9-1-2001

Fully Nonparametric Estimation of Scalar Diffusion Models

Federico M. Bandi

Peter C.B. Phillips

Fully Nonparametric Estimation of Scalar Diffusion Models

By

Federico M. Bandi and Peter C.B. Phillips

September 2001

COWLES FOUNDATION DISCUSSION PAPER NO. 1332

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY

Box 208281

New Haven, Connecticut 06520-8281

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

FULLY NONPARAMETRIC ESTIMATION OF SCALAR DIFFUSION MODELS*

FEDERICO M. BANDI
GSB, The University of Chicago
and
PETER C.B. PHILLIPS
Cowles Foundation, Yale University
University of Auckland & University of York

September 20, 2000
First Draft: November 1998

Abstract

We propose a functional estimation procedure for homogeneous stochastic differential equations based on a discrete sample of observations and with minimal requirements on the data generating process. We show how to identify the drift and diffusion function in situations where one or the other function is considered a nuisance parameter. The asymptotic behavior of the estimators is examined as the observation frequency increases and as the time span lengthens (that is, we implement both *infill* and *long span* asymptotics). We prove consistency and convergence to mixtures of normal laws, where the mixing variates depend on the *chronological local time* of the underlying process, that is the time spent by the process in the vicinity of a spatial point. The estimation method and asymptotic results apply to both stationary and nonstationary processes.

KEYWORDS: Diffusion, Drift, Infill asymptotics, Kernel density, Local time, Long span asymptotics, Martingale, Nonparametric estimation, Semimartingale, Stochastic differential equation.

JEL CLASSIFICATION: C14, C22

*We are grateful to the co-editor Alain Monfort and three anonymous referees for useful comments. Bandi thanks Mediocredito Centrale and the Sloan Foundation for Fellowship support and Phillips thanks the NSF for support under Grant No. SBR 97–30295. *Comments are welcome and can be sent to federico.bandi@gsb.uchicago.edu.*

1. INTRODUCTION

Many popular models in economics and finance, like those for pricing derivative securities, involve diffusion processes formulated in continuous-time as stochastic differential equations. These processes have been used to model options prices, the term structure of interest rates, exchange rates, and foreign currency interest rates, *inter alia*. A recent introduction to some of these applications is given in Baxter and Rennie (1996). Stochastic differential equations have also been used to model macroeconomic aggregates like consumption and investment, and systems of such equations have been used for many years to model economic activity at the national level, as described in Bergstrom (1988). In all these applications, statistical estimation involves the use of discrete data. It is then necessary to identify and estimate with discretely sampled observations the parameters and functionals of a process that is defined in continuous time.

The stochastic differential equation that defines a diffusion process, like X_t in (2.1) below, involves two components. These components measure the conditional drift, $\mu(X_t)$, and the conditional variation, $\sigma^2(X_t)$, of the process in the vicinity of each point visited by X_t. The most general approach to estimating stochastic differential equations is to avoid any functional form specification for the drift and the diffusion term. In some cases, attention may focus on one of the functions and it is then of interest to estimate it in the context of the other function being treated as a nuisance parameter. A substantial simplification to the estimation problem is obtained by the commonly made assumption of stationarity. Indeed, under stationarity and provided suitable regularity conditions are met, the marginal density of the process is fully characterized by the two functions of interest (e.g. see Karatzas and Shreve (1991) and Karlin and Taylor (1981)). This fact justifies some estimation methods that have appeared recently in the literature which exploit the restrictions imposed on the drift and diffusion function by virtue of the existence of a time-invariant density of the process (see, in particular, Aït-Sahalia (1996a,b) and Jiang and Knight (1997)). Notwithstanding the advantages of assuming stationarity, it would appear that, for many of the empirical applications mentioned in the preceding paragraph at least, it would be more appropriate to allow for martingale and other possible forms of nonstationary behavior in the process. In such cases, it becomes necessary to achieve identification without resorting to cross restrictions delivered from the existence of a time-invariant density and transitional density, and estimation and inference must be performed when such restrictions cannot be imposed, namely when the process is nonstationary. Of course, there may also be interest in testing either local or more general martingale behavior in the process.

The aim of the present paper is to construct a nonparametric estimation method for diffusion models without imposing a stationarity assumption. Recurrence, which is a substantially milder assumption than stationarity, is our identifying condition. In other words, we simply require the continuous trajectory of the process to visit any level in its range an infinity number of times over time. Our approach is a refined sample analog method, which builds local estimates of the drift and diffusion components from the local behavior of the process at each spatial point that the process visits. We assume that the process is discretely sampled, but we explore the limit theory of the proposed estimators as the sample frequency increases (i.e. as the interval between observations tends to zero, as in Florens-Zmirou (1993), Jacod (1997) and Jiang and Knight (1997)) and also as the total time span of observation lengthens. In technical terms this amounts to both *infill* and *long span* asymptotics. The twofold limit theory allows us to avoid the well-known *aliasing problem* (i.e. different continuous-time processes may be indistinguishable when sampled at discrete points in time) and be extremely general about the dynamic features of the underlying diffusion process (Phillips (1973) and Hansen and Sargent (1983) are early references on the aliasing phenomenom in the econometric literature on the identification of continuous-time Markov systems).

We give conditions for almost sure convergence of the proposed sample analog estimators to the theoretical functions and provide a limit distribution theory for the general case. The asymptotic distributions of the estimates are mixed normal and the mixture variates can be expressed in terms of the *chronological local time* (see Phillips and Park (1998)) of the underlying process, a random quantity that measures in chronological time units the amount of time the process spends in the vicinity of each spatial point. Our results also enable us to comment on the fixed time span situation. We confirm earlier findings that the diffusion term can be consistently estimated over a fixed time span (as in Florens-Zmirou (1993) and Jacod (1997), for example) and discuss the difference between this case and the long span situation. We also confirm that, in general, the drift term can not be identified nonparametrically on a fixed interval without cross-restrictions, no matter how frequently the data is sampled (c.f. Merton (1973), Aït-Sahalia (1996a) and Bandi (1998, theorem 2.1)). Despite this limitation, by letting the time span increase to infinity, the theoretical drift term can be recovered in the limit, provided the process continues to repeat itself, that is provided the process is recurrent. Geman (1979) utilized the same property but assumed the availability of a continuous record of observations. To our knowledge, our drift estimator is the first fully nonparametric estimator which permits identification of the drift function by use of discretely sampled data, without relying on cross-restrictions based on the existence of a time-invariant marginal density. It is therefore robust against deviations from stationarity.

Interestingly, both the nonparametric theory on the estimation of conditional expectations in the stationary discrete time framework (c.f. Pagan and Ullah (1999) for references) and the recent functional theory on the identification of conditional first moments in the unit root literature (c.f. Phillips and Park (1998)) are reflected in our general results which can be specialized to various forms of recurrent behavior and, in consequence, cover both the stationary case and the Brownian motion (unit root, that is) case in the existing nonparametric literature, *inter alia*.

Our work is presented as follows. Section 2 lays out the model and objects of interest. Section 3 gives some useful theoretical preliminaries. Section 4 contains a description of the methodology. Section 5 presents the main results and Section 6 concludes. Appendix A provides proofs and technicalities. Notation is laid out in Appendix B.

2. THE MODEL, ASSUMPTIONS AND OBJECTS OF INTEREST

The model we consider is the autonomous stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t, \tag{2.1}$$

with initial condition $X_0 = \overline{X}$ and where B_t is a standard Brownian motion defined on the filtered probability space $(\Omega, \Im^B, (\Im_t^B)_{t\geq 0}, P)$. The initial condition $\overline{X} \in L^2$ and is taken to be independent of $\{B_t : t \geq 0\}$. We define the left-continuous filtration

$$\overline{\Im}_t := \sigma(\overline{X}) \vee \Im_t^B = \sigma(\overline{X}, B_s; 0 \leq s \leq t) \qquad 0 \leq t < \infty$$

and the collection of null sets

$$\aleph := \{N \subseteq \Omega; \exists G \in \overline{\Im}_\infty \textit{ with } N \subseteq G \textit{ and } P(G) = 0\}.$$

We create the *augmented* filtration

$$\widetilde{\Im}_t^X := \sigma(\overline{\Im}_t \cup \aleph) \qquad 0 \leq t < \infty.$$

The following conditions will be used in the study of (2.1). They will assure the existence and pathwise uniqueness of a nonexplosive solution to (2.1) that is adapted to the *augmented* filtration $\{\widetilde{\Im}_t^X\}$.

Assumption 1:

(i) *$\mu(\cdot)$ and $\sigma(\cdot)$ are time-homogeneous, $\mathfrak{B}$-measurable functions on $\mathfrak{D} = (l, u)$ with $-\infty \leq l < u \leq \infty$ where $\mathfrak{B}$ is the σ-field generated by Borel sets on $\mathfrak{D}$. Both functions are at least once continuously differentiable. Hence, they satisfy local Lipschitz and growth conditions. Thus,* *for every compact subset J of the range of the process, there exist constants C_1 and C_2 such that, for all x and y in J,*

$$|\mu(x) - \mu(y)| + |\sigma(x) - \sigma(y)| \leq C_1|x - y|,$$

and

$$|\mu(x)| + |\sigma(x)| \leq C_2\{1 + |x|\}.$$

(ii) $\sigma^2(\cdot) > 0$ *on* $\mathfrak{D}$.

(iii) (Feller's (1952) necessary and sufficient condition for nonexplosion). *We define* $V(\alpha)$ *as*

$$\int_0^\alpha S'(y)\left\{\int_0^y \left[\frac{2}{S'(x)\sigma^2(x)}\right] dx\right\} dy$$

where $S'(x)$ *is the first derivative of the natural scale function,*

$$S(\alpha) = \int_0^\alpha \exp\left\{\int_0^y \left[-\frac{2\mu(x)}{\sigma^2(x)}\right] dx\right\} dy.$$

We require $V(\alpha)$ *to diverge at the boundaries of* $\mathfrak{D}$, *i.e.*

$$\lim_{\alpha \to l^+} V(\alpha) = \lim_{\alpha \to u^-} V(\alpha) = \infty.$$

Assumption (i) is sufficient for pathwise uniqueness of the solution to (2.1) (c.f. Karatzas and Shreve (1991, Theorem 5.2.5, page 287)). Assumptions (i) and (ii) assure the existence of a unique strong solution up to an explosion time (c.f. Karatzas and Shreve (1991, Theorem 5.5.15, page 341 and Corollary 5.3.23, page 310)). Assumption (iii) guarantees that neither l nor u are attained in finite time (c.f. Karatzas and Shreve (1991, Theorem 5.5.29, page 348)); and the same condition is necessary and sufficient for recurrence, meaning that, for each $c \in (l, u)$, there exist a sequence of times $\{t_i\}$ increasing to infinity such that $X_{t_i} = c$ for each i, almost surely.

Remark 1: Global Lipschitz and growth conditions are typically assumed to guarantee existence and uniqueness of a strong solution to (2.1) (c.f. Karatzas and Shreve (1991, Theorem 5.2.9, page 289), for example). We do not impose these conditions here because, as Aït-Sahalia (1996a,b) points out, they fail to be satisfied for interesting models in economics and finance.

Remark 2: Geman (1979) requires the natural scale measure $S(\alpha)$ to diverge to ∞ as $\alpha \to u$, and to $-\infty$ as $\alpha \to l$. Notice that this condition is only sufficient for nonexplosion and recurrence.

Feller's (1952) condition based on the function $V(\alpha)$ is necessary and sufficient. The following implications are easily derived (c.f. Karatzas and Shreve (1991, Problem 5.5.27, page 348)):

$$S(l^+) = -\infty \Rightarrow V(l^+) = \infty$$

and

$$S(u^-) = \infty \Rightarrow V(u^-) = \infty.$$

Thus, under conditions (i), (ii) and (iii), the stochastic differential equation has a strong solution X_t that is unique, recurrent and continuous in $t \in [0, T]$. X_t satisfies

$$X_t = X_0 + \int_0^t \mu(X_s)ds + \int_0^t \sigma(X_s)dB_s$$

a.s., with $\int_0^T \mathbf{E}[X_t^2]dt < \infty$.

The objects of econometric interest are the drift and diffusion terms in (2.1). These functions have the following definitions:

$$\mathbf{E}^x[X_t - x] = t\mu(x) + o(t) \tag{2.2}$$

$$\mathbf{E}^x[(X_t - x)^2] = t\sigma^2(x) + o(t) \tag{2.3}$$

where x is a generic initial condition and $\mathbf{E}^x$ is the expectation operator associated with the process started at x. Loosely speaking, (2.2) and (2.3) can be interpreted as representing the "instantaneous" conditional mean and the "instantaneous" conditional variance of the process when $X_t = x$. More precisely, (2.2) describes the conditional expected rate of change of the process for infinitesimal time changes, whereas (2.3) gives the conditional rate of change of volatility at x.

3. LOCAL TIME PRELIMINARIES

In what follows we introduce some preliminary results regarding the local or *sojourn* time of a continuous semimartingale (SMG). These results will be useful in the development of our analysis (Protter (1990) and Revuz and Yor (1998) are standard references).

DEFINITION 1: (CONTINUOUS SMG) *A continuous SMG is a continuous process M which can be written as $M = LM + A$ where LM is a continuous local martingale and A is a continuous adapted process of finite variation.*

Continuous-time stochastic differential equations like (2.1) are known to have solutions that are SMGs since $X_0 + \int_0^t \mu(X_s)ds$ is a continuous adapted process of finite variation and $\int_0^t \sigma(X_s)dB_s$ is a continuous local martingale. Hence, our theory comes within the ambit of SMG analysis. The local time of a continuous SMG M is defined as follows:

Definition 2: (The Tanaka Formula) *For any real number a, there exists a non-decreasing continuous process $L_M(.,a)$ called the local time of M at a, such that*

$$\begin{aligned} |M_t - a| &= |M_0 - a| + \int_0^t sgn(M_s - a)dM_s + L_M(t,a), \\ (M_t - a)^+ &= (M_0 - a)^+ + \int_0^t \mathbf{1}_{\{M_s > a\}} dM_s + \frac{1}{2} L_M(t,a), \\ (M_t - a)^- &= (M_0 - a)^- - \int_0^t \mathbf{1}_{\{M_s \le a\}} dM_s + \frac{1}{2} L_M(t,a). \end{aligned}$$

In particular, $|M_t - a|$, $(M_t - a)^+$ and $(M_t - a)^-$ are SMGs.

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

Lemma 2: (The Occupation Time Formula) *Let M be a continuous SMG with quadratic variation process $[M]_s$ and let L^a be the local time at a. Then,*

$$\int_0^t f(M_s, s)d[M]_s = \int_{-\infty}^{+\infty} da \int_0^t f(a,s)dL_M(s,a)$$

for every positive Borel measurable function f. If f is homogeneous, then the expression simplifies to

$$\int_0^t f(M_s)d[M]_s = \int_{-\infty}^{+\infty} f(a)L_M(t,a)da. \tag{3.1}$$

Lemma 3: *If M is a continuous SMG then, almost surely*

$$L_M(t,a) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^t \mathbf{1}_{[a,\ a+\varepsilon[}(M_s)d[M]_s \ \forall a, t. \tag{3.2}$$

If M is a continuous local martingale then, almost surely

$$L_M(t,a) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t \mathbf{1}_{]a-\varepsilon,\ a+\varepsilon[}(M_s)d[M]_s \ \forall a, t. \tag{3.3}$$

The process $L_M(t,a)$ is called the *local time* of M at the point a over the time interval $[0,t]$. It is measured in units of the quadratic variation process and gives the amount of time that the process spends in the vicinity of a. The *chronological local time* (terminology from Phillips and Park (1998)) is a standardized version of the conventional local time that is defined in terms of pure time units. It can be easily derived in the Brownian motion case. >From (3.3), the local time of a standard Brownian motion W is

$$L_W(t,a) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t \mathbf{1}_{(|W_s - a| < \varepsilon)} ds \text{ a.s. } \forall a, t.$$

Now, consider the Brownian motion $B = \sigma W$ with variance σ^2. We can write, as in Phillips and Park (1998),

$$L_B(t,a) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t \mathbf{1}_{(|B_s - a| < \varepsilon)} \sigma^2 ds = \sigma L_W\left(t, \frac{a}{\sigma}\right) \text{ a.s. } \forall a, t.$$

Since the quadratic variation of Brownian motion is deterministic, the chronological local time can be obtained as a scaled version of the conventional sojourn time as

$$\overline{L}_B(t,a) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t \mathbf{1}_{(|B_s - a| < \varepsilon)} ds = \sigma^{-2} L_B(t,a) \text{ a.s. } \forall a, t. \tag{3.4}$$

Equation (3.4) clarifies the sense in which $\overline{L}_B(t,a)$ measures the amount of time (out of t) that the process spends in the neighborhood of a generic spatial point a.

It turns out that a similar expression can be defined for more general processes such as those driven by stochastic differential equations like (2.1). In this case, the measure $d[X]_s$ is random and equal to $\sigma^2(X_s)ds$. Hence, given the limit operation, a natural way to define the *chronological local time* of a process like (2.1) is by

$$\overline{L}_X(t,a) = \frac{1}{\sigma^2(a)} \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^t \mathbf{1}_{[a,\ a+\varepsilon[}(X_s) \sigma^2(X_s) ds = \frac{1}{\sigma^2(a)} L_X(t,a) \text{ a.s. } \forall a, t. \tag{3.5}$$

This is the notion of local time that we will use extensively in what follows. It appears in other recent work on the nonparametric treatment of diffusion processes (Bosq (1998, p. 146) and Florens-Zmirou (1993)) where it is sometimes referred to simply as local time.

Lemma 4 and 5 below contain additional results that will be used in the development of our limit theory. Lemma 4 generalizes to diffusion processes the limit theory for Brownian local time (see Yor (1983), Revuz and Yor (1998) and Phillips and Park (1998)).

LEMMA 4: (LIMIT THEORY FOR THE LOCAL TIME OF A DIFFUSION) *Let X satisfy the properties in Section 2. Let r and $a > 0$ be fixed real numbers and treat $\{L_X(t, r + \frac{a}{\lambda}) - L_X(t,r)\}$ as a double indexed stochastic process in (t,a). Then, as $\lambda \to \infty$*

$$\frac{1}{2}\sqrt{\lambda}\left\{L_X(t, r + \frac{a}{\lambda}) - L_X(t,r)\right\} \Rightarrow \mathfrak{B}(L_X(t,r), a)$$

where $\mathfrak{B}(t,a)$ *is a standard Brownian sheet independent of* X. *If* $a<0$, *then*

$$\frac{1}{2}\sqrt{\lambda}\left\{L_X(t,r+\frac{a}{\lambda})-L_X(t,r)\right\}\Rightarrow\mathfrak{B}(L_X(t,r),-a).$$

Finally, Lemma 5 specializes to scalar diffusion processes a result that has wider applicability in the theory of occupation times for recurrent Markov processes (c.f. Revuz and Yor (1998)).

LEMMA 5: *Let* X *satisfy the properties in Section 2. Then, for any Borel measurable pair* $f(.)$ *and* $g(.)$ *that are integrable with respect to the speed measure* $s(dx)=\frac{2dx}{S'(x)\sigma^2(x)}$ *of* X *where* $S(x)$ *is the scale function (c.f. Assumption 1 (iii)), the ratio of the additive functionals* $\int_0^T f(X_s)ds$ *and* $\int_0^T g(X_s)ds$ *is such that*

$$\Pr\left[\lim_{T\to\infty}\frac{\int_0^T f(X_s)ds}{\int_0^T g(X_s)ds}=\frac{\int_{-\infty}^{\infty}f(x)s(dx)}{\int_{-\infty}^{\infty}g(x)s(dx)}\right]=1.$$

We now turn to the estimation method.

4. ECONOMETRIC ESTIMATION

Assume the process X_t is observed at $\{t=t_1,t_2,..,t_n\}$ in the time interval $[0,T]$, with $T\geq T_0>0$, where T_0 is a positive constant. Further assume that the observations are equispaced. Then, $\{X_t=X_{\Delta_{n,T}},X_{2\Delta_{n,T}},X_{3\Delta_{n,T}},...,X_{n\Delta_{n,T}}\}$ are n observations on the process X_t at $\{t_1=\Delta_{n,T},t_2=2\Delta_{n,T},t_3=3\Delta_{n,T},...,t_n=n\Delta_{n,T}\}$ where $\Delta_{n,T}=T/n$.

We want the number of sampled points (n) to increase as the time span (T) lengthens. We also want the frequency of observation to increase with n. Thus, we will explore the limit theory of the proposed estimators as $n\to\infty$, $T\to\infty$ and $\Delta_{n,T}=T/n\to 0$. We will also comment on the fixed T case where $T=\overline{T}$.

We propose the following estimators for (2.2) and (2.3).

$$\begin{aligned}\widehat{\mu}_{(n,T)}(x) &= \frac{\sum_{i=1}^n \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\left(\frac{1}{m_{n,T}(i\Delta_{n,T})\Delta_{n,T}}\sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1}[X_{t(i\Delta_{n,T})_j+\Delta_{n,T}}-X_{t(i\Delta_{n,T})_j}]\right)}{\sum_{i=1}^n \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}\\ &:= \frac{\sum_{i=1}^n \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\widetilde{\mu}_{n,T}(X_{i\Delta_{n,T}})}{\sum_{i=1}^n \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)},\end{aligned}\tag{4.1}$$

$$
\begin{aligned}
\widehat{\sigma}^2_{(n,T)}(x) &= \frac{\sum_{i=1}^n \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\left(\frac{1}{m_{n,T}(i\Delta_{n,T})\Delta_{n,T}}\sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1}[X_{t(i\Delta_{n,T})_j+\Delta_{n,T}} - X_{t(i\Delta_{n,T})_j}]^2\right)}{\sum_{i=1}^n \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} \\
&:= \frac{\sum_{i=1}^n \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\widetilde{\sigma}^2_{n,T}(X_{i\Delta_{n,T}})}{\sum_{i=1}^n \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}, \qquad (4.2)
\end{aligned}
$$

where $\mathbf{K}(.)$ is a standard kernel function whose properties are specified below. In the above formulae, $\{t(i\Delta_{n,T})_j\}$ is a sequence of random times defined in the following manner:

$$
t(i\Delta_{n,T})_0 = \inf\{t \geq 0 : |X_t - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}\},
$$

and

$$
t(i\Delta_{n,T})_{j+1} = \inf\{t \geq t(i\Delta_{n,T})_j + \Delta_{n,T} : |X_t - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}\}.
$$

The number $m_{n,T}(i\Delta_{n,T}) \leq n$ counts the stopping times associated with the value $X_{i\Delta_{n,T}}$ and is defined as

$$
m_{n,T}(i\Delta_{n,T}) = \sum_{j=1}^n \mathbf{1}_{\left[\left|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}\right|\leq\varepsilon_{n,T}\right]},
$$

where $\mathbf{1}_A$ denotes the indicator of A. The quantity $\varepsilon_{n,T}$ is a bandwidth-like parameter that is taken to depend on the time span and on the sample size. We call this parameter the spatial bandwidth. As usual, the random time $t(i\Delta_{n,T})$ is defined on Ω and takes values on $[0,\infty)$. Further, $\{t(i\Delta_{n,T}) < t^*\} \in \Im^X_{t+^*}$ where $\Im^X_{t+^*}$is a right-continuous filtration defined as $\bigcap_{u>t^*} \Im^X_u$.

The kernel $\mathbf{K}(\cdot)$ that appears in (4.1) and (4.2) is assumed to satisfy the following condition.

Assumption 2: *The kernel* $\mathbf{K}(.)$ *is a continuous differentiable, symmetric and nonnegative function whose derivative* $\mathbf{K}'$ *is absolutely integrable and for which*

$$
\int_{-\infty}^{\infty}\mathbf{K}(s)ds = 1, \quad \int_{-\infty}^{\infty}\mathbf{K}^2(s)ds < \infty, \quad \sup_s \mathbf{K}(s) < C_3,
$$

and

$$
\int_{-\infty}^{\infty} s^2\mathbf{K}(s)ds < \infty.
$$

The method hinges on the simultaneous operation of *infill* and *long span* asymptotics. The intuition underlying the construction of (4.1) and (4.2) is fairly clear. By using observations over a lengthening time span as well as of increasing frequency we aim to "reconstruct" as well as possible the path of the process in terms of the key objects of interest, the drift and diffusion functions, which vary over the path. The idea is twofold.

First, the use of local averaging and stopping times in the algorithm is designed to replicate as well as possible the instantaneous features of the actual functions. Notice, in fact, that the components $\widetilde{\sigma}^2_{n,T}(X_{i\Delta_{n,T}})$ and $\widetilde{\mu}_{n,T}(X_{i\Delta_{n,T}})$ in (4.1) and (4.2) are defined as empirical analogs to the true functions for all i. Further, the estimates $\widetilde{\sigma}^2_{n,T}(X_{i\Delta_{n,T}})$ and $\widetilde{\mu}_{n,T}(X_{i\Delta_{n,T}})$ are consistent for $\sigma^2(X_{i\Delta_{n,T}})$ and $\mu(X_{i\Delta_{n,T}})$ as the random quantity $m_{n,T}(i\Delta_{n,T})$ goes to infinity $\forall i$. Under suitable conditions on the bandwidths, $m_{n,T}(i\Delta_{n,T})$ diverges to infinity almost surely when $T \to \infty$. In particular, given appropriate choices of the smoothing sequences, divergence occurs when the process X_t is recurrent, as it is under Condition (iii) in Assumption 1. In this case, the process almost surely hits any point in its range an infinite number of times, i.e. $P_x\{X_t$ hits z at a sequence of times increasing to $\infty\} = 1$, $\forall x, z$ (here x represents possible initializations of the process X_t).

Second, we apply standard nonparametric smoothing to recover the two functions of interest from the crude estimates $\widetilde{\sigma}^2_{n,T}(X_{i\Delta_{n,T}})$ and $\widetilde{\mu}_{n,T}(X_{i\Delta_{n,T}})$ calculated at the sample points.

5. MAIN RESULTS

5.1. *Some Preliminary Theory*

We start with the following preliminary result. Throughout, we assume that Assumptions 1 and 2 hold.

THEOREM 1: (ALMOST SURE CONVERGENCE TO THE CHRONOLOGICAL LOCAL TIME) *Given $n \to \infty$, T fixed ($= \overline{T}$) and $h_{n,\overline{T}} \to 0$ (as $n \to \infty$) in such a way that $\frac{1}{h_{n,\overline{T}}}(\Delta_{n,\overline{T}})^{\alpha} = O(1)$ for some $\alpha \in (0, \frac{1}{2})$, the estimator $\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n} \mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)$ converges to $\overline{L}_X(\overline{T}, x)$ a.s.*

REMARK 3: Theorem 1 is general enough to be applicable to *transient* processes. The following Corollary illustrates the difference between the two cases when we let T go to infinity.

COROLLARY 1: *If $T \to \infty$ with n but $\frac{T}{n} = \Delta_{n,T} \to 0$ and $h_{n,T} \to 0$ (as $n \to \infty$) in such a way that $\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T})^{\alpha} = O_{a.s.}(1)$ for some $\alpha \in \left(0, \frac{1}{2}\right)$ $\forall x \in \mathfrak{D}$, then*

$$\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n} \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right) \overset{a.s.}{\to} \overline{L}_X(\sup\{t : X_t = x\}, x).$$

Further, if the process is recurrent, then $\overline{L}_X(\sup\{t : X_t = x\}, x) = \infty$ a.s.

5.2. *Function Estimation of the Drift*

We next develop the asymptotic theory for the drift estimator.

Theorem 2: (Almost Sure Convergence to the Drift Term) *Given $n \to \infty$, $T \to \infty$, $\Delta_{n,T} \to 0$ and $h_{n,T} \to 0$ (as $n,T \to \infty$) such that $\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T})^{\alpha} = O_{a.s.}(1)$ for some $\alpha \in \left(0,\frac{1}{2}\right)$ $\forall x \in \mathfrak{D}$, and provided $\varepsilon_{n,T} \to 0$ (as $n,T \to \infty$) such that $\frac{\overline{L}_X(T,x)}{\varepsilon_{n,T}}(\Delta_{n,T})^{\beta} = O_{a.s.}(1)$ for some $\beta \in \left(0,\frac{1}{2}\right)$ and $\varepsilon_{n,T}\overline{L}_X(T,x) \overset{a.s.}{\to} \infty$ $\forall x \in \mathfrak{D}$, the estimator (4.1) converges to the true function with probability one.*

Theorem 3: (The Limiting Distribution of the Drift Estimator) *Given $n \to \infty$, $T \to \infty$, $\Delta_{n,T} \to 0$, $h_{n,T} \to 0$ (as $n,T \to \infty$) such that $\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T})^{\alpha} = O_{a.s.}(1)$ for some $\alpha \in \left(0,\frac{1}{2}\right)$ $\forall x \in \mathfrak{D}$, and provided $\varepsilon_{n,T} \to 0$ (as $n,T \to \infty$) such that $\frac{\overline{L}_X(T,x)}{\varepsilon_{n,T}}(\Delta_{n,T})^{\beta} = O_{a.s.}(1)$ for some $\beta \in \left(0,\frac{1}{2}\right)$, $\varepsilon_{n,T}\overline{L}_X(T,x) \overset{a.s.}{\to} \infty$ and $\varepsilon_{n,T}^5\overline{L}_X(T,x) \overset{a.s.}{\to} 0$ $\forall x \in \mathfrak{D}$, then the asymptotic distribution of the drift function estimator is of the form*

$$\sqrt{\varepsilon_{n,T}\overline{L}_X(T,x)}\left\{\widehat{\mu}_{(n,T)}(x) - \mu(x)\right\} \Rightarrow \mathbf{N}\left(0, \mathbf{K}_2^{ind}\sigma^2(x)\right), \tag{5.1}$$

where $\mathbf{K}_2^{ind} = \frac{1}{4}\int_{-\infty}^{\infty}\mathbf{1}_{\{|a|\leq 1\}}^2 da = \frac{1}{2}$ if $h_{n,T} = o(\varepsilon_{n,T})$. If $h_{n,T} = O(\varepsilon_{n,T})$ with $h_{n,T}/\varepsilon_{n,T} \to \phi > 0$, then

$$\sqrt{\varepsilon_{n,T}\overline{L}_X(T,x)}\left\{\widehat{\mu}_{(n,T)}(x) - \mu(x)\right\} \Rightarrow \mathbf{N}\left(0, \frac{1}{2}\theta_{\phi}\sigma^2(x)\right), \tag{5.2}$$

where $\theta_{\phi} = \frac{1}{2}\int_{-\infty}^{\infty}\int_{(z-1)/\phi}^{(z+1)/\phi}\int_{(z-1)/\phi}^{(z+1)/\phi}\mathbf{K}(a)\mathbf{K}(e)dzdade$.

Under the same conditions, but provided $\varepsilon_{n,T}^5\overline{L}_X(T,x) = O_{a.s.}(1)$ $\forall x \in \mathfrak{D}$, the limiting distribution of the drift estimator displays an asymptotic bias term whose form is

$$\Gamma_{\mu}(x) = \varepsilon_{n,T}^2\mathbf{K}_1^{ind}\left[\mu^{'}(x)\frac{s^{'}(x)}{s(x)} + \frac{1}{2}\mu^{''}(x)\right] \tag{5.3}$$

with $\mathbf{K}_1^{ind} = \frac{1}{2}\int_{-\infty}^{\infty}a^2\mathbf{1}_{\{|a|\leq 1\}}da = \frac{1}{3}$, provided $h_{n,T} = o(\varepsilon_{n,T})$, and

$$\Gamma_{\mu}^{(\phi)}(x) = \varepsilon_{n,T}^2\left(\mathbf{K}_1\phi^2 + \mathbf{K}_1^{ind}\right)\left[\mu^{'}(x)\frac{s^{'}(x)}{s(x)} + \frac{1}{2}\mu^{''}(x)\right], \tag{5.4}$$

with $\mathbf{K}_1 = \int_{-\infty}^{\infty}a^2\mathbf{K}(a)da$, provided $h_{n,T} = O(\varepsilon_{n,T})$ with $h_{n,T}/\varepsilon_{n,T} \to \phi > 0$. The function $s(x)$ in (5.3) and (5.4) is the speed function of the process X, namely $s(x) = \frac{2}{S^{'}(x)\sigma^2(x)}$ (c.f. Lemma 5).

Remark 4: (The Fixed T Case) If we fix the time span T the drift function cannot be identified. In particular, the drift estimator would diverge at a speed equal to $\frac{1}{\sqrt{\varepsilon_{n,T}}}$ (c.f. Theorem

2.1 in Bandi (1998)). However, if we do not constrain the time span to be fixed, by virtue of recurrence, there are repeated visits to every level over time and this opens up the possibility of recovering the true function by using a single trajectory of the process over a long time, through a combination of *infill* and *long* span asymptotics. Since the local dynamics of the underlying continuous process reflect more of the features of the diffusion function than those of the drift, only the diffusion function estimator can be meaningfully defined over a fixed time span of observations as we will see in the sequel (c.f. Geman (1979) and Merton (1973), *inter alia*).

Remark 5: (The Rate of Convergence) The normalizations in (5.1) and (5.2) are random because of the presence of the local time factor $\left(\overline{L}_X(T,x)\right)^{1/2}$. In general, therefore, the rate of convergence will be path-dependent. The precise rate of convergence in (5.1) and (5.2) will depend on the asymptotic divergence characteristics of the chronological local time of the process $\{X_t; t \geq 0\}$. We consider the two cases for which closed-form expressions for the rates of convergence exist: Brownian motion and the wide class of stationary processes. First, assume X_t is a Brownian motion (i.e. $\mu(X) = 0$ and $\sigma(X) = \sigma$). Then,

$$\overline{L}_X(T,x) = \overline{L}_B(T,x) = T^{1/2}\frac{1}{\sigma}L_W\left(1, \frac{a}{T^{1/2}\sigma}\right) = O_{a.s.}(T^{1/2}).$$

In this case, the convergence rate of $\widehat{\mu}_{(n,T)}(x)$ is $\sqrt{\varepsilon_{n,T}T^{1/2}}$, the asymptotic distribution is mixed normal and the limiting variance depends inversely on the local time of the underlying standard Brownian motion at the origin and time 1. Now consider the class of stationary processes. For any strictly stationary real ergodic process, Bosq (Theorem 6.3, 1998, page 150) proves that

$$\frac{\overline{L}_X(T,x)}{T} \overset{a.s.}{\to} f(x),$$

where $f(x)$ is the time-invariant stationary distribution of the process at x. As expected, for stationary processes the rate of convergence is faster than in the Brownian motion case, i.e. $\sqrt{\varepsilon_{n,T}T}$, the distribution is normal and the limiting variance depends inversely on the marginal distribution function of X_t.

Remark 6: (Single Smoothing) We can simplify (4.1) above and write the estimator as a weighted average of differences with weights based on simple kernels. Consider

$$\overline{\mu}_{(n,T)}(x) = \frac{1}{\Delta_{n,T}}\frac{\sum_{i=1}^{n-1}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{g_{n,T}}\right)\left(X_{(i+1)\Delta_{n,T}} - X_{i\Delta_{n,T}}\right)}{\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{g_{n,T}}\right)}. \tag{5.5}$$

The limit theory in this paper and Bandi (2000) allows us to show that $\overline{\mu}_{(n,T)}(x)$ is consistent almost surely for the unknown drift function provided the window width $g_{n,T}$ is such that $\frac{\overline{L}_X(T,x)}{g_{n,T}}(\Delta_{n,T})^{\beta} =$

$O_{a.s.}(1)$ for some $\beta \in \left(0, \frac{1}{2}\right)$ and $g_{n,T}\overline{L}_X(T,x) \stackrel{a.s.}{\to} \infty \; \forall x \in \mathfrak{D}$ as $n, T \to \infty$ with $\frac{T}{n} \to 0$. Furthermore, if $g_{n,T}^5 \overline{L}_X(T,x) \stackrel{a.s.}{\to} 0 \; \forall x \in \mathfrak{D}$, then

$$\sqrt{g_{n,T}\overline{L}_X(T,x)} \left\{ \overline{\mu}_{(n,T)}(x) - \mu(x) \right\} \Rightarrow \mathbf{N}\left(0, \mathbf{K}_2 \sigma^2(x)\right),$$

where $\mathbf{K}_2 = \int_{-\infty}^{\infty} \mathbf{K}^2(s)ds$. Additionally, if $g_{n,T}^5 \overline{L}_X(T,x) = O_{a.s.}(1) \; \forall x \in \mathfrak{D}$, then

$$\sqrt{g_{n,T}\overline{L}_X(T,x)} \left\{ \overline{\mu}_{(n,T)}(x) - \mu(x) - \Gamma_\mu(x) \right\} \Rightarrow \mathbf{N}\left(0, \mathbf{K}_2 \sigma^2(x)\right),$$

where

$$\Gamma_\mu(x) = (g_{n,T})^2 \, \mathbf{K}_1 \left[\mu^{'}(x) \frac{s^{'}(x)}{s(x)} + \frac{1}{2} \mu^{''}(x) \right],$$

$s(x)$ is the speed function of the process X and $\mathbf{K}_1 = \int_{-\infty}^{\infty} s^2 \mathbf{K}(s) ds$.

It is noted that (5.5) behaves asymptotically like (4.1) in the case where $h_{n,T} = o(\varepsilon_{n,T})$ and (4.1) is originated from a smooth kernel convoluted with another smooth kernel rather than with an indicator function as in our original formulation. In other words, single-smoothing is the same as double-smoothing asymptotically if $h_{n,T}/\varepsilon_{n,T} \to \phi = 0$. If $h_{n,T}/\varepsilon_{n,T} \to \phi \geq 0$, then double-smoothing offers additional flexibility over its simple counterpart. In fact, the parameter θ_ϕ (which affects the asymptotic variance) is a decreasing function of the constant ϕ, whereas the parameter $\mathbf{K}_\phi = \mathbf{K}_1 \phi^2 + \mathbf{K}_1^{ind}$ (which affects the asymptotic bias) is an increasing function of the same constant. For some processes and some levels x, appropriate choice of the smoothing sequences (and, consequently, appropriate choice of ϕ) can improve the limiting trade-off between bias and variance delivering an asymptotic mean-squared error that is minimized at values ϕ that are strictly larger than 0 (as would be the case in the single-smoothing case). Notice that if $h_{n,T}/\varepsilon_{n,T} \to \phi = 0$ and $\varepsilon_{n,T}^5 \overline{L}_X(T,x) \stackrel{a.s.}{\to} 0$ (which implies undersmoothing with respect to the optimal bandwidth, i.e. $\varepsilon_{n,T}^5 \overline{L}_X(T,x) \stackrel{a.s.}{=} O(1)$), then the asymptotic bias of our double-smoothed estimator is zero, while the limiting variance is $\frac{1}{2}\sigma^2(x)$. These are the same limiting bias and variance of the single-smoothed estimator originated using an indicator kernel. If $h_{n,T}/\varepsilon_{n,T} \to \phi > 0$ and $\varepsilon_{n,T}^5 \overline{L}_X(T,x) \stackrel{a.s.}{\to} 0$, then the limiting bias remains zero but the limiting variance becomes $\frac{1}{2}\theta_\phi \sigma^2(x)$ which is strictly smaller than $\frac{1}{2}$. In other words, for suboptimal bandwidth choices, which are usually implemented to eliminate the bias term and center the limiting distribution around zero, double-smoothing guarantees a smaller asymptotic mean-squared error than single-smoothing for any processes and any level x.

The finite sample benefits of convoluted kernels for drift estimation are discussed in a recent simulation study by Bandi and Nguyen (2000).

5.3. *Function Estimation of the Diffusion*

We now turn to the asymptotic theory for the diffusion estimator (4.2).

Theorem 4: (Almost Sure Convergence of the Diffusion Estimator) *Given* $n \to \infty$, $T \to \infty$, $\Delta_{n,T} \to 0$ *and* $h_{n,T} \to 0$ *(as* $n, T \to \infty$*) such that* $\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T})^{\alpha} = O_{a.s.}(1)$ *for some* $\alpha \in \left(0, \frac{1}{2}\right)$ $\forall x \in \mathfrak{D}$*, and provided* $\varepsilon_{n,T} \to 0$ *(as* $n, T \to \infty$*) such that* $\frac{\overline{L}_X(T,x)}{\varepsilon_{n,T}}(\Delta_{n,T})^{\beta} = O_{a.s.}(1)$ *for some* $\beta \in \left(0, \frac{1}{2}\right)$ $\forall x \in \mathfrak{D}$*, the estimator (4.2) converges to the true function with probability one.*

Theorem 5: (Limiting Distribution of the Diffusion Estimator) *Assume* $n \to \infty$, $T \to \infty$, $\Delta_{n,T} \to 0$, $h_{n,T} \to 0$ *(as* $n, T \to \infty$*) such that* $\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T})^{\alpha} = O_{a.s.}(1)$ *for some* $\alpha \in \left(0, \frac{1}{2}\right)$ $\forall x \in \mathfrak{D}$*. Also, assume* $\varepsilon_{n,T} \to 0$ *(as* $n, T \to \infty$*) such that* $\frac{\overline{L}_X(T,x)}{\varepsilon_{n,T}}(\Delta_{n,T})^{\beta} = O_{a.s.}(1)$ *for some* $\beta \in \left(0, \frac{1}{2}\right)$, $\varepsilon_{n,T}\overline{L}_X(T,x) \overset{a.s.}{\to} 0$ *and* $\frac{\varepsilon^5_{n,T}\overline{L}_X(T,x)}{\Delta_{n,T}} \overset{a.s.}{\to} 0$ $\forall x \in \mathfrak{D}$*. Then, the asymptotic distribution of the diffusion function estimator is of the form*

$$\sqrt{\frac{\varepsilon_{n,T}\overline{L}_X(T,x)}{\Delta_{n,T}}}\left\{\widehat{\sigma}^2_{(n,T)}(x) - \sigma^2(x)\right\} \Rightarrow \mathbf{N}\left(0, 4\mathbf{K}^{ind}_2\sigma^4(x)\right), \tag{5.6}$$

where $\mathbf{K}^{ind}_2 = \frac{1}{4}\int_{-\infty}^{\infty}\mathbf{1}^2_{\{|a|\le 1\}}da = \frac{1}{2}$ *if* $h_{n,T} = o(\varepsilon_{n,T})$*. If* $h_{n,T} = O(\varepsilon_{n,T})$ *with* $h_{n,T}/\varepsilon_{n,T} \to \phi > 0$*, then*

$$\sqrt{\frac{\varepsilon_{n,T}\overline{L}_X(T,x)}{\Delta_{n,T}}}\left\{\widehat{\sigma}^2_{(n,T)}(x) - \sigma^2(x)\right\} \Rightarrow \mathbf{N}\left(0, 2\theta_{\phi}\sigma^4(x)\right), \tag{5.7}$$

where $\theta_{\phi} = \frac{1}{2}\int_{-\infty}^{\infty}\int_{(z-1)/\phi}^{(z+1)/\phi}\int_{(z-1)/\phi}^{(z+1)/\phi}\mathbf{K}(a)\mathbf{K}(e)dzdade.$

Under the same conditions, but provided $\frac{\varepsilon^5_{n,T}\overline{L}_X(T,x)}{\Delta_{n,T}} = O_{a.s.}(1)$ $\forall x \in \mathfrak{D}$*, the limiting distribution of the diffusion estimator displays an asymptotic bias term whose form is*

$$\Gamma_{\sigma^2}(x) = \varepsilon^2_{n,T}\mathbf{K}^{ind}_1\left[\left(\sigma^2(x)\right)'\frac{s'(x)}{s(x)} + \frac{1}{2}\left(\sigma^2(x)\right)''\right] \tag{5.8}$$

with $\mathbf{K}^{ind}_1 = \frac{1}{2}\int_{-\infty}^{\infty}a^2\mathbf{1}_{\{|a|\le 1\}}da = \frac{1}{3}$*, provided* $h_{n,T} = o(\varepsilon_{n,T})$*, and*

$$\Gamma^{(\phi)}_{\sigma^2}(x) = \varepsilon^2_{n,T}\left(\mathbf{K}_1\phi^2 + \mathbf{K}^{ind}_1\right)\left[\left(\sigma^2(x)\right)'\frac{s'(x)}{s(x)} + \frac{1}{2}\left(\sigma^2(x)\right)''\right], \tag{5.9}$$

with $\mathbf{K}_1 = \int_{-\infty}^{\infty}a^2\mathbf{K}(a)da$*, provided* $h_{n,T} = O(\varepsilon_{n,T})$ *with* $h_{n,T}/\varepsilon_{n,T} \to \phi > 0$*. The function* $s(x)$ *in (5.8) and (5.9) is the speed function of the process* X*, namely* $s(x) = \frac{2}{S'(x)\sigma^2(x)}$ *(c.f. Lemma 5).*

We now consider the fixed $T = \overline{T}$ case.

Theorem 6: (Limiting Distribution of the Diffusion Estimator For a Fixed Time Span T) *Given $n \to \infty$, $T = \overline{T}$ and $h_{n,\overline{T}} \to 0$ (as $n \to \infty$) such that $\frac{1}{h_{n,\overline{T}}}(\Delta_{n,\overline{T}})^{\alpha} = O(1)$ for some $\alpha \in \left(0, \frac{1}{2}\right)$, and provided $\varepsilon_{n,\overline{T}} \to 0$ (as $n \to \infty$) such that $\frac{1}{\varepsilon_{n,\overline{T}}}(\Delta_{n,\overline{T}})^{\beta} = O(1)$ for some $\beta \in \left(0, \frac{1}{2}\right)$, the estimator (4.2) converges to the true function with probability one.*

If $h_{n,\overline{T}} = o(\varepsilon_{n,\overline{T}})$ and $n\varepsilon^4_{n,\overline{T}} \to 0$, then the asymptotic distribution of the diffusion function estimator is driven by a "martingale" effect and has the form

$$\sqrt{n\varepsilon_{n,\overline{T}}}\left\{\widehat{\sigma}^2_{(n,\overline{T})}(x) - \sigma^2(x)\right\} \Rightarrow \mathbf{MN}\left(0, \frac{2\sigma^4(x)}{\overline{L}_X(\overline{T}, x)/\overline{T}}\right), \tag{5.10}$$

If $h_{n,\overline{T}} = o(\varepsilon_{n,\overline{T}})$ and $n\varepsilon^4_{n,\overline{T}} \to \infty$, then the asymptotic distribution of the diffusion function estimator is driven by a "bias" effect and has the form

$$\frac{1}{\varepsilon^{3/2}_{n,\overline{T}}}\left\{\widehat{\sigma}^2_{(n,\overline{T})}(x) - \sigma^2(x)\right\} \Rightarrow \mathbf{MN}\left(0, 16\varphi^{ind}\frac{\left(\sigma'(x)\right)^2}{\overline{L}_X(\overline{T}, x)}\right), \tag{5.11}$$

where $\varphi^{ind} = 2\int_0^\infty\int_0^\infty \left(\frac{1}{2}\mathbf{1}_{\{|a|\leq 1\}}\right)\left(\frac{1}{2}\mathbf{1}_{\{|b|\leq 1\}}\right)\min(a,b)dadb$.

If $h_{n,\overline{T}} = O(\varepsilon_{n,\overline{T}})$ with $h_{n,\overline{T}}/\varepsilon_{n,\overline{T}} \to \phi > 0$ and $n\varepsilon^4_{n,\overline{T}} \to 0$, then the asymptotic distribution of the diffusion function estimator is driven by a "martingale" effect and is of the form

$$\sqrt{n\varepsilon_{n,\overline{T}}}\left\{\widehat{\sigma}^2_{(n,\overline{T})}(x) - \sigma^2(x)\right\} \Rightarrow \mathbf{MN}\left(0, \frac{2\theta_\phi\sigma^4(x)}{\overline{L}_X(\overline{T}, x)/\overline{T}}\right), \tag{5.12}$$

where $\theta_\phi = \frac{1}{2}\int_{-\infty}^{\infty}\int_{(z-1)/\phi}^{(z+1)/\phi}\int_{(z-1)/\phi}^{(z+1)/\phi}\mathbf{K}(a)\mathbf{K}(e)dzdade$.

If $h_{n,\overline{T}} = O(\varepsilon_{n,\overline{T}})$ with $h_{n,\overline{T}}/\varepsilon_{n,\overline{T}} \to \phi > 0$ and $n\varepsilon^4_{n,\overline{T}} \to \infty$, then the asymptotic distribution of the diffusion function estimator is driven by a "bias" effect and is of the form

$$\frac{1}{\varepsilon^{3/2}_{n,\overline{T}}}\left\{\widehat{\sigma}^2_{(n,\overline{T})}(x) - \sigma^2(x)\right\} \Rightarrow \mathbf{MN}\left(0, 16\left(\varphi^{ind,\mathbf{K}}(\phi)\right)\frac{\left(\sigma'(x)\right)^2}{\overline{L}_X(\overline{T}, x)}\right), \tag{5.13}$$

where $\varphi^{ind,\mathbf{K}}(\phi)$ is a positive function of ϕ (c.f. Proof of Theorem 6) such that $\varphi^{ind,\mathbf{K}}(\phi) \to \varphi^{ind}$ as $\phi \to 0$.

Remark 7: The statement of Theorem 6 uses the terms 'bias' effect and 'martingale' effect to refer to the principal terms that govern the asymptotic distribution. These effects are revealed in the proof of the theorem. The essential factor governing the magnitude of the two effects is the relation of the observation rate, $\Delta_{n,\overline{T}}$, of the process to the spatial bandwidth parameter, $\varepsilon_{n,\overline{T}}$. If $\Delta_{n,\overline{T}}$ is small relative to $\varepsilon_{n,\overline{T}}$, so that $n\varepsilon^4_{n,\overline{T}} \to \infty$, then the bias effect dominates the asymptotics. In contrast to conventional nonparametric regression situations (Härdle (1990)) and to the long span case (c.f. Theorem 5 above), the bias effect turns out to be random, as it is in the nonstationary autoregressive case studied in Phillips and Park (1998). If the spatial bandwidth $\varepsilon_{n,\overline{T}}$ is small relative to the observation interval and $n\varepsilon^4_{n,\overline{T}} \to 0$, the bias effects are eliminated asymptotically and the martingale effect governs the limit theory. Due to the very slow rate of convergence of the variance term in the estimation error decomposition for the drift, the bias term never plays a role in the limit theory for the infinitesimal first moment.

Remark 8: When T is fixed as in Theorem 6 above, the admissible bandwith conditions can be easily written as a function of the number of observations. The variance term dominates if

$$\varepsilon_{n,\overline{T}} \propto n^{-k_1} \text{ with } k_1 \in \left(\frac{1}{4}, \frac{1}{2}\right)$$

and

$$h_{n,\overline{T}} \propto n^{-k_2} \text{ with } k_2 \in \left(0, \frac{1}{2}\right).$$

On the other hand, if

$$\varepsilon_{n,\overline{T}} \propto n^{-k_1} \text{ with } k_1 \in \left(0, \frac{1}{4}\right)$$

and

$$h_{n,\overline{T}} \propto n^{-k_2} \text{ with } k_2 \in \left(0, \frac{1}{2}\right),$$

then the "bias" term drives the limiting distribution.

Remark 9: (The rate of Convergence) The diffusion function estimator converges at a faster rate than the drift estimator $\left(\text{i.e. } \sqrt{\frac{\varepsilon_{n,T}\overline{L}_X(T,x)}{\Delta_{n,T}}} \text{ versus } \sqrt{\varepsilon_{n,T}\overline{L}_X(T,x)}\right)$. Using the results in Remark 5 above, in the Brownian motion and stationary case the normalizations in (5.6) and (5.7) are $\sqrt{\frac{\varepsilon_{n,T}T^{\frac{1}{2}}}{\Delta_{n,T}}} = \sqrt{\frac{n\varepsilon_{n,T}}{T^{1/2}}}$ and $\sqrt{\frac{\varepsilon_{n,T}T}{\Delta_{n,T}}} = \sqrt{n\varepsilon_{n,T}}$, respectively.

Remark 10: (Single Smoothing) As in the drift case, we consider a simpler version of our infinitesimal volatility estimator based on single smoothing. Define

$$\overline{\sigma^2_{(n,T)}}(x) = \frac{1}{\Delta_{n,T}} \frac{\sum_{i=1}^{n-1} \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{g_{n,T}}\right)\left(X_{(i+1)\Delta_{n,T}} - X_{i\Delta_{n,T}}\right)^2}{\sum_{i=1}^{n} \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{g_{n,T}}\right)}. \tag{5.14}$$

Following our derivations in the convoluted case (also, c.f. Bandi (2000)), we can prove that (5.14) is consistent almost surely for the unknown function provided the window width $g_{n,T}$ is such that $\frac{\overline{L}_X(T,x)}{g_{n,T}}(\Delta_{n,T})^{\beta} = O_{a.s.}(1)$ for some $\beta \in \left(0, \frac{1}{2}\right)$ as $n, T \to \infty$ with $\frac{T}{n} \to 0$. Furthermore, if $\frac{g_{n,T}^5 \overline{L}_X(T,x)}{\Delta_{n,T}} \stackrel{a.s.}{\to} 0 \; \forall x \in \mathfrak{D}$, then

$$\sqrt{\frac{g_{n,T}\overline{L}_X(T,x)}{\Delta_{n,T}}} \left\{\overline{\sigma}^2_{(n,T)}(x) - \sigma^2(x)\right\} \Rightarrow \mathbf{N}\left(0, 4\mathbf{K}_2\sigma^4(x)\right),$$

where $\mathbf{K}_2 = \int_{-\infty}^{\infty} \mathbf{K}^2(s)ds$. Additionally, if $\frac{g_{n,T}^5 \overline{L}_X(T,x)}{\Delta_{n,T}} = O_{a.s.}(1) \; \forall x \in \mathfrak{D}$, then

$$\sqrt{\frac{g_{n,T}\overline{L}_X(T,x)}{\Delta_{n,T}}} \left\{\overline{\sigma}^2_{(n,T)}(x) - \sigma^2(x) - \Gamma_{\sigma^2}(x)\right\} \Rightarrow \mathbf{N}\left(0, 4\mathbf{K}_2\sigma^4(x)\right),$$

where

$$\Gamma_{\sigma^2}(x) = (g_{n,T})^2 \, \mathbf{K}_1 \left[\left(\sigma^2(x)\right)' \frac{s'(x)}{s(x)} + \frac{1}{2}\left(\sigma^2(x)\right)''\right],$$

$s(x)$ is the speed measure of the process X and $\mathbf{K}_1 = \int_{-\infty}^{\infty} s^2 \mathbf{K}(s)ds$.

As in the case of drift estimation (c.f. Remark 6 above), double-smoothing can reduce the asymptotic mean-squared error of the diffusion estimator for some processes and some levels x, thus offerering increased flexibility over its simple counterpart. Contrary to drift estimation (c.f. Remark 6 above), the finite sample performance of alternative diffusion estimators based on simple and convoluted kernels is quite similar (c.f. Bandi and Nguyen (2000)).

5.4. *Relation to Florens-Zmirou (1993)*

There is an important similarity between (5.10) and the limiting distribution obtained in Florens-Zmirou (1993). It is useful to recall her results before commenting further.

Theorem 7: (Florens-Zmirou (1993)) *Assume we observe X_t at $\{t = t_1, t_2, ..., t_n\}$ in the time interval $[0, \overline{T}]$ where $\overline{T}$ can be normalized to* 1. *Also, the data is equispaced. Consequently, $\{X_t = X_{\Delta_n}, X_{2\Delta_n}, X_{3\Delta_n}, ..., X_{n\Delta_n}\}$ are n observations at points $\{t_1 = \Delta_n, t_2 = 2\Delta_n, ..., t_n = \Delta_n\}$, where $\Delta_n = 1/n$. The estimator*

$$\widehat{\sigma}^2_{(n)}(x) = \frac{1}{\Delta_n} \frac{\sum_{i=1}^{n-1} \mathbf{1}_{\{|X_{i/n} - x| \leq h_n\}} [X_{(i+1)/n} - X_{i/n}]^2}{\sum_{i=1}^{n} \mathbf{1}_{\{|X_{i/n} - x| \leq h_n\}}} \stackrel{L_2}{\to} \sigma^2(x)$$

provided the sequence h_n is such that $nh_n \to \infty$ and $nh_n^4 \to 0$. *Further, if $nh_n^3 \to 0$, then*

$$\sqrt{nh_n}\left\{\widehat{\sigma}^2_{(n)}(x) - \sigma^2(x)\right\} \Rightarrow \mathbf{MN}\left(0, 2\frac{\sigma^4(x)}{\overline{L}_X(1,x)}\right).$$

Provided $nh_n^4 \to 0$, the bias term disappears asymptotically and the limiting distribution is the normal distribution to which the 'martingale' term converges. It is not surprising that the limiting distribution in Florens-Zmirou (1993) resembles the limiting distribution of the estimator proposed here for choices of $\varepsilon_{n,\overline{T}}$ and $h_{n,\overline{T}}$ that make the bias term negligible (and provided $h_{n,\overline{T}} = o(\varepsilon_{n,\overline{T}})$). Note, in fact, that in the fixed T case the estimator that we suggest here can be interpreted as a convoluted version of the Florens-Zmirou's estimator. In particular, it can be written as a weighted average of estimates obtained using the Florens-Zmirou's method. In effect, $\widetilde{\sigma}^2_{n,\overline{T}}(X_{i\Delta_{n,\overline{T}}})$ can be rearranged as follows $\forall i$,

$$\begin{aligned}\widetilde{\sigma}^2_{n,\overline{T}}(X_{i\Delta_{n,\overline{T}}}) &= \frac{1}{m_{n,\overline{T}}(i\Delta_{n,\overline{T}})\Delta_{n,\overline{T}}} \sum_{j=0}^{m_{n,\overline{T}}(i\Delta_{n,\overline{T}})-1} [X_{t(i\Delta_{n,\overline{T}})_j+\Delta_{n,\overline{T}}} - X_{t(i\Delta_{n,\overline{T}})_j}]^2 \\ &= \frac{1}{\Delta_{n,\overline{T}}} \frac{\sum_{j=1}^{n-1} \mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}} - X_{i\Delta_{n,\overline{T}}}| \le \varepsilon_{n,\overline{T}}\}} [X_{(j+1)\Delta_{n,\overline{T}}} - X_{j\Delta_{n,\overline{T}}}]^2}{\sum_{j=1}^{n} \mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}} - X_{i\Delta_{n,\overline{T}}}| \le \varepsilon_{n,\overline{T}}\}}}.\end{aligned}$$

It is easy to prove that when $nh_n^4 \to \infty$ the Florens-Zmirou's estimator is still consistent but, in the same manner as our own limit theory, the "bias" term drives the asymptotic distribution, namely

$$\frac{1}{h_n^{3/2}} \left\{ \widehat{\sigma}^2_{(n)}(x) - \sigma^2(x) \right\} \Rightarrow \mathbf{MN} \left(0, 16\varphi^{ind} \frac{\left(\sigma'(x)\right)^2}{(\overline{L}_X(1,x))} \right)$$

where $\varphi^{ind} = 2\int_0^\infty \int_0^\infty \left(\frac{1}{2}\mathbf{1}_{\{|u|\le 1\}}\right)\left(\frac{1}{2}\mathbf{1}_{\{|e|\le 1\}}\right) \min(u,e)dude.$

Of course, the similarity between our approach to diffusion function estimation and the approach in Florens-Zmirou is even more striking when considering sample analogues to the unknown diffusion function based on single smoothing, as in Remark 10 above, for a fixed time span $\overline{T}$. Nonetheless, our limit theory presents important differences over the results in Florens-Zmirou. First, we extend her analysis to general smooth kernels (c.f. (5.14)). Second, we provide a proof of convergence with probability one and related conditions on the relevant bandwidth(s). Third, based on different bandwidth choices, we describe the potential limiting trade-off between bias (c.f. (5.11)) and variance (c.f. (5.10)) in the asymptotic distribution.

5.5. *Remarks on the Stationary Case*

When stationarity holds, our general theory reflects existing results in the estimation of conditional first moments for discrete time series (c.f. Pagan and Ullah (1999) for a recent discussion).

Corollary 2 (c.f. Theorem 3): *Assume X is stationary. Furthermore, assume $n \to \infty$, $T \to \infty$, $\Delta_{n,T} \to 0$, $h_{n,T} \to 0$ (as $n, T \to \infty$) such that $\frac{T}{h_{n,T}}(\Delta_{n,T})^\alpha = O(1)$ for some $\alpha \in \left(0, \frac{1}{2}\right)$, and $\varepsilon_{n,T} \to 0$ (as $n, T \to \infty$) such that $\frac{T}{\varepsilon_{n,T}}(\Delta_{n,T})^\beta = O(1)$ for some $\beta \in \left(0, \frac{1}{2}\right)$ and $\varepsilon_{n,T}T \to \infty$.*

Then, $\widehat{\mu}_{(n,T)}(x) \overset{a.s.}{\to} \mu(x)\ \forall x \in \mathfrak{D}$. *Additionally, the asymptotic distribution of the drift function estimator is of the form*

$$\sqrt{\varepsilon_{n,T}T}\left\{\widehat{\mu}_{(n,T)}(x) - \mu(x) - \Gamma_{\mu}(x)\right\} \Rightarrow \mathbf{N}\left(0, \frac{1}{2}\frac{\sigma^2(x)}{f(x)}\right), \tag{5.15}$$

if $h_{n,T} = o(\varepsilon_{n,T})$ *and* $\varepsilon_{n,T} = O(T^{-1/5})$ *where*

$$\Gamma_{\mu}(x) = \varepsilon_{n,T}^2 \frac{1}{3}\left[\mu^{'}(x)\frac{f^{'}(x)}{f(x)} + \frac{1}{2}\mu^{''}(x)\right], \tag{5.16}$$

and $f(x)$ *is the stationary distribution function of the process at* x.

Equivalently,

Corollary 3 (c.f. Theorem 5): *Assume* X *is stationary. Furthermore, assume* $n \to \infty$, $T \to \infty$, $\Delta_{n,T} \to 0$, $h_{n,T} \to 0$ *(as* $n, T \to \infty$*) such that* $\frac{T}{h_{n,T}}(\Delta_{n,T})^{\alpha} = O(1)$ *for some* $\alpha \in \left(0, \frac{1}{2}\right)$ *and* $\varepsilon_{n,T} \to 0$ *(as* $n, T \to \infty$*) such that* $\frac{T}{\varepsilon_{n,T}}(\Delta_{n,T})^{\beta} = O(1)$ *for some* $\beta \in \left(0, \frac{1}{2}\right)$. *Then,* $\widehat{\sigma}^2_{(n,T)}(x) \overset{a.s.}{\to} \sigma^2(x)\ \forall x \in \mathfrak{D}$. *Additionally, the asymptotic distribution of the diffusion function estimator is of the form*

$$\sqrt{n\varepsilon_{n,T}}\left\{\widehat{\sigma}^2_{(n,T)}(x) - \sigma^2(x) - \Gamma_{\sigma^2}(x)\right\} \Rightarrow \mathbf{N}\left(0, 2\frac{\sigma^4(x)}{f(x)}\right), \tag{5.17}$$

if $h_{n,T} = o(\varepsilon_{n,T})$, $\varepsilon_{n,T}T \to 0$ *and* $\varepsilon_{n,T} = O(n^{-1/5})$ *where*

$$\Gamma_{\sigma^2}(x) = \varepsilon_{n,T}^2 \frac{1}{3}\left[\left(\sigma^2(x)\right)^{'}\frac{f^{'}(x)}{f(x)} + \frac{1}{2}\left(\sigma^2(x)\right)^{''}\right], \tag{5.18}$$

and $f(x)$ *is the stationary distribution function of the process at* x.

Interestingly, Corollaries 2 and 3 apply to the stationary case as well as to the case where the process is not initialized at the stationary distribution, while being endowed with a time-invariant stationary density at least in the limit. The latter situation is known as positive-recurrence and is such that the speed measure of the process (from Lemma 5 above) is finite, i.e. $s(\mathfrak{D}) < \infty$. In particular, the normalized speed measure coincides with the limiting distribution of X (c.f. Pollack and Siegmund (1985)), that is

$$\lim_{t\to\infty} P^x(X_t < z) = \frac{s((l,z))}{s(\mathfrak{D})} \qquad \forall x, z \in \mathfrak{D}.$$

5.6. *Practical Implementation*

The estimators presented and discussed in this paper are sample analogues to the true theoretical functions. They are written as weighted averages based on convoluted smoothing functions. As shown in Remark 6 and 10 above our asymptotic results readily apply to weighted averages based on simple kernels. In this case, by virtue of the generality of our set-up, only straightforward modifications to the theory outlined in the convoluted case are needed.

In both the simple and the convoluted case, practical implementation of our methodology requires the choice of the kernel and relevant bandwidth(s) along with an appropriate specification for the local time factor estimator ($\overline{L}_X(T,x)$, that is) that drives the rates of convergence of the functional estimates.

We start with local time. Theorem 1 provides us with an easy way to estimate it consistently for every sample path using kernels. Note that in applications it is often conventional to normalize T to 1. This implies that the admissible bandwidth $h_{n,T}$ is proportional to n^{-k} with $k \in \left(0, \frac{1}{2}\right)$. Since the rate of convergence of the estimated local time to the true process is $\frac{1}{\sqrt{h}}$ (c.f. Bandi (1998)), it is convenient to set $h^{ltime}_{n,\overline{T}}$ equal to $c_{ltime}\frac{1}{\log(n)}n^{-\frac{1}{2}}$ where c_{ltime} is a constant of proportionality that might be chosen using automated methods for bandwidth selection in density estimation (c.f. Pagan and Ullah (1999)). >From a practical standpoint, functional estimation of the local time factor is perfectly analogous to functional estimation of a marginal density function. What changes with respect to the standard case that assumes stationarity is the broader interpretation of the proposed estimator (c.f. Bandi (1998) for additional discussion). We now turn to the functions of interest.

In the convoluted case two window widths (i.e. $h_{n,T}$ and $\varepsilon_{n,T}$) need to be chosen. In light of the asymptotic role played by the local time factor in the additional smoothing (see the proof of Theorem 3, for example), it is natural to choose $h_{n,T}$ equal to $h^{ltime}_{n,\overline{T}}$ both in the drift and in the diffusion case. The choice of the 'leading' (provided $h_{n,T} = o\left(\varepsilon_{n,T}\right)$) bandwidth $\varepsilon_{n,T}$ is more awkward. Consider the diffusion case and normalize T to 1. Remark 7 above illustrates the relationship between the rate of convergence of the leading bandwidth $\varepsilon_{n,\overline{T}}$ and the limiting trade-off between bias and variance effects for a fixed time span T. Based on the limit theory and Remark 8 it is convenient to set $\varepsilon^{diff}_{n,\overline{T}}$ equal to $c_{diff}\frac{1}{\log(n)}n^{-\frac{1}{4}}$. We undersmooth slightly with respect to the optimal rate $\left(\frac{1}{4}\right)$ to eliminate the influence of the bias term from the asymptotic distribution. The constant c_{diff} can be found using standard automated criteria (such as cross-validation) under the constraint that $h_{n,T} \leq \varepsilon^{diff}_{n,\overline{T}}$. Given that the drift cannot be identified consistently over a fixed span of data, the admissible condition that the 'leading' drift bandwidth ought to satisfy cannot be expressed in closed-form as a function of the number of observations. Nonetheless, since the feasible drift bandwidth vanishes at a slower pace than the feasible diffusion bandwidth, a simple rule-of-thumb can be applied: we can set $\varepsilon_{n,\overline{T}}^{drift} = c_{drift}\frac{1}{\log(n)}n^{-\frac{1}{4}}$ and choose c_{drift} using automated methods under the constraint that $\varepsilon_{n,\overline{T}}^{drift} > \varepsilon_{n,\overline{T}}^{diff}$. More rigorously, one could recognize the role played by local time in the functional estimation of the drift (c.f. Theorem 3) and set $\varepsilon_{n,\overline{T}}^{drift}(x) = c_{drift}\frac{1}{\log(n)}\widehat{\overline{L}}_X(T,x)^{-\frac{1}{5}}$. Again, we undersmooth slightly with respect to the optimal case ($\varepsilon_{n,\overline{T}}^{drift}(x) \propto \widehat{\overline{L}}_X(T,x)^{-\frac{1}{5}}$) in order to achieve a close-to-optimal rate, eliminate the influence of the bias term from the limiting distribution and center it around zero. This choice is level-specific and implies less smoothing in areas that are often visited. In other words, there is explicit scope for local adaptation of the leading drift bandwidth to the number of visits to the point at which estimation is performed. Being the diffusion function estimable over a fixed span of time, the need for level-dependent bandwidth choices appears to be less compelling. Nonetheless, standard arguments in favor of level-specific choices leading to bias reduction (c.f. Pagan and Ullah (1999), for example) can still be made in our framework, even in the diffusion case.

In light of the limiting results in Remarks 6 and 10 above, it is noted that bandwidth choice in the simple case entails the same procedures as in the convoluted case with the leading bandwidth $\varepsilon_{n,T}$ being replaced by $h_{n,T}$. As a caveat, the use of selection criteria designed for density estimation and/or standard regression analysis can only be considered a preliminary solution in our framework. Future research should focus on the design of automated criteria for window selection in the context of nonparametric diffusion model estimation.

We now turn to the kernel. It is well known that choosing the kernel is less crucial than choosing the optimal window width (also, see Bandi and Nguyen (2000) for simulations in the diffusion case). For the suggested bandwidth choices, what matters to determine the constant of proportionality in the asymptotic variance in the convoluted case is the kernel being used in the preliminary smoothing. We use an indicator function but the generality of the methodology makes it clear that any smooth kernel could have been used instead. In fact, if we had used a smooth kernel, the constants of proportionality in the limiting variances would be the same as in the single smoothing case $\left(\text{i.e. } \int \mathbf{K}^2(s)ds \text{ and } 4\int \mathbf{K}^2(s)ds\right)$. In consequence, the smoothing function can be chosen to minimize the asymptotic dispersion of the estimates. Coherently with the standard functional estimation of conditional expectations in the discrete-time context, the use of higher-order kernels is expected to improve the rate of convergence to zero of the bias term (c.f. the proof of Theorem 3) when T is not fixed. Interestingly, for a fixed T, the use of higher-order kernels does not increase the rate of convergence of the random bias term in the case of diffusion estimation.

Generally speaking, the analogy between our theory and the standard theory for estimating conditional expectations in discrete-time reveals that conventional methods for selecting the kernel function (Härdle (1990) and Pagan and Ullah (1999), for example) readily extend to the nonparametric estimation of diffusions in the presence of single-smoothing and an enlarging span of data.

6. CONCLUSION

This paper shows how to identify and consistently estimate both the drift and the diffusion terms of a general homogeneous stochastic differential equation under broad assumptions on the data generating process. The method relies on the construction of functional sample counterparts to conditional expectations and can be extended to multi-equation specifications. The definition of the estimators in the multivariate case is straightforward but important technical difficulties associated with the curse of dimensionality arise in that case when deriving a limit theory along the lines given here. In particular, important issues concerning the recurrence of the process and the existence of local time in higher dimensions make the problem especially challenging. For instance, Brownian motion is well known to be transient rather than recurrent in dimensions greater than two.

Nonetheless, Brugière (1993) extends the methods in Florens-Zmirou (1993) to prove a general limit theory for a matrix of diffusion functions based on a probabilistic tool that corresponds to a general version of the local time factor. Equivalently, at the natural cost of a reduction in the rates of convergence, we expect the techniques that we introduce in the present paper to be generalizable to permit the development of an asymptotic theory for nonparametric estimates of the drift and diffusion matrices of multivariate processes that might not possess a time-invariant density. Research on this topic is being conducted and will be reported in later work.

APPENDIX A: PROOFS

Proof of Lemma 1: See Revuz and Yor (1998), Corollary 1.8, page 226.

Proof of Lemma 2: See Revuz and Yor (1998), Exercise 1.15, page 232.

Proof of Lemma 3: See Revuz and Yor (1998), Corollary 1.9, page 227.

Proof of Lemma 4: The first part of the result is stated in Yor (1983). We prove the result in the second case ($a < 0$, that is). Start by considering a simple application of the *Tanaka formula* (c.f. Definition 1), namely

$$
\begin{aligned}
X_t^+ &= X_0^+ + \int_0^t \mathbf{1}_{(X_s>0)} dX_s + \frac{1}{2} L_X(t,0),\\
(X_t - a)^+ &= (X_0 - a)^+ + \int_0^t \mathbf{1}_{(X_s>a)} dX_s + \frac{1}{2} L_X(t,a).
\end{aligned}
$$

Subtract the second expression from the first expression, giving

$$
\begin{aligned}
& X_t^+ - (X_t - a)^+ \\
= \; & X_0^+ - (X_0 - a)^+ - \int_0^t \mathbf{1}_{(a \le X_s \le 0)} dX_s + \frac{1}{2}(L_X(t,0) - L_X(t,a)).
\end{aligned}
$$

Equivalently, we can write

$$
\begin{aligned}
& X_t^+ - (X_t - a/\lambda)^+ \\
= \; & X_0^+ - (X_0 - a/\lambda)^+ - \int_0^t \mathbf{1}_{(a/\lambda \le X_s \le 0)} dX_s + \frac{1}{2}(L_X(t,0) - L_X(t,a/\lambda)).
\end{aligned}
$$

Now, multiply through by $\sqrt{\lambda}$. This gives,

$$
\begin{aligned}
& \sqrt{\lambda}(X_t^+ - (X_t - a/\lambda)^+) \\
= \; & \sqrt{\lambda}(X_0^+ - (X_0 - a/\lambda)^+) - \sqrt{\lambda} \int_0^t \mathbf{1}_{(a/\lambda \le X_s \le 0)} dX_s + \\
& + \frac{1}{2}\sqrt{\lambda}(L_X(t,0) - L_X(t,a/\lambda)).
\end{aligned}
$$

Apparently,

$$
\sqrt{\lambda}|X_t^+ - (X_t - a/\lambda)^+| + \sqrt{\lambda}|X_0^+ - (X_0 - a/\lambda)^+| \le 2\frac{|a|}{\sqrt{\lambda}}.
$$

Hence, the asymptotic distribution of $\frac{1}{2}\sqrt{\lambda}(L_X(t,0) - L_X(t,a/\lambda))$ is driven by the term $\sqrt{\lambda}\int_0^t \mathbf{1}_{(a/\lambda \le X_s \le 0)} dX_s$ as $\lambda \to \infty$. Further,

$$
\sqrt{\lambda}\int_0^t \mathbf{1}_{(a/\lambda \le X_s \le 0)} dX_s = \sqrt{\lambda}\int_0^t \mathbf{1}_{(a/\lambda \le X_s \le 0)} \mu(X_s) ds + \sqrt{\lambda}\int_0^t \mathbf{1}_{(a/\lambda \le X_s \le 0)} \sigma(X_s) dB_s. \tag{7.1}
$$

Now notice that $\sqrt{\lambda}\int_0^t \mathbf{1}_{(a/\lambda \le X_s \le 0)} \mu(X_s) ds \overset{a.s.}{\to} 0$ as $\lambda \to \infty$. In fact, by the occupation time formula (c.f. Lemma 2) we can write

$$
\begin{aligned}
& \sqrt{\lambda}\int_0^t \mathbf{1}_{(a/\lambda \le X_s \le 0)} \mu(X_s) ds \\
= \; & \sqrt{\lambda}\int_{-\infty}^{\infty} \mathbf{1}_{(a/\lambda \le b \le 0)} \frac{\mu(b)}{\sigma^2(b)} L_X(t,b) db,
\end{aligned}
$$

and, setting $\lambda b = c$, this becomes

$$
\frac{1}{\sqrt{\lambda}}\int_{-\infty}^{\infty} \mathbf{1}_{(a \le c \le 0)} \frac{\mu(c/\lambda)}{\sigma^2(c/\lambda)} L_X(t,c/\lambda) dc.
$$

By the properties of the local time (in particular, the map $a \to L_X(t,a)$ is a.s. continuous and has compact support – c.f. Lemma 1) and the dominated convergence theorem, it follows that

$$
\int_{-\infty}^{\infty} \mathbf{1}_{(a \le c \le 0)} \frac{\mu(c/\lambda)}{\sigma^2(c/\lambda)} L_X(t,c/\lambda) dc \overset{a.s.}{\to} -a\frac{\mu(0)}{\sigma^2(0)} L_X(t,0),
$$

as $\lambda \to \infty$. In consequence,

$$\frac{1}{\sqrt{\lambda}} \int_{-\infty}^{\infty} \mathbf{1}_{(a \le c \le 0)} \frac{\mu(c/\lambda)}{\sigma^2(c/\lambda)} L_X(t, c/\lambda) dc \overset{a.s.}{\to} 0.$$

This, in turn, implies that the asymptotic behaviour of (7.1) is determined by $\sqrt{\lambda} \int_0^t \mathbf{1}_{(a/\lambda \le X_s \le 0)} \sigma(X_s) dB_s$. Now define

$$M^\lambda(t) := \sqrt{\lambda} \int_0^t \mathbf{1}_{(a/\lambda \le X_s \le 0)} \sigma(X_s) dB_s.$$

M^λ is a continuous martingale with quadratic variation process $\{[M^\lambda]_t : t \ge 0\}$ given by

$$\lambda \int_0^t \mathbf{1}_{(a/\lambda \le X_s \le 0)} \sigma^2(X_s) ds.$$

Again, by the occupation time formula, the properties of local time and dominated convergence, we get

$$[M^\lambda]_t \overset{a.s.}{\to} -a L_X(t, 0).$$

Setting

$$T_t^\lambda = \inf\{s : [M^\lambda]_s > t\},$$

$\widetilde{B}_t = M^\lambda_{T_t^\lambda}$ is a Brownian motion and $M_t^\lambda = \widetilde{B}_{[M^\lambda]_t}$. In fact, $\widetilde{B}_t$ is the so-called *Dambis, Dubins-Schwarz* Brownian motion of M_t^λ (c.f. Revuz and Yor (1998, Theorem 1.6, page 173 and, for an asymptotic version, Theorem 2.3, page 496)). It follows that

$$\begin{aligned} M_t^\lambda &= \sqrt{\lambda} \int_0^t \mathbf{1}_{(a \le \lambda X_s \le 0)} \sigma(X_s) dB_s \\ &\underset{\lambda \to \infty}{\overset{d}{\to}} \widetilde{B}_{-a L_X(t,0)} \\ &\overset{d}{=} \sqrt{-a} \widetilde{B}_{L_X(t,0)} \\ &\overset{d}{=} \mathfrak{B}_{(L_X(t,0), -a)}, \end{aligned}$$

where $L_X(t, 0) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^t 1_{[0,\ \varepsilon[} \sigma^2(X_s) ds$ a.s. $\forall x, t$ and $\mathfrak{B}$ is a standard Brownian sheet. So far, we have proved convergence of the marginals of a generic family $\mathfrak{P}_\lambda$ of probability measures to corresponding marginal limit distributions. It is easy to verify the compactness of $\mathfrak{P}_\lambda$. The proof follows standard arguments and is omitted here for brevity (see Billingsley (1968)). Weak convergence then follows. In particular, as $\lambda \to \infty$, the process (indexed by $(t, a) \in \mathfrak{R}_+ \times \mathfrak{R}_-$)

$$\left(X_t \,;\, L_X(t, a) \,;\, \frac{\sqrt{\lambda}}{2}\{L_X\left(t, \frac{a}{\lambda}\right) - L_X(t, 0)\}\right)$$

converges weakly to

$$(X_t \,;\, L_X(t, a) \,;\, \mathfrak{B}(L_X(t, 0), -a),$$

where $(\mathfrak{B}(s, c) \,;\, (s, c) \in \mathfrak{R}_+^2)$ is a standard Brownian sheet independent of X. (For the independence property, see Revuz and Yor (1998, Exercise 2.12, Chapter XIII).) Then, a simple generalization of the previous finding to the spatial location $r \ne 0$ gives

$$\frac{1}{2}\sqrt{\lambda}\{L_X\left(t, r + \frac{a}{\lambda}\right) - L_X(t, r)\} \overset{d}{\to} \mathfrak{B}(L_X(t, r), -a),$$

as $\lambda \to \infty$, and this proves the stated result.

Proof of Lemma 5: Immediate given the *limit-quotient theorem* in Revuz and Yor (1998, Theorem 3.12, page 408) and the observation that any invariant measure for scalar diffusions has to be equal (up to multiplication by a constant) to the speed measure.

Proof of Theorem 1: See Florens-Zmirou (1993) for the case involving a discontinuous kernel function. For full derivations in the case of a continuous kernel, see Bandi and Phillips (1998).

Proof of Corollary 1: If $T \to \infty$ and $\frac{T}{n} = \Delta_{n,T} \to 0$, then $\frac{\Delta_{n,T}}{h_{n,T}} \sum_{i=1}^{n} \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)$ converges to $\overline{L}_X(\infty, x)$ provided $h_{n,T} \to 0$ (as $n \to \infty$) in such a way that $\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T})^{\alpha} = O_{a.s.}(1)$ for some $\alpha \in \left(0, \frac{1}{2}\right)$. But $\overline{L}_X(\infty, x) = \overline{L}_X(\sup\{t : X_t = x\}, x)$ a.s. (c.f. Revuz and Yor (1998, page 223, Proposition 1.3, Remark 2)). And, if the process is recurrent, then $\overline{L}_X((\sup\{t : X_t = x\}), x) = \infty$ a.s.

Proof of Theorem 2: We start by considering the expression

$$\frac{\frac{\Delta_{n,T}}{h_{n,T}} \sum_{i=1}^{n} \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right) \left(\widetilde{\mu}_{n,T}(X_{i\Delta_{n,T}}) - \mu(X_{i\Delta_{n,T}})\right)}{\frac{\Delta_{n,T}}{h_{n,T}} \sum_{i=1}^{n} \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} \tag{7.2}$$

$$+\frac{\frac{\Delta_{n,T}}{h_{n,T}} \sum_{i=1}^{n} \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right) \mu(X_{i\Delta_{n,T}})}{\frac{\Delta_{n,T}}{h_{n,T}} \sum_{i=1}^{n} \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}. \tag{7.3}$$

First, we examine (7.3). We want to prove that for some $\varepsilon > 0$

$$\frac{\frac{\Delta_{n,T}}{h_{n,T}} \sum_{i=1}^{n} \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right) \mu(X_{i\Delta_{n,T}})}{\frac{\Delta_{n,T}}{h_{n,T}} \sum_{i=1}^{n} \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}$$

$$= \frac{\int_0^T \frac{1}{h_{n,T}} \mathbf{K}\left(\frac{X_s-x}{h_{n,T}}\right) \mu(X_s) ds + o_{a.s.}\left(\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T})^{1/2-\varepsilon}\right)}{\int_0^T \frac{1}{h_{n,T}} \mathbf{K}\left(\frac{X_s-x}{h_{n,T}}\right) ds + o_{a.s.}\left(\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T})^{1/2-\varepsilon}\right)}.$$

We begin with the numerator and look at the quantity

$$\frac{\Delta_{n,T}}{h_{n,T}} \sum_{i=1}^{n} \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right) \mu(X_{i\Delta_{n,T}}) - \int_0^T \frac{1}{h_{n,T}} \mathbf{K}\left(\frac{X_s-x}{h_{n,T}}\right) \mu(X_s) ds. \tag{7.4}$$

Given the properties of $\mathbf{K}(.)$ and the assumptions on $\mu(.)$, (7.4) is seen to be bounded as follows

$$\frac{1}{h_{n,T}} \left| \sum_{i=0}^{n-1} \int_{iT/n}^{(i+1)T/n} \left[\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right) \mu(X_{i\Delta_{n,T}}) - \mathbf{K}\left(\frac{X_s-x}{h_{n,T}}\right) \mu(X_s)\right] ds \right|$$

$$+\left|\frac{\Delta_{n,T}}{h_{n,T}} \mathbf{K}\left(\frac{X_0-x}{h_{n,T}}\right) \mu(X_0)\right| + \left|\frac{\Delta_{n,T}}{h_{n,T}} \mathbf{K}\left(\frac{X_{n\Delta_{n,T}}-x}{h_{n,T}}\right) \mu(X_{n\Delta_{n,T}})\right|$$

$$\leq \frac{1}{h_{n,T}} \left| \sum_{i=0}^{n-1} \int_{i\Delta_{n,T}}^{(i+1)\Delta_{n,T}} \left[\mathbf{K}\left(\frac{X_s-x}{h_{n,T}}\right) \mu(X_{i\Delta_{n,T}}) - \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right) \mu(X_{i\Delta_{n,T}})\right] ds \right|$$

$$+\frac{1}{h_{n,T}} \left| \sum_{i=0}^{n-1} \int_{i\Delta_{n,T}}^{(i+1)\Delta_{n,T}} \left[\mathbf{K}\left(\frac{X_s-x}{h_{n,T}}\right) \mu(X_s) - \mathbf{K}\left(\frac{X_s-x}{h_{n,T}}\right) \mu(X_{i\Delta_{n,T}})\right] ds \right| + 2C_3 O_{a.s.}\left(\frac{\Delta_{n,T}}{h_{n,T}}\right)$$

$$\leq \frac{1}{h_{n,T}} \sum_{i=0}^{n-1} \int_{i\Delta_{n,T}}^{(i+1)\Delta_{n,T}} \left|\mathbf{K}'\left(\frac{\widetilde{X}_{is}-x}{h_{n,T}}\right)\right| \left|\left(\frac{X_s - X_{\Delta_{n,T}}}{h_{n,T}}\right)\right| \mu(X_{i\Delta_{n,T}}) ds \tag{7.5}$$

$$+\frac{1}{h_{n,T}} \left| \sum_{i=0}^{n-1} \int_{i\Delta_{n,T}}^{(i+1)\Delta_{n,T}} \mathbf{K}\left(\frac{X_s-x}{h_{n,T}}\right) \left(\mu(X_s) - \mu(X_{i\Delta_{n,T}})\right) ds \right| \tag{7.6}$$

$$+2C_3 O_{a.s.}\left(\frac{\Delta_{n,T}}{h_{n,T}}\right),$$

where $\widetilde{X}_{is}$ in (7.5) is on the line segment connecting X_s to $X_{i\Delta_{n,T}}$. Now define

$$\kappa_{n,T} = \max_{i\leq n} \sup_{i\Delta_{n,T}\leq s\leq (i+1)\Delta_{n,T}} |X_s - X_{i\Delta_{n,T}}|. \tag{7.7}$$

By the *Hölder property* for continuous SMGs (e.g. Revuz and Yor (1998, Exercise 1.20, Chapter V))

$$\mathfrak{P}\left(\left[t\geq 0 : \limsup_{\varepsilon\to 0} \frac{|X_{t+\varepsilon}-X_t|}{\varepsilon^{\alpha}} > 0\right]\right) = 0 \quad a.s.\ , \tag{7.8}$$

where $\mathfrak{P}$ is the Lebesgue measure on $\Re_+$ and (7.8) holds for every $\alpha < \frac{1}{2}$. In turn, (7.8) implies that

$$\frac{\kappa_{n,T}}{\Delta_{n,T}^{\alpha}} = o_{a.s.}(1) \tag{7.9}$$

for every $\alpha < \frac{1}{2}$. Hence, if $h_{n,T}$ is such that $\frac{1}{h_{n,T}}(\Delta_{n,T})^{\alpha} = O(1)$ for some $\alpha \in \left(0, \frac{1}{2}\right)$, then

$$\frac{\kappa_{n,T}}{h_{n,T}} = \frac{\kappa_{n,T}}{\Delta_{n,T}^{\alpha}} \frac{\Delta_{n,T}^{\alpha}}{h_{n,T}} = o_{a.s.}(1) \tag{7.10}$$

as $n, T \to \infty$. In view of (7.10) we have

$$\mathbf{K}'\left(\frac{\widetilde{X}_{is}-x}{h_{n,T}}\right) = \mathbf{K}'\left(\frac{X_s - x}{h_{n,T}} + o_{a.s}(1)\right), \tag{7.11}$$

uniformly over $i = 1, ..., n$. It follows from (7.7) and (7.11) that (7.5) is bounded by

$$\begin{aligned}
& \left(\frac{\kappa_{n,T}}{h_{n,T}}\right)\frac{1}{h_{n,T}}\sum_{i=0}^{n-1}\int_{i\Delta_{n,T}}^{(i+1)\Delta_{n,T}}\left|\mathbf{K}'\left(\frac{X_s-x}{h_{n,T}}+o_{a.s}(1)\right)\right|\mu(X_s+o_{a.s.}(1))ds \\
\leq\ & \left(\frac{\kappa_{n,T}}{h_{n,T}}\right)\frac{1}{h_{n,T}}\int_0^T\left|\mathbf{K}'\left(\frac{X_s-x}{h_{n,T}}+o_{a.s}(1)\right)\right|\mu(X_s+o_{a.s.}(1))ds \\
=\ & \left(\frac{\kappa_{n,T}}{h_{n,T}}\right)\frac{1}{h_{n,T}}\int_{-\infty}^{\infty}\left|\mathbf{K}'\left(\frac{p-x}{h}+o_{a.s}(1)\right)\right|\mu(p)\overline{L}_X\left(T,p\right)dp \\
=\ & \left(\frac{\kappa_{n,T}}{h_{n,T}}\right)\int_{-\infty}^{\infty}\left|\mathbf{K}'\left(q+o_{a.s}(1)\right)\right|\mu(qh_{n,T}+x)\overline{L}_X\left(T,qh_{n,T}+x\right)dq \\
\leq\ & C_4\left(\frac{\kappa_{n,T}}{h_{n,T}}\right)O_{a.s}\left(\overline{L}_X(T,x)\right),
\end{aligned}$$

for some constant C_4, by virtue of the integrability of $\mathbf{K}'$ and the continuity of $\overline{L}_X$ and μ. Employing similar methods we can prove that (7.6) is bounded by

$$C_5\left(\kappa_{n,T}\right)O_{a.s}\left(\overline{L}_X(T,x)\right).$$

In consequence, the formula for the numerator (7.4) holds for some $\varepsilon > 0$ such that $\alpha \leq \frac{1}{2} - \varepsilon$. As for the denominator of (7.3), we can show the stated result using the same steps as for (7.5) above. Next, we prove that

$$\frac{\int_0^T \frac{1}{h_{n,T}}\mathbf{K}\left(\frac{X_s-x}{h_{n,T}}\right)\mu(X_s)ds + o_{a.s.}\left(\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T})^{1/2-\varepsilon}\right)}{\int_0^T \frac{1}{h_{n,T}}\mathbf{K}\left(\frac{X_s-x}{h_{n,T}}\right)ds + o_{a.s.}\left(\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T})^{1/2-\varepsilon}\right)} = \frac{\mu(x)s(x)+o_{a.s.}(1)}{s(x)+o_{a.s.}(1)} + o_{a.s.}(1) \overset{a.s.}{\to} \mu(x) \tag{7.12}$$

where $s(x)$ is the speed measure of the process. By virtue of Lemma 5, for a fixed $h_{n,T}$, we can write

$$\frac{\int_0^T \frac{1}{h_{n,T}}\mathbf{K}\left(\frac{X_s-x}{h_{n,T}}\right)\mu(X_s)ds}{\int_0^T \frac{1}{h_{n,T}}\mathbf{K}\left(\frac{X_s-x}{h_{n,T}}\right)ds} \overset{a.s.}{\to} \frac{\int_{-\infty}^{\infty} \frac{1}{h_{n,T}}\mathbf{K}\left(\frac{a-x}{h_{n,T}}\right)\mu(a)s(a)ds}{\int_{-\infty}^{\infty} \frac{1}{h_{n,T}}\mathbf{K}\left(\frac{a-x}{h_{n,T}}\right)s(a)ds},$$

which becomes

$$\begin{aligned}\frac{\int_{-\infty}^{\infty} \frac{1}{h_{n,T}}\mathbf{K}\left(\frac{a-x}{h_{n,T}}\right)\mu(a)s(a)ds}{\int_{-\infty}^{\infty} \frac{1}{h_{n,T}}\mathbf{K}\left(\frac{a-x}{h_{n,T}}\right)s(a)ds} &= \frac{\int_{-\infty}^{\infty}\mathbf{K}(u)\mu(x+h_{n,T}u)s(x+h_{n,T}u)du}{\int_{-\infty}^{\infty}\mathbf{K}(u)s(x+h_{n,T}u)du} \\ &\to \frac{\mu(x)s(x)}{s(x)} = \mu(x)\end{aligned}$$

by the continuity of $s \circ \mu$ and dominated convergence as $h_{n,T} \to 0$ with $n, T \to \infty$ so that $\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T})^{\alpha} = O_{a.s.}(1)$ for some $\alpha \in \left(0, \frac{1}{2}\right)$ $\forall x \in \mathfrak{D}$. This establishes (7.12). We now turn to the analysis of (7.2). It is sufficient to prove that

$$\widetilde{\mu}(X_{i\Delta_{n,T}}) = \mu(X_{i\Delta_{n,T}}) + o_{a.s.}(1) \tag{7.13}$$

in order to verify the stated result. To do so, we bound

$$\frac{1}{m_{n,T}(i\Delta_{n,T})}\sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1}\frac{[X_{t(i\Delta_{n,T})_j+\Delta_{n,T}} - X_{t(i\Delta_{n,T})_j}]}{\Delta_{n,T}} - \mu(X_{i\Delta_{n,T}})$$

using the Lipschitz property of μ as follows:

$$\begin{aligned}&\frac{1}{m(i\Delta_{n,T})}\sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1}\frac{[X_{t(i\Delta_{n,T})_j+\Delta_{n,T}} - X_{t(i\Delta_{n,T})_j}]}{\Delta_{n,T}} - \mu(X_{i\Delta_{n,T}}) \\ = \; & \frac{1}{m_{n,T}(i\Delta_{n,T})\Delta_{n,T}}\sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1}\int_{t(i\Delta_{n,T})_j}^{t(i\Delta_{n,T})_j+\Delta_{n,T}}(\mu(X_s)-\mu(X_{i\Delta_{n,T}}))ds \\ & + \frac{1}{m_{n,T}(i\Delta_{n,T})\Delta_{n,T}}\sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1}\int_{t(i\Delta_{n,T})_j}^{t(i\Delta_{n,T})_j+\Delta_{n,T}}\sigma(X_s)dB_s \\ = \; & C_6 O_{a.s.}(\kappa_{n,T}) + \frac{1}{m(i\Delta_{n,T})\Delta_{n,T}}\sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1}\int_{t(i\Delta_{n,T})_j}^{t(i\Delta_{n,T})_j+\Delta_{n,T}}\sigma(X_s)dB_s, \end{aligned} \tag{7.14}$$

where $\kappa_{n,T}$ has its usual definition. Define $y_{t(i\Delta_{n,T})_j+\Delta_{n,T}} = \int_{t(i\Delta_{n,T})_j}^{t(i\Delta_{n,T})_j+\Delta_{n,T}}\sigma(X_s)dB_s$, which is measurable with respect to $\Im_{t(i\Delta_{n,T})_j+\Delta_{n,T}}$ where $\Im_{t(i\Delta_{n,T})_j+\Delta_{n,T}} = \{A \in \Im : A\{t(i\Delta_{n,T})_j + \Delta_{n,T} \leq t^*\} \in \Im_{t^+} \forall t \geq 0\}$ for all $j \leq m_{n,T}$. Furthermore,

$$\mathbf{E}\left(y_{t(i\Delta_{n,T})_j+\Delta_{n,T}}\right) = 0,$$

and, by the Ito isometry,

$$\theta_{t(i\Delta_{n,T})_j+\Delta_{n,T}} = var\left(y_{t(i\Delta_{n,T})_j+\Delta_{n,T}}\right) = \mathbf{E}\left(\int_{t(i\Delta_{n,T})_j}^{t(i\Delta_{n,T})_j+\Delta}\sigma^2(X_s)ds\right) < \infty,$$

for all $j \leq m_{n,T}$. So, $\left(y_{t(i\Delta_{n,T})_j+\Delta_{n,T}}, \Im_{t(i\Delta_{n,T})_j+\Delta_{n,T}}\right)$ is a martingale difference array with zero mean and variance $\theta_{t(i\Delta_{n,T})_j+\Delta_{n,T}}$. Invoking a strong law of large numbers for martingale differences (e.g. Hall and Heyde (1980, Theorem 2.19, page 36)), we have

$$
\begin{aligned}
& \frac{1}{m(i\Delta_{n,T})} \sum_{j=0}^{m(i\Delta_{n,T})-1} y_{t(i\Delta_{n,T})j+\Delta_{n,T}} \\
= \; & \frac{1}{m(i\Delta_{n,T})} \sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1} \int_{t(i\Delta_{n,T})j}^{t(i\Delta_{n,T})j+\Delta_{n,T}} \sigma(X_s) dB_s \overset{a.s.}{\to} 0 \text{ as } n, T \to \infty,
\end{aligned}
$$

as $m_{n,T} \to \infty$ $(\forall i)$. We now explore the rate of convergence. Consider,

$$
\begin{aligned}
& \frac{1}{m_{n,T}(i\Delta_{n,T})\Delta_{n,T}} \sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1} \int_{t(i\Delta_{n,T})j}^{t(i\Delta_{n,T})j+\Delta_{n,T}} \sigma(X_s) dB_s \\
= \; & \frac{1}{\Delta_{n,T}} \frac{\frac{\Delta_{n,T}}{\varepsilon_{n,T}} \sum_{j=1}^{n-1} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}\}} \int_{j\Delta_{n,T}}^{(j+1)\Delta_{n,T}} \sigma(X_s) dB_s}{\frac{\Delta_{n,T}}{\varepsilon_{n,T}} \sum_{j=1}^{n} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}\}}} \\
= \; & \frac{\frac{1}{2\varepsilon_{n,T}} \sum_{j=1}^{n-1} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}\}} \int_{j\Delta_{n,T}}^{(j+1)\Delta_{n,T}} \sigma(X_s) dB_s}{\frac{\Delta_{n,T}}{2\varepsilon_{n,T}} \sum_{j=1}^{n} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}\}}}.
\end{aligned}
$$

First, analyze the numerator of this expression. Write,

$$
\begin{aligned}
\mathbf{U}_{n,T}^{X_{i\Delta_{n,T}}}(r) &= \sqrt{\varepsilon_{n,T}} \left(\frac{1}{2\varepsilon_{n,T}} \sum_{j=1}^{[nr]-1} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}\}} \int_{j\Delta_{n,T}}^{(j+1)\Delta_{n,T}} \sigma(X_s) dB_s \right) \\
&= \frac{1}{2\sqrt{\varepsilon_{n,T}}} \sum_{j=1}^{[nr]-1} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}\}} \int_{j\Delta_{n,T}}^{(j+1)\Delta_{n,T}} \sigma(X_s) dB_s.
\end{aligned}
$$

$\mathbf{U}_{n,T}^{X_{i\Delta_{n,T}}}$ is a martingale whose quadratic variation process $\left[\mathbf{U}_{n,T}^{X_{i\Delta_{n,T}}}\right]_r$ is

$$
\begin{aligned}
\left[\mathbf{U}_{n,T}^{X_{i\Delta_{n,T}}}\right]_r &= \frac{1}{4\varepsilon_{n,T}} \sum_{j=1}^{[nr]-1} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}\}} \int_{j\Delta_{n,T}}^{(j+1)\Delta_{n,T}} \sigma^2(X_s) ds \\
&= \frac{1}{4\varepsilon_{n,T}} \sum_{j=1}^{[nr]-1} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}\}} \sigma^2(X_{j\Delta_{n,T}} + o_{a.s.}(1)) \Delta_{n,T} \\
&= \frac{1}{4\varepsilon_{n,T}} \int_0^{rT} \mathbf{1}_{\{|X_s - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}\}} \sigma^2(X_s + o_{a.s.}(1)) ds + o_{a.s.}(1) \\
&= \frac{1}{2} L_X(rT, X_{i\Delta_{n,T}}) + o_{a.s.}(1) \\
&= \frac{1}{2} \sigma^2(X_{i\Delta_{n,T}}) \overline{L}_X(rT, X_{i\Delta_{n,T}}) + o_{a.s.}(1),
\end{aligned}
$$

by virtue of (3.5). Now, as in Theorem 3.4 in Phillips and Ploberger (1996), expanding the probability space as needed, we have

$$
\left(\mathbf{U}_{n,T}^{X_{i\Delta_{n,T}}}\right)^2 / \left[\mathbf{U}_{n,T}^{X_{i\Delta_{n,T}}}\right]_1 = O_p(1),
$$

and then it follows that

$$
\sqrt{\overline{L}_X(T, X_{i\Delta_{n,T}})\varepsilon_{n,T}} \left(\frac{\frac{1}{\Delta_{n,T}} \frac{\Delta_{n,T}}{2\varepsilon_{n,T}} \sum_{j=1}^{[nr]-1} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}\}} \int_{j\Delta_{n,T}}^{(j+1)\Delta_{n,T}} \sigma(X_s) dB_s}{\frac{\Delta_{n,T}}{2\varepsilon_{n,T}} \sum_{j=1}^{n-1} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}\}}} \right) = O_p(1).
$$

This result implies that the bound (7.14) becomes

$$C_6(\kappa_{n,T}) + O_{a.s.}\left(\sqrt{\frac{1}{\overline{L}_X(T, X_{i\Delta_{n,T}})\varepsilon_{n,T}}}\right) \overset{a.s}{\to} 0.$$

In fact, $\sqrt{\varepsilon_{n,T}\overline{L}_X(T, X_{i\Delta_{n,T}})} \overset{a.s.}{\to} \infty$ as $n, T \to \infty$ since we control $\varepsilon_{n,T}$ to ensure that this property holds. This proves the stated result.

Proof of Theorem 3: Write the estimation error in two components as follows

$$\begin{aligned}
&\frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\widetilde{\mu}_{n,T}(X_{i\Delta_{n,T}})}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} - \mu(x) \\
= \; & \underbrace{\frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\widetilde{\mu}_{n,T}(X_{i\Delta_{n,T}})}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} - \frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\mu(X_{i\Delta_{n,T}})}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}}_{\text{term } V} \\
& + \underbrace{\frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\mu(X_{i\Delta_{n,T}})}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} - \frac{\mu(x)\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}}_{\text{term } B} \\
= \; & \text{term } V + \text{term } B.
\end{aligned}$$

Roughly speaking, this is a decomposition into a bias term B and a second effect, V. We start with the bias term B. Combining the two fractions constituting B, we obtain

$$\frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\left(\mu(X_{i\Delta_{n,T}})-\mu(x)\right)}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}.$$

By Lemma 5 (c.f. the proof of Theorem 2), we find that

$$\begin{aligned}
&\frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\left(\mu(X_{i\Delta_{n,T}})-\mu(x)\right)}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} \\
= \; & \frac{\frac{1}{h_{n,T}}\int_0^T \mathbf{K}\left(\frac{X_s-x}{h_{n,T}}\right)\left(\mu(X_s)-\mu(x)\right)ds + o_{a.s.}(1)}{\frac{1}{h_{n,T}}\int_0^T \mathbf{K}\left(\frac{X_s-x}{h_{n,T}}\right)ds + o_{a.s.}(1)} \\
= \; & \frac{\frac{1}{h_{n,T}}\int_{-\infty}^{\infty} \mathbf{K}\left(\frac{a-x}{h_{n,T}}\right)\left(\mu(a)-\mu(x)\right)s(a)da + o_{a.s.}(1)}{\frac{1}{h_{n,T}}\int_{-\infty}^{\infty} \mathbf{K}\left(\frac{a-x}{h_{n,T}}\right)s(a)da + o_{a.s.}(1)} + o_{a.s.}(1).
\end{aligned}$$

Neglecting the smaller orders of magnitude we can write

$$\begin{aligned}
&\frac{\frac{1}{h_{n,T}}\int_{-\infty}^{\infty} \mathbf{K}\left(\frac{a-x}{h_{n,T}}\right)\left(\mu(a)-\mu(x)\right)s(a)da}{\frac{1}{h_{n,T}}\int_{-\infty}^{\infty} \mathbf{K}\left(\frac{a-x}{h_{n,T}}\right)s(a)da} \\
= \; & \frac{\int_{-\infty}^{\infty} \mathbf{K}(u)\left(\mu(x+uh_{n,T})-\mu(x)\right)s(x+uh_{n,T})du}{\int_{-\infty}^{\infty} \mathbf{K}(u)\, s(x+uh_{n,T})du}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\int_{-\infty}^{\infty} \mathbf{K}(u)\left(uh_{n,T}\mu^{'}(x)+(uh_{n,T})^2\frac{1}{2}\mu^{''}(x)+o(u^2h_{n,T}^2)\right)}{\int_{-\infty}^{\infty}\mathbf{K}(u)\,s(x+uh_{n,T})du}\times \\
&\quad \times\left(s(x)+uh_{n,T}s^{'}(x)+(uh_{n,T})^2\frac{1}{2}s^{''}(x)+o(u^2h_{n,T}^2)\right)du \\
&= h_{n,T}^2\mathbf{K}_1\left(\frac{1}{2}\mu^{''}(x)+\frac{\mu^{'}(x)s^{'}(x)}{s(x)}\right)+o(h_{n,T}^2)
\end{aligned}
$$

where $\mathbf{K}_1=\int u^2\mathbf{K}(u)du$ by the symmetry of $\mathbf{K}(.)$ (c.f. Assumption 2) and dominated convergence. Now consider the term V, viz.

$$
\begin{aligned}
& \frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\widetilde{\mu}_{n,T}(X_{i\Delta_{n,T}})}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}-\frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\mu(X_{i\Delta_{n,T}})}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} \\
&= \frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\left(\widetilde{\mu}_{n,T}(X_{i\Delta_{n,T}})-\mu(X_{i\Delta_{n,T}})\right)}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}.
\end{aligned}
$$

The numerator can be written as

$$
\begin{aligned}
& \frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\left(\widetilde{\mu}_{n,T}(X_{i\Delta_{n,T}})-\mu(X_{i\Delta_{n,T}})\right) \\
&= \frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\leq\varepsilon_{n,T}\}}\frac{n}{T}\left[\left(X_{(j+1)T/n}-X_{jT/n}\right)-\mu(X_{iT/n})\Delta_{n,T}\right]}{\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\leq\varepsilon_{n,T}\}}} \\
&\quad -\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\frac{\left[\mu(X_{iT/n})-\mu(X_{i\Delta_{n,T}})\right]\frac{\Delta_{n,T}}{2\varepsilon_{n,T}}\mathbf{1}_{\{|X_{n\Delta_{n,T}}-X_{i\Delta_{n,T}}|\leq\varepsilon_{n,T}\}}}{\frac{\Delta_{n,T}}{2\varepsilon_{n,T}}\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\leq\varepsilon_{n,T}\}}} \\
&= \mathbf{V}_1^{num}+\mathbf{V}_2^{num}.
\end{aligned}
$$

It is immediate to prove that $\mathbf{V}_2^{num}=O\left(\frac{\Delta_{n,T}}{\varepsilon_{n,T}}\right)$. Notice that

$$
X_{(j+1)T/n}-X_{jT/n}=\int_{jT/n}^{(j+1)T/n}\mu(X_s)ds+\int_{jT/n}^{(j+1)T/n}\sigma(X_s)dB_s.
$$

Hence,

$$
\begin{aligned}
&\mathbf{V}_1^{num}= \\
&\underbrace{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\leq\varepsilon_{n,T}\}}\frac{n}{T}\left[\int_{jT/n}^{(j+1)T/n}\left(\mu(X_s)-\mu(X_{iT/n})\right)ds\right]}{\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\leq\varepsilon_{n,T}\}}}}_{(\mathbf{A}_{n,T})} \\
&+\underbrace{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\leq\varepsilon_{n,T}\}}\frac{n}{T}\left[\int_{jT/n}^{(j+1)T/n}\sigma(X_s)dB_s\right]}{\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\leq\varepsilon_{n,T}\}}}}_{(\mathbf{B}_{n,T}(1))}.
\end{aligned}
$$

These two components comprise an additional bias effect, $\mathbf{A}_{n,T}$, and a variance effect, $\mathbf{B}_{n,T}(1)$. First, examine $\sqrt{\varepsilon_{n,T}}\mathbf{B}_{n,T}(r)$, viz.

$$\sqrt{\varepsilon_{n,T}}\mathbf{B}_{n,T}(r) = \frac{\Delta_{n,T}}{h_{n,T}} \sum_{i=1}^{[nr]} \mathbf{K}\left(\frac{X_{i\Delta_{n,T}} - x}{h_{n,T}}\right) \frac{\frac{1}{2\sqrt{\varepsilon_{n,T}}} \sum_{j=1}^{n-1} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{i\Delta_{n,T}}| \le \varepsilon_{n,T}\}} \left[\int_{jT/n}^{(j+1)T/n} \sigma(X_s) dB_s\right]}{\frac{\Delta_{n,T}}{2\varepsilon_{n,T}} \sum_{j=1}^{n} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{i\Delta_{n,T}}| \le \varepsilon_{n,T}\}}}.$$

The term $\mathbf{B}_{n,T}(r)$ has a quadratic variation which can be analysed as follows:

$$\begin{aligned}
& [\mathbf{B}_{n,T}]_r \\
= \; & \left(\frac{\Delta_{n,T}}{h_{n,T}}\right)^2 \sum_{i=1}^{[nr]} \sum_{k=1}^{[nr]} \mathbf{K}\left(\frac{X_{i\Delta_{n,T}} - x}{h_{n,T}}\right) \mathbf{K}\left(\frac{X_{k\Delta_{n,T}} - x}{h_{n,T}}\right) \times \\
& \times \frac{\frac{1}{4}\left(\frac{1}{\sqrt{\varepsilon_{n,T}}}\right)^2 \sum_{j=1}^{n-1} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{i\Delta_{n,T}}| \le \varepsilon_{n,T}\}} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{k\Delta_{n,T}}| \le \varepsilon_{n,T}\}} \left[\int_{jT/n}^{(j+1)T/n} \sigma^2(X_s) ds\right]}{\left(\frac{\Delta_{n,T}}{2\varepsilon_{n,T}} \sum_{j=1}^{n} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{i\Delta_{n,T}}| \le \varepsilon_{n,T}\}}\right) \left(\frac{\Delta_{n,T}}{2\varepsilon_{n,T}} \sum_{j=1}^{n} \mathbf{1}_{\{|X_{j\Delta_{n,T}} - X_{k\Delta_{n,T}}| \le \varepsilon_{n,T}\}}\right)} \\
= \; & \left(\frac{1}{h_{n,T}}\right)^2 \int_0^{[Tr]} ds \int_0^{[Tr]} du \mathbf{K}\left(\frac{X_s - x}{h_{n,T}}\right) \mathbf{K}\left(\frac{X_u - x}{h_{n,T}}\right) \times \\
& \times \frac{\frac{1}{4\varepsilon_{n,T}} \int_0^T db \mathbf{1}_{\{|X_b - X_s| \le \varepsilon_{n,T}\}} \mathbf{1}_{\{|X_b - X_u| \le \varepsilon_{n,T}\}} \sigma^2(X_b + o_{a.s}(1))}{\left(\frac{1}{2\varepsilon_{n,T}} \int_0^T \mathbf{1}_{\{|b-s| \le \varepsilon_{n,T}\}} db\right) \left(\frac{1}{2\varepsilon_{n,T}} \int_0^T \mathbf{1}_{\{|b-u| \le \varepsilon_{n,T}\}} db\right)} + o_{a.s}(1) \\
= \; & \left(\frac{1}{h_{n,T}}\right)^2 \int_{-\infty}^{\infty} ds \int_{-\infty}^{+\infty} du \mathbf{K}\left(\frac{s - x}{h_{n,T}}\right) \mathbf{K}\left(\frac{u - x}{h_{n,T}}\right) \times \\
& \times \frac{\frac{1}{4\varepsilon_{n,T}} \int_{-\infty}^{\infty} db \mathbf{1}_{\{|b-s| \le \varepsilon_{n,T}\}} \mathbf{1}_{\{|b-u| \le \varepsilon_{n,T}\}} \sigma^2(b) \overline{L}_X(T,b) \overline{L}_X(rT,s) \overline{L}_X(rT,u)}{\left(\frac{1}{2\varepsilon_{n,T}} \int_{-\infty}^{\infty} \mathbf{1}_{\{|b-s| \le \varepsilon_{n,T}\}} \overline{L}(T,b) db\right) \left(\frac{1}{2\varepsilon_{n,T}} \int_{-\infty}^{\infty} \mathbf{1}_{\{|b-u| \le \varepsilon_{n,T}\}} \overline{L}(T,b) db\right)} + o_{a.s}(1).
\end{aligned}$$

provided $h_{n,T}$ and $\varepsilon_{n,T}$ are such that $\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T})^{\alpha} = O_{a.s.}(1)$ for some $\alpha \in \left(0, \frac{1}{2}\right)$ and $\frac{\overline{L}_X(T,x)}{\varepsilon_{n,T}}(\Delta_{n,T})^{\beta} = O_{a.s.}(1)$ for some $\beta \in \left(0, \frac{1}{2}\right)$ as $n, T \to \infty$. Let

$$\frac{s - x}{h_{n,T}} = a \text{ and } \frac{u - x}{h_{n,T}} = e.$$

Then,

$$\begin{aligned}
& \frac{1}{4\varepsilon_{n,T}} \int_{-\infty}^{\infty} da \int_{-\infty}^{+\infty} de \mathbf{K}(a) \mathbf{K}(e) \times \\
& \times \frac{\int_{-\infty}^{\infty} db \mathbf{1}_{\{|b-x-ah_{n,T}| \le \varepsilon_{n,T}\}} \mathbf{1}_{\{|b-x-eh_{n,T}| \le \varepsilon_{n,T}\}} \sigma^2(b) \overline{L}_X(T,b) \overline{L}_X(rT, x+ah_{n,T}) \overline{L}_X(rT, x+eh_{n,T})}{\left(\frac{1}{2\varepsilon_{n,T}} \int_{-\infty}^{\infty} \mathbf{1}_{\{|b-x-ah_{n,T}| \le \varepsilon_{n,T}\}} \overline{L}_X(T,b) db\right) \left(\frac{1}{2\varepsilon_{n,T}} \int_{-\infty}^{\infty} \mathbf{1}_{\{|b-x-eh_{n,T}| \le \varepsilon_{n,T}\}} \overline{L}_X(T,b) db\right)} \\
& + o_{a.s}(1) \\
= \; & \frac{1}{4\varepsilon_{n,T}} \int_{-\infty}^{\infty} da \int_{-\infty}^{+\infty} de \mathbf{K}(a) \mathbf{K}(e) \times \\
& \times \frac{\int_{-\infty}^{\infty} db \mathbf{1}_{\{|\frac{b-x}{\varepsilon_{n,T}} - a\frac{h_{n,T}}{\varepsilon_{n,T}}| \le 1\}} \mathbf{1}_{\{|\frac{b-x}{\varepsilon_{n,T}} - e\frac{h_{n,T}}{\varepsilon_{n,T}}| \le 1\}} \sigma^2(b) \overline{L}_X(T,b) \overline{L}_X(rT, x+ah_{n,T}) \overline{L}_X(rT, x+eh_{n,T})}{\left(\frac{1}{2\varepsilon_{n,T}} \int_{-\infty}^{\infty} \mathbf{1}_{\{|\frac{b-x}{\varepsilon_{n,T}} - a\frac{h_{n,T}}{\varepsilon_{n,T}}| \le 1\}} \overline{L}_X(T,b) db\right) \left(\frac{1}{2\varepsilon_{n,T}} \int_{-\infty}^{\infty} \mathbf{1}_{\{|\frac{b-x}{\varepsilon_{n,T}} - e\frac{h_{n,T}}{\varepsilon_{n,T}}| \le 1\}} \overline{L}_X(T,b) db\right)}.
\end{aligned}$$

Setting,

$$\frac{b - x}{\varepsilon_{n,T}} = z,$$

this last expression becomes

$$\int_{-\infty}^{\infty} da \int_{-\infty}^{+\infty} de \mathbf{K}(a)\mathbf{K}(e) \times$$

$$\frac{\frac{1}{4}\int_{-\infty}^{\infty} dz \mathbf{1}_{\{|z-a\frac{h_{n,T}}{\varepsilon_{n,T}}|\leq 1\}}\mathbf{1}_{\{|z-e\frac{h_{n,T}}{\varepsilon_{n,T}}|\leq 1\}}\sigma^2(x)\overline{L}_X(T,x+z\varepsilon_{n,T})\overline{L}_X(rT,x+ah_{n,T})\overline{L}_X(rT,x+eh_{n,T})}{\left(\frac{1}{2}\int_{-\infty}^{\infty}\mathbf{1}_{\{|z-a\frac{h_{n,T}}{\varepsilon_{n,T}}|\leq 1\}}\overline{L}_X(T,x+z\varepsilon_{n,T})dz\right)\left(\frac{1}{2}\int_{-\infty}^{\infty}\mathbf{1}_{\{|z-e\frac{h_{n,T}}{\varepsilon_{n,T}}|\leq 1\}}\overline{L}_X(T,x+z\varepsilon_{n,T})dz\right)}$$

$$+o_{a.s}(1).$$

Now, if $h_{n,T} = o(\varepsilon_{n,T})$, then

$$[\mathbf{B}_{n,T}]_r \overset{a.s.}{\to} \frac{1}{2}\sigma^2(x)\frac{\left(\overline{L}_X(rT,x)\right)^2}{\overline{L}_X(T,x)},$$

whereas if $h_{n,T} = O(\varepsilon_{n,T})$ with $h_{n,T}/\varepsilon_{n,T} \to \phi > 0$, then

$$[\mathbf{B}_{n,T}]_r \overset{a.s.}{\to} \frac{1}{2}\theta_\phi\sigma^2(x)\frac{\left(\overline{L}_X(rT,x)\right)^2}{\overline{L}_X(T,x)},$$

where

$$\begin{aligned}
\theta_\phi &= \int_{-\infty}^{\infty} da \int_{-\infty}^{+\infty} de \mathbf{K}(a)\mathbf{K}(e)\left(\frac{\frac{1}{2}\int_{-\infty}^{\infty} dz \mathbf{1}_{\{|z-\phi a|\leq 1\}}\mathbf{1}_{\{|z-\phi e|\leq 1\}}}{\left(\frac{1}{2}\int_{-\infty}^{\infty}\mathbf{1}_{\{|z-\phi a|\leq 1\}}dz\right)\left(\frac{1}{2}\int_{-\infty}^{\infty}\mathbf{1}_{\{|z-\phi e|\leq 1\}}dz\right)}\right) \\
&= \int_{-\infty}^{\infty} da \int_{-\infty}^{+\infty} de \mathbf{K}(a)\mathbf{K}(e)\frac{1}{2}\int_{-\infty}^{\infty} dz \mathbf{1}_{\{|z-\phi a|\leq 1\}}\mathbf{1}_{\{|z-\phi e|\leq 1\}} \\
&= \frac{1}{2}\int_{-\infty}^{\infty} da \int_{-\infty}^{+\infty} de \int_{-\infty}^{\infty} dz \mathbf{K}(a)\mathbf{K}(e)\mathbf{1}_{\{|z-\phi a|\leq 1\}}\mathbf{1}_{\{|z-\phi e|\leq 1\}} \\
&= \frac{1}{2}\int_{-\infty}^{\infty} dz \int_{(z-1)/\phi}^{(z+1)/\phi} de \int_{(z-1)/\phi}^{(z+1)/\phi} da \mathbf{K}(a)\mathbf{K}(e).
\end{aligned}$$

By earlier arguments (e.g. the proof of Lemma 4) the above results imply that

$$\sqrt{\varepsilon_{n,T}}\left(\frac{\mathbf{B}_{n,T}(1)}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}\right) \Rightarrow \mathbf{MN}\left(0,\frac{1}{2}\frac{\sigma^2(x)}{\overline{L}_X(T,x)}\right)$$

and

$$\sqrt{\varepsilon_{n,T}\overline{L}_X(T,x)}\left(\frac{\mathbf{B}_{n,T}(1)}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}\right) \Rightarrow \mathbf{N}\left(0,\frac{1}{2}\sigma^2(x)\right)$$

provided $h_{n,T} = o(\varepsilon_{n,T})$. If $h_{n,T} = O(\varepsilon_{n,T})$ and $h_{n,T}/\varepsilon_{n,T} \to \phi > 0$, then

$$\sqrt{\varepsilon_{n,T}\overline{L}_X(T,x)}\left(\frac{\mathbf{B}_{n,T}(1)}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}\right) \Rightarrow \mathbf{N}\left(0,\frac{1}{2}\theta_\phi\sigma^2(x)\right).$$

Next, examine the additional bias term $\mathbf{A}_{n,T}$. We have

$$\frac{\mathbf{A}_{n,T}}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} = \frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\leq\varepsilon_{n,T}\}}\left(\mu(X_{jT/n})-\mu(X_{iT/n})\right)}{\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\leq\varepsilon_{n,T}\}}}}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n \mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}$$

$$+\frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\leq\varepsilon_{n,T}\}}\frac{n}{T}\left[\int_{jT/n}^{(j+1)T/n}\mu(X_s)-\mu(X_{jT/n})ds\right]}{\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\leq\varepsilon_{n,T}\}}}}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}$$

$$=\frac{\frac{1}{h_{n,T}}\int_{-\infty}^{\infty}\mathbf{K}\left(\frac{a-x}{h_{n,T}}\right)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|b-a|\leq\varepsilon_{n,T}\}}(\mu(b)-\mu(a))s(b)db}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|b-a|\leq\varepsilon_{n,T}\}}s(b)db}s(a)da}{\frac{1}{h_{n,T}}\int_{-\infty}^{\infty}\mathbf{K}\left(\frac{a-x}{h_{n,T}}\right)s(a)da}+O_{a.s}\left(\Delta_{n,T}^{1/2}\right).$$

Neglecting the order term we can write

$$=\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\frac{1}{\varepsilon_{n,T}}\int_{-\infty}^{\infty}\mathbf{1}_{\left\{\left|\frac{b-x-h_{n,T}c}{\varepsilon_{n,T}}\right|\leq 1\right\}}(\mu(b)-\mu(x+h_{n,T}c))s(b)db}{\frac{1}{\varepsilon_{n,T}}\int_{-\infty}^{\infty}\mathbf{1}_{\left\{\left|\frac{b-x-h_{n,T}c}{\varepsilon_{n,T}}\right|\leq 1\right\}}s(b)db}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}$$

$$=\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\left\{\left|a-\frac{h_{n,T}c}{\varepsilon_{n,T}}\right|\leq 1\right\}}(\mu(x+\varepsilon_{n,T}a)-\mu(x))s(x+\varepsilon_{n,T}a)da}{\int_{-\infty}^{\infty}\mathbf{1}_{\left\{\left|a-\frac{h_{n,T}c}{\varepsilon_{n,T}}\right|\leq 1\right\}}s(x+\varepsilon_{n,T}a)da}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}$$

$$+\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\left\{\left|a-\frac{h_{n,T}c}{\varepsilon_{n,T}}\right|\leq 1\right\}}(\mu(x)-\mu(x+h_{n,T}c))s(x+\varepsilon_{n,T}a)da}{\int_{-\infty}^{\infty}\mathbf{1}_{\left\{\left|a-\frac{h_{n,T}c}{\varepsilon_{n,T}}\right|\leq 1\right\}}s(x+\varepsilon_{n,T}a)da}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}$$

$$=\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\left\{\left|a-\frac{h_{n,T}c}{\varepsilon_{n,T}}\right|\leq 1\right\}}\left(\mu^{'}(x)\varepsilon_{n,T}a+\frac{1}{2}\mu^{''}(x)(\varepsilon_{n,T}a)^2+o\right)\left(s(x)+\varepsilon_{n,T}as^{'}(x)+\frac{1}{2}s^{''}(x)(\varepsilon_{n,T}a)^2+o\right)da}{\int_{-\infty}^{\infty}\mathbf{1}_{\left\{\left|a-\frac{h_{n,T}c}{\varepsilon_{n,T}}\right|\leq 1\right\}}s(x+\varepsilon_{n,T}a)da}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}$$

$$-\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\left\{\left|a-\frac{h_{n,T}c}{\varepsilon_{n,T}}\right|\leq 1\right\}}\left(\mu^{'}(x)h_{n,T}c+\frac{1}{2}\mu^{''}(x)(h_{n,T}c)^2+o\right)\left(s(x)+\varepsilon_{n,T}as^{'}(x)+\frac{1}{2}s^{''}(x)(\varepsilon_{n,T}a)^2+o\right)da}{\int_{-\infty}^{\infty}\mathbf{1}_{\left\{\left|a-\frac{h_{n,T}c}{\varepsilon_{n,T}}\right|\leq 1\right\}}s(x+\varepsilon_{n,T}a)da}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}.$$

Hence, if $h_{n,T}=o(\varepsilon_{n,T})$, then

$$\frac{\mathbf{A}_{n,T}}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}$$

$$=\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a|\leq 1\}}\left(\mu^{'}(x)\varepsilon_{n,T}a+\frac{1}{2}\mu^{''}(x)(\varepsilon_{n,T}a)^2+o\right)\left(s(x)+\varepsilon_{n,T}as^{'}(x)+\frac{1}{2}s^{''}(x)(\varepsilon_{n,T}a)^2+o\right)da}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a|\leq 1\}}s(x+\varepsilon_{n,T}a)da}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}$$

$$-\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a|\leq 1\}}\left(\mu^{'}(x)h_{n,T}c+\frac{1}{2}\mu^{''}(x)(h_{n,T}c)^2+o\right)\left(s(x)+\varepsilon_{n,T}as^{'}(x)+\frac{1}{2}s^{''}(x)(\varepsilon_{n,T}a)^2+o\right)da}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a|\leq 1\}}s(x+\varepsilon_{n,T}a)da}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}$$

$$=\left(\varepsilon_{n,T}\right)^2\mathbf{K}_1^{ind}\left[\mu^{'}(x)\frac{s^{'}(x)}{s(x)}+\frac{1}{2}\mu^{''}(x)\right]+o\left(\varepsilon_{n,T}^2\right)+o\left(h_{n,T}^2\right),$$

where $\mathbf{K}_1^{ind}=\frac{1}{2}\int_{-\infty}^{\infty}a^2\mathbf{1}_{\{|a|\leq 1\}}da=\frac{1}{3}$. Now suppose $h_{n,T}=O(\varepsilon_{n,T})$ and $\frac{h_{n,T}}{\varepsilon_{n,T}}\to\phi$. Then,

$$\frac{\mathbf{A}_{n,T}}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}=$$

$$\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\le 1\}}\left(\mu^{'}(x)\varepsilon_{n,T}a+\frac{1}{2}\mu^{''}(x)(\varepsilon_{n,T}a)^2+o\right)\left(s(x)+\varepsilon_{n,T}as^{'}(x)+\frac{1}{2}s^{''}(x)(\varepsilon_{n,T}a)^2+o\right)da}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\le 1\}}s(x+\varepsilon_{n,T}a)da}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)s(x+h_{n,T}c)dc} \tag{7.15}$$

$$-\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\le 1\}}\left(\mu^{'}(x)h_{n,T}c+\frac{1}{2}\mu^{''}(x)(h_{n,T}c)^2+o\right)\left(s(x)+\varepsilon_{n,T}as^{'}(x)+\frac{1}{2}s^{''}(x)(\varepsilon_{n,T}a)^2+o\right)da}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\le 1\}}s(x+\varepsilon_{n,T}a)da}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)s(x+h_{n,T}c)dc} \tag{7.16}$$

We can use the transformation $g = a - \phi c$ and write (7.15) as

$$\begin{aligned}
&\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\mu^{'}(x)\varepsilon_{n,T}(g+\phi c)+\frac{1}{2}\mu^{''}(x)(\varepsilon_{n,T}(g+\phi c))^2+o\right)\left(s(x)+\varepsilon_{n,T}(g+\phi c)s^{'}(x)+\frac{1}{2}s^{''}(x)(\varepsilon_{n,T}(g+\phi c))^2+o\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}}{\int_{-\infty}^{\infty}\mathbf{K}(c)s(x+h_{n,T}c)dc}\times \\
&\times s(x+h_{n,T}c)dc \\
=&\ \frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\mu^{'}(x)s(x)\varepsilon_{n,T}(g+\phi c)+\frac{1}{2}\mu^{''}(x)s(x)(\varepsilon_{n,T}(g+\phi c))^2+\varepsilon_{n,T}^2(g+\phi c)^2s^{'}(x)\mu^{'}(x)\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)s(x+h_{n,T}c)dc}+o(\varepsilon_{n,T}^2) \\
=&\ \frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\mu^{'}(x)s(x)\varepsilon_{n,T}(\phi c)+\frac{1}{2}\mu^{''}(x)s(x)\varepsilon_{n,T}^2(g^2+\phi^2c^2)+\varepsilon_{n,T}^2(g^2+\phi^2c^2)s^{'}(x)\mu^{'}(x)\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)s(x+h_{n,T}c)dc}+o(\varepsilon_{n,T}^2).
\end{aligned}$$

by the symmetry of the indicator kernel. Then,

$$\begin{aligned}
&\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\mu^{'}(x)s(x)\varepsilon_{n,T}(\phi c)+\frac{1}{2}\mu^{''}(x)s(x)\varepsilon_{n,T}^2(g^2+\phi^2c^2)+\varepsilon_{n,T}^2(g^2+\phi^2c^2)s^{'}(x)\mu^{'}(x)\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)s(x+h_{n,T}c)dc} \\
=&\ \frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\mu^{'}(x)s(x)\varepsilon_{n,T}(\phi c)\frac{1}{\frac{1}{2}\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)s(x+h_{n,T}c)dc} \\
&+\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\frac{1}{2}\mu^{''}(x)s(x)\varepsilon_{n,T}^2\phi^2c^2\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)s(x+h_{n,T}c)dc} \\
&+\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\frac{1}{2}\mu^{''}(x)s(x)\varepsilon_{n,T}^2g^2\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)s(x+h_{n,T}c)dc} \\
&+\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\varepsilon_{n,T}^2g^2s^{'}(x)\mu^{'}(x)\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)s(x+h_{n,T}c)dc} \\
&+\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\varepsilon_{n,T}^2\phi^2c^2s^{'}(x)\mu^{'}(x)\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)s(x+h_{n,T}c)dc}.
\end{aligned}$$

But,

$$\begin{aligned}
&\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\mu^{'}(x)s(x)\varepsilon_{n,T}(\phi c)\frac{1}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)s(x+h_{n,T}c)dc} \\
=&\ \mathbf{K}_1\phi^2\varepsilon_{n,T}^2\frac{\mu^{'}(x)}{s(x)}s^{'}(x),
\end{aligned}$$

$$\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\frac{1}{2}\mu^{''}(x)s(x)\varepsilon_{n,T}^2\phi^2c^2\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}$$

$$= \mathbf{K}_1\phi^2\varepsilon_{n,T}^2\frac{1}{2}\mu^{''}(x),$$

$$\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\frac{1}{2}\mu^{''}(x)s(x)\varepsilon_{n,T}^2g^2\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}$$

$$= \mathbf{K}_1^{ind}\varepsilon_{n,T}^2\frac{1}{2}\mu^{''}(x),$$

$$\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\varepsilon_{n,T}^2g^2s^{'}(x)\mu^{'}(x)\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}$$

$$= \mathbf{K}_1^{ind}\varepsilon_{n,T}^2\frac{s^{'}(x)\mu^{'}(x)}{s(x)},$$

$$\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\varepsilon_{n,T}^2\phi^2c^2s^{'}(x)\mu^{'}(x)\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}$$

$$= \mathbf{K}_1\phi^2\varepsilon_{n,T}^2\frac{s^{'}(x)\mu^{'}(x)}{s(x)}.$$

Now, write (7.16) as

$$\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\mu^{'}(x)h_{n,T}c+\frac{1}{2}\mu^{''}(x)(h_{n,T}c)^2+o\right)\left(s(x)+\varepsilon_{n,T}(g+\phi c)s^{'}(x)+\frac{1}{2}s^{''}(x)(\varepsilon_{n,T}(g+\phi c))^2+o\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}$$

$$= \frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\mu^{'}(x)s(x)h_{n,T}c+\frac{1}{2}\mu^{''}(x)s(x)(h_{n,T}c)^2+\varepsilon_{n,T}h_{n,T}(g+\phi c)cs^{'}(x)\mu^{'}(x)\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}+o(\varepsilon_{n,T}^2)$$

$$= \frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\mu^{'}(x)s(x)h_{n,T}c+\frac{1}{2}\mu^{''}(x)s(x)h_{n,T}^2c^2+\varepsilon_{n,T}h_{n,T}\phi c^2s^{'}(x)\mu^{'}(x)\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}+o(\varepsilon_{n,T}^2),$$

by the symmetry of the indicator kernel. Then,

$$\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\mu^{'}(x)s(x)h_{n,T}c+\frac{1}{2}\mu^{''}(x)s(x)h_{n,T}^2c^2+\varepsilon_{n,T}h_{n,T}\phi c^2s^{'}(x)\mu^{'}(x)\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}s(x+h_{n,T}c)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}$$

$$= \frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\,\mu^{'}(x)s(x)h_{n,T}c\frac{1}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}$$

$$
+\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\leq 1\}}\left(\frac{1}{2}\mu^{''}(x)s(x)h_{n,T}^{2}c^{2}\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\leq 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}
+\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\leq 1\}}\left(h_{n,T}\varepsilon_{n,T}\phi c^{2}s^{'}(x)\mu^{'}(x)\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\leq 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}.
$$

But,

$$
\begin{aligned}
&\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\,\mu^{'}(x)s(x)h_{n,T}c\frac{1}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\leq 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}\\
&= \mathbf{K}_{1}\phi^{2}\varepsilon_{n,T}^{2}\frac{\mu^{'}(x)}{s(x)}s^{'}(x),
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\leq 1\}}\left(\frac{1}{2}\mu^{''}(x)s(x)h_{n,T}^{2}c^{2}\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\leq 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}\\
&= \mathbf{K}_{1}\phi^{2}\varepsilon_{n,T}^{2}\frac{1}{2}\mu^{''}(x),
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\leq 1\}}\left(h_{n,T}\varepsilon_{n,T}\phi c^{2}s^{'}(x)\mu^{'}(x)\right)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\leq 1\}}s(x+\varepsilon_{n,T}(g+\phi c))dg}\left(s(x)+s^{'}(x)h_{n,T}c+o\right)dc}{\int_{-\infty}^{\infty}\mathbf{K}(c)\,s(x+h_{n,T}c)dc}\\
&= \mathbf{K}_{1}\phi^{2}\varepsilon_{n,T}^{2}\frac{\mu^{'}(x)}{s(x)}s^{'}(x).
\end{aligned}
$$

To conclude, when $h_{n,T}=O(\varepsilon_{n,T})$ with $\frac{h_{n,T}}{\varepsilon_{n,T}}\to\phi$, then

$$
\begin{aligned}
&\frac{\mathbf{A}_{n,T}}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}\\
&= \varepsilon_{n,T}^{2}\left(\mathbf{K}_{1}\phi^{2}+\mathbf{K}_{1}^{ind}\right)\left(\frac{1}{2}\mu^{''}(x)+\frac{s^{'}(x)}{s(x)}\mu^{'}(x)\right).
\end{aligned}
$$

In consequence, defining the estimation error decomposition as

$$
E=B+\frac{\mathbf{A}_{n,T}}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}+\frac{\mathbf{B}_{n,T}(1)}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)},
$$

we can write

$$
\begin{aligned}
&\sqrt{\varepsilon_{n,T}\overline{L}_{X}(T,x)}\left(\frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\widetilde{\mu}_{n,T}(X_{i\Delta_{n,T}})}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}-\mu(x)\right)\\
&= \sqrt{\varepsilon_{n,T}\overline{L}_{X}(T,x)}\Bigg(\frac{\mathbf{B}_{n,T}(1)}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}\\
&\quad +O\left(\Delta_{n,T}^{1/2}\right)+O\left(\frac{\Delta_{n,T}}{\varepsilon_{n,T}}\right)+O(h_{n,T}^{2})+O(\varepsilon_{n,T}^{2})\Bigg)\\
&\Rightarrow \mathbf{N}\left(0,\frac{1}{2}\sigma^{2}(x)\right),
\end{aligned}
$$

if $h_{n,T}=o(\varepsilon_{n,T})$, $\varepsilon^5_{n,T}\overline{L}_X(T,x)\overset{a.s.}{\to}0$ and $\varepsilon_{n,T}\overline{L}_X(T,x)\overset{a.s.}{\to}\infty$. If $h_{n,T}=o(\varepsilon_{n,T})$, $\varepsilon_{n,T}\overline{L}_X(T,x)\overset{a.s.}{\to}\infty$ and $\varepsilon^5_{n,T}=O_p\left(\overline{L}_X(T,x)\right)$, then

$$\sqrt{\varepsilon_{n,T}\overline{L}_X(T,x)}\left(\frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\widetilde{\mu}_{n,T}(X_{i\Delta_{n,T}})}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}-\mu(x)-\Gamma_\mu(x)\right)\Rightarrow\mathbf{N}\left(0,\frac{1}{2}\sigma^2(x)\right),$$

where

$$\Gamma_\mu(x)=(\varepsilon_{n,T})^2\,\mathbf{K}_1^{ind}\left[\frac{1}{2}\mu^{''}(x)+\mu^{'}(x)\frac{s^{'}(x)}{s(x)}\right],$$

with $\mathbf{K}_1^{ind}=\frac{1}{2}\int_{-\infty}^{\infty}\mathbf{1}_{\{|a|\le 1\}}a^2da=\frac{1}{3}$. If $h_{n,T}=O(\varepsilon_{n,T})$ with $h_{n,T}/\varepsilon_{n,T}\to\phi>0$, $\varepsilon^5_{n,T}\overline{L}_X(T,x)\overset{a.s.}{\to}0$ and $\varepsilon_{n,T}\overline{L}_X(T,x)\overset{a.s.}{\to}\infty$, then

$$\sqrt{\varepsilon_{n,T}\overline{L}_X(T,x)}\left(\frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\widetilde{\mu}_{n,T}(X_{i\Delta_{n,T}})}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}-\mu(x)\right)\Rightarrow\mathbf{N}\left(0,\frac{1}{2}\theta_\phi\sigma^2(x)\right),$$

where $\theta_\phi=\frac{1}{2}\int_{-\infty}^{\infty}\int_{(z-1)/\phi}^{(z+1)/\phi}\int_{(z-1)/\phi}^{(z+1)/\phi}\mathbf{K}(a)\mathbf{K}(e)dzdade$. If $h_{n,T}=O(\varepsilon_{n,T})$ with $h_{n,T}/\varepsilon_{n,T}\to\phi>0$, $\varepsilon_{n,T}\overline{L}_X(T,x)\overset{a.s.}{\to}\infty$ and $\varepsilon^5_{n,T}=O_p\left(\overline{L}_X(T,x)\right)$, then

$$\sqrt{\varepsilon_{n,T}\overline{L}_X(T,x)}\left(\frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\widetilde{\mu}_{n,T}(X_{i\Delta_{n,T}})}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}-\mu(x)-\Gamma_\mu^{(\phi)}(x)\right)\Rightarrow\mathbf{N}\left(0,\frac{1}{2}\theta_\phi\sigma^2(x)\right),$$

where

$$\Gamma_\mu^{(\phi)}(x)=\varepsilon^2_{n,T}\left(\mathbf{K}_1\phi^2+\mathbf{K}_1^{ind}\right)\left(\frac{1}{2}\mu^{''}(x)+\frac{s^{'}(x)}{s(x)}\mu^{'}(x)\right).$$

This proves the stated result.

Proof of Theorem 4: The proof follows that of Theorem 2 and is omitted here for brevity. See Bandi and Phillips (1998) for full derivations.

Proof of Theorem 5: C.f. the proof of Theorem 3.

Proof of Theorem 6: Fix $T=\overline{T}$. We write the estimation error as follows:

$$\begin{aligned}
&\frac{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)\widetilde{\sigma}^2_{n,\overline{T}}(X_{i\Delta_{n,\overline{T}}})}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}-\sigma^2(x)\\
=\;&\underbrace{\frac{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)\widetilde{\sigma}^2_{n,\overline{T}}(X_{i\Delta_{n,\overline{T}}})}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}-\frac{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)\sigma^2(X_{i\Delta_{n,\overline{T}}})}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}}_{\text{term }V}\\
&+\underbrace{\frac{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)\sigma^2(X_{i\Delta_{n,\overline{T}}})}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}-\frac{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)\sigma^2(x)}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}}_{\text{term }B}\\
=\;&\text{term }V+\text{term }B.
\end{aligned}$$

As earlier in the drift case, we have a bias term B and a second effect, V. We start with the bias term B. Using Lemma 2, the numerator of B can be written as

$$\underbrace{\int_{-\infty}^{\infty}\mathbf{K}(c)\left(\sigma^2(x)\right)' ch_{n,\overline{T}}\left(\frac{L_X(\overline{T},x+ch_{n,T})-L_X(\overline{T},x)}{\sigma^2(x+ch_{n,\overline{T}})}\right)dc}_{B^1_{num}}$$

$$+\underbrace{\int_{-\infty}^{\infty}\mathbf{K}(c)\left(\sigma^2(x)\right)' ch_{n,\overline{T}}\frac{\left(\sigma^2(x)-\sigma^2(x+ch_{n,\overline{T}})\right)}{\sigma^2(x+ch_{n,\overline{T}})\sigma^2(x)}L_X(\overline{T},x)dc}_{B^2_{num}}$$

$$+\underbrace{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{1}{2}\left(\sigma^2(x^*)\right)''\left(ch_{n,\overline{T}}\right)^2\frac{L_X(\overline{T},x+ch_{n,\overline{T}})}{\sigma^2(x+ch_{n,\overline{T}})}dc}_{B^3_{num}}.$$

By Lemma 4 the first term has the following limiting form as a functional of a Brownian sheet $\mathfrak{B}$,

$$\begin{aligned}
B^1_{num} &= 2h^{3/2}_{n,\overline{T}}\frac{\left(\sigma^2(x)\right)'}{\sigma^2(x)}\int_{-\infty}^{\infty}c\mathbf{K}(c)\frac{1}{2\sqrt{h_{n,\overline{T}}}}\left(L_X(\overline{T},x+h_{n,T}c)-L_X(\overline{T},x)\right)dc \qquad (7.17)\\
&= 2h^{3/2}_{n,\overline{T}}\frac{\left(\sigma^2(x)\right)'}{\sigma^2(x)}\int_{0}^{\infty}c\mathbf{K}(c)\frac{1}{2\sqrt{h_{n,\overline{T}}}}\left(L_X(\overline{T},x+h_{n,T}c)-L_X(\overline{T},x)\right)dc\\
&\quad +2h^{3/2}_{n,\overline{T}}\frac{\left(\sigma^2(x)\right)'}{\sigma^2(x)}\int_{-\infty}^{0}c\mathbf{K}(c)\frac{1}{2\sqrt{h_{n,\overline{T}}}}\left(L_X(\overline{T},x+h_{n,\overline{T}}c)-L_X(\overline{T},x)\right)dc\\
&\Rightarrow 2h^{3/2}_{n,\overline{T}}\frac{\left(\sigma^2(x)\right)'}{\sigma^2(x)}\int_{0}^{\infty}c\mathbf{K}(c)\mathfrak{B}^{\oplus}(L_X(\overline{T},x),c)dc\\
&\quad +2h^{3/2}_{n,\overline{T}}\frac{\left(\sigma^2(x)\right)'}{\sigma^2(x)}\int_{-\infty}^{0}c\mathbf{K}(c)\mathfrak{B}^{\otimes}(L_X(\overline{T},x),-c)dc\\
&\overset{d}{=} 2h^{3/2}_{n,\overline{T}}\frac{\left(\sigma^2(x)\right)'}{\sigma(x)}\sqrt{L_X(\overline{T},x)}\left(\int_{0}^{\infty}c\mathbf{K}(c)\mathfrak{B}^{\oplus}(1,c)dc+\int_{-\infty}^{0}c\mathbf{K}(c)\mathfrak{B}^{\otimes}(1,-c)dc\right).
\end{aligned}$$

where $\mathfrak{B}^{\oplus}$ and $\mathfrak{B}^{\otimes}$ are independent Brownian sheets (c.f. Revuz and Yor (1994)). It follows that,

$$\begin{aligned}
&\int_{0}^{\infty}c\mathbf{K}(c)\mathfrak{B}^{\oplus}(1,c)dc\\
&\overset{d}{=}\int_{0}^{\infty}c\mathbf{K}(c)\mathbf{B}^{\oplus}(c)dc\\
&\overset{d}{=}\mathbf{N}\left(0,\int_{0}^{\infty}\int_{0}^{\infty}us\mathbf{K}(u)\mathbf{K}(s)\min(u,s)duds\right).
\end{aligned}$$

Analogously,

$$\begin{aligned}
&\int_{-\infty}^{0}c\mathbf{K}(c)\mathfrak{B}^{\otimes}(1,-c)dc\\
&\overset{d}{=}\int_{-\infty}^{0}c\mathbf{K}(c)\mathbf{B}^{\otimes}(-c)dc\\
&\overset{d}{=}\mathbf{N}\left(0,-\int_{-\infty}^{0}\int_{-\infty}^{0}us\mathbf{K}(u)\mathbf{K}(s)\max(u,s)duds\right).
\end{aligned}$$

In consequence,

$$
\begin{aligned}
&2h_{n,\overline{T}}^{3/2}\frac{\left(\sigma^2(x)\right)^{'}}{\sigma(x)}\sqrt{\overline{L}_X(\overline{T},x)}\left(\int_0^\infty c\mathbf{K}(c)\mathfrak{B}^{\oplus}(1,c)dc+\int_{-\infty}^0 c\mathbf{K}(c)\mathfrak{B}^{\otimes}(1,-c)dc\right)\\
&\overset{d}{=}\left(2h_{n,\overline{T}}^{3/2}\frac{\left(\sigma^2(x)\right)^{'}}{\sigma(x)}\sqrt{\overline{L}_X(\overline{T},x)}\right)\mathbf{N}\left(0,2\int_0^\infty\int_0^\infty us\mathbf{K}(u)\mathbf{K}(s)\min(u,s)duds\right)\\
&\overset{d}{=}h_{n,\overline{T}}^{3/2}\mathbf{N}\left(0,16\varphi\left(\sigma^{'}(x)\right)\overline{L}_X(\overline{T},x)\right)
\end{aligned}
$$

where $\varphi=2\int_0^\infty\int_0^\infty ue\mathbf{K}(u)\mathbf{K}(e)\min(u,e)dude$. Then,

$$
\frac{1}{(h_{n,\overline{T}})^{3/2}}\left(\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\left(\sigma^2(x)\right)^{'}ch_{n,\overline{T}}\left(\frac{L_X(\overline{T},x+ch_{n,T})-L_X(\overline{T},x)}{\sigma^2(x+ch_{n,\overline{T}})}\right)dc}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}\right)\Rightarrow\mathbf{N}\left(0,\frac{16\varphi\left(\sigma^{'}(x)\right)^2}{\overline{L}_X(\overline{T},x)}\right).
$$

As for $\frac{B_{num}^2}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}$ and $\frac{B_{num}^3}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}$, we can write

$$
\begin{aligned}
&\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\left(\sigma^2(x)\right)^{'}ch_{n,\overline{T}}\frac{\left(\sigma^2(x)-\sigma^2(x+ch_{n,\overline{T}})\right)}{\sigma^2(x+ch_{n,\overline{T}})\sigma^2(x)}L_X(\overline{T},x)dc}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}\\
&+\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{1}{2}\left(\sigma^2(x^*)\right)^{''}\left(ch_{n,\overline{T}}\right)^2\frac{L_X(T,x+ch_{n,\overline{T}})}{\sigma^2(x+ch_{n,\overline{T}})}dc}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}=O_{a.s.}\left(h_{n,\overline{T}}^2\right).
\end{aligned}
$$

But,

$$
\begin{aligned}
\frac{1}{(h_{n,\overline{T}})^{3/2}}(B)&=\frac{1}{(h_{n,\overline{T}})^{3/2}}\left(\frac{\int_{-\infty}^{\infty}\mathbf{K}(c)\left(\sigma^2(x)\right)^{'}ch_{n,\overline{T}}\left(\frac{L_X(\overline{T},x+ch_{n,T})-L_X(\overline{T},x)}{\sigma^2(x+ch_{n,\overline{T}})}\right)dc}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}+O_{a.s.}\left(h_{n,\overline{T}}^2\right)\right)\\
&\overset{d}{=}\mathbf{MN}\left(0,\frac{16\varphi\left(\sigma^{'}(x)\right)^2}{\overline{L}_X(\overline{T},x)}\right).
\end{aligned}
$$

Next, consider the numerator of the term V which can be written as

$$
V=\mathbf{V}_1^{num}+\mathbf{V}_2^{num}
$$

where $\mathbf{V}_2^{num}=O\left(\frac{\Delta_{n,\overline{T}}}{\varepsilon_{n,\overline{T}}}\right)$ and

$$
\begin{aligned}
&\mathbf{V}_1^{num}\\
=&\underbrace{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}}-X_{i\Delta_{n,\overline{T}}}|\leq\varepsilon_{n,\overline{T}}\}}\frac{n}{\overline{T}}\left[\int_{j\overline{T}/n}^{(j+1)\overline{T}/n}\left(\sigma^2(X_s)-\sigma^2(X_{i\overline{T}/n})\right)ds\right]}{\sum_{j=1}^n\mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}}-X_{i\Delta_{n,\overline{T}}}|\leq\varepsilon_{n,\overline{T}}\}}}}_{(\mathbf{A}_{n,\overline{T}})}
\end{aligned}
$$

$$+\underbrace{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}}-X_{i\Delta_{n,\overline{T}}}|\leq\varepsilon_{n,\overline{T}}\}}\frac{n}{\overline{T}}\left[\int_{j\overline{T}/n}^{(j+1)\overline{T}/n}2\left(X_s-X_{j\overline{T}/n}\right)\sigma(X_s)dB_s\right]}{\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}}-X_{i\Delta_{n,\overline{T}}}|\leq\varepsilon_{n,\overline{T}}\}}}}_{(\mathbf{B}_{n,\overline{T}}(1))}$$

$$+\underbrace{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}}-X_{i\Delta_{n,\overline{T}}}|\leq\varepsilon_{n,\overline{T}}\}}\frac{n}{\overline{T}}\left[\int_{j\overline{T}/n}^{(j+1)\overline{T}/n}2\left(X_s-X_{j\overline{T}/n}\right)\mu(X_s)ds\right]}{\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}}-X_{i\Delta_{n,\overline{T}}}|\leq\varepsilon_{n,\overline{T}}\}}}}_{(\mathbf{C}_{n,\overline{T}})}$$

$$=\ \mathbf{A}_{n,\overline{T}}+\mathbf{B}_{n,\overline{T}}(1)+\mathbf{C}_{n,\overline{T}}. \tag{7.18}$$

The three terms comprise an additional bias effect, $\mathbf{A}_{n,\overline{T}}$, a martingale effect, $\mathbf{B}_{n,\overline{T}}(1)$, and a residual effect, $\mathbf{C}_{n,\overline{T}}$. As we shall see, depending on the bandwidth choices, either $\mathbf{A}_{n,\overline{T}}$ or $\mathbf{B}_{n,\overline{T}}$ may dominate the asymptotic distribution. Using the results in Theorem 3 we can show that

$$\sqrt{\frac{\varepsilon_{n,\overline{T}}}{\Delta_{n,\overline{T}}}}\left(\frac{\mathbf{B}_{n,\overline{T}}(1)}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}\right)\Rightarrow\mathbf{MN}\left(0,\frac{2\sigma^4(x)}{\overline{L}_X(\overline{T},x)}\right) \tag{7.19}$$

if $h_{n,\overline{T}}=o(\varepsilon_{n,\overline{T}})$, and

$$\sqrt{\frac{\varepsilon_{n,\overline{T}}}{\Delta_{n,\overline{T}}}}\left(\frac{\mathbf{B}_{n,\overline{T}}(1)}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}\right)\Rightarrow\mathbf{MN}\left(0,\frac{2\theta_\phi\sigma^4(x)}{\overline{L}_X(\overline{T},x)}\right) \tag{7.20}$$

if $h_{n,\overline{T}}=O(\varepsilon_{n,\overline{T}})$ and $h_{n,\overline{T}}/\varepsilon_{n,\overline{T}}\to\phi>0$. Next, examine $\mathbf{A}_{n,\overline{T}}$. If $h_{n,\overline{T}}=o(\varepsilon_{n,\overline{T}})$, then

$$\begin{aligned}
\mathbf{A}_{n,\overline{T}}=&\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\frac{1}{2}\int_{-\infty}^{\infty}\mathbf{1}_{\{|a|\leq1\}}\left(\left(\sigma^2(x)\right)^{'}\varepsilon_{n,\overline{T}}a\right)\left(\frac{L_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)-L_X(\overline{T},x)}{\sigma^2(x+\varepsilon_{n,\overline{T}}a)}\right)da}{\frac{1}{2}\int_{-\infty}^{\infty}\mathbf{1}_{\{|a|\leq1\}}\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}\overline{L}_X(\overline{T},x+h_{n,\overline{T}}c)dc\\
&+\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\frac{1}{2}\int_{-\infty}^{\infty}\mathbf{1}_{\{|a|\leq1\}}\left(\left(\sigma^2(x)\right)^{'}\varepsilon_{n,\overline{T}}a\right)\left(\frac{\sigma^2(x)-\sigma^2(x+\varepsilon_{n,\overline{T}}a)}{\sigma^2(x+\varepsilon_{n,\overline{T}}a)\sigma^2(x)}\right)L_X(\overline{T},x)da}{\frac{1}{2}\int_{-\infty}^{\infty}\mathbf{1}_{\{|a|\leq1\}}\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}\overline{L}_X(\overline{T},x+h_{n,\overline{T}}c)dc\\
&+\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\frac{1}{2}\int_{-\infty}^{\infty}\mathbf{1}_{\{|a|\leq1\}}\left(\frac{1}{2}\left(\sigma^2(x^*)\right)^{''}\left(\varepsilon_{n,\overline{T}}a\right)^2\right)\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}{\frac{1}{2}\int_{-\infty}^{\infty}\mathbf{1}_{\{|a|\leq1\}}\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}\overline{L}_X(\overline{T},x+h_{n,\overline{T}}c)dc\\
&+\int_{-\infty}^{\infty}\mathbf{K}(c)\left(\left(\sigma^2(x)\right)^{'}h_{n,\overline{T}}c+\frac{1}{2}\left(\sigma^2(x^*)\right)^{''}\left(h_{n,\overline{T}}c\right)^2\right)\overline{L}_X(\overline{T},x+h_{n,\overline{T}}c)dc.
\end{aligned}$$

Using Lemma 4 and proceeding as above, we find that

$$\frac{1}{\varepsilon_{n,\overline{T}}^{3/2}}\left(\frac{\mathbf{A}_{n,\overline{T}}}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}\right)\Rightarrow\mathbf{MN}\left(0,16\varphi^{ind}\frac{\left(\sigma^{'}(x)\right)^2}{\overline{L}_X(\overline{T},x)}\right) \tag{7.21}$$

with $\varphi^{ind}=2\int_0^\infty\int_0^\infty dadb\left(\frac{1}{2}\mathbf{1}_{\{|a|\leq1\}}a\right)\left(\frac{1}{2}\mathbf{1}_{\{|b|\leq1\}}b\right)\min(a,b)dadb$. Next, if $h_{n,\overline{T}}=O(\varepsilon_{n,\overline{T}})$ and $h_{n,\overline{T}}/\varepsilon_{n,\overline{T}}\to\phi>0$, then

$$\mathbf{A}_{n,\overline{T}}=\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\leq1\}}\left(\left(\sigma^2(x)\right)^{'}\varepsilon_{n,\overline{T}}a\right)\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\leq1\}}\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}\overline{L}_X(\overline{T},x+\phi\varepsilon_{n,\overline{T}}c)dc$$

$$+\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\le 1\}}\left(\frac{1}{2}\left(\sigma^2(x^*)\right)''\left(\varepsilon_{n,\overline{T}}a\right)^2\right)\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\le 1\}}\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}\overline{L}_X(\overline{T},x+\phi\varepsilon_{n,T}c)dc$$

$$+\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\le 1\}}\left(\left(\sigma^2(x)\right)'h_{n,\overline{T}}c\right)\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\le 1\}}\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}\overline{L}_X(\overline{T},x+h_{n,\overline{T}}c)dc.$$

$$+\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\le 1\}}\left(\frac{1}{2}\left(\sigma^2(x^*)\right)''\left(h_{n,\overline{T}}c\right)^2\right)\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\le 1\}}\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}\overline{L}_X(\overline{T},x+h_{n,\overline{T}}c)dc.$$

Only the first and the third term can affect the asymptotic distribution of $\mathbf{A}_{n,T}$. Write the first term as follows,

$$\begin{aligned}
&\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\le 1\}}\left(\left(\sigma^2(x)\right)'\varepsilon_{n,\overline{T}}a\right)\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\le 1\}}\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}\overline{L}_X(\overline{T},x+\phi\varepsilon_{n,\overline{T}}c)dc\\
=\;&\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\left(\sigma^2(x)\right)'\varepsilon_{n,\overline{T}}(g+\phi c)\right)\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}(g+\phi c))dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|s|\le 1\}}\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}(s+\phi c))ds}\overline{L}_X(\overline{T},x+\phi\varepsilon_{n,\overline{T}}c)dc\\
=\;&\phi\varepsilon_{n,\overline{T}}\left(\sigma^2(x)\right)'\int_{-\infty}^{\infty}c\mathbf{K}(c)\left(\overline{L}_X(\overline{T},x+\phi\varepsilon_{n,\overline{T}}c)-\overline{L}_X(\overline{T},x)\right)dc\\
&+\left(\sigma^2(x)\right)'\varepsilon_{n,\overline{T}}\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}g\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}(g+\phi c))dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|s|\le 1\}}\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}(s+\phi c))ds}\overline{L}_X(\overline{T},x+\phi\varepsilon_{n,\overline{T}}c)dc\\
=\;&\phi\varepsilon_{n,\overline{T}}\left(\sigma^2(x)\right)'\int_{-\infty}^{\infty}c\mathbf{K}(c)\left(\overline{L}_X(\overline{T},x+\phi\varepsilon_{n,\overline{T}}c)-\overline{L}_X(\overline{T},x)\right)dc\\
&+\left(\sigma^2(x)\right)'\varepsilon_{n,\overline{T}}\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{1}{2}\int_{-\infty}^{\infty}\mathbf{1}_{\{|u-\phi c|\le 1\}}(u-\phi c)\left(\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}u)-\overline{L}_X(\overline{T},x)\right)dudc.
\end{aligned}$$

As for the third term, write

$$\begin{aligned}
&\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\le 1\}}\left(\left(\sigma^2(x)\right)'h_{n,\overline{T}}c\right)\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|a-\phi c|\le 1\}}\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}a)da}\overline{L}_X(\overline{T},x+h_{n,\overline{T}}c)dc\\
=\;&\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\left(\left(\sigma^2(x)\right)'h_{n,\overline{T}}c\right)\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}(g+\phi c)dg}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|g|\le 1\}}\overline{L}_X(\overline{T},x+\varepsilon_{n,\overline{T}}(g+\phi c))dg}\overline{L}_X(\overline{T},x+h_{n,\overline{T}}c)dc\\
=\;&\phi\varepsilon_{n,\overline{T}}\left(\sigma^2(x)\right)'\int_{-\infty}^{\infty}c\mathbf{K}(c)\left(\overline{L}_X(\overline{T},x+h_{n,\overline{T}}c)-\overline{L}_X(\overline{T},x)\right)dc.
\end{aligned}$$

Then,

$$\begin{aligned}
&\mathbf{A}_{n,\overline{T}}\\
=\;&2\phi^{3/2}\varepsilon_{n,\overline{T}}^{3/2}\frac{\left(\sigma^2(x)\right)'}{\sigma^2(x)}\frac{1}{\sqrt{h_{n,\overline{T}}}}\int_{-\infty}^{\infty}c\mathbf{K}(c)\left(L_X(\overline{T},x+h_{n,\overline{T}}c)-L_X(\overline{T},x)\right)dc\\
&+\frac{\left(\sigma^2(x)\right)'}{\sigma^2(x)}\varepsilon_{n,\overline{T}}^{3/2}\frac{1}{\sqrt{\varepsilon_{n,\overline{T}}}}\int_{-\infty}^{\infty}\mathbf{K}(c)\frac{1}{2}\int_{-\infty}^{\infty}\mathbf{1}_{\{|u-\phi c|\le 1\}}(u-\phi c)\left(L_X(\overline{T},x+\varepsilon_{n,T}u)-L_X(\overline{T},x)\right)dudc\\
=\;&2\phi^{3/2}\varepsilon_{n,\overline{T}}^{3/2}\frac{\left(\sigma^2(x)\right)'}{\sigma^2(x)}\frac{1}{\sqrt{h_{n,\overline{T}}}}\int_{-\infty}^{\infty}c\mathbf{K}(c)\left(L_X(\overline{T},x+h_{n,\overline{T}}c)-L_X(\overline{T},x)\right)dc
\end{aligned}$$

$$+\frac{\left(\sigma^2(x)\right)^{'}}{\sigma^2(x)}\varepsilon_{n,\overline{T}}^{3/2}\frac{1}{\sqrt{\varepsilon_{n,\overline{T}}}}\frac{1}{2}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\mathbf{K}\left(c\right)\mathbf{1}_{\{|u-\phi c|\leq 1\}}\left(u-\phi c\right)dc\right)\left(L_X(\overline{T},x+\varepsilon_{n,T}u)-L_X(\overline{T},x)\right)du$$

$$= 2\phi^{3/2}\varepsilon_{n,\overline{T}}^{3/2}\frac{\left(\sigma^2(x)\right)^{'}}{\sigma^2(x)}\frac{1}{\sqrt{h_{n,\overline{T}}}}\int_{-\infty}^{\infty}c\mathbf{K}\left(c\right)\left(L_X(\overline{T},x+h_{n,\overline{T}}c)-L_X(\overline{T},x)\right)dc$$

$$+\frac{\left(\sigma^2(x)\right)^{'}}{\sigma^2(x)}\varepsilon_{n,\overline{T}}^{3/2}\frac{1}{\sqrt{\varepsilon_{n,\overline{T}}}}\frac{1}{2}\int_{-\infty}^{\infty}\Lambda(\phi,u)\left(L_X(\overline{T},x+\varepsilon_{n,\overline{T}}u)-L_X(\overline{T},x)\right)du$$

$$= 2\phi\varepsilon_{n,\overline{T}}^{3/2}\frac{\left(\sigma^2(x)\right)^{'}}{\sigma^2(x)}\frac{1}{\sqrt{\varepsilon_{n,\overline{T}}}}\int_{-\infty}^{\infty}c\mathbf{K}\left(c\right)\left(L_X(\overline{T},x+\phi\varepsilon_{n,\overline{T}}c)-L_X(\overline{T},x)\right)dc$$

$$+\frac{\left(\sigma^2(x)\right)^{'}}{\sigma^2(x)}\varepsilon_{n,\overline{T}}^{3/2}\frac{1}{\sqrt{\varepsilon_{n,\overline{T}}}}\frac{1}{2}\int_{-\infty}^{\infty}\Lambda(\phi,u)\left(L_X(\overline{T},x+\varepsilon_{n,\overline{T}}u)-L_X(\overline{T},x)\right)du. \tag{7.22}$$

where $\Lambda(\phi,u)=\int_{-\infty}^{\infty}\mathbf{K}\left(c\right)\mathbf{1}_{\{|u-\phi c|\leq 1\}}\left(u-\phi c\right)dc$. Now, we note that $\frac{\mathbf{C}_{n,\overline{T}}}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}$ is $o_{a.s.}\left(\sqrt{\frac{\Delta_{n,\overline{T}}}{\varepsilon_{n,\overline{T}}}}\right)$.

Then, defining the overall estimation error as

$$E = B+\frac{\mathbf{A}_{n,\overline{T}}}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}+\frac{\mathbf{C}_{n,\overline{T}}}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}+\frac{\mathbf{B}_{n,\overline{T}}(1)}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}$$

and scaling by $\sqrt{\frac{\varepsilon_{n,\overline{T}}}{\Delta_{n,\overline{T}}}}$, we have

$$\sqrt{\frac{\varepsilon_{n,\overline{T}}}{\Delta_{n,\overline{T}}}}\left(\frac{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)\widetilde{\sigma}_{n,\overline{T}}^2(X_{i\Delta_{n,\overline{T}}})}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}-\sigma^2(x)\right)$$

$$= \sqrt{\frac{\varepsilon_{n,\overline{T}}}{\Delta_{n,\overline{T}}}}\left\{O_p\left(h_{n,\overline{T}}^{3/2}\right)+o_{a.s.}\left(\sqrt{\Delta_{n,\overline{T}}}\right)+O\left(\frac{\Delta_{n,\overline{T}}}{\varepsilon_{n,\overline{T}}}\right)+O_p\left(h_{n,\overline{T}}^2\right)\right.$$

$$\left.+o_{a.s.}\left(\sqrt{\frac{\Delta_{n,\overline{T}}}{\varepsilon_{n,\overline{T}}}}\right)+O_p\left(\varepsilon_{n,\overline{T}}^{3/2}\right)+\frac{\mathbf{B}_{n,\overline{T}}(1)}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}\right\}$$

$$\Rightarrow \mathbf{MN}\left(0,\frac{2\sigma^4(x)}{\overline{L}_X(\overline{T},x)}\right)$$

from (7.19), for choices of $\varepsilon_{n,\overline{T}}$ such that $\frac{\varepsilon_{n,\overline{T}}^4}{\Delta_{n,\overline{T}}}\to 0$ and $h_{n,\overline{T}}=o(\varepsilon_{n,\overline{T}})$. If $\frac{\varepsilon_{n,\overline{T}}^4}{\Delta_{n,\overline{T}}}\to 0$ and $h_{n,\overline{T}}=O(\varepsilon_{n,\overline{T}})$ with $h_{n,\overline{T}}/\varepsilon_{n,\overline{T}}\to\phi>0$, then

$$\sqrt{\frac{\varepsilon_{n,\overline{T}}}{\Delta_{n,\overline{T}}}}\left(\frac{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)\widetilde{\sigma}_{n,\overline{T}}^2(X_{i\Delta_{n,\overline{T}}})}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}-\sigma^2(x)\right)\Rightarrow\mathbf{MN}\left(0,\frac{2\theta_{\phi}\sigma^4(x)}{\overline{L}_X(\overline{T},x)}\right)$$

from (7.20), where $\theta_{\phi}=\frac{1}{2}\int_{-\infty}^{\infty}\int_{(z-1)/\phi}^{(z+1)/\phi}\int_{(z-1)/\phi}^{(z+1)/\phi}\mathbf{K}(a)\mathbf{K}(e)dzdade$. Finally, provided $\frac{\varepsilon_{n,\overline{T}}^4}{\Delta_{n,\overline{T}}}\to\infty$ and $h_{n,\overline{T}}=o(\varepsilon_{n,\overline{T}})$, then the $\mathbf{A}_{n,\overline{T}}$ term dominates, leading to

$$\frac{1}{\varepsilon_{n,\overline{T}}^{3/2}}\left\{O_p\left(\sqrt{\frac{\Delta_{n,\overline{T}}}{\varepsilon_{n,\overline{T}}}}\right)+O_p\left(h_{n,\overline{T}}^{3/2}\right)+o_{a.s.}\left(\sqrt{\Delta_{n,\overline{T}}}\right)\right.$$

$$\left.+O\left(\frac{\Delta_{n,\overline{T}}}{\varepsilon_{n,\overline{T}}}\right)+\frac{\mathbf{A}_{n,\overline{T}}}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}\right\}\Rightarrow \mathbf{MN}\left(0,16\varphi^{ind}\frac{\left(\sigma^{'}(x)\right)^2}{\overline{L}_X(\overline{T},x)}\right)$$

from (7.21), where $\varphi^{ind}=2\int_0^\infty\int_0^\infty\left(\frac{1}{2}\mathbf{1}_{\{|a|\leq 1\}}a\right)\left(\frac{1}{2}\mathbf{1}_{\{|b|\leq 1\}}b\right)\min(a,b)dadb$. Under the same conditions, but when $h_{n,T}=O(\varepsilon_{n,T})$ with $h_{n,T}/\varepsilon_{n,T}\to\phi>0$, we have

$$\frac{1}{\varepsilon_{n,\overline{T}}^{3/2}}\left\{O_p\left(\sqrt{\frac{\Delta_{n,\overline{T}}}{\varepsilon_{n,\overline{T}}}}\right)+\frac{B(1)}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}+o_{a.s.}\left(\Delta_{n,\overline{T}}{}^{1/2-\varepsilon}\right)\right.$$

$$\left.+\frac{\mathbf{A}_{n,\overline{T}}}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}\mathbf{K}\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}\right\}\Rightarrow \mathbf{MN}\left(0,16\left(\sigma^{'}(x)\right)^2\frac{\left(\varphi^{*}+9\phi^3\varphi+6\phi\varphi^{\hat{}}\right)}{\overline{L}_X(\overline{T},x)}\right)$$

$$\overset{d}{=}\mathbf{MN}\left(0,16\varphi^{ind,\mathbf{K}}\frac{\left(\sigma^{'}(x)\right)^2}{\overline{L}_X(\overline{T},x)}\right),$$

from (7.17) and (7.22), where

$$\begin{aligned}\varphi^{*} &= \frac{1}{4}\int_0^\infty\int_0^\infty\Lambda(\phi,a)\Lambda(\phi,b)\min(a,b)dadb\\ &\quad -\frac{1}{4}\int_{-\infty}^0\int_{-\infty}^0\Lambda(\phi,a)\Lambda(\phi,b)\max(a,b)dadb\\ &= \frac{1}{2}\int_0^\infty\frac{1}{2}\int_0^\infty\left(\int_{-\infty}^\infty\mathbf{K}(c)\,\mathbf{1}_{\{|a-\phi c|\leq 1\}}(a-\phi c)\,dc\right)\left(\int_{-\infty}^\infty\mathbf{K}(c)\,\mathbf{1}_{\{|b-\phi c|\leq 1\}}(b-\phi c)\,dc\right)\min(a,b)dadb\\ &\quad -\frac{1}{2}\int_{-\infty}^0\frac{1}{2}\int_{-\infty}^0\left(\int_{-\infty}^\infty\mathbf{K}(c)\,\mathbf{1}_{\{|a-\phi c|\leq 1\}}(a-\phi c)\,dc\right)\left(\int_{-\infty}^\infty\mathbf{K}(c)\,\mathbf{1}_{\{|b-\phi c|\leq 1\}}(b-\phi c)\,dc\right)\max(a,b)dadb\end{aligned}$$

and

$$\begin{aligned}\varphi^{\hat{}} &= \frac{1}{2}\int_0^\infty\int_0^\infty a\mathbf{K}(a)\Lambda(\phi,b)\min(\phi a,b)dadb\\ &\quad -\frac{1}{2}\int_{-\infty}^0\int_{-\infty}^0 a\mathbf{K}(a)\Lambda(\phi,b)\max(\phi a,b)dadb\\ &= \frac{1}{2}\int_0^\infty\int_0^\infty a\mathbf{K}(a)\left(\int_{-\infty}^\infty\mathbf{K}(c)\,\mathbf{1}_{\{|b-\phi c|\leq 1\}}(b-\phi c)\,dc\right)\min(\phi a,b)dadb\\ &\quad -\frac{1}{2}\int_{-\infty}^0\int_{-\infty}^0 a\mathbf{K}(a)\left(\int_{-\infty}^\infty\mathbf{K}(c)\,\mathbf{1}_{\{|b-\phi c|\leq 1\}}(b-\phi c)\,dc\right)\max(\phi a,b)dadb.\end{aligned}$$

This concludes the proof of the stated result.

Proof of Corollary 2: Immediate after noticing that under stationarity (or positive recurrence)

$$\frac{\widehat{\overline{L}}_X(T,x)}{T}\overset{a.s.}{\to}f(x),$$

$$\overline{L}_X(T,x) = O_{a.s.}(T),$$

and

$$\frac{s^{'}(x)}{s(x)} = \frac{s^{'}(x)/s(\mathfrak{D})}{s(x)/s(\mathfrak{D})} = \frac{f^{'}(x)}{f(x)}$$

where $f(x)$ is the time invariant density of the process at x.

PROOF OF COROLLARY 3: C.f. the proof of corollary 2.

APPENDIX B: NOTATION

$\rightarrow_{a.s.}$	almost sure convergence
$\rightarrow_p$	convergence in probability
$\Rightarrow, \rightarrow_d$	weak convergence
$:=$	definitional equality
$o_p(1)$	tends to zero in probability
$O_p(1)$	bounded in probability
$o_{a.s.}(1)$	tends to zero almost surely
$O_{a.s.}(1)$	bounded almost surely
$=_d$	distributional equivalence
$\sim_d$	asymptotically distributed as
$\mathrm{MN}(0,V)$	mixed normal distribution with variance V
$\mathbf{1}_A$	indicator function for the set A
$a \vee b$	$\max\{a,b\}$
$C_k, \quad k=1,2,...$	constants

REFERENCES

AÏT-SAHALIA, Y. (1996a): "Nonparametric Pricing of Interest Rate Derivative Securities," *Econometrica*, 64, 527-560.

AÏT-SAHALIA, Y. (1996b): "Testing Continuous-Time Models of the Spot Interest Rate," *The Review of Financial Studies,* 2, 385-426.

BANDI, F. (1998): "Short-Term Interest Rate Dynamics: A Spatial Approach," *The Journal of Financial Economics, Forthcoming.*

BANDI, F. (2000): "Nonparametric Fixed-Income Pricing: Theoretical Issues," *Unpublished paper, The University of Chicago.*

BANDI, F. and T. NGUYEN (2000): "Fully Nonparametric Estimators for Diffusions: a Small Sample Analysis," *Unpublished paper, The University of Chicago.*

BANDI, F. and P. C. B. PHILLIPS (1998): "Econometric Estimation of Diffusion Models," *Unpublished paper, Yale University (http://cowles.econ.yale.edu/).*

BAXTER, M. and A. RENNIE, (1996): *Financial Calculus. An Introduction to Derivative Pricing*, Cambridge University Press: Cambridge.

BERGSTROM, A. R. (1988): "A History of Continuous Time Econometric Models," *Econometric Theory*, 4, 365-383.

BRUGIÈRE, P. (1993): "Théorém de Limite Centrale pour un Estimateur Non Paramétrique de la Variance d'un Processus de Diffusion Multidimensionnelle," *Ann. Inst. Henri Poincaré.*

BOSQ, D. (1998): *Nonparametric Statistics for Stochastic Processes.* Second Edition. Springer-Verlag: New York.

DACUNHA-CASTELLE, D. and D. FLORENS-ZMIROU (1986): "Estimation of the Coefficients of a Diffusion from Discrete Observations," *Stochastics,* 19, 263-284.

DUFFIE, D. (1996): *Dynamic Asset Pricing Theory.* Princeton University Press: Princeton.

DURRETT, R. (1996): *Stochastic Calculus. A Practical Introduction.* CRC Press: New York.

FLORENS-ZMIROU, D. (1993): "On Estimating the Diffusion Coefficient from Discrete Observations," *Journal of Applied Probability*, 30, 790-804.

GEMAN, S.A. (1979): "On a Common Sense Estimator for the Drift of a Diffusion," *Working Paper, Brown University.*

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

JACOD, J. (1997): "Nonparametric Kernel Estimation of the Diffusion Coefficient of a Diffusion," *Prépublication No. 405. du Laboratoire de Probabilités de l'Université Paris VI.*

JIANG G.J. AND J. 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.

KARATZAS, I. AND S. E. SHREVE (1988): *Brownian Motion and Stochastic Calculus.* Springer-Verlag: New York.

KARLIN, S. AND H. M. TAYLOR (1981): *A Second Course in Stochastic Processes,* Academic Press: New York.

PAGAN, A. and A. ULLAH (1999): *Nonparametric Statistics.* Cambridge University Press.

PHILLIPS, P.C.B. (1998): "Econometric Analysis of Fisher's Equation," *Cowles Foundation Discussion Paper, No. 1180. Presented at the Irving Fisher Conference, Yale University, 1998.*

PHILLIPS, P.C.B. AND J. PARK (1998): "Nonstationary Density Estimation and Kernel Autoregression," *Cowles Foundation Discussion Paper, No. 1181, Yale University.*

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

POLLACK, M. and D. SIEGMUND (1985): "A diffusion process and its applications to detecting a change in the drift of Brownian motion," *Biometrica,* 72, 267-280.

PROTTER, P. (1990): *Stochastic Integration and Differential Equations.* Springer-Verlag: New York.

REVUZ, D. AND M. YOR (1998): *Continuous Martingales and Brownian Motion.* Third Edition, Springer-Verlag: New York.

STANTON, R. (1998): "A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk," *Journal of Finance,* 5, 1973-2002.

YOR, M. (1983): "Le Drap Brownien Comme Limite en Loi de Temps Locaux Linéaires," *Sém. Prob. XVII. Lecture Notes in Mathematics, vol 986, Springer, Berlin Heidelberg New York 1983, pp. 89-105.*