NONSTATIONARY DENSITY ESTIMATION AND KERNEL AUTOREGRESSION

Peter C. B. Phillips*
Cowles Foundation for Research in Economics
Yale University

and

Joon Y. Park
Economics Division
Seoul National University

March 1998

Abstract

An asymptotic theory is developed for the kernel density estimate of a random walk and the kernel regression estimator of a nonstationary first order autoregression. The kernel density estimator provides a consistent estimate of the local time spent by the random walk in the spatial vicinity of a point that is determined in part by the argument of the density and in part by initial conditions. The kernel regression estimator is shown to be consistent and to have a mixed normal limit theory. The limit distribution has a mixing variate that is given by the reciprocal of the local time of a standard Brownian motion. The permissible range for the bandwidth parameter h_n includes rates which may increase as well as decrease with the sample size n, in contrast to the case of a stationary autoregression. However, the convergence rate of the kernel regression estimator is at most $n^{1/4}$, and this is slower than that of a stationary kernel autoregression, in contrast to the parametric case. In spite of these differences in the limit theory and the rates of convergence between the stationary and nonstationary cases, it is shown that the usual formulae for confidence intervals for the regression function still apply when $h_n \to 0$.

Key words and phrases: Brownian sheet, kernel regression, local time, martingale embedding, mixture normal, nonstationary density, occupation time, quadratic variation, unit root autoregression.

*This paper grew out of a preliminary draft by Phillips that was written in July 1994, became a joint project and went through successive further drafts while Park visited the Cowles Foundation for Research in Economics during the fall of 1995. It was finally completed in March 1998. Phillips thanks the NSF for research support under Grant No. SBR 94–22922.

1 Introduction

The asymptotic behavior of parametric autoregression has been extensively researched in time series models that include both stationary and nonstationary (unit root) cases. The properties of nonparametric kernel methods in time series applications have also been studied, but attention has so far been limited to the case of stationary models. Convergence results for kernel regression estimators are particularly well developed for dependent data under weak dependence conditions, following work by Robinson (1983), Collomb (1985a), and Bierens (1983). The asymptotic results are analogous to those with independent and identically distributed data and include consistency and asymptotic normality of the kernel regression estimator under standard regularity conditions that involve a shrinking bandwidth as the sample size (n) tends to infinity. An optimal data-driven bandwidth selector is also available for kernel regression with stationary time series using cross validation methods (Hardle and Vieu, 1992). The literature is reviewed in Collomb (1985b), Bierens (1987), Hardle (1990), and Hardle and Linton (1995), and these methods are now used extensively in empirical econometric research.

As yet, there appear to be no asymptotic results for kernel regression estimators in models with unit roots or integrated processes. The random walk model is an important special case that is uncovered by existing theory and yet is empirically relevant in econometric work. The goal of the present paper is to begin the study of nonparametric estimation of nonstationary regressions by developing an asymptotic theory of kernel density estimation and kernel regression in the special case of a random walk. Although this model is very simple, it serves to illustrate the methods that are needed for an asymptotic theory in this class of problem and yields some interesting results on convergence rates that are indicative of the effects of nonstationarity on kernel regression. The methods and results of this paper should be useful for nonparametric inference about nonstationary time series and are also likely to be useful in the field of nonparametric estimation of nonlinear diffusion processes (Aït Sahalia, 1996 & 1997, and Jiang and Knight, 1997).

Specifically, this paper considers the case of a bivariate times series $(y_t,\, u_t)$ on a probability space $(\Omega,\, \mathcal{F},\, P)$ adapted to a filtration $\mathcal{F}_t$ and satisfying the unit root autoregression

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

with $\alpha = 1$ and $u_t \equiv iid(0, \sigma^2)$. The observed process y_t is initialized at $t = 0$ and we will allow for various possibilities regarding the stochastic properties of the initial value y_0. The regression function in (1) is given by the conditional mean function $m(x) = E(y_t|y_{t-1} = x)$ at x and our primary interest is in the asymptotic properties of the Nadaraya–Watson kernel estimate of $m(x)$ given by

$$m_n(x) = \Sigma_{t=1}^n K_n(x - y_{t-1})y_t / \Sigma_{t=1}^n K_n(x - y_{t-1}), \tag{2}$$

where $K_n(\cdot) = h_n^{-1}K(\cdot/h_n)$, h_n is the bandwidth parameter and $K(\cdot)$ is the kernel function.

We are also interested in the quantity

$$\widehat{f}_n(x) = (nh_n)^{-1} \sum_{t=1}^n K\left(\frac{x - y_{t-1}}{h_n}\right), \tag{3}$$

a scaled version of the denominator of (2), which is the usual kernel estimate of what would be the probability density of y_t at x if y_t were strictly stationary (i.e. if $|\alpha| < 1$, and $y_t \overset{d}{=} y_0 \overset{d}{=} N\left(0, \sigma^2/(1-\alpha^2)\right)$). It will be shown that the quantity $\sqrt{n}\widehat{f}_n(x)$ still has a meaning as a type of density estimate even in the nonstationary case. In fact, in the nonstationary case, $\sqrt{n}\widehat{f}_n(x)$ tells us how dense the process is about a particular spatial point, and in this sense can be interpreted as a form of 'density' estimate. The point where this density is being measured turns out to depend in part on initial conditions. Such estimates and the asymptotic theory associated with them will be useful in assessing the spatial characteristics of nonstationary time series like stock prices, interest rates and exchange rates.

It is further shown that the kernel estimator of the conditional mean function $m_n(x)$ is a consistent estimator of $m(x)$, as in the case of a stationary autoregression. However, unlike the stationary case ($|\alpha| < 1$), the limit distribution of the kernel estimator in the unit root case turns out to be mixed normal, rather than normal, and the mixture variate can be expressed in terms of the local time of a Brownian motion in the neighborhood of the origin. Also, the rate of convergence to this limit distribution is shown to be slower than that of the kernel estimator in the stationary case. These latter two results are very different from the corresponding asymptotics for parametric autoregressions in the stationary and nonstationary cases.

2 Some Preliminary Theory

Our development relies on the local time of a Brownian motion. Using the Tanaka formula and following Revuz and Yor (1991, pp. 207–216), we define local time as follows and make specific some of the properties that are needed in our development.

2.1 Definition *For a continuous local martingale M, there exists an increasing process $L_M(\cdot, s)$ called the local time of M at s such that*

$$|M(r)-s| = |M(0)-s| + \int_0^r \operatorname{sgn}(M(t)-s)dM(t) + L_M(r,s)$$

$$(M(r)-s)^+ = (M(0)-s)^+ + \int_0^r 1(M(t) > s)dM(t) + (1/2)L_M(r,s)$$

$$(M(r)-s)^- = (M(0)-s)^- - \int_0^r 1(M(t) \leq s)dM(t) + (1/2)L_M(r,s)$$

where $\operatorname{sgn}(x) = 1, -1$ *if* $x > 0$, $x \leq 0$, *and* $1(A)$ *is the indicator of* A.

2.2. Lemma (Continuity of Martingale local times) *For any continuous local martingale M, there exists a version of the local time such that $(r, s) \mapsto L_M(r, s)$ is a.s. continuous in both r and s. Moreover, it can be chosen so that $s \mapsto L_M(r, s)$ is Hölder continuous of order k for every $k < 1/2$ uniformly in r on every compact interval.*

2.3. Lemma (Occupation times formula) *Let M be a continuous local martingale with quadratic variation process $[M]_r$ and let $L_M(\cdot, s)$ be its local time at s. Then*

$$\int_0^t f(M(r), r)d[M]_r = \int_{-\infty}^{\infty} ds \int_0^t f(s,r)dL_M(r,s) \tag{4}$$

for every Borel function f. When f has only a single argument we have the simpler formula

$$\int_0^t f(M(r))d[M]_r = \int_{-\infty}^{\infty} f(s)L_M(t,s)ds. \tag{5}$$

Setting f to be the indicator function $f(M) = 1(|M(r) - p| < \varepsilon)$ in formula (5), we deduce that

$$L_M(t,p) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t 1(|M(r) - p| < \varepsilon) d[M]_r \ , \tag{6}$$

a representation which explains why $L_M(t,p)$ is called the local time of M at the point p. For standard Brownian motion W we have

$$L_W(t,p) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t 1(|W(r) - p| < \varepsilon) dr \ ,$$

which we denote by $L(t,p)$. The local time of standard Brownian motion at the origin will be denoted simply by $L(t)$. If $B = \sigma W$ is Brownian motion with variance σ^2, then by a simple calculation we have $L_B(t,p) = \sigma L(t, p/\sigma)$, and so $L_B(t) = \sigma L(t)$. As is clear from formula (6), local time is measured in units of the quadratic variation process, which we can think of as information units because they reflect the amount of information that is being accumulated about the process. In the case of the Brownian motion B, this is simply a scaled version of chronological time since $d\,[B]_r = \sigma^2 dr$.

We define the *chronological local time* of B directly as

$$\overline{L}_B(t,p) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t 1(|B(r) - p| < \varepsilon) dr \ = \sigma^{-2} L_B(t,p). \tag{7}$$

This measures the local time that B spends at p in the chronological units dr (as distinct from information units measured in $d\,[B]_r = \sigma^2 dr$). For reasons that will become apparent later on, the chronological local time of Brownian motion plays an important role in kernel regression asymptotics.

We make the following assumptions about the kernel function $K(\cdot)$, the bandwidth h_n, the initial condition y_0 and the errors u_j in (1).

2.4. Assumption. *The kernel* $K(\cdot)$ is a *symmetric and nonnegative density with integrable characteristic function* φ_K *and satisfies the following conditions for some* $r > 2$*:*

$$\int_{-\infty}^{\infty} K(s) = 1, \ \int_{-\infty}^{\infty} K(s)^2 ds < \infty, \ \int_{-\infty}^{\infty} s^{2r} K(s)\, ds < \infty, \ \sup_s K(s) < \infty;$$

2.5. Assumption.

(a) $n^{1-\delta} h_n^2 \to \infty$, *and* $h_n / n^{(1-\delta)/12} \to 0$ *for some* $\delta > 0$.

(b) $n^{1-\delta} h_n^4 \to \infty$, *and* $h_n / n^{(1-\delta)/12} \to 0$ *for some* $\delta > 0$.

2.6. Assumption. *The initial conditions of (1) are set at* $t = 0$ *and* y_0 *has the following general form*

$$y_0 = u + \sum_{j=0}^{[n\kappa]} u_{-j}, \ \textit{for some } \kappa \geq 0 \tag{8}$$

where u *is an* $O_{a.s.}(1)$ *random variable with* $E(|u|^p) < \infty$ *for some* $p > 2$.

2.7. Assumption. *The errors* $\{u_j\}_{j=-\infty}^{\infty}$ in (1) and (8) are $iid(0, \sigma^2)$ *with* $E(|u_j|^p) < \infty$, *for some* $p > 2$.

2.8. Remarks

(i) Condition 2.4 is analogous to conditions that are commonly made in kernel density and kernel regression asymptotics for stationary processes. The moment condition implies that the kernel has tail behaviour of the form $K(x) = o(|x|^{-2r})$ as $|x| \to \infty$. In the proof of Theorem 3.1 below and some later results, it will be convenient to place stronger conditions on the moment exponent r than $r > 2$. These stronger conditions are made explicit as we need them.

(ii) Condition 2.5 allows for bandwidth parameters of the form $h_n = cn^k$, $-1/2 + \delta < k < 1/12 - \delta$ for some constant c and some $\delta > 0$ in case (a), and $-1/4 + \delta < k < 1/12 - \delta$ in case (b). Thus, the bandwidth h_n may increase, as well as decrease, as $n \to \infty$. However, it should not decrease too fast nor increase too fast with n. More specific bandwidth rates will be given in our asymptotic development.

(iii) Assumption 2.6 permits the initial value y_0 to be zero, a constant, a fixed random variable (all of these are obtained by setting $\kappa = 0$ and by making the appropriate assumption about u), or a random variable that is dependent on the sample size and the parameter κ. The latter condition allows for the variance of y_0 to grow with n, so that the behaviour of y_0 more closely resembles that of the sample data.

The following results are useful in our main development.

2.9. Lemma (Limit theory for functionals of Brownian motion) *Let $B(s)$ be a Brownian motion with variance σ^2.*

(a) *If $f(\cdot)$ is a piecewise continuous and integrable function with $\overline{f} = \int_{-\infty}^{\infty} f(x)dx \neq 0$, then as $\lambda \to \infty$*

$$\lambda^{-1/2} \int_0^{\lambda t} f(B(s))\, ds \xrightarrow{d} \overline{f}\sigma^{-2} L_B(t).$$

(b) *Suppose $f(\cdot)$ is piecewise continuous and integrable with $\overline{f} = 0$, and $\int_{-\infty}^{\infty} x^2 f(x)dx < \infty$. Then, as $\lambda \to \infty$*

$$\lambda^{-1/4} \int_0^{\lambda t} f(B(s))\, ds \xrightarrow{d} \{2\langle f, f\rangle\}^{1/2} \sigma^{-2} U(L_B(t)),$$

where $\langle f, f\rangle = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x-y| f(x)f(y)dxdy$ is the 'energy' of f and U is a standard Brownian motion independent of B.

(c) *Let $(s_1, ..., s_k)$ be real numbers with $s_i \neq s_j$, $\forall i \neq j$,* and let $a > 0$. *Then, as $\lambda \to \infty$*

$$\left[\lambda^{1/2}\{L_B(t, s_i + a/\lambda) - L_B(t, s_i)\}, i = 1, ..., k\right] \xrightarrow{d} \left[2a^{1/2} U_i(L_B(t, s_i)), i = 1, ..., k\right],$$

where $L_B(t,s)$ is the local time at t of B at s, and $\{U_i, i = 1, ..., k\}$ are independent standard Brownian motions that are independent of B.

(d) *Let r be a fixed real number and treat $\left\{L_B(t, r + \frac{s}{\lambda}) - L_B(t, r)\right\}$ as a double indexed stochastic process in the arguments (t,s). Then, as $\lambda \to \infty$*

$$2^{-1}\lambda^{1/2}\left\{L_B(t, r + \frac{s}{\lambda}) - L_B(t, r)\right\} \xrightarrow{d} Q(L_B(t, r), s),$$

where $Q(t, s)$ is a standard Brownian sheet.

2.10. Lemma (Limit theory for kernel integrals of local time) *If $B(s)$ is a Brownian motion with variance σ^2, and $K(s)$ is a kernel function satisfying* Assumption 2.4, *then as $\lambda \to \infty$*

$$\lambda^{-1/2}\int_{-\infty}^{\infty}K(s)L_B(\lambda t,s)ds\xrightarrow{d}L_B(t),\tag{9}$$

$$\lambda^{-1/4}\int_{-\infty}^{\infty}sK(s)L_B(\lambda t,s)ds\xrightarrow{d}\varphi^{1/2}U(L_B(t)),\tag{10}$$

$$\lambda^{1/2}\int_{-\infty}^{\infty}sK(s)L_B(t,r+\frac{s}{\lambda})ds\xrightarrow{d}\varphi^{1/2}U(L_B(t,r)),\tag{11}$$

where $L_B(t)$ and $L_B(t,r)$ are the local times at t of B at the origin and at r, respectively, U is a standard Brownian motion independent of B, and $\varphi=-2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|a-b|abK(a)K(b)dadb$.

3 Main Results

We start with a construction that enables us to deal in a convenient way with sample moments of kernel functions of the data. In the form given here, the construction uses an expanded probability space in which the data can be represented almost surely and up to a negligible error in terms of a Brownian motion that is defined on the same space. This argument relies on an almost sure invariance principle and an embedding argument like that used in Phillips and Ploberger (1996).

More specifically, define the partial sum process $S_k=\sum_{j=1}^{k}u_j$ for $k\geq 1$, and $S_0=0$, for $k=0$. Since u_j has finite moments of order $p>2$, we can expand the probability space as necessary to set up a partial sum process that is distributionally equivalent to B_{nt} and a Brownian motion $B(\cdot)$ with variance σ^2 on the same space for which

$$\sup_{0\leq k\leq n}|S_k-B(k)|=o_{a.s}(n^{\frac{1}{p}}).\tag{12}$$

As p becomes large, the error in this approximation becomes smaller and it becomes bounded, or $O_{a.s}(1)$, when S_k is Gaussian. Almost sure invariance principles or strong approximations of the type (12) have been proved by many authors using a variety of techniques, a popular recent approach being the Hungarian construction, e.g. see Shorack and Wellner (1986) and Csörgõ, M. and L. Horváth (1993). Setting $B_{nk}=n^{-1/2}\Sigma_{j=1}^{k}u_j$, we can write this approximation in the form

$$\sup_{0\leq k\leq n}\left|B_{nk}-B(\frac{k}{n})\right|=o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right).$$

Let $[nr]$ be the integer part of nr. Then, by adjusting the space as needed, we have

$$n^{-1/2}y_t=n^{-1/2}y_{[nr]}=n^{-1/2}y_0+B\left(\frac{[nr]}{n}\right)+o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right)$$

for $(t-1)/n\leq r<t/n$. Further, by the local Hölder continuity of Brownian motion $B\left(\frac{[nr]}{n}\right)=B(r)+o_{a.s}\left(n^{-\frac{1}{2}+\delta}\right)$ for $\delta>0$, so we have

$$n^{-1/2}y_t=n^{-1/2}y_{[nr]}=n^{-1/2}y_0+B\left(r\right)+o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right)\tag{13}$$

In a similar way, when Assumption 2.6 applies to the initialization at y_0, we can expand the probability space as needed to apply the strong approximation and use Hölder continuity of the Brownian motion to write

$$n^{-1/2}y_0 = n^{-1/2}u + n^{-1/2}\Sigma_{j=0}^{[n\kappa]}u_{-j} = n^{-1/2}u + B_0(\kappa) + o_{a.s}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right), \tag{14}$$

where B_0 is a Brownian motion (independent of B) with variance σ^2.

In what follows, we will proceed as if the probability space and variates on them have all been adjusted so that relations (13) and (14) hold for the original data. As usual, the results imply distributional convergence in place of almost sure convergence in the original random coordinates.

Using these strong approximations, we can establish the following asymptotic result for the density estimate $\widehat{f}_n(x)$ in (3).

3.1. Theorem *Let* Assumptions 2.4, 2.5(a), 2.6 and 2.7 *hold. Define $c_n = \sqrt{n}/h_n$. Then, for some $\gamma \in (1/2, 1)$ and $\varepsilon > 0$, as $n \to \infty$*

$$\begin{aligned} &\frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right) \\ &\quad - c_n \int_0^1 K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right) 1\left(\frac{x-y_0}{\sqrt{n}} - \frac{1}{c_n^{\gamma}} \le B(r) \le \frac{x-y_0}{\sqrt{n}} + \frac{1}{c_n^{\gamma}}\right) dr \\ &= o_{a.s.}\left(\frac{1}{n^{1/4+\gamma/6-\varepsilon}}\right) \end{aligned} \tag{15}$$

For $x = x_0 + \sqrt{n}a$, with x_0 and a fixed, we have

$$\sqrt{n}\widehat{f}_n(x) = \frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right) \overset{a.s.}{\to} \overline{L}_B^{a,\kappa}, \tag{16}$$

where $\overline{L}_B^{a,\kappa} = \overline{L}_B(1, a - B_0(\kappa))$, and *$\overline{L}_B$ is the chronological local time of the Brownian motion B.*

3.2. Remarks

(i) As remarked in the introduction, the quantity $\widehat{f}_n(x) = (nh_n)^{-1}K\left(\frac{x-y_{t-1}}{h_n}\right)$ is the usual kernel estimate of what would be the density of y_{t-1} at x if the process y_{t-1}were strictly stationary. Theorem 3.1 shows that the quantity $\sqrt{n}\widehat{f}_n(x)$ still has a meaning as a type of density estimate even in the nonstationary case where $\alpha = 1$ in (1). In effect, $\sqrt{n}\widehat{f}_n(x)$ estimates the local time, $\overline{L}_B(1, a + B_0(\kappa))$, that y_{t-1} spends in the immediate vicinity of a point that is determined by the value of x and the initial value of the process. Thus, even in the nonstationary case, $\sqrt{n}\widehat{f}_n(x)$ still tells us how dense the process is about a particular point, and in this sense can be interpreted as a form of 'density' estimate for a nonstationary time series.

(ii) The limit behaviour in (16) of Theorem 3.1 depends on the precise assumptions about x and the initial condition. The cases given are probably the most important in practice. Thus, when $a = 0$ and the density ordinate $x = x_0$ is fixed, the limit behaviour of the density estimate $\widehat{f}_n(x)$ is essentially determined by the local time of the underlying Brownian motion process at the origin. In effect, since y_t is of order $O_p(\sqrt{t})$, the data tend to drift away from a fixed point like x_0, and, in consequence, the corresponding estimate is of the density or local time at the origin. But, when $x = \sqrt{n}a$, $\sqrt{n}\widehat{f}_n(x)$ estimates the local time spent by the process y_t at the point a.

In a similar way, when the initial conditions are fixed or bounded, then they do not influence the asymptotic behaviour of $\widehat{f}_n(x)$. But, when the initial observation is of order $O_p(\sqrt{n})$, then $\sqrt{n}\widehat{f}_n(x)$ estimates the local time spent by the process at a point which is, in part, determined by the initial conditions.

(iii) We see from Theorem 3.1 that for the random walk $\{y_t\}$

$$\sum_{t=1}^{n} K\left(\frac{x-y_t}{h_n}\right) = O_p(\sqrt{n}h_n),$$

in contrast to the case of a stationary autoregression, where

$$\sum_{t=1}^{n} K\left(\frac{x-y_t}{h_n}\right) = O_p(nh_n)\ .$$

Thus, the order of magnitude of the density estimate $\widehat{f}_n(x)$ in the integrated case is smaller than in the stationary case when $n \to \infty$. This is explained by the fact that an integrated process like y_t eventually (as $t \to \infty$) has a bigger probability of being away from a given point x than a stationary process and the kernel function $K(\cdot)$ assigns smaller values to the more distant points. This has important implications for kernel regression with nonstationary time series. In effect, the use of the kernel function in a kernel regression on (1) reduces the strength of the signal in the lagged regressor y_{t-1}. This, in turn, reduces the rate of convergence of the kernel estimate of the regression function.

(iv) Akonom (1993) proved a result like (15) in the case of a uniform kernel and with a zero initialization. Our derivations in the Appendix (for Lemmas 5.5 and 5.7) draw substantially on the line of argument developed in his paper.

3.3. Lemma *Let* Assumptions 2.4 - 2.6 *hold. Then, as $n \to \infty$ we have the following limits:*

(a) *For $x = x_0$ fixed and $y_0 = u$ (i.e. $a = 0,\ \kappa = 0$)*

$$\frac{1}{\sqrt{nh_n^2}} \sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right) y_{t-1} \overset{a.s.}{\to} x_0 \overline{L}_B^{0,0}, \tag{17}$$

(b) *For $x = x_0 + \sqrt{n}a$,* and *$y_0 = u + \Sigma_{j=0}^{[n\kappa]} u_{-j}$* with *$a \neq 0,\ \kappa > 0$*

$$\frac{1}{nh_n} \sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right) y_{t-1} \overset{a.s.}{\to} a\overline{L}_B^{a,\kappa}, \tag{18}$$

where $\overline{L}_B^{a,\kappa}$ is defined in Theorem 3.1.

3.4. Lemma *Let* Assumptions 2.4, 2.5(b), 2.6 *and* 2.7 *hold. Define $c_n = \sqrt{n}/h_n$. Then, for $x = x_0 + \sqrt{n}a$, with x_0 and a fixed, and for $y_0 = u + \sum_{j=0}^{[n\kappa]} u_{-j}$, with $u = O_{a.s.}(1)$, we have, as $n \to \infty$*

$$\frac{\sqrt{c_n}}{h_n\sqrt{nh_n^2}} \sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right) (y_{t-1} - x) \overset{d}{\to} \varphi^{1/2}\sigma^{-1} U\left(\overline{L}_B^{a,\kappa}\right), \tag{19}$$

where U is a standard Brownian motion independent of B, $\overline{L}_B^{a,\kappa}$ is defined in Theorem 3.1, *and $\varphi = -2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |a-b| abK(a)K(b)dadb$.*

3.5. Lemma *Under the conditions of* Lemma 3.4 *as* $n \to \infty$

$$\frac{1}{\sqrt[4]{nh_n^2}} \sum_{t=1}^{n} K\left(\frac{x - y_{t-1}}{h_n}\right) u_t \xrightarrow{d} \left(k_2 L_B^{a,\kappa}\right)^{1/2} V(1) = \left\{k_2 \sigma^2 \overline{L}_B^{a,\kappa}\right\}^{1/2} V(1),$$

where V *is a standard Brownian motion independent of* B, and $k_2 = \int_{-\infty}^{\infty} K(s)^2 ds$.

We are now in a position to develop asymptotics for the kernel estimator $m_n(x)$ in (2). First, we give conditions for the consistency of $m_n(x)$, and then give its limiting distribution.

3.6. Theorem *Under the conditions of* Lemma 3.4 *as* $n \to \infty$,

$$\begin{aligned} m_n(x) &\xrightarrow{p} x \quad \text{for } x = x_0 = \text{fixed} \\ n^{-1/2} m_n(x) &\xrightarrow{p} a \quad \text{for } x = x_0 + \sqrt{n}a. \end{aligned}$$

.

3.7. Theorem *Let the conditions of* Lemma 3.4 *hold and let* $x = x_0 + \sqrt{n}a$, *with* x_0 and a *fixed. The asymptotic distribution of the kernel estimate* $m_n(x)$ as $n \to \infty$ is as follows.

(i) *if* $h_n \to 0$

$$\sqrt[4]{nh_n^2}(m_n(x) - x) \xrightarrow{d} \left\{\frac{k_2 \sigma^2}{\overline{L}_B^{a,\kappa}}\right\}^{1/2} V(1) \equiv MN\left(0,\, k_2 \sigma^2 \overline{L}_B^{a,\kappa\,-1}\right) \tag{20}$$

(ii) *if* $h_n = h = constant$

$$\sqrt[4]{n}(m_n(x) - x) \xrightarrow{d} \left\{\frac{h^3 \varphi \sigma^{-2}}{\overline{L}_B^{a,\kappa}}\right\}^{1/2} U(1) + \left\{\frac{k_2 \sigma^2}{h \overline{L}_B^{a,\kappa}}\right\}^{1/2} V(1) \tag{21}$$

$$\equiv MN\left(0,\, (h\varphi\sigma^{-2} + h^{-1} k_2 \sigma^2) \overline{L}_B^{a,\kappa\,-1}\right) \tag{22}$$

(iii) *if* $h_n \to \infty$

$$\sqrt[4]{\frac{n}{h_n^6}}(m_n(x) - x) \xrightarrow{d} \left\{\frac{\varphi \sigma^{-2}}{\overline{L}_B^{a,\kappa}}\right\}^{1/2} U(1) \equiv MN\left(0,\, \varphi \sigma^{-2} \overline{L}_B^{a,\kappa\,-1}\right) \tag{23}$$

where $MN(0, \cdot)$ *signifies a mixed normal distribution. In the above formulae,* U *and* V *are independent standard Brownian motions that are independent of* B, *and* $k_2 = \int_{-\infty}^{\infty} K(s)^2 ds$.

3.8. Remarks

(a) Theorem 3.6 shows that the kernel estimator $m_n(x)$ is consistent, as in the standard case of stationary autoregression. However, consistency holds for increasing as well as decreasing bandwidths.

(b) Theorem 3.7 shows that the asymptotic distribution of $m_n(x)$ is mixed normal. The mixing parameter depends on the chronological local time of the Brownian motion B in which the time series is embedded. When x is fixed, the local time is measured at the origin. When $x = \sqrt{n}a$, the local time is measured at a.

(c) The form of the limit distribution depends on the bandwidth expansion or contraction rate. When $h_n \to 0$, there is no 'bias' term in the limit, *i.e.* the component

$$\frac{\frac{1}{\sqrt[4]{nh_n^2}} \sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right)(y_{t-1}-x)}{\frac{1}{\sqrt{nh_n^2}} \sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right)}$$

(see (47)) does not contribute to the limiting distribution. In fact, the limit has a mixed normal form with a mixing parameter that is given by the expression $k_2\sigma^2 \overline{L}_B^{a,\kappa-1}$. This limiting 'variance' is the analogue in the unit root case of the expression for the variance of the kernel estimator that applies in stationary kernel regression, *viz.* $k_2\sigma^2 f(x)^{-1}$, where $f(x)$ is the probability density of the stationary process at x. For the nonstationary case, the probability density in this expression is replaced by the chronological local time of the process. Note that this correspondence implies that confidence interval formulae for the regression function $m(x)$ are the same for the two cases, *viz.*

$$\left[m_n(x) \pm 1.96 \left\{\frac{k_2\widehat{\sigma}^2}{\sqrt{nh_n^2}\widehat{L}_B^{a,\kappa}}\right\}^{1/2}\right] \equiv \left[m_n(x) \pm 1.96 \left\{\frac{k_2\widehat{\sigma}^2}{nh_n\widehat{f}_n(x)}\right\}^{1/2}\right].$$

Thus, in spite of the major differences in the limit theory and the rates of convergence between the two cases, the usual formulae for confidence intervals for the regression function still apply, at least when $h_n \to 0$.

(d) Theorem 3.7 shows that the maximum achievable rate of convergence in the case of a unit root autoregression for the kernel estimator of the regression function is $n^{1/4}$. This rate is achieved when the bandwidth parameter is a constant as $n \to \infty$. The limit distribution in this case depends on the limit of the bias term - the first component of (21). Since this component is also mixed normal in form, the asymptotic distribution is still mixed normal but the mixing parameter is now more complex.

(e) Unlike stationary kernel regression, the kernel estimator $m_n(x)$ is consistent even when the bandwidth $h_n \to \infty$, provided that $h_n/n^{\frac{1}{12}-\delta} \to 0$, for some $\delta > 0$. In this event, the rate of convergence is slower and the limit distribution is dominated by the effects of the bias term. Again, note that the bias term effects are random in the limit in the nonstationary case.

4 Conclusion

The asymptotic theory for the first order unit root autoregression given in this paper shows that kernel estimation of this nonstationary autoregression is consistent, has a slower rate of convergence than that for stationary autoregressions, and the limit theory is mixture normal rather than normal. However, the usual formulae for confidence intervals for the regression function still apply provided the bandwidth $h_n \to 0$.

The methods involve the use of kernel density estimates for nonstationary time series and show that there is an interesting interpretation to such density estimates in terms of the local time of the Brownian motion that is the weak limit process of the standardised time series $\frac{1}{\sqrt{n}}y_{[nr]}$. Such density estimates can be used in empirical work to quantify the amount of time that a nonstationary process like a random walk spends in the vicinity of a certain point. Such quantification can be expected to be useful in studying the empirical properties of trajectories of economic time series like exchange rates and interest rates.

The methods in this paper will also be useful in extending the regression theory to more complex nonstationary time series models, including nonlinear models of integrated processes. They can further be used in analysing the nonparametric estimation of nonlinear diffusion models. Some extensions of the methods of the paper along these lines are being conducted and will be reported by the authors in later work.

5 Proofs

5.1 Proof of Lemma 2.2

See Corollary 1.8, p. 211 of Revuz and Yor (1991).

5.2 Proof of Lemma 2.3

See exercise 1.15, p. 216 of Revuz and Yor (1991).

5.3 Proof of Lemma 2.9

Theorem 4.4, p.146-147 of Ikeda and Watanabe (1989), gives the corresponding formulae for parts (a) and (b) of the lemma for standard Brownian motion. Note that Ikeda and Watanabe define the local time of standard Brownian motion as $(1/2)L(t,p)$, and their formulae are adjusted accordingly in what follows. Also, Ikeda and Watanabe assume that $f(\cdot)$ is continuous and has compact support. The results given here continue to hold because, under the stated conditions, $F(x) = \int_{-\infty}^{x} f(s)ds$ has compact support and $G(x) = \int_0^x F(s)ds$ is bounded. Together with piecewise continuity of $f(\cdot)$, these two conditions are sufficient for Ikeda and Watanabe's theorem to hold.

Using the formulae for standard Brownian motion and rescaling delivers the stated results. Thus, defining $f_\sigma(\cdot) = f(\sigma\cdot)$, we have $f_\sigma(W(s)) = f(\sigma W(s)) = f(B(s))$, and then

$$\lambda^{-1/2}\int_0^{\lambda t} f\left(B(s)\right)ds \overset{d}{=} \lambda^{-1/2}\int_0^{\lambda t} f_\sigma\left(W(s)\right)ds \overset{d}{\longrightarrow} \overline{f_\sigma}L(t) = \overline{f}\sigma^{-1}L(t) \overset{d}{=} \overline{f}\sigma^{-2}L_B(t)$$

as required for part (a). In the same way, we find for part(b), where $\overline{f} = \overline{f}_\sigma = 0$,

$$\begin{aligned} \lambda^{-1/4}\int_0^{\lambda t} f\left(B(s)\right)ds &\overset{d}{=} \lambda^{-1/4}\int_0^{\lambda t} f_\sigma\left(W(s)\right)ds \overset{d}{\longrightarrow} \left\{2\left\langle f_\sigma, f_\sigma\right\rangle\right\}^{1/2} U\left(L(t)\right) \\ &= \left\{2\sigma^{-3}\left\langle f,f\right\rangle\right\}^{1/2} U\left(L(t)\right) \overset{d}{=} \left\{2\left\langle f,f\right\rangle\right\}^{1/2}\sigma^{-2}U\left(L_B(t)\right), \end{aligned}$$

giving the stated result (b).

For part (c), we use the following result for standard Brownian motion from Revuz and Yor (1991, exercise 2.14, p. 486).

$$\left[\varepsilon^{-1/2}\left\{L(t,s_i+\varepsilon) - L(t,s_i)\right\}, i=1,...,k\right] \overset{d}{\to} \left[2U_i\left(L(t,s_i)\right), i=1,...,k\right] \quad as\ \varepsilon \to 0,$$

where U_i are standard Brownian motions that are independent of L. From this result, we obtain

$$\begin{aligned}
&\left[\lambda^{1/2}\left\{L_B(t,s_i+a/\lambda)-L_B(t,s_i)\right\},i=1,...,k\right]\\
&\overset{d}{=}\left[\lambda^{1/2}\sigma\left\{L(t,\sigma^{-1}(s_i+a/\lambda))-L(t,\sigma^{-1}s_i)\right\},i=1,...,k\right]\\
=\;&\left[\varepsilon^{-1/2}\sigma^{1/2}a^{1/2}\left\{L(t,s_i/\sigma+\varepsilon)-L(t,s_i/\sigma)\right\},i=1,...,k\right]
\end{aligned}$$

$$\overset{d}{\to}\left[2(\sigma a)^{1/2}U_i\left(L(t,s_i/\sigma)\right),i=1,...,k\right]\overset{d}{=}\left[2a^{1/2}U_i\left(L_B(t,s_i)\right),i=1,...,k\right],$$

as $\lambda\to\infty$, giving the required result.

For part (d), we use the following result for standard Brownian motion from Revuz and Yor (1991, exercise 2.12, p. 486).

$$2^{-1}\lambda^{1/2}\left\{L(t,r+\frac{s}{\lambda})-L(t,r)\right\}\overset{d}{\to}Q(L(t,r),s),$$

where Q is a standard Brownian sheet. It follows that

$$2^{-1}\lambda^{1/2}\left\{L_B(t,r+\frac{s}{\lambda})-L_B(t,r)\right\}\overset{d}{=}2^{-1}\lambda^{1/2}\sigma\left\{L_B(t,\sigma^{-1}(r+s/\lambda))-L_B(t,\sigma^{-1}r)\right\}$$

$$\overset{d}{\to}\sigma Q(L(t,r/\sigma),s/\sigma)\overset{d}{=}Q(\sigma L(t,r/\sigma),s)\overset{d}{=}Q(L_B(t,r),s),$$

as required. The penultimate distributional equivalence in the above argument uses the fact that $abQ(t,s)\overset{d}{=}Q(a^2t,,b^2s)$ - e.g Revuz and Yor (1991, exercise 3.11 p.37)-and sets $a=b=\sigma^{1/2}$.

5.4 Proof of Lemma 2.10

To prove (9), we use the occupation formula (5) of Lemma 2.3 to write the integral in the form

$$\lambda^{-1/2}\int_{-\infty}^{\infty}K(s)L_B(\lambda t,s)ds=\lambda^{-1/2}\sigma^2\int_0^{\lambda t}K(B(s))ds$$

as a continuous additive linear functional of the Brownian motion B. Then, applying Lemma 2.9(a) gives

$$\lambda^{-1/2}\sigma^2\int_0^{\lambda t}K(B(s))ds\overset{d}{\to}\sigma^2\int_{-\infty}^{\infty}K(s)ds\sigma^{-2}L_B(t)=L_B(t),$$

as required. Result (10) is proved in a similar way, using the occupation formula (5) to write

$$\lambda^{-1/4}\int_{-\infty}^{\infty}sK(s)L_B(\lambda t,s)ds=\lambda^{-1/4}\sigma^2\int_0^{\lambda t}B(s)K(B(s))ds$$

Now use Lemma 2.9(b), giving

$$\lambda^{-1/4}\sigma^2\int_0^{\lambda t}B(s)K(B(s))ds\overset{d}{\to}\varphi^{1/2}U(L_B(t)),$$

as required for (10).

To prove (11) we proceed as follows. First let Q_1 and Q_2 be two independent standard Brownian sheets and define $Q(t,s) = Q_1(t,s)1(s \geq 0) + Q_2(t,-s)1(s \leq 0)$. Next, from Lemma 2.9(d) we deduce that as $\lambda \to \infty$

$$2^{-1}\lambda^{1/2}\left\{L_B(t, r + \frac{s}{\lambda}) - L_B(t,r)\right\} \xrightarrow{d} Q(L_B(t,r), s),$$

where Q is independent of L_B. It now follows by the continuous mapping theorem that

$$\begin{aligned}
\lambda^{1/2}\int_{-\infty}^{\infty} sK(s)L_B(t, r+\frac{s}{\lambda})ds &= \int_{-\infty}^{\infty} sK(s)\lambda^{1/2}\left\{L_B(t, r+\frac{s}{\lambda}) - L_B(t,r)\right\}ds \\
&\xrightarrow{d} 2\int_{-\infty}^{\infty} sK(s)Q(L_B(t,r), s)ds \\
&\stackrel{d}{=} 2\left\{L_B(t,r)\right\}^{1/2}\int_{-\infty}^{\infty} sK(s)Q(1,s)ds.
\end{aligned}$$

Next let $F(t) = \int_{-\infty}^{t} sK(s)ds$. Direct calculation shows that $\int_{-\infty}^{\infty} F(s)^2 ds = \varphi/4$, and that as $t \to \infty$, $F(t)Q(1,t) \to 0 \quad a.s.$. Then, integration by parts yields

$$\int_{-\infty}^{\infty} sK(s)Q(1,s)ds = -\int_{-\infty}^{\infty} F(s)dQ(1,s) \stackrel{d}{=} N\left(0, \int_{-\infty}^{\infty} F(s)^2 ds\right),$$

from which the stated result follows immediately.

5.5 Lemma

Let $M_n(a+kh, a+(k+1)h; z) = \sum_{j=1}^{n} \exp\left(-\frac{iz}{h}\{a - S_j\}\right) 1\left[a + kh \leq S_j \leq a + (k+1)h\right]$ *for some integer* $k \geq 0$ *and* $h > 0$, and where $S_j = \sum_{t=1}^{j} u_s$. *Let* $r \geq 1$ *be any positive integer. Then, for any* $a \in \mathbb{R}$, $h \geq \frac{1}{\sqrt{n}}$, *and any positive integer* k, *there exists a constant* C *such that*

$$\begin{aligned}
& E\left|M_n(a, a+h; z) - M_n(a+kh, a+(k+1)h; z)\right|^{2r} \\
\leq\ & C\left(\frac{h}{\sqrt{n}}\right)^r \left[1 + kh^2 \log n\right]^r
\end{aligned}$$

uniformly in $z \in \mathbb{R}$.

5.6 Proof

The result follows from Akonom (1993, Lemma 1). Note that

$$\begin{aligned}
& M_n(a, a+h; z) - M_n(a+kh, a+(k+1)h; z) \\
=\ & \sum_{j=1}^{n} \exp\left(-\frac{iz}{h}\{a - S_j\}\right) U(S_j)
\end{aligned} \tag{24}$$

where

$$U(S_j) = 1\left[a \leq S_j \leq a+h\right] - 1\left[a + kh \leq S_j \leq a + (k+1)h\right].$$

Akonom (1993, Lemma 1) proved the stated result for

$$N_n(a,a+h)-N_n(a+kh,a+(k+1)h)=\sum_{j=1}^{n}U\left(S_j\right), \tag{25}$$

where $N_n(a,a+h)=\sum_{j=1}^{n}1\left[a\leq S_j\leq a+h\right]$ is the number of visits of S_j to the interval $[a,a+h]$. Now

$$\begin{aligned}&\left|M_n(a,a+h;z)-M_n(a+kh,a+(k+1)h;z)\right|^2\\ =&\sum_{j,l=1}^{n}\exp\left(-\frac{iz}{h}\left\{S_l-S_j\right\}\right)U\left(S_j\right)\overline{U}\left(S_l\right).\end{aligned}$$

The factor $\exp\left(-\frac{iz}{h}\left\{S_k-S_j\right\}\right)$ in this expression has modulus unity, and does not affect the bound on the r'th moment of the expression which can be calculated in the same manner as in Akonom (1993) for (25). Thus, the bound is the same as for the $2r'$th moment of (25) and this bound holds uniformly in z.

5.7 Lemma

Let $a\in\mathbb{R}$, let h_n be a sequence of positive numbers satisfying $h_n\geq\frac{1}{\sqrt{n}}$, and let k_n be be a sequence of positive integers with $k_n\leq n$. Then, for all $\varepsilon>0$

$$\begin{aligned}&M_n(a,a+h_n;z)-\frac{1}{k_n}M_n(a+h_n,a+(k_n+1)h_n;z)\\ =\;&o_{a.s.}\left(h_n^{\frac{1}{2}}\left(1+k_nh_n^2\right)^{\frac{1}{2}}n^{\frac{1}{4}+\varepsilon}\right),\end{aligned} \tag{26}$$

uniformly in z as $n\to\infty$.

5.8 Proof

The result follows as in Akonom (1993, Proposition 1). Let $\beta>0$, and define the events

$$A_{n,k}(\beta)=\left\{\begin{array}{c}\left|M_n(a,a+h_n;z)-M_n(a+kh_n,a+(k+1)h_n;z)\right|\\ \geq\beta h_n^{\frac{1}{2}}\left(1+kh_n^2\right)^{\frac{1}{2}}n^{\frac{1}{4}+\varepsilon}\end{array}\right\},$$

$$A_n(\beta)=\left\{\begin{array}{c}\left|M_n(a,a+h_n;z)-\frac{1}{k_n}M_n(a+h_n,a+(k_n+1)h_n;z)\right|\\ \geq\beta h_n^{\frac{1}{2}}\left(1+k_nh_n^2\right)^{\frac{1}{2}}n^{\frac{1}{4}+\varepsilon}\end{array}\right\}.$$

and note that $A_n(\beta)\subset\cup_{k\leq k_n}A_{n,k}(\beta)$. Using the Markov inequality of order $2r$, Lemma 5.5, and the fact that $A_n(\beta)\subset\cup_{k\leq k_n}A_{n,k}(\beta)$, we obtain

$$P\left[A_n(\beta)\right]\leq k_n\frac{C\left(\log n\right)^r}{\beta^{2r}n^{2r\varepsilon}},$$

uniformly in z. Since $k_n\leq n$, the result then follows from the Borel-Cantelli lemma by suitable choices of r and a sequence β_n for which $\beta_n\to0$ and

$$\sum_{n=1}^{\infty}\beta_n^{-2r}n^{1-2r\varepsilon}\left(\log n\right)^r<\infty.$$

5.9 Proof of Theorem 3.1

Let $\gamma \in (1/2, 1)$. Using the strong approximation to the partial sum process $S_k = \sum_{i=1}^{k} u_j$ given in (13) we have the representation

$$y_{t-1} = B\left(t-1\right) + y_0 + o_{a.s.}(n^{1/p}) = \sqrt{n}B(r) + y_0 + o_{a.s.}(n^{1/p}),$$

for $(t-2)/n \leq r < (t-1)/n,\ t \geq 1$. Write the scaled density estimate as follows:

$$\begin{aligned} \frac{1}{\sqrt{nh_n^2}} \sum_{i=1}^{n} K\left(\frac{x - y_{i-1}}{h_n}\right) &= \frac{n}{\sqrt{nh_n^2}} \frac{1}{n} \sum_{i=1}^{n} K\left(\frac{x - y_0 - S_{i-1}}{h_n}\right) \\ &= c_n \frac{1}{n} \sum_{i=1}^{n} K\left(c_n \left\{\frac{x - y_0 - S_{i-1}}{\sqrt{n}}\right\}\right) 1\left[\left|\frac{x - y_0 - S_{i-1}}{\sqrt{n}}\right| \leq \frac{1}{c_n^{\gamma}}\right] \\ &\quad + c_n \frac{1}{n} \sum_{i=1}^{n} K\left(c_n \left\{\frac{x - y_0 - S_{i-1}}{\sqrt{n}}\right\}\right) 1\left[\left|\frac{x - y_0 - S_{i-1}}{\sqrt{n}}\right| > \frac{1}{c_n^{\gamma}}\right] \\ &= c_n \frac{1}{n} \sum_{i=1}^{n} K\left(c_n \left\{\frac{x - y_0 - S_{i-1}}{\sqrt{n}}\right\}\right) 1\left[\left|\frac{x - y_0 - S_{i-1}}{\sqrt{n}}\right| \leq \frac{1}{c_n^{\gamma}}\right] \\ &\quad + o_{a.s}(c_n^{1-2r(1-\gamma)}), \end{aligned} \tag{27}$$

since $K\left(x\right) = o\left(\left|x\right|^{-2r}\right)$ as $|x| \to \infty$. The second term of (27) is negligible when $1-2r\left(1-\gamma\right) < 0$, which will be assumed in what follows.

The next part of the proof follows a line of argument developed in Akonom (1993, Theorem 2) that applies for the case of a uniform kernel. The idea is to show that the first part of (27) can be replaced with a corresponding integral expression that involves the Brownian motion B. The final part of the argument then makes use of the occupation formula.

Let k_n be an integer for which $1 \leq k_n/c_n^{\gamma} \leq \sqrt{n}/log(n)$. The role of k_n is to widen the bands in the formulae below so that the embedding of S_{i-1} in the Brownian motion B can be used. We seek to majorize the difference

$$\begin{aligned} & c_n \left| \frac{1}{n} \sum_{i=1}^{n} K\left(c_n \left\{\frac{x - y_0 - S_{i-1}}{\sqrt{n}}\right\}\right) 1\left(\frac{x - y_0}{\sqrt{n}} - \frac{1}{c_n^{\gamma}} \leq \frac{S_{i-1}}{\sqrt{n}} \leq \frac{x - y_0}{\sqrt{n}} + \frac{1}{c_n^{\gamma}}\right) \right. \\ & \left. - \int_0^1 K\left(c_n \left\{\frac{x - y_0}{\sqrt{n}} - B(r)\right\}\right) 1\left(\frac{x - y_0}{\sqrt{n}} - \frac{1}{c_n^{\gamma}} \leq B(r) \leq \frac{x - y_0}{\sqrt{n}} + \frac{1}{c_n^{\gamma}}\right) dr \right| \\ \leq\ & A_1(n) + \frac{1}{2k_n} A_2(n) + A_3(n), \end{aligned}$$

using the terms

$$\begin{aligned} A_1(n) &= c_n \left| \frac{1}{n} \sum_{j=1}^{n} K\left(c_n \left\{\frac{x - y_0 - S_{j-1}}{\sqrt{n}}\right\}\right) 1\left(\frac{x - y_0}{\sqrt{n}} - \frac{1}{c_n^{\gamma}} \leq \frac{S_{j-1}}{\sqrt{n}} \leq \frac{x - y_0}{\sqrt{n}} + \frac{1}{c_n^{\gamma}}\right) \right. \\ & \left. - \frac{1}{2k_n} \frac{1}{n} \sum_{j=1}^{n} K\left(c_n \left\{\frac{x - y_0 - S_{j-1}}{\sqrt{n}}\right\}\right) 1\left(\frac{x - y_0}{\sqrt{n}} - \frac{k_n}{c_n^{\gamma}} \leq \frac{S_{j-1}}{\sqrt{n}} \leq \frac{x - y_0}{\sqrt{n}} + \frac{k_n}{c_n^{\gamma}}\right) \right|, \end{aligned}$$

$$
\begin{aligned}
A_2(n) &= c_n \left| \frac{1}{n} \sum_{j=1}^{n} K\left(c_n\left\{\frac{x-y_0-S_{j-1}}{\sqrt{n}}\right\}\right) 1\left(\frac{x-y_0}{\sqrt{n}}-\frac{k_n}{c_n^\gamma} \le \frac{S_{j-1}}{\sqrt{n}} \le \frac{x-y_0}{\sqrt{n}}+\frac{k_n}{c_n^\gamma}\right) \right. \\
&\quad \left. - \int_0^1 K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}}-B(r)\right\}\right) 1\left(\frac{x-y_0}{\sqrt{n}}-\frac{k_n}{c_n^\gamma} \le B(r) \le \frac{x-y_0}{\sqrt{n}}+\frac{k_n}{c_n^\gamma}\right) dr \right|, \\
A_3(n) &= c_n \left| \frac{1}{2k_n} \int_0^1 K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}}-B(r)\right\}\right) 1\left(\frac{x-y_0}{\sqrt{n}}-\frac{k_n}{c_n^\gamma} \le B(r) \le \frac{x-y_0}{\sqrt{n}}+\frac{k_n}{c_n^\gamma}\right) dr \right. \\
&\quad \left. - \int_0^1 K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}}-B(r)\right\}\right) 1\left(\frac{x-y_0}{\sqrt{n}}-\frac{1}{c_n^\gamma} \le B(r) \le \frac{x-y_0}{\sqrt{n}}+\frac{1}{c_n^\gamma}\right) dr \right|.
\end{aligned}
$$

Take each of these in turn in what follows. Let $\varphi_K(z)$ be the characteristic function of the density $K(\cdot)$. Because φ_K is absolutely integrable, we have

$$
\begin{aligned}
A_1(n) &= \frac{c_n}{2\pi} \left| \frac{1}{n} \sum_{j=1}^{n} \int_{-\infty}^{\infty} e^{-izc_n\left\{\frac{x-y_0}{\sqrt{n}}-\frac{S_{j-1}}{\sqrt{n}}\right\}} \varphi_K(z) dz \left[1\left(\frac{x-y_0}{\sqrt{n}}-\frac{1}{c_n^\gamma} \le \frac{S_{j-1}}{\sqrt{n}} \le \frac{x-y_0}{\sqrt{n}}+\frac{1}{c_n^\gamma}\right) \right. \right. \\
&\quad \left. \left. - \frac{1}{k_n} 1\left(\frac{x-y_0}{\sqrt{n}}-\frac{k_n}{c_n^\gamma} \le \frac{S_{j-1}}{\sqrt{n}} \le \frac{x-y_0}{\sqrt{n}}+\frac{k_n}{c_n^\gamma}\right) \right] \right| \\
&\le \frac{c_n}{2\pi} \int_{-\infty}^{\infty} |\varphi_K(z)| \frac{1}{n} \left| \sum_{j=1}^{n} e^{-izc_n\left\{\frac{x-y_0}{\sqrt{n}}-\frac{S_{j-1}}{\sqrt{n}}\right\}} \left[1\left(\frac{x-y_0}{\sqrt{n}}-\frac{1}{c_n^\gamma} \le \frac{S_{j-1}}{\sqrt{n}} \le \frac{x-y_0}{\sqrt{n}}+\frac{1}{c_n^\gamma}\right) \right. \right. \\
&\quad \left. \left. - \frac{1}{k_n} 1\left(\frac{x-y_0}{\sqrt{n}}-\frac{k_n}{c_n^\gamma} \le \frac{S_{j-1}}{\sqrt{n}} \le \frac{x-y_0}{\sqrt{n}}+\frac{k_n}{c_n^\gamma}\right) \right] \right| dz.
\end{aligned}
$$

By Lemma 5.7

$$
\begin{aligned}
&\left| \frac{c_n}{n} \sum_{j=1}^{n} e^{-izc_n\left\{\frac{x-y_0}{\sqrt{n}}-\frac{S_{j-1}}{\sqrt{n}}\right\}} \left\{ 1\left(\frac{x-y_0}{\sqrt{n}}-\frac{1}{c_n^\gamma} \le \frac{S_{j-1}}{\sqrt{n}} \le \frac{x-y_0}{\sqrt{n}}+\frac{1}{c_n^\gamma}\right) \right. \right. \\
&\qquad \left. \left. - \frac{1}{k_n} 1\left(\frac{x-y_0}{\sqrt{n}}-\frac{k_n}{c_n^\gamma} \le \frac{S_{j-1}}{\sqrt{n}} \le \frac{x-y_0}{\sqrt{n}}+\frac{k_n}{c_n^\gamma}\right) \right\} \right| \\
&\quad = o_{a.s.}\left(\frac{c_n}{n} \frac{1}{c_n^{\gamma/2}} \left[1+\frac{k_n}{c_n^{2\gamma}}\right]^{1/2} n^{1/4+\varepsilon} \right),
\end{aligned}
$$

uniformly in z, thereby giving the order of magnitude of $A_1(n)$ as

$$
A_1(n) = o_{a.s.}\left(\frac{c_n}{n} \frac{1}{c_n^{\gamma/2}} \left[1+\frac{k_n}{c_n^{2\gamma}}\right]^{1/2} n^{1/4+\varepsilon} \right). \tag{28}
$$

To bound $A_2(n)$ we set up intervals $I_n' \subset I_n \subset I_n''$ defined as

$$
\begin{aligned}
I_n &= \left[\frac{x-y_0}{\sqrt{n}}-\frac{k_n}{c_n^\gamma}, \frac{x-y_0}{\sqrt{n}}+\frac{k_n}{c_n^\gamma}\right], \\
I_n' &= \left[\frac{x-y_0}{\sqrt{n}}-\frac{k_n}{c_n^\gamma}+\frac{n^{1/p}}{\sqrt{n}}, \frac{x-y_0}{\sqrt{n}}+\frac{k_n}{c_n^\gamma}-\frac{n^{1/p}}{\sqrt{n}}\right], \\
I_n'' &= \left[\frac{x-y_0}{\sqrt{n}}-\frac{k_n}{c_n^\gamma}-\frac{n^{1/p}}{\sqrt{n}}, \frac{x-y_0}{\sqrt{n}}+\frac{k_n}{c_n^\gamma}+\frac{n^{1/p}}{\sqrt{n}}\right],
\end{aligned}
$$

and let G_n be the event

$$\begin{aligned} G_n &= \left\{ \int_0^1 \left[K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right)\right] 1_{I_n'}[B(r)]\,dr \right. \\ &\le \frac{1}{n}\sum_{i=1}^n K\left(c_n\left\{\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right\}\right) 1\left(\frac{S_{i-1}}{\sqrt{n}} \in I_n\right) \\ &\left. \le \int_0^1 \left[K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right)\right] 1_{I_n''}[B(r)]\,dr \right\} \\ &= \left\{ -\int_0^1 \left[K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right)\right] 1_{I_n - I_n'}[B(r)]\,dr \right. \\ &\le \frac{1}{n}\sum_{i=1}^n K\left(c_n\left\{\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right\}\right) 1\left(\frac{S_{i-1}}{\sqrt{n}} \in I_n\right) \\ &\quad - \int_0^1 \left[K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right)\right] 1_{I_n}[B(r)]\,dr \\ &\left. \le \int_0^1 \left[K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right)\right] 1_{I_n'' - I_n}[B(r)]\,dr \right\} \end{aligned}$$

In view of the strong approximation (12) , $S_{i-1}/\sqrt{n}$ is almost surely of distance less than $n^{1/p}/\sqrt{n}$ from $B(r)$ uniformly over $(i-1)/n \le r < i/n$ as $i, n \to \infty$. Also, the kernel function $K(\cdot)$ is nonnegative and bounded. It follows that the event G_n holds eventually as $n \to \infty$, so that

$$P\left(\lim\inf_{n\to\infty} G_n\right) = 1.$$

Hence

$$P\left(\lim\sup_{n\to\infty}\left\{A_2(n) \ge c_n\int_0^1 K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right) 1_{I_n''-I_n'}[B(r)]\,dr\right\}\right) = 0.$$

Using the occcupation formula (5) we have the equivalence

$$\begin{aligned} \int_0^1 K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right) 1_{I_n''-I_n'}[B(r)]\,dr &= \sigma^{-2}\int_{-\infty}^{\infty} K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - p\right\}\right) 1_{I_n''-I_n'}[p]\,L(1,p)dp \\ &= \sigma^{-2}\int_{I_n''-I_n'} K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - p\right\}\right) L(1,p)dp, \end{aligned}$$

whose modulus is bounded by a constant multiple of

$$\int_{I_n''-I_n'} L(1,p)dp = O_{a.s.}\left(\frac{n^{1/p}}{\sqrt{n}}\right).$$

It follows that $A_2(n) = O_{a.s.}\left(c_n n^{1/p}/\sqrt{n}\right)$, and

$$\frac{1}{k_n}A_2(n) = O_{a.s.}\left(\frac{c_n n^{1/p}}{k_n\sqrt{n}}\right). \tag{29}$$

To find the order of magnitude of $A_3(n)$, we can use alternate versions of Lemmas 5.5 and 5.7 in which the partial sum process S_k is replaced by the Brownian motion $\sqrt{n}B(r)$ - see Akonom

(1993, p. 73) - and then the derivation of the order of magnitude follows in the same way as that of $A_1(n)$.

We deduce from (27) (28) and (29) that

$$\begin{aligned}
&\frac{1}{\sqrt{nh_n^2}}\sum_{i=1}^{n} K\left(\frac{x-y_{i-1}}{h_n}\right) \\
&\qquad - c_n \int_0^1 K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right) 1\left(\frac{x-y_0}{\sqrt{n}} - \frac{1}{c_n^{\gamma}} \le B(r) \le \frac{x-y_0}{\sqrt{n}} + \frac{1}{c_n^{\gamma}}\right) dr \\
&= o_{a.s.}\left(\frac{c_n}{n}\frac{1}{c_n^{\gamma/2}}\left[1+\frac{k_n}{c_n^{2\gamma}}\right]^{1/2} n^{1/4+\varepsilon}\right) + O_{a.s.}\left(\frac{c_n n^{1/p}}{k_n\sqrt{n}}\right) + o_{a.s}(c_n^{1-2r(1-\gamma)}) \\
&= o_{a.s.}\left(\frac{c_n}{n}\frac{1}{c_n^{\gamma/2}} n^{1/4+\varepsilon}\right) + O_{a.s.}\left(\frac{c_n n^{\varepsilon}}{k_n\sqrt{n}}\right) + o_{a.s}(c_n^{1-2r(1-\gamma)}) \\
&= o_{a.s.}\left(\frac{n^{\varepsilon} h_n^{1+\gamma/2}}{n^{1/4} n^{\gamma/4}}\right) + O_{a.s.}\left(\frac{n^{\varepsilon}}{n^{\gamma} h_n^{1-2\gamma}}\right) + o_{a.s}(c_n^{1-2r(1-\gamma)})
\end{aligned}$$

when k_n is chosen in such a way that $k_n = O(c_n^{2\gamma})$, p is sufficiently large that $1/p \le \varepsilon$ and r is large enough that the term $o_{a.s}(c_n^{1-2r(1-\gamma)})$ is negligible in relation to the other two terms, which can be arranged by modifying Assumption 2.4 as needed. By Assumption 2.6 $n^{1-\delta}h_n^2 \to \infty$, and $h_n/n^{(1-\delta)/12} \to 0$ for some $\delta > 0$, so that

$$\begin{aligned}
& o_{a.s.}\left(\frac{n^{\varepsilon} h_n^{1+\gamma/2}}{n^{1/4} n^{\gamma/4}}\right) + O_{a.s.}\left(\frac{n^{\varepsilon}}{n^{\gamma} h_n^{1-2\gamma}}\right) \\
&= o_{a.s.}\left(\frac{n^{\varepsilon} n^{(1-\delta)(1+\gamma/2)/12}}{n^{1/4} n^{\gamma/4}}\right) + o_{a.s.}\left(\frac{n^{\varepsilon} n^{(1-2\gamma)(1-\delta)/2}}{n^{\gamma}}\right) \\
&= o_{a.s.}\left(\frac{n^{\varepsilon} n^{(1+\gamma/2)/12}}{n^{1/4} n^{\gamma/4}}\right) + o_{a.s.}\left(\frac{n^{\varepsilon} n^{(1-2\gamma)/2}}{n^{\gamma}}\right) \\
&= o_{a.s.}\left(\frac{1}{n^{\frac{1}{4}+\frac{5}{24}\gamma-\frac{1}{12}-\varepsilon}}\right) + o_{a.s.}\left(\frac{n^{\varepsilon}}{n^{2\gamma-1/2}}\right) \\
&= o_{a.s.}\left(\frac{1}{n^{\frac{1}{4}+\frac{5}{24}\gamma-\frac{1}{6}\gamma-\varepsilon}}\right) + o_{a.s.}\left(\frac{n^{\varepsilon}}{n^{2\gamma-1/2}}\right) \\
&= o_{a.s.}\left(\frac{1}{n^{\frac{1}{4}+\frac{1}{6}\gamma-\varepsilon}}\right)
\end{aligned}$$

which gives the required result (15).

To prove (16) we need to show that

$$c_n \int_0^1 K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right) 1\left(\frac{x-y_0}{\sqrt{n}} - \frac{1}{c_n^{\gamma}} \le B(r) \le \frac{x-y_0}{\sqrt{n}} + \frac{1}{c_n^{\gamma}}\right) dr \overset{a.s.}{\to} \overline{L}_B^{a,\kappa}, \qquad (30)$$

for then

$$\sqrt{n}\widehat{f}_n(x) = \frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right)$$

$$
\begin{aligned}
&= c_n \int_0^1 K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right) 1\left(\frac{x-y_0}{\sqrt{n}} - \frac{1}{c_n^{\gamma}} \le B(r) \le \frac{x-y_0}{\sqrt{n}} + \frac{1}{c_n^{\gamma}}\right) \\
&\quad + o_{a.s.}\left(\frac{1}{n^{1/4+\gamma/6-\varepsilon}}\right) \\
&\overset{a.s.}{\to} \overline{L}_B^{a,\kappa}.
\end{aligned} \tag{31}
$$

To show (30), note from the occupation times formula that the left side of (30) can be written as

$$
\begin{aligned}
& c_n \int_{-\infty}^{\infty} K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - p\right\}\right) 1\left(\frac{x-y_0}{\sqrt{n}} - \frac{1}{c_n^{\gamma}} \le p \le \frac{x-y_0}{\sqrt{n}} + \frac{1}{c_n^{\gamma}}\right) \overline{L}_B(1,p)dp \qquad (32) \\
= & \int_{-\infty}^{\infty} c_n K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - p\right\}\right) \overline{L}_B(1,p)dp + o_{a.s}(c_n^{1-2r(1-\gamma)})
\end{aligned}
$$

as in (27) above. However, according to (14)

$$
\frac{x-y_0}{\sqrt{n}} \overset{a.s.}{\to} a - B_0(\kappa) \tag{33}
$$

and, since $c_n = \sqrt{n}/h_n \to \infty$ as $n \to \infty$, $\overline{L}_B(1,p)$ is continuous in p, and $\int_{-\infty}^{\infty} K(s)ds = 1$, we have

$$
c_n \int_{-\infty}^{\infty} K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - p\right\}\right) \overline{L}_B(1,p)dp \to_{a.s.} \overline{L}_B\left(1, a - B_0(\kappa)\right) = \overline{L}_B^{a,\kappa},
$$

as required for (30), and (31) then follows.

5.10 Proof of Lemma 3.3

Start by considering the case where $a \neq 0$ and $\kappa > 0$. Let $\gamma \in (1/2, 1)$. Then, following the same line of argument as in the first part of the proof of Theorem 3.1, we find

$$
\begin{aligned}
\frac{1}{nh_n}\sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right) y_{t-1} &= \frac{1}{\sqrt{n}h_n}\sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right)\frac{y_{t-1}}{\sqrt{n}} \\
&= c_n \int_0^1 K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right)\left[B(r) + \frac{y_0}{\sqrt{n}}\right] dr \\
&\quad + o_{a.s.}\left(\frac{1}{n^{1/4+\gamma/6-\varepsilon}}\right) \qquad (34) \\
&= c_n \int_{-\infty}^{\infty}\left[p + \frac{y_0}{\sqrt{n}}\right]\left[K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - p\right\}\right)\overline{L}_B(1,p)\right] dp \\
&\quad + o_{a.s.}\left(\frac{1}{n^{1/4+\gamma/6-\varepsilon}}\right).
\end{aligned}
$$

Since $\frac{y_0}{\sqrt{n}} \to_{a.s.} B_0(\kappa)$, and in view of (33) and the continuity of $p\overline{L}_B(1,p)$, we obtain

$$
\frac{1}{nh_n}\sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right) y_{t-1} \to_{a.s.} a\overline{L}_B\left(1, a - B_0(\kappa)\right) = a\overline{L}_B^{a,\kappa},
$$

as required for (18).

When $a = 0$, and $\kappa = 0$, we proceed in the same manner, but use the normalization $1/\sqrt{nh_n^2}$, rather than $1/nh_n$ in (34) and let $\gamma \in (1/2, 1)$. Then, following the same approach, we have

$$\frac{1}{\sqrt{nh_n^2}} \sum_{t=1}^{n} K\left(\frac{x - y_{t-1}}{h_n}\right) y_{t-1} = c_n \int_0^1 K\left(c_n \left\{\frac{x - y_0}{\sqrt{n}} - B(r)\right\}\right) \left\{\sqrt{n}B(r) + y_0\right\} dr + o_{a.s.}\left(\frac{n^{1/p}}{n^{1/4+\gamma/6-\varepsilon}}\right). \tag{35}$$

For large enough p the remainder in (35) is negligible and we are left with

$$\begin{aligned} & c_n \int_0^1 K\left(c_n \left\{\frac{x - y_0}{\sqrt{n}} - B(r)\right\}\right) \left\{\sqrt{n}B(r) + y_0\right\} dr \\ = \ & c_n \int_{-\infty}^{\infty} \left(\sqrt{n}p + y_0\right) \left\{K\left(\frac{x - \sqrt{n}p - y_0}{h_n}\right) \overline{L}_B(1, p)\right\} dp \\ = \ & \int_{-\infty}^{\infty} \{x - h_n s\} K(s) \overline{L}_B\left(1, \frac{x - y_0}{\sqrt{n}} - \frac{s}{c_n}\right) ds \\ = \ & x \int_{-\infty}^{\infty} K(s)\overline{L}_B\left(1, \frac{x - y_0}{\sqrt{n}} - \frac{s}{c_n}\right) ds - h_n \int_{-\infty}^{\infty} sK(s) \overline{L}_B\left(1, \frac{x - y_0}{\sqrt{n}} - \frac{s}{c_n}\right) ds \end{aligned} \tag{36}$$

Now, since $x = x_0$, the first term in (36) is

$$x \int_{-\infty}^{\infty} K(s)\overline{L}_B\left(1, \frac{x - y_0}{\sqrt{n}} - \frac{s}{c_n}\right) ds \overset{a.s}{\to} x\overline{L}_B(1, 0) = x_0 \bar{L}_B^{0,0}, \tag{37}$$

which gives the required limit (17) in this case. The second term in (36) is negligible, as we now show. First, $\int_{-\infty}^{\infty} sK(s)ds\overline{L}_B\left(1, n^{-1/2}(x - y_0)\right) = 0$ by virtue of the fact that $\int_{-\infty}^{\infty} sK(s) = 0$ by the symmetry condition in Assumption 2.4. We may therefore write

$$\begin{aligned} & \left| h_n \int_{-\infty}^{\infty} sK(s)\overline{L}_B(1, n^{-1/2}(x - y_0) - s/c_n) ds \right| \\ = \ & h_n \left| \int_{-\infty}^{\infty} sK(s) \left\{\overline{L}_B\left(1, n^{-1/2}(x - y_0) - s/c_n\right) - \overline{L}_B\left(1, n^{-1/2}(x - y_0)\right)\right\} ds \right| \\ \leq \ & const.\ h_n \left(\frac{h_n}{\sqrt{n}}\right)^{1/2-\eta} \int_{-\infty}^{\infty} |s|^{3/2-\eta} |K(s)| ds, \quad a.s. \end{aligned} \tag{38}$$

for any $\eta > 0$, due to the Hölder continuity of $L_B(1, \cdot)$. The right side (38) tends to zero by virtue of Assumption 2.4(a) provided $n^{-1/2} h_n^{1+2/(1-2\eta)} \to 0$, which holds whenever $h_n = cn^k$ and $k < 1/6 - \delta$, since η can be arbitrarily small. Thus, from (36), (37), and (38) we get the convergence

$$\int_{-\infty}^{\infty} \{x - h_n s\} K(s)\overline{L}_B\left(1, \frac{x - y_0}{\sqrt{n}} - \frac{s}{c_n}\right) ds \overset{a.s}{\to} x_0 \bar{L}_B^{0,0}$$

for the same range of bandwidths h_n. This proves the stated result (17) for $x = x_0$ fixed and $y_0 = u$.

5.11 Proof of Lemma 3.4

Let $\gamma \in (1/2, 1)$ and following the same line of argument as in the proof of Theorem 3.1, we find

$$
\begin{aligned}
& \frac{\sqrt{c_n}}{h_n\sqrt{nh_n^2}} \sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right)(y_{t-1}-x) \\
= & \frac{c_n^{3/2}}{n} \sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right)\left(\frac{y_{t-1}-x}{h_n}\right) \\
= & c_n^{3/2}\frac{1}{n}\sum_{i=1}^{n} K\left(c_n\left\{\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right\}\right)\left(c_n\left\{\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right\}\right) 1\left[\left|\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right| \le \frac{1}{c_n^{\gamma}}\right] \\
& + c_n^{3/2}\frac{1}{n}\sum_{i=1}^{n} K\left(c_n\left\{\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right\}\right)\left(c_n\left\{\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right\}\right) 1\left[\left|\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right| > \frac{1}{c_n^{\gamma}}\right] \\
= & c_n^{3/2}\frac{1}{n}\sum_{i=1}^{n} K\left(c_n\left\{\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right\}\right)\left(c_n\left\{\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right\}\right) 1\left[\left|\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right| \le \frac{1}{c_n^{\gamma}}\right] \\
& + o_{a.s.}(c_n^{3/2+(1-\gamma)-2r(1-\gamma)}),
\end{aligned}
$$

and the error is arbitrarily small as $n \to \infty$ provided $r > \frac{1}{2} + \frac{3}{4(1-\gamma)}$, which can be arranged by suitable restriction of the kernel function by modifying Assumption 2.4 and this is assumed in what follows.

Proceeding as in first part of the proof of Theorem 3.1, we have

$$
\begin{aligned}
& \frac{\sqrt{c_n}}{h_n\sqrt{nh_n^2}} \sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right)(y_{t-1}-x) \\
= & c_n^{3/2}\int_0^1 K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right)\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right) \\
& \times 1\left(\frac{x-y_0}{\sqrt{n}} - \frac{1}{c_n^{\gamma}} \le B(r) \le \frac{x-y_0}{\sqrt{n}} + \frac{1}{c_n^{\gamma}}\right) dr + o_{a.s.}\left(\frac{c_n^{3/2}}{n}\frac{c_n^{1-\gamma}}{c_n^{\gamma/2}} n^{1/4+\varepsilon}\right) + o_{a.s.}(1) \\
= & c_n^{3/2}\int_0^1 K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right)\left(c_n\left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right) dr \\
& + o_{a.s.}\left(\frac{c_n^{\frac{5}{2}-\frac{3}{2}\gamma}}{n^{\frac{3}{4}}} n^{\varepsilon}\right) \qquad (39)
\end{aligned}
$$

Observe that

$$\frac{c_n^{\frac{5}{2}-\frac{3}{2}\gamma}}{n^{\frac{3}{4}}} n^{\varepsilon} = \frac{n^{\frac{5}{4}-\frac{3}{4}\gamma+\varepsilon}}{n^{\frac{3}{4}} h_n^{\frac{5}{2}-\frac{3}{2}\gamma}} = \frac{n^{\frac{1}{2}-\frac{3}{4}\gamma+\varepsilon}}{h_n^{\frac{5}{2}-\frac{3}{2}\gamma}}$$

and let $\gamma = 1-\delta$ for $\delta > 0$ small. Then

$$\frac{n^{\frac{1}{2}-\frac{3}{4}\gamma+\varepsilon}}{h_n^{\frac{5}{2}-\frac{3}{2}\gamma}} = \frac{n^{\frac{3}{4}\delta+\varepsilon}}{n^{\frac{1}{4}} h_n^{1+\frac{3}{2}\delta}}$$

Since ε is arbitrarily small and δ can be chosen arbitrarily small, we deduce that

$$\frac{n^{\frac{3}{4}\delta+\varepsilon}}{n^{\frac{1}{4}} h_n^{1+\frac{3}{2}\delta}} = o(1)$$

provided h_n is not too small, in particular for

$$h_n > \frac{1}{n^{\frac{1}{4}-\mu}}$$

for some arbirtrarily small $\mu > 0$. It follows that (39) holds with a negligible error as $n \to \infty$.

From the occupation formula we deduce that

$$\begin{aligned} & c_n^{3/2} \int_0^1 K\left(c_n \left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right)\left(c_n \left\{\frac{x-y_0}{\sqrt{n}} - B(r)\right\}\right) dr \\ = \; & c_n^{3/2} \int_{-\infty}^{\infty} K\left(c_n \left\{\frac{x-y_0}{\sqrt{n}} - p\right\}\right)\left(c_n \left\{\frac{x-y_0}{\sqrt{n}} - p\right\}\right) \overline{L}_B(1,p) dp, \end{aligned}$$

and transforming $p \mapsto s = c_n \left[(x-y_0)/\sqrt{n} - p\right]$, this expression becomes

$$-c_n^{1/2} \int_{-\infty}^{\infty} sK(s)\, \overline{L}_B\left(1, \left[\frac{x-y_0}{\sqrt{n}} - \frac{s}{c_n}\right]\right) ds. \tag{40}$$

Since $\int_{-\infty}^{\infty} sK(s)\, ds = 0$, (40) equals

$$\begin{aligned} & -\int_{-\infty}^{\infty} sK(s) c_n^{1/2} \left[\overline{L}_B\left(1, \left[\frac{x-y_0}{\sqrt{n}} - \frac{s}{c_n}\right]\right) - \overline{L}_B\left(1, \frac{x-y_0}{\sqrt{n}}\right)\right] ds \\ & \xrightarrow{d} -\sigma^{-1}\varphi^{1/2} U\left(\overline{L}_B\left(1, a - B_0(\kappa)\right)\right), \end{aligned} \tag{41}$$

in view of (11) of Lemma 2.10 and the fact that $n^{-1/2}(x-y_0) \xrightarrow{a.s.} a - B_0(\kappa)$. In the limit (41), φ is defined by $\varphi = -2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |a-b| abK(a)K(b) da db$, and $U(\cdot)$ is a standard Brownian motion independent of B. We deduce that

$$\frac{\sqrt{c_n}}{h_n\sqrt{nh_n^2}} \sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right)(y_{t-1}-x) \xrightarrow{d} -\varphi^{1/2} U\left(L_B^{a,\kappa}\right) \overset{d}{=} \varphi^{1/2}\sigma^{-1} U\left(\overline{L}_B^{a,\kappa}\right), \tag{42}$$

as required for (19).

5.12 Proof of Lemma 3.5

Following the line of argument used in the first part of the proof of Theorem 3.1, we find

$$\begin{aligned} & \frac{1}{\sqrt[4]{nh_n^2}} \sum_{t=1}^{n} K\left(\frac{x-y_{t-1}}{h_n}\right) u_t \\ = \; & c_n^{1/2} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} K\left(c_n\left\{\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right\}\right) 1\left[\left|\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right| \le \frac{1}{c_n^{\gamma}}\right] u_i \\ & + c_n^{1/2} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} K\left(c_n\left\{\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right\}\right) 1\left[\left|\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right| > \frac{1}{c_n^{\gamma}}\right] u_i \\ = \; & c_n^{1/2} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} K\left(c_n\left\{\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right\}\right) 1\left[\left|\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right| \le \frac{1}{c_n^{\gamma}}\right] u_i \\ & + O_p\left(c_n^{1/2} c_n^{1-2r(1-\gamma)}\right) \end{aligned}$$

$$= c_n^{1/2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}K\left(c_n\left\{\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right\}\right)1\left[\left|\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right|\leq\frac{1}{c_n^{\gamma}}\right]u_i$$
$$+O_p\left(c_n^{1/2}c_n^{1-2r(1-\gamma)}\right). \tag{43}$$

Now $c_n^{\frac{3}{2}-2r(1-\gamma)}=o(1)$ as $n\to\infty$, for $r>3/4(1-\gamma)$, which is hereby assumed, and then the last term of (43) can be neglected. The first term is a martingale with quadratic variation

$$\sigma^2 c_n\frac{1}{n}\sum_{i=1}^{n}K\left(c_n\left\{\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right\}\right)^2 1\left[\left|\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right|\leq\frac{1}{c_n^{\gamma}}\right].$$

As in the proof of Theorem 3.1 we can determine the asymptotic behaviour of this conditional variance function using the strong approximation. In particular, we have

$$\sigma^2 c_n\frac{1}{n}\sum_{i=1}^{n}K\left(c_n\left\{\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right\}\right)^2 1\left[\left|\frac{x-y_0-S_{i-1}}{\sqrt{n}}\right|\leq\frac{1}{c_n^{\gamma}}\right]$$
$$= \sigma^2 c_n\int_0^1 K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}}-B(r)\right\}\right)^2 1\left[\left|\frac{x-y_0}{\sqrt{n}}-B(r)\right|\leq\frac{1}{c_n^{\gamma}}\right]dr$$
$$+o_{a.s.}\left(\frac{1}{n^{1/4+\gamma/6-\varepsilon}}\right)$$
$$= c_n\int_{-\infty}^{\infty}K\left(c_n\left\{\frac{x-y_0}{\sqrt{n}}-p)\right\}\right)^2 1\left[\left|\frac{x-y_0}{\sqrt{n}}-p\right|\leq\frac{1}{c_n^{\gamma}}\right]L_B(1,p)\,dp+o_{a.s.}(1)$$
$$= \int_{-\infty}^{\infty}K(s)^2\,1\left[|s|\leq c_n^{1-\gamma}\right]L_B\left(1,\frac{x-y_0}{\sqrt{n}}-\frac{s}{c_n}\right)ds+o_{a.s.}(1)$$
$$\to_{a.s.}\left(\int_{-\infty}^{\infty}K(s)^2\,ds\right)L_B(1,a-B_0(\kappa))$$
$$= k_2 L_B^{a,\kappa}(1), \tag{44}$$

where $L_B^{a,\kappa}(r)=L_B(r,a-B_0(\kappa))$ and $k_2=\int_{-\infty}^{\infty}K(s)^2\,ds$.

It now follows from the generalized martingale central limit theorem (e.g Hall and Heyde, 1980, p.58) that the first term of (43) has a mixed normal limit distribution with conditional variance (44). Hence,

$$\frac{1}{\sqrt[4]{nh_n^2}}\sum_{t=1}^{n}K\left(\frac{x-y_{t-1}}{h_n}\right)u_t\to_d MN\left(0,k_2L_B^{a,\kappa}(1)\right),$$

or, equivalently,

$$\frac{1}{\sqrt[4]{nh_n^2}}\sum_{t=1}^{n}K\left(\frac{x-y_{t-1}}{h_n}\right)u_t\to_d V\left(k_2L_B^{a,\kappa}(1)\right),$$

where V is a Brownian motion that is independent of B. This gives the stated result.

Proof of Theorem 3.6 Write

$$m_n(x)=\frac{\sum_{t=1}^{n}K\left(\frac{x-y_{t-1}}{h_n}\right)y_{t-1}}{\sum_{t=1}^{n}K\left(\frac{x-y_{t-1}}{h_n}\right)}+\frac{\sum_{t=1}^{n}K\left(\frac{x-y_{t-1}}{h_n}\right)u_t}{\sum_{t=1}^{n}K\left(\frac{x-y_{t-1}}{h_n}\right)}. \tag{45}$$

From Theorem 3.1 and Lemma 3.3, we deduce that

$$\frac{\frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)y_{t-1}}{\frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)} \overset{a.s.}{\to} \frac{x\overline{L}_B^{0,\kappa}}{\overline{L}_B^{0,\kappa}} = x \quad \text{for } x = fixed, \tag{46}$$

and

$$\frac{\frac{1}{nh_n}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)y_{t-1}}{\frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)} \overset{a.s.}{\to} \frac{a\overline{L}_B^{a,\kappa}}{\overline{L}_B^{a,\kappa}} = a \quad \text{for } x = \sqrt{n}a.$$

Moreover, from Lemma 3.4 it follows that

$$\frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)u_t \overset{p}{\to} 0 \quad \text{for } x = \begin{cases} fixed \\ \sqrt{n}a \end{cases}$$

We deduce that

$$m_n(x) \overset{p}{\to} x \qquad \text{for } x = fixed,$$

and

$$\frac{1}{\sqrt{n}}m_n(x) \overset{p}{\to} a \quad \text{for } x = \sqrt{n}a,$$

as required.

Proof of Theorem 3.7 Start with the case where x is fixed. From (45) we have

$$\sqrt[4]{nh_n^2}\left\{m_n(x) - x\right\} = \frac{\frac{1}{\sqrt[4]{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)(y_{t-1}-x)}{\frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)} + \frac{\frac{1}{\sqrt[4]{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)u_t}{\frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)}. \tag{47}$$

In view of Lemma 3.4, the first term on the right side of (47) is

$$\frac{\sqrt[4]{nh_n^2}\left\{\frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)(y_{t-1}-x)\right\}}{\frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)} = O_p\left(\frac{h_n\sqrt[4]{nh_n^2}}{\sqrt{c_n}}\right) = O_p\left(h_n^2\right) \to 0.$$

when $h_n \to 0$. On the other hand, the second term of (47) has the following limit by virtue of Theorem 3.1 and Lemma 3.5

$$\frac{\frac{1}{\sqrt[4]{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)u_t}{\frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)} \overset{d}{\to} \frac{\left\{k_2\sigma^2\overline{L}_B^{0,\kappa}\right\}^{1/2}V(1)}{\overline{L}_B^{0.\kappa}} \overset{d}{=} \left\{k_2\sigma^2/\overline{L}_B^{0,\kappa}\right\}^{1/2}V(1).$$

Thus, when $h_n \to 0$ we have

$$\sqrt[4]{nh_n^2}\left\{m_n(x) - x\right\} \overset{d}{\to} \left\{k_2\sigma^2/L_B^{0,\kappa}\right\}^{1/2}V(1).$$

which proves (20).

When $h_n \to \infty$ the first term in (47) dominates and we have using (17) of Lemma 3.3 and (19) of Lemma 3.4

$$\frac{\sqrt{c_n}}{h_n}\{m_n(x)-x\} = \frac{\frac{\sqrt{c_n}}{h_n\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)(y_{t-1}-x)}{\frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)} + \frac{\frac{\sqrt{c_n}}{h_n\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)u_t}{\frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)} \tag{48}$$

$$= \frac{\frac{\sqrt{c_n}}{h_n\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)(y_{t-1}-x)}{\frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)} + O_p\left(\frac{1}{h_n^2}\right) \tag{49}$$

$$\xrightarrow{d} \frac{\left\{\varphi\sigma^{-2}\overline{L}_B^{0,\kappa}\right\}^{1/2}U(1)}{\overline{L}_B^{0,\kappa}} = \left\{\varphi\sigma^{-2}/\overline{L}_B^{0,\kappa}\right\}^{1/2}U(1),$$

which completes the proof of (23).

When $h_n = h =$ constant, both terms of (47) affect the limit distribution and we get

$$\sqrt{c_n}\{m_n(x)-x\} = \frac{\frac{\sqrt{c_n}}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)(y_{t-1}-x)}{\frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)} + \frac{\frac{\sqrt{c_n}}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)u_t}{\frac{1}{\sqrt{nh_n^2}}\sum_{t=1}^n K\left(\frac{x-y_{t-1}}{h_n}\right)}$$

$$\xrightarrow{d} \frac{h\left\{\varphi\sigma^{-2}\overline{L}_B^{0,\kappa}\right\}^{1/2}U(1)}{\overline{L}_B^{0,\kappa}} + \frac{h^{-1}\left\{k_2\sigma^2\overline{L}_B^{0,\kappa}\right\}^{1/2}V(1)}{\overline{L}_B^{0.\kappa}}$$

$$= h\left\{\varphi\sigma^{-2}/\overline{L}_B^{0,\kappa}\right\}^{1/2}U(1) + h^{-1}\left\{k_2\sigma^2/\overline{L}_B^{0,\kappa}\right\}^{1/2}V(1),$$

leading to (22).

The case where $x = \sqrt{n}a$ is proved in the same way and the limits involve $\overline{L}_B^{a,\kappa}$ rather than $\overline{L}_B^{0,\kappa}$.

6 References

Aït Sahalia, Y. (1996) "Nonparametric pricing of interest rate derivative securities". *Econometrica*, 64, 527-560.

Aït Sahalia, Y. (1997) "Maximum likelihood estimation of discretely sampled diffusions: a closed form approach". University of Chicago, mimeographed.

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

Bierens, H. (1978). "Kernel Estimators of Regression Functions," in T. Bewley (ed.), *Advances in Econometrics, Fifth World Congress, Vol. 1.* Cambridge: Cambridge University Press, pp. 99–144.

Collomb, G. (1985a). "Nonparametric time series analysis and prediction: uniform almost sure convergence of the window and k-NN autoregression estimates." *Statistics*, 16, 297-307.

Collomb, G. (1985b). "Nonparametric regression: an up-to-date bibliography." *Statistics*, 16, 309-324.

Collomb, G. and W. Hardle (1986). "Strong uniform convergence rates in robust nonparametric time series analysis and prediction: Kernel regression estimation from dependent observations," *Stochastic Processes and their Applications*, 23, 77–89.

Csörgõ, M. and L. Horváth (1993). *Weighted Approximations in Probability and Statistics.* Wiley: New York.

Hardle, W. (1990). *Applied Nonparametric Regression.* Cambridge: Cambridge University Press.

Hardle, W. and O. B. Linton (1995). "Applied nonparametric methods," in D.F. McFadden and R.F. Engle (eds.) *The Handbook of Econometrics Vol IV*, Amsterdam: North Holland.

Hardle, W. and P. Vieu (1992). "Kernel regression smoothing of time series," *Journal of Time Series Analysis*, 13, 209–232.

Ikeda, N. and S. Watanabe (1989). *Stochastic Differential Equations and Diffusion Processes.* Amsterdam: North Holland.

Jiang G. J. and J. L. Knight (1997) "A nonparametric approach to the estimation of diffusion processes, with an application to a short-term interest rate model". *Econometric Theory*, 13, 615-645.

Phillips P. C. B. and W. Ploberger (1996). "An Asymptotic Theory of Bayesian Inference for Time Series", *Econometrica*, 64, 381-413.

Revuz, D. and M. Yor (1991). *Continuous Martingale and Brownian Motion.* New York: Springer–Verlag.

Robinson, P. M. (1983). "Nonparametric estimators for time series," *Journal of Time Series Analysis*, 4, 185–207.

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

White, J. S. (1958). "The limiting distribution of the serial correlation coefficient in the explosive case," *Annals of Mathematical Statistics,* 29, 1188–1197.