Econometric Theory, **27**, 2011, 235–259.
doi:10.1017/S0266466610000277

ASYMPTOTIC THEORY FOR ZERO ENERGY FUNCTIONALS WITH NONPARAMETRIC REGRESSION APPLICATIONS

QIYING WANG
University of Sydney

PETER C.B. PHILLIPS
Yale University
University of Auckland
University of Southampton
and
Singapore Management University

A local limit theorem is given for the sample mean of a zero energy function of a nonstationary time series involving twin numerical sequences that pass to infinity. The result is applicable in certain nonparametric kernel density estimation and regression problems where the relevant quantities are functions of both sample size and bandwidth. An interesting outcome of the theory in nonparametric regression is that the linear term is eliminated from the asymptotic bias. In consequence and in contrast to the stationary case, the Nadaraya–Watson estimator has the same limit distribution (to the second order including bias) as the local linear nonparametric estimator.

1. INTRODUCTION

Consider an array $x_{k,n}, 1 \leq k \leq n, n \geq 1$ constructed from some underlying nonstationary time series and assume that there is a continuous limiting Gaussian process $G(t), 0 \leq t \leq 1$, to which $x_{[nt],n}$ converges weakly, where $[a]$ denotes the integer part of a. For instance, in many applications we encounter quantities such as $x_{k,n} = d_n^{-1} x_k$ where x_k is a nonstationary time series, such as a unit root or long memory process, for which d_n is an appropriate standardization factor. A common functional of interest S_n of $x_{k,n}$ is defined by

Our thanks to the co-editor and two referees for helpful comments on the original version of this paper. Wang acknowledges partial research support from the Australian Research Council. Phillips acknowledges partial research support from the NSF under grant SES 06-47086. Address correspondence to Peter C.B. Phillips, Cowles Foundation, Yale University, P.O. Box 208281, New Haven, CT 06520, USA; e-mail: peter.phillips@yale.edu.

the sample quantity

$$S_n = \sum_{k=1}^{n} g(c_n x_{k,n}), \tag{1.1}$$

where c_n is a certain sequence of positive constants and g is a real integrable function on R. Such functionals arise in nonparametric estimation problems, particularly those involving nonlinear cointegration models, where the underlying time series x_k are nonstationary, g is a kernel function, and the secondary sequence c_n depends on the bandwidth used in the nonparametric regression.

The limit behavior of S_n in the situation where $\int_{-\infty}^{\infty} g(s)\,ds \neq 0$ was studied in Wang and Phillips (2009a), where it was shown that when $c_n \to \infty$ and $n/c_n \to \infty$,

$$\frac{c_n}{n} S_n \to_D \int_{-\infty}^{\infty} g(x)\,dx\, L_G(1,0), \tag{1.2}$$

where $L_G(t,s)$ is the local time of the process $G(t)$ at the spatial point s, defined in the following section. When the function g is a kernel density, the limit (1.2) is simply the local time of G at the origin. This limit may be recentered at an arbitrary spatial point s by using $g\left(c_n\left(x_{k,n}-s\right)\right)$ in place of $g(c_n x_{k,n})$ in (1.1). Jeganathan (2004) 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 independent and identically distributed (i.i.d.) random variables, related results were given in Borodin and Ibragimov (1995), Akonom (1993), and Phillips and Park (1998). Results of the type (1.2) have many statistical applications, especially in nonparametric estimation—see Wang and Phillips (2009a, 2009b).

The present work is concerned with developing a limit theory for the sample function S_n in the zero energy case where $\int_{-\infty}^{\infty} g(s)\,ds = 0$. Such cases are important in nonparametric regression and appear in the analysis of bias and in derivative estimation problems. In bias analysis, e.g., we need to consider functions of the form $g(s) = sK(s)$, where $K(s)$ is the kernel function used in nonparametric estimation, and then $\int g(s)\,ds = 0$ when K is a symmetric function. Interestingly, in this case it turns out that for nonstationary time series, the expression for the bias in the limit theory involves no linear term in the bandwidth, in contrast to the stationary case. One consequence of this change in the limit theory is that the local level (Nadaraya–Watson) estimator has the same asymptotic distribution including the bias correction as that of the local linear estimator in nonstationary cointegrating regression. These issues are explored in Section 2 (see Remarks 2.5 and 2.6 for details). Similarly, in nonparametric derivative estimation, we need to deal with functions such as the kernel derivative $g(s) = K'(s)$, which again have zero energy when K is symmetric. Theorem 2.1 shows that the limit theory for S_n in (1.1) differs from (1.2) when g has zero energy in terms of both rate of convergence and the limiting process.

2. MAIN RESULTS

Let $\{\xi_j, j \geq 1\}$ be a linear process defined by

$$\xi_j = \sum_{k=0}^{\infty} \phi_k \epsilon_{j-k}, \tag{2.1}$$

where $\{\epsilon_j, -\infty < j < \infty\}$ is a sequence of i.i.d. random variables with $\mathrm{E}\epsilon_0 = 0$, $\mathrm{E}\epsilon_0^2 = 1$, and characteristic function $\varphi(t)$ of ϵ_0 satisfying $\int_{-\infty}^{\infty} |\varphi(t)| dt < \infty$. Throughout the paper, the coefficients $\phi_k, k \geq 0$, are assumed to satisfy one of the following conditions.

C1. $\phi_k \sim k^{-\mu} \rho(k)$, *where* $1/2 < \mu < 1$ *and* $\rho(k)$ *is a function slowly varying at* ∞.

C2. $\sum_{k=0}^{\infty} |\phi_k| < \infty$ *and* $\phi \equiv \sum_{k=0}^{\infty} \phi_k \neq 0$.

Put $x_i = \sum_{j=1}^{i} \xi_j$ and let $g(x)$ be a Borel measurable function on R. As discussed earlier, the present paper is concerned with the limit behavior of sample functions of the form $\sum_{k=1}^{n} g\left(x_k/h\right)$, when $n \to \infty$, $h \equiv h_n \to 0$, and g is an integrable zero energy function for which $\int_{-\infty}^{\infty} g(x)\, dx = 0$.

We start with the following notation. A fractional Brownian motion with $0 < \beta < 1$ on $D[0,1]$ is defined by

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

where

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

$W(s), 0 \leq s < \infty$ is a standard Brownian motion, and for $-\infty < s \leq 0$, $W(s)$ is taken to be $W^*(-s)$, where $W^*(s), 0 \leq s < \infty$ is an independent copy of $W(s), 0 \leq s < \infty$. It is readily seen that $W_{1/2}(t) = W(s)$ 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, e.g., Theorem 22.1 of Geman and Horowitz (1980).

Here and subsequently, 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 $T(x)$,

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

with probability one.

We now develop a limit theory for the sample function (1.1) in the zero energy case. Write $d_n^2 = \mathrm{E}x_n^2$. It is well known (see, e.g., Wang, Lin, and Gulati, 2003)

https://doi.org/10.1017/S0266466610000277 Published online by Cambridge University Press

that

$$d_n^2 \sim \begin{cases} c_\mu\, n^{3-2\mu}\, \rho^2(n), & \text{under C1,} \\ \phi^2\, n, & \text{under C2,} \end{cases} \tag{2.2}$$

where $c_\mu = 1/((1-\mu)(3-2\mu)) \int_0^\infty x^{-\mu}(x+1)^{-\mu} dx$. Setting $c_n = d_n/h$, we consider the standardized version

$$\left(\frac{c_n}{n}\right)^{1/2} \sum_{k=1}^n g(c_n\, x_{k,n}) = \left(\frac{d_n}{nh}\right)^{1/2} \sum_{k=1}^n g\,(x_k/h).$$

Our main result is as follows.

THEOREM 2.1. *Assume that* $\int |g(t)|dt < \infty$, $\int |\hat{g}(t)|dt < \infty$ *and* $|\hat{g}(t)| \le C \min\{|t|, 1\}$, *where* $\hat{g}(x) = \int e^{itx} g(t)dt$ *and* C *is a positive constant. Then, for any* $h \to 0$ *(*$h^2 \log n \to 0$ *under C2) and* $nh/d_n \to \infty$, *we have*

$$\left(\frac{d_n}{nh}\right)^{1/2} \sum_{k=1}^n g\left(x_k/h\right) \to_D \tau\, N\, \psi^{1/2}(1), \tag{2.3}$$

where $\tau^2 = \int g^2(s)ds$, N *is a standard normal variate independent of* $\psi(t)$, *and for* $0 \le t \le 1$, *the process* $\psi(t)$ *is defined by*

$$\psi(t) = \begin{cases} L_{W_{3/2-\mu}}(t,0), & \textit{under C1,} \\ L_W(t,0), & \textit{under C2.} \end{cases}$$

Remark 2.1. The conditions on $g(x)$ imply $\int g(x)\,dx = 0$ and $\int g^2(x)\,dx < \infty$. Indeed it follows by dominated convergence that

$$\int g(x)dx = \int \lim_{t\to 0} e^{itx} g(x)\,dx = \lim_{t\to 0} \hat{g}(t) = 0.$$

On the other hand, $\int g^2(x)\,dx = (2\pi)^{-1} \int \hat{g}^2(x)\,dx \le (2\pi)^{-1} \int |g(x)|\,dx \int |\hat{g}(x)|\,dx < \infty$. This fact will be used in the proof without further explanation. Integrability of $\hat{g}(x)$ is a mild condition, and $|\hat{g}(t)| \le C \min\{|t|, 1\}$ is implied by $\int (1+|x|)|g(x)|\,dx < \infty$. Many commonly used functions, such as the normal kernel function or functions having a compact support with $\int g(x)\,dx = 0$, satisfy the conditions on $g(x)$ in Theorem 2.1. These conditions are particularly convenient for our proofs. More direct conditions such as $\int g(x)\,dx = 0$, $\int (1+|x|)|g(x)|\,dx < \infty$, and $\int g^2(x)\,dx < \infty$ might be imposed on g, but it is not clear whether these are sufficient for our results.

Remark 2.2. If $\int g(t)dt \ne 0$, the limit behavior of $\sum_{k=1}^n g\left(x_k/h\right)$ is quite different and involves a different rate of convergence. It has been proved as a corollary of a more general result in Wang and Phillips (2009a) that

$$\frac{d_n}{nh} \sum_{k=1}^n g\left(x_k/h\right) \to_D \psi(1) \int g(x)dx.$$

Jeganathan (2004) and Borodin and Ibragimov (1995) provide related results for such sample functions. The latter monograph investigated the limit behavior of $\sum_{k=1}^{n} g\left(x_k/h\right)$ under more general settings on $g(x)$, but required x_k to be a partial sum of i.i.d. random variables.

Remark 2.3. Assume that $\phi_0 = 1$ and $\phi_j = 0$, $j \geq 1$. In this setting, $x_i = \sum_{j=1}^{i} \epsilon_j$ is a partial sum of i.i.d. random variables, and $d_n^2 = n$. Under some conditions on $g(x)$ that are similar to those in Theorem 2.1, Theorem 4.3.3 of Borodin and Ibragimov (1995) established that

$$\left(\frac{d_n}{n}\right)^{1/2} \sum_{k=1}^{n} g\left(x_k\right) \to_D \tau' N L_W^{1/2}(1,0), \tag{2.4}$$

where $\tau'^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{g}(x)|^2 \left[1 + 2\sum_{k=1}^{\infty} \varphi^k(x)\right] dx$ with $\varphi(t) = \mathrm{E}e^{it\epsilon_0}$. Note that $\tau^2 = \int g^2(x)\,dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{g}(x)|^2 dx$ in (2.3), which is related to τ'^2. But there is an essential difference between (2.3) and (2.4). In particular, (2.4) is only a partial invariance principle because the limit involves the characteristic function $\varphi(t) = \mathrm{E}e^{it\epsilon_0}$ of the innovations in x_k and so the constant τ' in (2.4) is dependent on this distribution. The reason underlying the difference between (2.3) and (2.4) is that the sample autocovariances of the summand in (2.3) satisfy

$$J_n \equiv \frac{d_n}{nh} \sum_{1 \leq k < l \leq n} g\left(x_k/h\right) g\left(x_l/h\right) = O_P(h).$$

See the proof of Proposition 3.3. Hence $J_n = o_P(1)$, when $h \to 0$, and so J_n does not contribute to the limit behavior of $\left(\frac{d_n}{nh}\right)^{1/2} \sum_{k=1}^{n} g\left(x_k/h\right)$. The extension of (2.4) to linear processes can be found in Jeganathan (2008). Our proof is different from Jeganathan (2008), and the presence of the bandwidth sequence h seems to simplify the limit theory.

Remark 2.4. If $|f_j(x)|$ and $f_j^2(x)$, $j = 1, 2$, are Lebesgue integrable functions on R with $\tau_1 = \int f_1(x)\,dx \neq 0$ and $\tau_2 = \int f_2(x)\,dx \neq 0$, in addition to the result (2.3), we have

$$\left\{ \left(\frac{d_n}{nh}\right)^{1/2} \sum_{k=1}^{n} g\left(x_k/h\right),\ \frac{d_n}{nh} \sum_{k=1}^{n} f_1\left(x_k/h\right),\ \frac{d_n}{nh} \sum_{k=1}^{n} f_2\left(x_k/h\right) \right\}$$
$$\to_D \left\{ \tau N \psi^{1/2}(1),\ \tau_1 \psi(1),\ \tau_2 \psi(1) \right\}, \tag{2.5}$$

where the notation $\to_D$ is defined as in Section 3.2. As a direct consequence of (2.5), we have the following corollary, which provides a self-normalized result for additive functionals of random sums.

COROLLARY 2.1. *Assume that* $\int [|g(t)| + g^4(t)]\,dt < \infty$, $\int |\hat{g}(t)|\,dt < \infty$, *and* $|\hat{g}(t)| \leq C \min\{|t|, 1\}$, *where* $\hat{g}(x) = \int e^{itx} g(t)\,dt$ *and* C *is a positive constant.*

Then, for any $h \to 0$ ($h^2 \log n \to 0$ under C2) and $nh/d_n \to \infty$, we have

$$\frac{\sum_{k=1}^{n} g\left(x_k/h\right)}{\sqrt{\sum_{k=1}^{n} g^2\left(x_k/h\right)}} \to_D N(0,1). \tag{2.6}$$

Remark 2.5. Result (2.5) is also useful in nonparametric bias analysis related to nonstationary cointegration regression. To illustrate, consider the following nonlinear structural model of cointegration:

$$y_t = f(x_t) + u_t, \quad t = 1, 2, \ldots, n, \tag{2.7}$$

where u_t is a zero mean stationary equilibrium error and f is an unknown function to be estimated with the observed data $\{y_t, x_t\}_{t=1}^n$. The conventional kernel estimate of $f(x)$ in model (2.7) is given by

$$\hat{f}(x) = \frac{\sum_{t=1}^{n} y_t K_h(x_t - x)}{\sum_{t=1}^{n} K_h(x_t - x)}, \tag{2.8}$$

where $K_h(s) = (1/h)K(s/h)$, $K(x)$ is a nonnegative real function, and the bandwidth parameter $h \equiv h_n \to 0$ as $n \to \infty$. Under certain conditions on $f(x)$, u_t, and h, it is shown in Wang and Phillips (2009a) that

$$(nh^2)^{1/4}\,(\hat{f}(x) - f(x)) \to_D C_0\, N\, L_W^{-1/2}(1,0), \tag{2.9}$$

where C_0 is a constant related to the kernel $K(x)$ and the moment $\mathrm{E}u_t^2$. By making use of the result (2.5), together with some additional smoothness conditions on $f(x)$, an explicit bias term may be incorporated into the limit theory (2.9). To do this, we use the following assumptions in the asymptotic development. The assumptions are similar to those in Wang and Phillips (2009b), but Assumption 3 allows for higher order kernel functions $K(y)$.

Assumption 1. $x_t = \sum_{j=1}^{t} \xi_j$, where ξ_j is defined as in (2.1) with ϕ_k satisfying C2.

Assumption 2. $(\epsilon_i, \eta_i), i \geq 1$, is assumed to be a sequence of i.i.d. random vectors. The equation error $u_t = u(\eta_t, \eta_{t-1}, \ldots, \eta_{t-m_0+1})$ in (2.7) satisfies $\mathrm{E}u_t = 0$ and $\mathrm{E}u_t^4 < \infty$ for $t \geq m_0$, where $u(y_1, \ldots, y_{m_0})$ is a real measurable function on R^{m_0}. We define $u_t = 0$ for $1 \leq t \leq m_0 - 1$.

Assumption 3.

(a) $K(x)$ satisfies that $\int K(y)\,dy = 1$ and for some $p \geq 2$,

$$\int y^p K(y)\,dy \neq 0, \qquad \int y^i K(y)\,dy = 0, \qquad i = 1, 2, \ldots, p-1.$$

(b) $K(x)$ has compact support.

https://doi.org/10.1017/S0266466610000277 Published online by Cambridge University Press

Assumption 4. For given fixed x, $f(x)$ has a continuous $p+1$ derivative in a small neighborhood of x, where $p \geq 2$ is defined as in Assumption 3.

THEOREM 2.2. *Under Assumptions 1–4, we have*

$$\left(nh^2\right)^{1/4}\left[\hat{f}(x) - f(x) - \frac{h^p f^{(p)}(x)}{p!}\int_{-\infty}^{\infty} y^p K(y)dy\right] \to_D \sigma_u N L_W^{-1/2}(1,0), \tag{2.10}$$

provided $nh^2 \to \infty$ *and* $nh^{2+4(p+1)} \to 0$, *where* $\sigma_u^2 = |\phi|^{-1} \mathrm{E}u_{m0}^2 \int_{-\infty}^{\infty} K^2(s)\,ds$.

Remark 2.6. An important distinction between (2.10) and the limit theory for the case of stationary x_t is that the expression for the bias in (2.10) involves only a term that depends on $f^{(p)}(x)$. In particular, in the important case where $p = 2$, there is no linear term (involving $f'(x)$) in the bias expression. The reason for this simplification in the limit theory is that in the usual Taylor development for the bias, the linear term takes the form

$$I_a = h\, f'(x) \sum_{t=1}^{n} H_1\left(\frac{x_t - x}{h}\right), \tag{2.11}$$

in which $H_1(s) = sK(s)$ is a zero energy function. It follows from Theorem 2.1 that $I_a = O_p\left(n^{1/4}h^{3/2}\right)$ when x_t is unit root nonstationary and $d_n = \sqrt{n}$ as occurs under Assumption 1. On the other hand, the quadratic term in the Taylor development of the bias has the form

$$I_b = \frac{h^2}{2} f''(x) \sum_{t=1}^{n} H_2\left(\frac{x_t - x}{h}\right),$$

where $H_2(x) = x^2 K(x)$, which is $O_p\left(n^{1/2}h^3\right)$ from (2.5). Thus, I_a is dominated by I_b as $n \to \infty$ provided $nh^6 \to \infty$. On the other hand, when $nh^6 = O(1)$, both I_a and I_b do not affect the limit theory. Details are given in the proof of Theorem 2.2 in Section 4. By contrast, in the stationary case both I_a and I_b are $O\left(nh^2\right)$, and then both terms contribute to the bias in the limit theory.

Remark 2.7. The result (2.10) implies that

$$\hat{f}(x) - f(x) = O_P\{h^2 + a_n(nh^2)^{-1/4}\}, \tag{2.12}$$

where a_n diverges to infinity as slowly as required. This indicates that a possible "optimal" bandwidth h that yields the best rate in (2.12) or the minimal $\mathrm{E}(\hat{f}(x) - f(x))^2$ satisfies

$$h^* \sim a\,\mathrm{argmin}_h\left\{h^2 + \left(nh^2\right)^{-1/4}\right\} \sim a' n^{-1/10},$$

https://doi.org/10.1017/S0266466610000277 Published online by Cambridge University Press

where a and a' are positive constants. This result is different from that of nonparametric regression with a stationary regressor, which typically requires $h = o\left(n^{-1/5}\right)$ for undersmoothing. Investigation of an "optimal" bandwidth h for nonstationary cointegration regression therefore involves different criteria from that of stationary regression. We leave this topic for future work.

Remark 2.8. Interestingly, the fact that the linear term in the bias is eliminated in (2.10) means that in the nonstationary case the Nadaraya–Watson estimator $\hat{f}(x)$ defined by (2.8), under Assumptions 3 and 4 with $p = 2$, has the same limit distribution (to the second order including bias) as the local linear nonparametric estimator (e.g., Fan and Gijbels, 1996), defined by

$$\hat{f}^L(x) = \sum_{i=1}^{n} w_i Y_i \Big/ \sum_{i=1}^{n} w_i, \qquad w_i = K_h(x_i - x)\{V_{n,2} - (x_i - x)V_{n,1}\}, \tag{2.13}$$

where $V_{n,j} = \sum_1^n K_h(x_i - x)(x_i - x)^j$.

Indeed, we have the following theorem.

THEOREM 2.3. *Theorem 2.2 (with $p = 2$) still holds if we replace $\hat{f}(x)$ by $\hat{f}^L(x)$.*

Remark 2.9. The local linear nonparametric estimator is popular partly because of its bias reducing properties in comparison with the Nadaraya–Watson estimator $\hat{f}(x)$ defined by (2.8). The present finding shows that this particular advantage is lost when x_t is nonstationary. The other main advantage of the local linear smoother is the absence of boundary effects when the distribution of x_t has bounded support. However, in the present case, x_t is recurrent with unbounded support, and so this second advantage also does not apply.

Remark 2.10. As pointed out by a referee, a comparison between the Nadaraya–Watson estimator $\hat{f}(x)$ defined by (2.8) and the local polynomial estimator with order $p > 2$ would be interesting. The asymptotics related to the local polynomial estimator with order $p > 2$ require more precise results than those of (2.5). This kind of extra precision in the limit theory is not available at the present time. So we leave this topic for future work.

3. PROOF OF THEOREM 2.1

Section 3.1 provides some preliminary lemmas. Section 3.2 outlines the proof of Theorem 2.1. In fact, we provide the proof of the more general joint convergence result (2.5). Some useful propositions are given in Section 3.3. These propositions are interesting in their own right. Throughout the section we denote constants by $C, C_1, \ldots,$ which may differ at each appearance.

3.1. Preliminaries

Write $\varphi_i = \sum_{j=0}^{i} \phi_j$, $S_k = \sum_{i=0}^{k} \varphi_i \epsilon_i$, $\Lambda_k^2 = \sum_{i=0}^{k} \varphi_i^2$, and $f_k(t) = \mathrm{E}e^{itS_k/\Lambda_k}$. Recalling the properties of ϕ_j, together with (2.2), simple calculations show that

$$d_k^2/\Lambda_k^2 \sim \begin{cases} (1-\mu)\int_0^\infty x^{-\mu}(x+1)^{-\mu}dx, & \text{under C1,} \\ 1, & \text{under C2.} \end{cases} \tag{3.1}$$

Next, because $\mathrm{E}\epsilon_0 = 0$, $\mathrm{E}\epsilon_0^2 = 1$, and the characteristic function $\varphi(t)$ of ϵ_0 satisfies $\int_{-\infty}^{\infty} |\varphi(t)|dt < \infty$, it follows that, $\forall \epsilon > 0$, we may choose A sufficiently large such that

$$\int_{|t|\geq A} |f_k(t)|dt < \epsilon, \tag{3.2}$$

uniformly on k. See, e.g., the proof of Corollary 2.2 of Wang and Phillips (2009a). Result (3.2) implies the following fact.

F. *S_k/Λ_k has a density $\nu_k(x)$, and the $\nu_k(x)$ are uniformly bounded on k and x by a constant C.*

See, e.g., Lukács (1970, Thm. 3.2.2). Note that, for any $s < m$,

$$\begin{aligned} x_m &= \sum_{j=1}^{m} \sum_{i=-\infty}^{j} \epsilon_i \phi_{j-i} \\ &= x_s + \sum_{j=s+1}^{m} \sum_{i=-\infty}^{s} \epsilon_i \phi_{j-i} + \sum_{j=s+1}^{m} \sum_{i=s+1}^{j} \epsilon_i \phi_{j-i} \\ &:= x_{s,m}^* + x_{s,m}', \end{aligned} \tag{3.3}$$

where $x_{s,m}^*$ depends only on $(\ldots, \epsilon_{s-1}, \epsilon_s)$ and

$$x_{s,m}' = \sum_{j=1}^{m-s} \sum_{i=1}^{j} \epsilon_{i+s} \phi_{j-i} = \sum_{i=s+1}^{m} \epsilon_i \sum_{j=0}^{m-i} \phi_j =_d S_{m-s-1},$$

where $=_d$ denotes equivalence in distribution. By virtue of (3.3), results (3.1) and (3.2) also imply the following lemma.

LEMMA 3.1. *x_k/d_k has a density $g_k(x)$ in which the $g_k(x)$ are uniformly bounded over k and x by a constant C, and as $k \to \infty$,*

$$\sup_x |g_k(x) - n(x)| \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{g}_k(t) - e^{-t^2/2}|dt \to 0, \tag{3.4}$$

where $\hat{g}_k(t) = \mathrm{E}e^{itx_k/d_k}$ and $n(x) = e^{-x^2/2}/\sqrt{2\pi}$.

https://doi.org/10.1017/S0266466610000277 Published online by Cambridge University Press

Proof. By virtue of (3.1) and (3.2), it follows from (3.3) with $s=-1$ and the independence of ϵ_j that

$$\int_{-\infty}^{\infty} |\hat{g}_k(t)|\,dt \le \int_{-\infty}^{\infty} |\mathrm{E}e^{itS_k/d_k}|\,dt \le C \max_k \int_{-\infty}^{\infty} |f_k(t)|\,dt < \infty, \tag{3.5}$$

uniformly on k. This proves that x_k/d_k has a density $g_k(x)$, and the $g_k(x)$ are uniformly bounded on k and x by a constant C. As for (3.4), for any $\epsilon > 0$, by noting that we may choose A sufficiently large such that

$$\int_{|t|\ge A} |\hat{g}_k(t)|\,dt + \int_{|t|\ge A} e^{-t^2/2}dt \le C \int_{|t|\ge A} |f_k(t)|\,dt + \int_{|t|\ge A} e^{-t^2/2}dt < \epsilon,$$

uniformly on k because of (3.2), we have

$$\begin{aligned} 2\pi \sup_x |g_k(x) - n(x)| &\le \int_{-\infty}^{\infty} |\hat{g}_k(t) - e^{-t^2/2}|\,dt \\ &\le \int_{|t|\le A} |\hat{g}_k(t) - e^{-t^2/2}|\,dt + \int_{|t|\ge A} |\hat{g}_k(t)|\,dt + \int_{|t|\ge A} e^{-t^2/2}dt \le 2\epsilon, \end{aligned}$$

when $k \to \infty$, where we have used the fact that $\int_{|t|\le A} |f_k(t) - e^{-t^2/2}|\,dt \to 0$, for any $A > 0$, as $x_k/d_k \to_d N(0,1)$. This proves (3.4) and also completes the proof of Lemma 3.1. ■

To introduce the next two lemmas, let $r(x)$ be a real function such that $\int_{-\infty}^{\infty} |r(x)|\,dx < \infty$. Define

$$I_{k,l}^{(s)} = \mathrm{E}\left[r(x'_{s,k}/h)\, r(x'_{s,l}/h) \exp\left\{ i\mu \sum_{j=1}^{l} \epsilon_j/\sqrt{n} \right\} \right],$$

$$II_k^{(s)} = \mathrm{E}\left[r(x'_{s,k}/h) \exp\left\{ i\mu \sum_{j=1}^{k} \epsilon_j/\sqrt{n} \right\} \right],$$

where $x'_{s,k}$ is defined as in (3.3) and μ is a constant.

LEMMA 3.2.

(i) $\mathrm{E}|r(x'_{s,k}/h)| \le C\, h/d_{k-s}$ *and*

$$\mathrm{E}r(x_k/h) - \frac{h\, d_k^{-1}}{\sqrt{2\pi}} \int r(x)\,dx = o(h/d_k). \tag{3.6}$$

(ii) Suppose that $|\hat{r}(t)| \le C \min\{|t|, 1\}$ *and* $\int |\hat{r}(t)|\,dt < \infty$, *where* $\hat{r}(t) = \int e^{itx} r(x)\,dx$. *Then, for all* $l-k \ge 1$ *and all* $k \ge s+1$,

$$|II_k^{(s)}| \le C\, h \left[(k-s)^{-2} + h/d_{k-s}^2 \right], \tag{3.7}$$

$$|I_{k,l}^{(s)}| \le C\, h \left[(l-k)^{-2} + h/d_{l-k}^2 \right] \left[(k-s)^{-2} + h/d_{k-s} \right]. \tag{3.8}$$

Remark 3.1. The constant C in Lemma 3.2 depends on $r(x)$ through $\int |r(x)|\,dx$ and $\int |\hat{r}(x)|\,dx$. This implies that, if

$$\sup_{r(\cdot)\in\Omega}\left(\int |r(x)|\,dx + \int |\hat{r}(x)|\,dx\right) < \infty,$$

where Ω is a set of functionals, then Lemma 3.2 holds true uniformly on $r(\cdot)\in G$.

Proof. The first part of result (i) follows from fact F. It follows from Lemma 3.1 that

$$\begin{aligned}\left|\mathrm{E}r(x_k/h) - \frac{h\,d_k^{-1}}{\sqrt{2\pi}}\int r(x)\,dx\right| &\le h\,d_k^{-1}\int |r(x)|\left|g_k(xh/d_k) - 1/\sqrt{2\pi}\right|dx\\ &\le h\,d_k^{-1}\int |r(x)|\Big(\left|g_k(xh/d_k) - n(xh/d_k)\right|\\ &\qquad + \left|n(xh/d_k) - n(0)\right|\Big)dx = o(h/d_k),\end{aligned}$$

which gives the second part of result (i).

We next prove result (ii). We prove (3.8) with $s = 0$ because the proofs of (3.7) and (3.8) with $s \neq 0$ are the same and so the details are omitted. For convenience of notation, write $x_k'' = x_{0,k}'$ and $I_{k,l} = I_{k,l}^{(0)}$. As $\int |\hat{r}(t)|\,dt < \infty$, we have $r(x) = \frac{1}{2\pi}\int e^{-ixt}\hat{r}(t)\,dt$. This yields

$$\begin{aligned} I_{k,l} &= \mathrm{E}\left[r(x_k''/h)\,r(x_l''/h)\,\exp\left\{i\mu\sum_{j=1}^{l}\epsilon_j/\sqrt{n}\right\}\right]\\ &= \frac{1}{(2\pi)^2}\int\int \mathrm{E}\left\{e^{-itx_k''/h}\,e^{i\lambda x_l''/h}\,e^{i\mu\sum_{j=1}^{l}\epsilon_j/\sqrt{n}}\right\}\hat{r}(t)\,\overline{\hat{r}(\lambda)}\,dt\,d\lambda.\end{aligned}$$

Define $\sum_{j=k}^{l} = 0$ if $l < k$ and put $a_{s,q} = \sum_{j=0}^{s-q}\phi_j$. Without loss of generality, assume that $\phi_0 \neq 0$. Indeed, if $\phi_0 = 0$, we may use ϕ_1, etc. Because

$$x_l'' = \sum_{q=1}^{l}\epsilon_q\sum_{j=0}^{l-q}\phi_j = \left(\sum_{q=1}^{k} + \sum_{q=k+1}^{l-1}\right)\epsilon_q\,a_{l,q} + \epsilon_l\phi_0,$$

it follows from independence of the ϵ_k's that

$$|I_{k,l}| \le \frac{1}{(2\pi)^2}\int\left|\mathrm{E}e^{i\epsilon_l(\lambda\phi_0 + uh/\sqrt{n})/h}\right|\left|\mathrm{E}\left\{e^{iz^{(2)}/h}\right\}\right||\hat{r}(\lambda)|\,\Lambda(\lambda,k)\,d\lambda, \tag{3.9}$$

where $\Lambda(\lambda,k) = \int\left|\mathrm{E}\{e^{iz^{(1)}/h}\}\right||\hat{r}(t)|\,dt$,

$$z^{(1)} = \sum_{q=1}^{k}\epsilon_q\left(\lambda a_{l,q} - t\,a_{k,q} + uh/\sqrt{n}\right),$$

$$z^{(2)} = \sum_{q=k+1}^{l-1}\epsilon_q\left(\lambda\,a_{l,q} + uh/\sqrt{n}\right).$$

As n can be taken sufficiently large so that $u/\sqrt{n}$ is as small as required, we assume $u = 0$ in the following proof for convenience. We first show that, for all λ,

$$\Lambda(\lambda, k) \ \le \ C\left(k^{-2} + h/d_k\right). \tag{3.10}$$

To estimate (3.10), write Ω_1 (Ω_2, respectively) for the set of $1 \le q \le k/2$ such that $|\lambda\, a_{l,q} - t\, a_{k,q}| \ge h$ ($|\lambda\, a_{l,q} - t\, a_{k,q}| < h$, respectively), and

$$B_1 = \sum_{q \in \Omega_2} a_{k,q}^2, \qquad B_2 = \sum_{q \in \Omega_2} a_{l,q} a_{k,q} \quad \text{and} \quad B_3 = \sum_{q \in \Omega_2} a_{l,q}^2.$$

By noting that

$$\left|a_{s,q}\right| = \left|\sum_{j=1}^{s-q} \phi_j\right| \sim \begin{cases} C\,(s-q)^{1-u} \rho(s-q), & \text{under C1} \\ \phi, & \text{under C2,} \end{cases} \tag{3.11}$$

as $s - q$ sufficiently large, it is readily seen that

$$B_1 \ge \begin{cases} C\,k^{3-2u} \rho^2(k), & \text{under C1} \\ C\,k, & \text{under C2,} \end{cases}$$

whenever $\#(\Omega_1) \le \sqrt{k}$ and k is sufficiently large, where $\#(A)$ denotes the number of elements in A. On the other hand, there exist constants $\gamma_1 > 0$ and $\gamma_2 > 0$ such that

$$\left|\mathrm{E} e^{i\,\epsilon_1 t}\right| \le \begin{cases} e^{-\gamma_1} & \text{if } |t| \ge 1, \\ e^{-\gamma_2 t^2} & \text{if } |t| \le 1, \end{cases} \tag{3.12}$$

because $\mathrm{E}\epsilon_1 = 0$, $\mathrm{E}\epsilon_1^2 = 1$, and ϵ_1 has a density. See, e.g., Chapter 1 of Petrov (1995). Also note that

$$\sum_{q \in \Omega_2} (\lambda\, a_{l,q} - t a_{k,q})^2 = \lambda^2 B_3 - 2\lambda\, t\, B_2 + t^2 B_1 = B_1 (t - \lambda B_2/B_1)^2 + \lambda^2 (B_3 - B_2^2/B_1) \ge B_1 (t - \lambda B_2/B_1)^2,$$

because $B_2^2 \le B_1 B_3$, by Hölder's inequality. By virtue of these facts, it follows from the independence of ϵ_t that

$$\begin{aligned} \left|\mathrm{E} e^{i\, W^{(1)}/h}\right| &\le \prod_{q=1}^{k/2} \left|\mathrm{E} e^{i \epsilon_1 (\lambda\, a_{l,q} - t\, a_{k,q})}\right| \\ &\le \exp\left\{ -\gamma_1 \#(\Omega_{1n}) - \gamma_2\, h^{-2} \sum_{q \in \Omega_2} (\lambda\, a_{l,q} - t a_{k,q})^2 \right\} \\ &\le \exp\left\{ -\gamma_1 \#(\Omega_{1n}) - \gamma_2\, B_1\, h^{-2} (t - \lambda B_2/B_1)^2 \right\}, \end{aligned}$$

where $W^{(1)} = \sum_{q=1}^{k/2} \epsilon_q (\lambda a_{l,q} - t a_{k,q})$. This, together with the fact that $z^{(1)} = W^{(1)} + \sum_{q=k/2+1}^{k} \epsilon_q (\lambda a_{l,q} - t a_{k,q})$, yields that, for all λ,

$$\begin{aligned}\Lambda(\lambda, k) &\le \int \left| \mathrm{E}\{e^{iW^{(1)}/h}\}\right| |\hat{r}(t)|\,dt \\ &\le \int_{\#(\Omega_1)\ge\sqrt{k}} e^{-\gamma_1 \#(\Omega_1)} |\hat{r}(t)|\,dt + \int_{\#(\Omega_1)\le\sqrt{k}} e^{-\gamma_2 B_1 h^{-2}(t-\lambda B_2/B_1)^2}\,dt \\ &\le C k^{-2} \int |\hat{r}(t)|\,dt + \int e^{-\gamma_2 B_1 h^{-2} t^2}\,dt \\ &\le C\,(k^{-2} + h/d_k),\end{aligned}$$

as required.

We now turn back to the proof of (3.8) for $s = 0$. Recall that we may assume $u = 0$ for convenience as earlier. By virtue of (3.9) and (3.10), it suffices to show that

$$\begin{aligned}\tilde{I}_{k,l} &:= \int \left| \mathrm{E} e^{i\lambda\phi_0\epsilon_1/h} \right| \left| \mathrm{E}\left\{ e^{i\lambda \sum_{q=k+1}^{l-1} \epsilon_q a_{l,q}/h} \right\} \right| |\hat{r}(\lambda)|\,d\lambda \\ &\le C h \left[(l-k)^{-2} + h/d_{l-k}^2 \right], \end{aligned} \tag{3.13}$$

for $l - k \ge 1$. First notice that, for any $\delta > 0$, there exist constants $\gamma_3 > 0$, $\gamma_4 > 0$, and k_0 sufficiently large such that, for all $s \ge k_0$ and $q \le s/2$,

$$\left| \mathrm{E} e^{i\epsilon_1 \lambda a_{s,q}/h} \right| \le \begin{cases} e^{-\gamma_3 s^{1-u} \rho(s)}, & \text{if } |\lambda| \ge \delta h, \\ e^{-\gamma_4 s^{2(1-u)} \rho^2(s) \lambda^2/h^2}, & \text{if } |\lambda| \le \delta h, \end{cases}$$

under C1, and

$$\left| \mathrm{E} e^{i\epsilon_1 \lambda a_{s,q}/h} \right| \le \begin{cases} e^{-\gamma_3}, & \text{if } |\lambda| \ge \delta h, \\ e^{-\gamma_4 \lambda^2/h^2}, & \text{if } |\lambda| \le \delta h, \end{cases}$$

under C2. These facts follow from (3.11) and (3.12) with a simple calculation. Hence, because $\rho(.)$ is a slowly varying function, whenever $l - k \ge k_0$,

$$\begin{aligned}\left| \mathrm{E}\left\{ e^{i\lambda \sum_{q=k+1}^{l-1} \epsilon_q a_{l,q}/h} \right| &\le \Pi_{q=k}^{(l+k)/2} \left| \mathrm{E} e^{i\epsilon_q \lambda a_{l,q}/h} \right| \\ &\le \begin{cases} e^{-\gamma_3 (l-k)} & \text{if } |\lambda| \ge \delta h, \\ e^{-\gamma_4 d_{l-k}^2 \lambda^2/h^2} & \text{if } |\lambda| \le \delta h. \end{cases}\end{aligned} \tag{3.14}$$

Now, using $|\hat{r}(t)| \le C \min\{|t|, 1\}$, we obtain that, whenever $l - k \ge k_0$,

$$\begin{aligned}\tilde{I}_{k,l} &\le C e^{-\gamma_3 (l-k)^{2-u} h(l-k)} \int_{|\lambda|\ge\delta h} \left| \mathrm{E} e^{i\lambda\phi_0\epsilon_1/h} \right| d\lambda + C \int_{|\lambda|\le\delta h} |\lambda| e^{-\gamma_4 d_{l-k}^2 \lambda^2/h^2} d\lambda \\ &\le C \left[h(l-k)^{-2} + h^2/d_{l-k}^2 \right],\end{aligned}$$

https://doi.org/10.1017/S0266466610000277 Published online by Cambridge University Press

where we have used the fact that $\int |\mathrm{E}e^{i\lambda\epsilon_1}|d\lambda < \infty$. This gives (3.13) for $l-k \geq k_0$. The result (3.13) for $l-k \leq k_0$ is obvious, because, in this case,

$$\tilde{I}_{k,l} \leq C \int \left|\mathrm{E}e^{i\lambda\phi_0\epsilon_1/h}\right| d\lambda \leq C\, k_0^2\, h\, (l-k)^{-2}.$$

The proof of Lemma 3.2 is now complete. ■

3.2. Proof of (2.5)

First, it is convenient to introduce the following definitions and notation. If $\alpha_n^{(1)}$, $\alpha_n^{(2)}, \ldots, \alpha_n^{(k)}$ $(1 \leq n \leq \infty)$ are random elements of $D[0,1]$, we will understand the condition

$$(\alpha_n^{(1)}, \alpha_n^{(2)}, \ldots, \alpha_n^{(k)}) \to_D (\alpha_\infty^{(1)}, \alpha_\infty^{(2)}, \ldots, \alpha_\infty^{(k)})$$

to mean that for all $\alpha_\infty^{(1)}, \alpha_\infty^{(2)}, \ldots, \alpha_\infty^{(k)}$-continuity sets $A_1, A_2, \ldots, A_k$

$$P\left(\alpha_n^{(1)} \in A_1, \alpha_n^{(2)} \in A_2, \ldots, \alpha_n^{(k)} \in A_k\right) \to P\left(\alpha_\infty^{(1)} \in A_1, \alpha_\infty^{(2)} \in A_2, \ldots, \alpha_\infty^{(k)} \in A_k\right)$$

(see Billingsley, 1968, Thm. 3.1; Hall, 1977). The term $D[0,1]^k$ will be used to denote $D[0,1] \times \cdots \times D[0,1]$, the k-times coordinate product space of $D[0,1]$. We still use $\Rightarrow$ to denote weak convergence on $D[0,1]$.

To prove (2.5), we use the following lemma, whose proof is the same as in Wang and Phillips (2009b). Also see Borodin and Ibragimov (1995).

LEMMA 3.3. *Suppose that* $\{\mathcal{F}_t\}_{t\geq 0}$ *is an increasing sequence of* σ*-fields,* $q(t)$ *is a process that is* $\mathcal{F}_t$*-measurable for each t and continuous with probability one,* $\mathrm{E}q^2(t) < \infty$*, and* $q(0) = 0$*. Let* $\psi(t), t \geq 0$*, be a process that is nondecreasing and continuous with probability one and satisfies* $\psi(0) = 0$ *and* $\mathrm{E}\psi^2(t) < \infty$*. Let* $\xi_1, \ldots, \xi_m$ *be random variables that are* $\mathcal{F}_t$*-measurable for each* $t \geq 0$*. If, for any* $\gamma_j \geq 0$*,* $j = 1, 2, \ldots, r$*, and any* $0 \leq s < t \leq t_0 < t_1 < \cdots < t_r < \infty$*,*

$$\mathrm{E}\left(e^{-\sum_{j=1}^r \gamma_j[\psi(t_j)-\psi(t_{j-1})]}\left[q(t)-q(s)\right] \mid \mathcal{F}_s\right) = 0, \quad a.s.,$$

$$\mathrm{E}\left(e^{-\sum_{j=1}^r \gamma_j[\psi(t_j)-\psi(t_{j-1})]}\left\{[q(t)-q(s)]^2 - [\psi(t)-\psi(s)]\right\} \mid \mathcal{F}_s\right) = 0, \quad a.s.$$

then the finite-dimensional distributions of the process $(q(t), \xi_1, \ldots, \xi_m)_{t\geq 0}$ *coincide with those of the process* $(W[\psi(t)], \xi_1, \ldots, \xi_m)_{t\geq 0}$*, where* $W(s)$ *is a standard Brownian motion with* $\mathrm{E}W^2(s) = s$ *independent of* $\psi(t)$*.*

By virtue of Lemma 3.3, we now obtain the proof of (2.5). Technical details of some subsidiary results that are used in this proof are given in the next section. Set

$$\zeta_n(t,l) = \frac{1}{\sqrt{n}} \sum_{k=-[nl]}^{[nt]} \epsilon_k, \qquad \psi_{1n}(t) = \frac{d_n}{nh} \sum_{k=1}^{[nt]} f_1\big(x_k/h\big),$$

$$\psi_{2n}(t) = \frac{d_n}{nh} \sum_{k=1}^{[nt]} f_2\big(x_k/h\big),$$

$$\eta_n(t) = \left(\frac{d_n}{nh}\right)^{1/2} \sum_{k=1}^{[nt]} g(x_k/h), \qquad \psi_n(t) = \frac{d_n}{nh} \sum_{k=1}^{[nt]} g^2(x_k/h),$$

for $0 \le t \le 1$ and $0 \le l < \infty$.

We will prove in Propositions 3.1 and 3.2 that $\zeta_n(t,l) \Rightarrow \zeta(t,l)$, for each $0 \le l < \infty$, where $\zeta(t,l) = W(t) - W(-l)$, $\psi_n(t) \Rightarrow \tau^2 \psi(t)$, and $\psi_{jn}(t) \Rightarrow \tau_j \psi(t)$, $j = 1,2$, on $D[0,1]$. Furthermore we will prove in Proposition 3.4 that $\{\eta_n(t)\}_{n\ge 1}$ is tight on $D[0,1]$. These facts imply that, for any $0 \le l_0 < l_1 < \cdots < l_{r'} < \infty$,

$$\{\eta_n(t),\ \psi_n(t),\ \psi_{1n}(t),\ \psi_{2n}(t),\ \zeta_n(t,l_0),\ldots,\zeta_n(t,l_{r'})\}_{n\ge 1}$$

is tight on $D[0,1]^{r'+4}$. Hence, by Prohorov's theorem (see Billingsley, 1968, Sect. 6), for each $\{n'\} \subseteq \{n\}$, there exists a subsequence $\{n''\} \subseteq \{n'\}$ such that

$$\big\{\eta_{n''}(t), \psi_{n''}(t), \psi_{1n''}(1), \psi_{2n''}(1), \zeta_{n''}(t,l_0),\ldots,\zeta_{n''}(t,l_{r'})\big\}$$
$$\to_d \big\{\eta(t), \tau^2\psi(t), \tau_1\psi(1), \tau_2\psi(1), \zeta(t,l_0),\ldots,\zeta(t,l_{r'})\big\} \qquad \textbf{(3.15)}$$

on $D[0,1]^{r'+4}$, where $\eta(t)$ is a process continuous with probability one by noting (3.28) later in this section. Write $\mathcal{F}_s = \sigma\{\zeta(t,l), 0 \le t \le 1, 0 \le l < \infty; \eta(t), 0 \le t \le s\}$. It is readily seen that $\mathcal{F}_s \uparrow$ and $\eta(s)$ is $\mathcal{F}_s$-measurable for each $0 \le s \le 1$. Also note that $\psi(t)$ (for any fixed $t \in [0,1]$) is $\mathcal{F}_s$-measurable for each $0 \le s \le 1$. If we prove that for any $0 \le s < t \le 1$,

$$\mathrm{E}\Big(\big[\eta(t) - \eta(s)\big] \mid \mathcal{F}_s\Big) = 0, \quad a.s., \qquad \textbf{(3.16)}$$

$$\mathrm{E}\Big(\big\{[\eta(t) - \eta(s)]^2 - [\psi(t) - \psi(s)]\big\} \mid \mathcal{F}_s\Big) = 0, \quad a.s., \qquad \textbf{(3.17)}$$

then it follows from Lemma 3.3 that the finite-dimensional distributions of $\big\{\eta(t), \tau_1\psi(1), \tau_2\psi(1)\big\}$ coincide with those of $\big\{\tau N \psi^{1/2}(t), \tau_1\psi(1), \tau_2\psi(1)\big\}$, where N is a normal variate independent of $\psi^{1/2}(t)$. The result (2.5) therefore follows because $\big\{\eta(t), \tau_1\psi(1), \tau_2\psi(1)\big\}$ does not depend on the choice of the subsequence. See, e.g., Theorem 2.3 of Billingsley (1968).

Let $0 = t_0 < t_1 < \cdots < t_r = 1$ and $0 = l_0 < l_1, \ldots, l_{r'} < \infty$, where r and r' are arbitrary integers, and $G(\bullet)$ be an arbitrary bounded measurable function on $R^{j+(r+1)(r'+1)}$. To prove (3.16) and (3.17), it suffices to show that

$$\mathrm{E}[\eta(t_j) - \eta(t_{j-1})]\, G(\cdots) = 0, \qquad \textbf{(3.18)}$$

$$\mathrm{E}\big\{[\eta(t_j) - \eta(t_{j-1})]^2 - [\psi(t_j) - \psi(t_{j-1})]\big\}\, G(\cdots) = 0, \qquad \textbf{(3.19)}$$

where $G(\cdots) = G[\eta(t_0),\ldots,\eta(t_{j-1});\zeta(t_0,l_0),\ldots,\zeta(t_0,l_{r'});\ldots;\zeta(t_r,l_0),\ldots,\zeta(t_r,l_{r'})]$.

The result (3.15) (for convenience of notation, we assume that the sequence $\{n''\}$ in (3.15) is just $\{n\}$ itself), together with the uniform integrability of $\eta_n(t)$, $\eta_n^2(t)$, and $\psi_n(t)$ for each $0 \le t \le 1$ (see Proposition 3.3), implies that the statements (3.18) and (3.19) will follow if we prove

$$\mathrm{E}[\eta_n(t_j)-\eta_n(t_{j-1})]\,G_n[\cdots]\to 0, \tag{3.20}$$

$$\mathrm{E}\big\{[\eta_n(t_j)-\eta_n(t_{j-1})]^2-[\psi_n(t_j)-\psi_n(t_{j-1})]\big\}\,G_n[\cdots]\to 0, \tag{3.21}$$

where $G_n[\cdots] = G[\eta_n(t_0),\ldots,\eta_n(t_{j-1});\zeta_n(t_0,l_0),\ldots,\zeta_n(t_0,l_{r'});\ldots;\zeta_n(t_r,l_0),\ldots,\zeta_n(t_r,l_{r'})]$ (see, e.g., Billingsley, 1968, Thm. 5.4).

Note that, by using similar arguments to those in the proofs of Lemmas 5.4 and 5.5 in Borodin and Ibragimov (1995), we may choose

$$G(\bullet) = \exp\left\{i\left(\sum_{k=0}^{j-1}\lambda_k y_k + \sum_{k=0}^{r}\sum_{s=0}^{r'}\mu_{ks}z_{ks}\right)\right\},$$

and simple calculations show that

$$\begin{aligned}
&\sum_{k=0}^{j-1}\lambda_k\eta_n(t_k)+\sum_{k=0}^{r}\sum_{s=0}^{r'}\mu_{ks}\zeta_n(t_k,l_s)\\
&\quad=\sum_{k=0}^{j-1}\lambda_k\eta_n(t_k)+\sum_{k=0}^{j-1}\sum_{s=0}^{r'}\mu_{ks}\zeta_n(t_k,l_s)\\
&\qquad+\frac{1}{\sqrt{n}}\sum_{k=j}^{r}\sum_{s=0}^{r'}\mu_{ks}\left(\sum_{u=-[nl_s]}^{[nt_{j-1}]}\epsilon_u+\sum_{u=[nt_{j-1}]+1}^{[nt_k]}\epsilon_u\right)\\
&\quad=\chi(t_{j-1})+\frac{1}{\sqrt{n}}\mu_j^*\sum_{u=[nt_{j-1}]+1}^{[nt_j]}\epsilon_u+\frac{1}{\sqrt{n}}\sum_{k=j+1}^{r}\sum_{s=0}^{r'}\mu_{ks}\sum_{u=[nt_j]+1}^{[nt_k]}\epsilon_u,
\end{aligned}$$

where $\chi(s)=\chi(\ldots,\epsilon_{s-1},\epsilon_s)$, a functional of $\ldots,\epsilon_{s-1},\epsilon_s$, and $\mu_j^*=\sum_{k=j}^{r}\sum_{s=0}^{r'}\mu_{ks}$. By independence of ϵ_k, we now only need to show that

$$\mathrm{E}\left\{\sum_{k=[nt_{j-1}]+1}^{[nt_j]} g(x_k/h)\,e^{i\mu_j^*(1/\sqrt{n})\sum_{k=[nt_{j-1}]+1}^{[nt_j]}\epsilon_k+i\chi(t_{j-1})}\right\}
= o\left[\left(\frac{nh}{d_n}\right)^{1/2}\right], \tag{3.22}$$

$$\mathrm{E}\left\{\left[\sum_{k=[nt_{j-1}]+1}^{[nt_j]} g(x_k/h)\right]^2-\sum_{k=[nt_{j-1}]+1}^{[nt_j]} g^2(x_k/h)\right\}e^{i\mu_j^*(1/\sqrt{n})\sum_{k=[nt_{j-1}]+1}^{[nt_j]}\epsilon_k+i\chi(t_{j-1})}
= o\left(\frac{nh}{d_n}\right). \tag{3.23}$$

https://doi.org/10.1017/S0266466610000277 Published online by Cambridge University Press

Furthermore, by independence of ϵ_k again and conditioning arguments, it suffices to show that, for any μ,

$$\sup_{y,0\le s<m\le n} \mathrm{E}\left\{\sum_{k=s+1}^{m} g(y+x'_{s,k}/h)\, e^{i\mu\sum_{i=1}^{m}\epsilon_i/\sqrt{n}}\right\} = o\left[\left(\frac{nh}{d_n}\right)^{1/2}\right], \tag{3.24}$$

$$\sup_{y,0\le s<m\le n} \mathrm{E}\left(\left\{\sum_{k=s+1}^{m} g(y+x'_{s,k}/h)\right\}^2 - \sum_{k=s+1}^{m} g^2(y+x'_{s,k}/h)\right) e^{i\mu\sum_{i=1}^{m}\epsilon_i/\sqrt{n}} = o\left(\frac{nh}{d_n}\right), \tag{3.25}$$

where $x'_{s,k}$ is defined as in (3.3). This follows from Proposition 3.5. The proof of Theorem 2.1 is now complete.

3.3. Some Useful Propositions

In this section we will prove the following propositions required in the proof of theorem 2.1. Our notation will be the same as in the previous sections except when explicitly mentioned.

PROPOSITION 3.1. *We have, for each* $0 \le l < \infty$,

$$\zeta_n(t,l) \Rightarrow \zeta(t,l) \quad \text{and} \quad \zeta'_n(t) := \frac{1}{d_n}\sum_{k=1}^{[nt]} x_k \Rightarrow \widetilde{W}(t) \quad \text{on } D[0,1], \tag{3.26}$$

where $\widetilde{W}(t) = W_{3/2-u}(t)$ *under C1 and* $\widetilde{W}(t) = W(t)$ *under C2.*

Proof. The first result of (3.26) is well known. The second result in (3.26) can be found in, e.g., Wang et al. (2003). ■

PROPOSITION 3.2. *For any* $h \to 0$ *and* $nh/d_n \to \infty$, *we have*

$$\psi_n(t) \Rightarrow \tau^2\,\psi(t) \quad \text{on } D[0,1]. \tag{3.27}$$

Similarly, we also have

$$\psi_{1n}(t) \Rightarrow \tau_1\,\psi(t), \qquad \psi_{2n}(t) \Rightarrow \tau_2\,\psi(t) \quad \text{on } D[0,1].$$

Proof. We only prove (3.27). It suffices to show that

(a) the finite-dimensional distributions of $\psi_n(t)$ converge to those of $\tau^2\,\psi(t)$;
(b) $\{\psi_n(t)\}_{n\ge 1}$ is tight on $D[0,1]$.

https://doi.org/10.1017/S0266466610000277 Published online by Cambridge University Press

Statement (a) has been established in Jeganathan (2004) (also see Wang and Phillips, 2009a). We will use Theorem 4 of Billingsley (1974) to establish statement (b). According to this theorem, we only need to show that

$$\max_{1\le k\le n} g^2(x_k/h) = o_P(nh/d_n), \tag{3.28}$$

and there exists a sequence of $\alpha_n(\epsilon,\delta)$ satisfying $\lim_{\delta\to 0}\limsup_{n\to\infty}\alpha_n(\epsilon,\delta)=0$ for each $\epsilon>0$ such that, for

$$0\le t_1\le t_2\le\cdots\le t_m\le t\le 1, \qquad t-t_m\le\delta,$$

we have

$$P\left[|\psi_n(t)-\psi_n(t_m)|\ge\epsilon \mid \psi_n(t_1),\psi_n(t_2),\ldots,\psi_n(t_m)\right]\le\alpha_n(\epsilon,\delta), \quad a.s. \tag{3.29}$$

To prove (3.28), by noting that, for all $\epsilon>0$,

$$P\left(\max_{1\le k\le n} g^2(x_k/h)\ge\epsilon^2 nh/d_n\right)$$
$$= P\left(\sum_{k=1}^n g^2(x_k/h) I_{g^2(x_k/h)\ge\epsilon^2 nh/d_n}\ge\epsilon^2 nh/d_n\right),$$

it suffices to show that, for all $\epsilon>0$,

$$J\equiv\frac{d_n}{nh}\sum_{k=1}^n \mathrm{E}g^2(x_k/h) I_{(g^2(x_k/h)\ge\epsilon^2 nh/d_n)}=o(1). \tag{3.30}$$

In fact, by recalling that x_k/d_k has a uniformly bounded density $g_k(x)$ by Lemma 3.1, we have

$$J=\frac{d_n}{nh}\sum_{k=1}^n\int g^2(xd_k/h) I_{(g^2(xd_k/h)\ge\epsilon^2 nh/d_n)}\, g_k(x)\,dx$$

$$\le C\,\frac{d_n}{n}\sum_{k=1}^n\frac{1}{d_k}\int g^2(x) I_{g^2(x)\ge\epsilon^2 nh/d_n}\,dx=o(1),$$

where we have used the fact that $\frac{d_n}{n}\sum_{k=1}^n\frac{1}{d_k}=O(1)$ and $\int g^2(x) I_{g^2(x)\ge\epsilon^2 nh/d_n}$ $dx=o(1)$.[1]

We next prove (3.29). It follows from the independence of ϵ_k and (3.3) that

$$\sup_{|t-s|\le\delta} P\left(\left|\sum_{k=[ns]+1}^{[nt]} g^2(x_k/h)\right|\ge\epsilon\, nh/d_n \mid \epsilon_{[ns]},\epsilon_{[ns]-1},\ldots\right)\le\alpha_n(\epsilon,\delta), \tag{3.31}$$

where

$$\alpha_n(\epsilon,\delta)=\epsilon^{-1}\left[d_n/(nh)\right]\sup_{y,0\le t\le\delta}\mathrm{E}\sum_{k=1}^{[nt]} g^2[(y+x'_{0,k})/h].$$

https://doi.org/10.1017/S0266466610000277 Published online by Cambridge University Press

The result (3.29) will follow if we prove $\lim_{\delta\to 0}\limsup_{n\to\infty}\alpha_n(\epsilon,\delta)=0$ for each $\epsilon>0$. In fact, by letting $r(x)=g^2(y/h+x)$, we have $\int r(x)\,dx=\int g^2(x)\,dx<\infty$ uniformly on $y\in R$ and h. Hence it follows from part (i) of Lemma 3.2 (also recall Remark 3.1) that, for all $\epsilon>0$,

$$\alpha_n(\epsilon,\delta)\le C\,\epsilon^{-1}\,\frac{d_n}{n}\sum_{k=1}^{[n\delta]}d_k^{-1}\to 0, \tag{3.32}$$

first $n\to\infty$ and then $\delta\to 0$, as required. The proof of Proposition 3.2 is complete. ■

PROPOSITION 3.3. *For any fixed* $0\le t\le 1$, $\eta_n(t)$, $\eta_n^2(t)$, *and* $\psi_n(t)$, $n\ge 1$, *are uniformly integrable.*

Proof. We first claim that, for each fixed t,

$$\mathrm{E}\psi_n(t)\to\tau^2\,\mathrm{E}\psi(t),\quad \text{as } n\to\infty. \tag{3.33}$$

In fact it follows from (3.6) with $r(x)=g^2(x)$ that, for each fixed t,

$$\begin{aligned}\mathrm{E}\psi_n(t)&=\frac{d_n}{nh}\sum_{k=1}^{[nt]}\mathrm{E}g^2(x_k/h)\ \sim\ \frac{\tau^2}{\sqrt{2\pi}}\,\frac{d_n}{n}\sum_{k=1}^{[nt]}d_k^{-1}\\ &\sim\frac{\tau^2}{\sqrt{2\pi}}\begin{cases}\frac{1}{u-1/2}\,t^{u-1/2}, & \text{under C1,}\\ \frac{1}{2}\,t^{1/2}, & \text{under C2,}\end{cases}\\ &=\tau^2\,\mathrm{E}\psi(t).\end{aligned}$$

By virtue of (3.33), together with Proposition 3.2 and the fact that $\psi_k(t)$ is positive, it follows from Theorem 5.4 of Billingsley (1968) that $\psi_k(t)$ are uniformly integrable for each fixed t.

To prove the uniform integrability of $\eta_n^2(t)$ for each fixed t, we first show that

$$\sup_{0\le t\le 1}\mathrm{E}|\psi_n(t)-\eta_n^2(t)|=o(1). \tag{3.34}$$

To prove (3.34), let $r(x)=g(y/h+x)$ and $\hat r(t)=\int e^{itx}r(x)\,dx$. It is readily seen that $\hat r(t)=\int e^{itx}g(y/h+x)\,dx=e^{-ity/h}\hat g(t)$ and $\int|r(x)|\,dx=\int|g(x)|\,dx<\infty$. Furthermore, $\int|\hat r(\lambda)|\,d\lambda\le\int|\hat g(\lambda)|\,d\lambda<\infty$ and

$$|\hat r(t)|\le|\hat g(t)|\le C\min\{|t|,1\}.$$

That is, the conditions on $r(t)$ in part (ii) of Lemma 3.2 hold true uniformly for all $y\in R$ and h. It now follows from (3.8) (recall Remark 3.1) with $u=0$ and $s=0$ that, for all $l-k\ge 1$,

$$\begin{aligned}&\sup_y\left|\mathrm{E}\left\{g\left[(y+x'_{0,k})/h\right]g\left[(y+x'_{0,l})/h\right]\right\}\right|\\ &\quad\le\ C\,h\left[(l-k)^{-2}+h/d_{l-k}^2\right]\left(k^{-2}+h/d_k\right).\end{aligned} \tag{3.35}$$

Hence, by noting that

$$\mathrm{E}|\eta_n^2(t) - \psi_n(t)| \le \frac{2d_n}{nh} \sum_{1\le k<l\le[nt]} |\mathrm{E}\big[g(x_k/h)\, g(x_l/h)\big]|,$$

and recalling (3.3), we obtain that

$$\begin{aligned}\sup_{0\le t\le 1} \mathrm{E}|\psi_n(t) - \eta_n^2(t)| &\le \frac{d_n}{nh} \sum_{1\le k<l\le n} \sup_y \big|\mathrm{E}\big\{g\big[(y+x'_{0,k})/h\big]\, g\big[(y+x'_{0,l})/h\big]\big\}\big| \\ &\le \frac{d_n}{n}\left(C + h\sum_{k=1}^n d_k^{-2}\right)\left(C + h\sum_{k=1}^n d_k^{-1}\right) \\ &\le C \begin{cases} h, & \text{under C1,} \\ h + h^2\log n, & \text{under C2,}\end{cases}\end{aligned}$$

which yields (3.34), because $h \to 0$ ($h^2 \log n \to 0$ under C2).

By virtue of (3.34), for any $A > 0$ and fixed t, we have

$$|\mathrm{E}\eta_n^2(t) I_{\eta_n^2(t)\ge A} - \mathrm{E}\psi_n(t) I_{\eta_n^2(t)\ge A}| \le \sup_{0\le t\le 1} \mathrm{E}|\psi_n(t) - \eta_n^2(t)| = o(1).$$

This, together with the fact that

$$\begin{aligned}\mathrm{E}\psi_n(t) I_{\eta_n^2(t)\ge A} &\le \mathrm{E}\psi_n(t) I_{\psi_n(t)\ge\sqrt{A}} + \sqrt{A}\, P(\eta_n^2(t) \ge A) \\ &\le \mathrm{E}\psi_n(t) I_{\psi_n(t)\ge\sqrt{A}} + A^{-1/2}\mathrm{E}\psi_n(t) + o(1),\end{aligned}$$

implies that

$$\lim_{A\to\infty} \sup_n \mathrm{E}\eta_n^2(t) I_{\eta_n^2(t)\ge A} \le \lim_{A\to\infty} \sup_n \big[\mathrm{E}\psi_n(t) I_{\psi_n(t)\ge\sqrt{A}} + A^{-1/2}\mathrm{E}\psi_n(t)\big] = 0,$$

where we have used the uniform integrability of $\psi_n(t)$. That is, $\eta_n^2(t)$ is uniformly integrable. The integrability of $\eta_n(t)$ follows from that of $\eta_n^2(t)$. The proof of Proposition 3.3 is now complete. ■

PROPOSITION 3.4. *$\{\eta_n(t)\}_{n\ge1}$ is tight on $D[0,1]$.*

Proof. As in the proof of Proposition 3.2, we will use Theorem 4 of Billingsley (1974) to establish the tightness of $\eta_n(t)$ on $D[0,1]$. According to the theorem, we only need to show that

$$\max_{1\le k\le n} |g(x_k/h)| = o_P[(d_n/nh)^{1/2}], \qquad \textbf{(3.36)}$$

and there exists a sequence of $\alpha'_n(\epsilon,\delta)$ satisfying $\lim_{\delta\to0}\limsup_{n\to\infty}\alpha'_n(\epsilon,\delta) = 0$ for each $\epsilon > 0$ such that, for

$$0 \le t_1 \le t_2 \le \cdots \le t_m \le t \le 1, \quad t - t_m \le \delta,$$

https://doi.org/10.1017/S0266466610000277 Published online by Cambridge University Press

we have

$$P\left[|\eta_n(t)-\eta_n(t_m)| \geq \epsilon \mid \eta_n(t_1), \eta_n(t_2), \ldots, \eta_n(t_m)\right] \leq \alpha_n'(\epsilon,\delta), \quad a.s. \tag{3.37}$$

The result (3.36) has been proved in (3.28). To prove (3.37), we choose

$$\alpha_n'(\epsilon,\delta) = \epsilon^{-2}\frac{d_n}{nh} \sup_{y, 0\leq t\leq \delta} \mathrm{E}\left\{\sum_{k=1}^{[nt]} g[(y+x_{0,k}')/h]\right\}^2.$$

It follows from (3.32) and (3.35) that

$$\begin{aligned}\alpha_n'(\epsilon,\delta) &\leq \epsilon^{-1}\alpha_n(\epsilon,\delta) + 2\epsilon^{-2}\frac{d_n}{nh}\sup_y \sum_{1\leq k<l\leq [n\delta]} \\ &\quad \times \left|\mathrm{E}\{g[(y+x_{0,k}')/h]g[(y+x_{0,l}')/h]\}\right| \\ &= \epsilon^{-1}\alpha_n(\epsilon,\delta) + 2\epsilon^{-2}\frac{d_n}{n}\left(C+h\sum_{k=1}^{[n\delta]} d_k^{-2}\right)\left(C+h\sum_{k=1}^{[n\delta]} d_k^{-1}\right) \\ &\leq \epsilon^{-1}\alpha_n(\epsilon,\delta) + C\epsilon^{-2}\delta \begin{cases} h, & \text{under C1,} \\ h+h^2\log n, & \text{under C2,} \end{cases} \\ &\to 0,\end{aligned}$$

first $n\to\infty$ and then $\delta\to 0$, as $h\to 0$ ($h^2\log n\to 0$ under C2). Now, by noting that

$$\begin{aligned}&\sup_{|t-s|\leq\delta} P\left(\left|\sum_{k=[ns]+1}^{[nt]} g(x_k/h)\right| \geq \epsilon\,(d_n/nh)^{1/2} \mid \epsilon_{[ns]}, \epsilon_{[ns]-1}, \ldots; \eta_{[ns]}, \ldots, \eta_1\right) \\ &\quad \leq \alpha_n'(\epsilon,\delta),\end{aligned}$$

by using Markov's inequality and the independence of ϵ_k, we obtain the required (3.37). The proof of Proposition 3.4 is complete. ■

PROPOSITION 3.5. *Results (3.24) and (3.25) hold true for any constant $u \in R$.*

Proof. Let $r(t) = g(y/h+t)$. It has been proved in Proposition 3.3 that $r(x)$ satisfies the conditions required in part (ii) of Lemma 3.2 (also recall Remark 3.1), uniformly on y and h. Hence it follows from (3.8) that, uniformly on y,

$$\begin{aligned}\sum_{1\leq k<l\leq n} |I_{k,l}^s| &\leq C\sum_{1\leq k<l\leq n}\left[h\,(l-k)^{-2} + h^2/d_{l-k}^2\right]\left(k^{-2}+h/d_k\right) \\ &\leq C\,(1+nh/d_n)\begin{cases} h+h^2, & \text{under C1,} \\ h+h^2\log n, & \text{under C2.}\end{cases}\end{aligned}$$

This implies (3.25) because $h\to 0$ ($h^2\log n\to 0$ under C2) and $nh/d_n\to\infty$. The proof of (3.24) is similar, and the details are omitted. ■

https://doi.org/10.1017/S0266466610000277 Published online by Cambridge University Press

4. PROOF OF THEOREM 2.2

We may write

$$\hat{f}(x) - f(x) = \frac{\sum_{t=1}^{n} \{f(x_t) - f(x)\} K\left(\frac{x_t - x}{h}\right)}{\sum_{t=1}^{n} K\left(\frac{x_t - x}{h}\right)} + \frac{\sum_{t=1}^{n} u_t K\left(\frac{x_t - x}{h}\right)}{\sum_{t=1}^{n} K\left(\frac{x_t - x}{h}\right)}$$
$$= \Lambda_{1n} + \Lambda_{2n}, \quad \text{say.} \tag{4.1}$$

It is readily seen that Assumptions 1–4 match those used in Theorem 3.1 of Wang and Phillips (2009b) except Assumption 2. The current Assumption 2 seems to be more natural and clearly does not affect the result and the proof of Theorem 3.1 in Wang and Phillips (2009b). It follows from (3.8) of Wang and Phillips (2009b) that

$$(nh^2)^{1/4} \Lambda_{2n} \to_D \sigma_u N L_W^{-1/2}(1,0). \tag{4.2}$$

We next consider Λ_{1n}. The numerator of Λ_{1n} involves

$$\sum_{t=1}^{n} \{f(x_t) - f(x)\} K\left(\frac{x_t - x}{h}\right) = \sum_{j=1}^{p+1} I_j, \tag{4.3}$$

where

$$I_j = \frac{f^{(j)}(x)}{j!} \sum_{t=1}^{n} (x_t - x)^j \; K\left(\frac{x_t - x}{h}\right), \qquad j = 1, 2, \ldots, p,$$

$$I_{p+1} = \sum_{t=1}^{n} \left\{ f(x_t) - \sum_{j=0}^{p} \frac{f^{(j)}(x)}{j!} (x_t - x)^j \right\} K\left(\frac{x_t - x}{h}\right).$$

Write $H_j(x) = x^j K(x)$, $j = 1, 2, \ldots, p$. Recall that $K(x)$ has a compact support (Ω, say), $\int K(x)\,dx = 1$, $\int H_j(x)\,dx = 0$, $j = 1, \ldots, p-1$, and $\int H_p(x)\,dx \neq 0$ by Assumption 3. It is readily seen from (2.5) that

$$\frac{h^{-j}(nh^2)^{1/4} I_j}{\sum_{t=1}^{n} K\left(\frac{x_t - x}{h}\right)} \to_D \sigma_j N L_W^{-1/2}(1,0), \qquad j = 1, 2, \ldots, p-1, \tag{4.4}$$

where $\sigma_j^2 = \left[\frac{f^{(j)}(x)}{j!}\right]^2 |\phi|^{-1} \int H_j^2(x)\,dx$, and

$$\frac{h^{-p} I_p}{\sum_{t=1}^{n} K\left(\frac{x_t - x}{h}\right)} \to_P \frac{f^{(p)}(x)}{p!} \int H_p(x)\,dx. \tag{4.5}$$

On the other hand, by noting that $\lim_{h \to 0} \sup_{y \in \Omega} |f^{(p+1)}(yh + x)| \leq C$ by Assumption 4, Taylor expansion yields

$$|I_{p+1}| \leq C \sum_{t=1}^{n} |x_t - x|^{p+1} K\left(\frac{x_t - x}{h}\right),$$

and hence

$$\frac{h^{-(p+1)}\,|I_{p+1}|}{\sum_{t=1}^n K\left(\frac{x_t-x}{h}\right)} \le C\,\frac{\sum_{t=1}^n H_{p+1}\left(\frac{x_t-x}{h}\right)}{\sum_{t=1}^n K\left(\frac{x_t-x}{h}\right)} \to_P C\int H_{p+1}(x)\,dx, \tag{4.6}$$

where $H_{p+1}(x) = |x|^{p+1}K(x)$.

Combining (4.3)–(4.6), simple calculations show that

$$\begin{aligned}
&(nh^2)^{1/4}\left[\Lambda_{1n} - \frac{h^p f^{(p)}(x)}{p!}\int_{-\infty}^{\infty} y^p K(y)\,dy\right] \\
&\quad\le \frac{(nh^2)^{1/4}}{\sum_{t=1}^n K\left(\frac{x_t-x}{h}\right)}\sum_{j=1}^{p-1}|I_j| + \frac{(nh^2)^{1/4}|I_{p+1}|}{\sum_{t=1}^n K\left(\frac{x_t-x}{h}\right)} \\
&\quad= O_P\left[h^{1/2} + (nh^2)^{1/4}h^{p+1}\right] = o_P(1),
\end{aligned}$$

whenever $nh^2 \to \infty$ and $nh^{2+4(p+1)} \to 0$. This, together with (4.1) and (4.2), yields (2.10). The proof of Theorem 2.2 is now complete. ■

5. PROOF OF THEOREM 2.3

We may write

$$\begin{aligned}
\hat f^L(x) &= \frac{V_{n,2}\sum_{i=1}^n K_h(x_i-x)Y_i - V_{n,1}\sum_{i=1}^n K_h(x_i-x)(x_i-x)Y_i}{V_{n,2}\sum_{i=1}^n K_h(x_i-x) - V_{n,1}^2} \\
&= \frac{\sum_{i=1}^n K[(x_i-x)/h]}{\sum_{i=1}^n K[(x_i-x)/h] - h\,V_{n,1}^2/V_{n,2}}\,\hat f(x) \\
&\quad - \frac{(h\,V_{n,1}/V_{n,2})\sum_{i=1}^n H_1[(x_i-x)/h]\,Y_i}{\sum_{i=1}^n K[(x_i-x)/h] - h\,V_{n,1}^2/V_{n,2}},
\end{aligned}$$

where $H_1(x) = xK(x)$. As in the proof of Theorem 2.2, it follows easily from (2.5) that

$$h\,\frac{V_{n,1}^2}{V_{n,2}} \to_D C_0 N, \qquad h\,(\sqrt{n}h)^{1/2}\,\frac{V_{n,1}}{V_{n,2}} \to_D C_1 N\,L_W^{-1/2}(1,0),$$

where C_0 and C_1 are constants. Also recall that

$$\frac{1}{\sqrt{n}h}\sum_{i=1}^n K[(x_i-x)/h] \to_D |\psi|^{-1}L_W(1,0).$$

By virtue of these facts, together with (2.10), to prove

$$\left(nh^2\right)^{1/4}\left[\hat f^L(x) - f(x) - \frac{h^2}{2}f''(x)\int_{-\infty}^{\infty} y^2 K(y)\,dy\right] \to_D \sigma_u\,N\,L_W^{-1/2}(1,0),$$

it suffices to show that

$$\Delta_{3n} := \frac{\sum_{i=1}^{n} H_1[(x_i - x)/h] Y_i}{\sum_{i=1}^{n} K[(x_i - x)/h]} = o_P(1), \qquad \textbf{(5.1)}$$

provided that $nh^{14} \to 0$ and $nh^2 \to \infty$. This follows from some arguments similar to those in the proof of Theorem 2.2. To see why, we may split the numerator of Δ_{3n} as

$$\begin{aligned} f(x)\sum_{i=1}^{n} H_1[(x_i - x)/h] + \sum_{i=1}^{n} H_1[(x_i - x)/h][f(x_i) - f(x)] \\ + \sum_{i=1}^{n} H_1[(x_i - x)/h] u_i \\ := \Delta_{4n} + \Delta_{5n} + \Delta_{6n}. \end{aligned}$$

Noting $H_{1(x)} \le CK(x)$ as $K(x)$ has a compact support, as in (4.2),

$$\frac{\Delta_{6n}}{\sum_{i=1}^{n} K[(x_i - x)/h]} = O_P[(nh^2)^{-1/4}] = o_P(1).$$

As in (4.4) (also see Theorem 2.1),

$$\frac{\Delta_{4n}}{\sum_{i=1}^{n} K[(x_i - x)/h]} = O_P[(nh^2)^{-1/4}] = o_P(1).$$

By noting that

$$|\Delta_{5n}| \le C \sum_{i=1}^{n} |H_1[(x_i - x)/h]||x_i - x| = Ch \sum_{i=1}^{n} H_2[(x_i - x)/h],$$

where $H_2(x) = x^2 K(x)$, as in (4.6),

$$\frac{\Delta_{5n}}{\sum_{i=1}^{n} K[(x_i - x)/h]} = O_P(h) = o_P(1).$$

Combining all these estimates, we obtain (5.1), and the proof of Theorem 2.3 is complete. ■

NOTE

1. Assuming that Y has a density $|g(x)|/\int |g(x)|dx$, we have $\mathrm{E}|g(Y)|I_{g^2(Y)\ge \epsilon^2 nh/d_n} = \int g^2(x) I_{g^2(x)\ge \epsilon^2 nh/d_n} dx/\int |g(x)|dx$. The fact that $\int g^2(x) I_{g^2(x)\ge \epsilon^2 nh/d_n} dx = o(1)$ follows from $\mathrm{E}|g(Y)| = \int g^2(x)dx/\int |g(x)|dx < \infty$ and $P(g^2(Y) \ge \epsilon^2 nh/d_n) \le \epsilon^{-1}[d_n/(nh)]^{1/2}\mathrm{E}|g(Y)| = o(1)$.

REFERENCES

Akonom, J. (1993) Comportement asymptotique du temps d'occupation du processus des sommes partielles. *Annals of the Institute of Henri Poincaré* 29, 57–81.

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

Billingsley, P. (1974) Conditional distributions and tightness. *Annals of Probability* 2, 480–485.

Borodin, A.N. & I.A. Ibragimov (1995) *Limit Theorems for Functionals of Random Walks.* Proceedings of the Steklov Institute of Mathematics,Vol.195 (V.N. Sudakov, ed.). American Mathematical Society.

Fan, J. & I. Gijbels (1996) *Local Polynomial Modeling and Its Applications.* Chapman and Hall.

Geman, D. & J. Horowitz (1980) Occupation densities. *Annals of Probability* 8, 1–67.

Hall, P. (1977) Martingale invariance principles. *Annals of Probability* 5, 875–887.

Jeganathan, P (2004) Convergence of functionals of sums of r.v.s to local times of fractional stable motions. *Annals of Probability* 32, 1771–1795.

Jeganathan, P. (2008) Limit Theorems for Functionals of Sums That Converge to Fractional Brownian and Stable Motions. Cowles Foundation Discussion Paper 1949.

Luklacs, E. (1970) *Characteristic Functions*. Hafner.

Petrov, V.V. (1995) *Limit Theorems of Probability Theory.* Oxford University Press.

Phillips, P.C.B. & J.Y. Park (1998) Nonstationary Density Estimation and Kernel Autoregression. Cowles Foundation Discussion Paper 1181.

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

Wang, Q. & P.C.B. Phillips (2009a) Asymptotic theory for local time density estimation and nonparametric cointegrating regression. *Econometric Theory* 25, 710–738.

Wang, Q. & P.C.B. Phillips (2009b) Structural Nonparametric Cointegrating Regression. *Econometrica* 77, 1901–1948.