frac{\tau_{n,k}}{n} - \frac{\tau_{n,k-1}}{n}\Big)\, \sigma_{uv} \, + Y_{n,k-1}\Big(r - \frac{\tau_{n,j-1}}{n}\Big)\sigma_{uv} \\
&= \sigma_{uv}\, \big(\frac{d_n}{h}\big)^{1/2} \int_0^r K\big[\frac{d_n}{h}\, U_n^0(s) - \frac{x}{h}\big]\, ds\, \big[1 + o_P(1)\big] \\
&\to_P \; 0, \qquad (5.25)
\end{aligned}$$

where $\sigma_{uv} = cov(V, U)$, since $(h/d_n)^{1/2} \to 0$ and (5.21) with $\lambda = 1$. Now, following the proof of Theorem 3.2 in Park and Phillips (2001), we obtain that

$$\sum_{t=1}^{j-1} Y_{nt}\big[V^0(\frac{\tau_{n,t}}{n}) - V^0(\frac{\tau_{n,t-1}}{n})\big] \;=\; M_n(\frac{\tau_{n,n}}{n}) \to_D \eta\, N,$$

which yields (5.22). This completes the proof of Theorem 3.2.

REFERENCES

Akonom, J. (1993). Comportement asymptotique du temps d'occupation du processus des sommes partielles. *Ann. Inst. H. Poincaré Probab. Statist.* **29**, 57–81.

Bandi, F. (2004). On Persistence and nonparametric estimation (with an application to stock return predictability). *Unpublished manuscript.*

Bandi, F. and P. C. B. Phillips (2003). Fully Nonparametric Estimation of Scalar Diffusion Models. *Econometrica*, **71(1)** 241-283.

Berkes, I.; Horváth, L. (2006). Convergence of integral functionals of stochastic processes. *Econometric Theory*, **22**, no. 2, 304–322.

Borodin, A. N.; Ibragimov, I. A. (1995). *Limit theorems for functionals of random walks.* Proc. Steklov Inst. Math., no. 2.

de Jong, R. (2004). Addendum to: "Asymptotics for nonlinear transformations of integrated time series". *Econometric Theory*, **20** , 627–635.

de Jong, R.; Wang, Chien-Ho (2005). Further results on the asymptotics for nonlinear transformations of integrated time series. *Econometric Theory*, **21**, 413–430.

Feller, W. (1971). *An introduction to probability theory and its applications.* Vol. II. Second edition John Wiley & Sons, Inc.

Geman, D, J. Horowitz (1980). Occupation densities. *Ann. Probab.*, **8** , 1–67.

Gorodeckiĭ, V. V. (1977). Convergence to semistable Gaussian processes. *Teor. Verojatnost. i Primenen.* **22**, 513–522.

Guerre, E. (2004). Design-Adaptive pointwise nonparametric regression estimation for recurrent Markov time series. *Unpublished pamnuscript.*

Hall, P.; Heyde, C. C. (1980). *Martingale limit theory and its application,* Probability and Mathematical Statistics. Academic Press, Inc.

Hannan, E. J.(1979). The central limit theorem for time series regression. *Stochastic Process. Appl.* **9**, 281–289.

Hong, S. H. and P. C. B. Phillips (2005). Testing Linearity in Cointegrating Relations with an Application to Purchasing Power Parity. *Cowles foundation discussion paper No. 1541.*

Hu, L. and P. C. B. Phillips (2004). Dynamics of the Federal Funds Target Rate: A Nonstationary Discrete Choice Approach. *Journal of Applied Econometrics*, **19**, 851-867.

Jeganathan, P. (2004). Convergence of functionals of sums of r.v.s to local times of fractional stable motions. *Ann. Probab.* **32** , no. 3A, 1771–1795.

Karlsen, H. A.; Tjøstheim, D. (2001). Nonparametric estimation in null recurrent time series. *Ann. Statist.* **29**, no. 2, 372–416.

Karlesn, H. A., Myklebust, T. and Tjostheim, D. (2007). Nonparametric estimation in a nonlinear cointegration model, *Ann. Statist.*, to appear.

Kasparis, I. (2004). Detection of Functional from Misspecification in cointegration relations. *(unpublished manuscript)*

Lukacs, E. (1970). *Characteristic functions.* Hafner Publishing Co., New York.

Park, J. Y. (2006). Spatial analysis of time series, *FEMES Meetings, Beijing.*

Park, J. Y. and P. C. B. Phillips (1999). Asymptitocs for nonlinear transformation of integrated time series, *Econometric Theory*, **15**, 269-298.

Park, J. Y. and P. C. B. Phillips (2001). Nonlinear regressions with integrated time series, *Econometrica*, **69**, 117-161.

Phillips, P. C. B. (2001). Descriptive Econometrics for Non-Stationary Time Series with Empirical Applications. *Journal of Applied Econometrics*, **16**, 389413.

Phillips, P. C. B. (2005). Econometric Analysis of Fisher's equation, *American Journal of Economics and Sociology*, **64(1)**, 125-168.

Phillips, P. C. B. and J. Y. Park (1998). Nonstationary Density Estimation and Kernel

Autoregression. *Cowles Foundation discuss paper No. 1181.*

Pötscher, B. M. (2004). Nonlinear functions and convergence to Brownian motion: beyond the continuous mapping theorem, *Econometric Theory*, **20**, no. 1, 1–22.

Revuz, D.; Yor, M. (1999). *Continuous martingales and Brownian motion.* 293. Springer-Verlag, Berlin.

Taqqu, M. S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrsch. Verw. Geb.* **31** , 287–302.

Wang, Q., Lin, Y.-X., Gulati, C. M. (2003a). Strong approximation for long memory processes with applications. *J. Theoret. Probab.* **16**, 377–389.

Wang, Q., Lin, Y.-X., Gulati, C. M. (2003b). Asymptotics for general fractionally integrated processes with applications to unit root tests. *Econometric Theory*, **19**, 143–164.

Shorack, G. R. And J. A. Wellner (1986). *Empirical Processes with Applications to Statistics.* New York: Wiley.