Econometrica, Vol. 71, No. 1 (January, 2003), 241–283

FULLY NONPARAMETRIC ESTIMATION OF SCALAR DIFFUSION MODELS

BY FEDERICO M. BANDI AND PETER C. B. PHILLIPS[1]

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. We prove almost sure consistency and weak convergence to mixtures of normal laws, where the mixing variates depend on the chronological local time of the underlying diffusion process, that is the random time spent by the process in the vicinity of a generic spatial point. The estimation method and asymptotic results apply to both stationary and nonstationary recurrent processes.

KEYWORDS: Diffusion, drift, local time, martingale, nonparametric estimation, semimartingale, stochastic differential equation.

1. INTRODUCTION

MANY POPULAR MODELS in economics and finance, like those for pricing derivative securities, involve diffusion processes formulated in continuous-time as solutions to stochastic differential equations. These processes have been used to model options prices, the term structure of interest rates, and exchange 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 employed 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 whose evolution is defined continuously in time.

The stochastic differential equation that defines a diffusion process, like X_t in (2.1) below, involves two components. These components denote the (infinitesimal) conditional drift, $\mu(.)$, and the (infinitesimal) conditional variation, $\sigma^2(.)$, of the process in the vicinity of each spatial level 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

[1] We are grateful to the co-editor and three anonymous referees for useful comments. Bandi thanks Mediocredito Centrale and the Sloan Foundation for Fellowship support. Phillips thanks the NSF for support under Grant Nos. SBR 97-30295 and SES SES-00-92509.

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 (see Karatzas and Shreve (1991) and Karlin and Taylor (1981), for example). 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 distribution function of the process of interest (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 first 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 marginal data density. In consequence, 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 scalar diffusion models without imposing a stationarity assumption. Recurrence, which is a substantially milder assumption than stationarity, is our identifying condition. Specifically, we require the continuous trajectory of the process to visit any level in its range an infinite 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), among others) and also as the time span of observation lengthens. In technical terms this amounts to both *infill* and *long span* asymptotics. As clarified below, 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 very general about the dynamic features of the underlying diffusion process (Phillips (1973) and Hansen and Sargent (1983) are early references on the aliasing phenomenon 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 of the underlying diffusion, i.e., a random quantity that measures in real time units the amount of time that the process spends in the local neighborhood of each spatial point in its admissible range.

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 cannot be identified nonparametrically on a fixed interval without cross-restrictions (as in Jiang and Knight (1997), for instance), no matter how frequently the data are sampled (see Merton (1980) and Bandi (2002, Theorem 2.1), among others). 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, as implied by our assumptions (see Section 2). 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 that 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 distribution function for the process of interest. It is therefore robust against deviations from stationarity in the wide class of recurrent processes.

Not surprisingly, both the nonparametric theory on the estimation of conditional expectations in the stationary discrete time framework (see Pagan and Ullah (1999), for example, and the references therein) and the recent functional theory on the identification of conditional first moments in the unit root literature (Phillips and Park (1998)) are reflected in the general results given here. These results can in turn 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, without being limited to them.

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. The Appendix provides proofs and technicalities.

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 $X_0 = \overline{X}$, where $\{B_t : t \geq 0\}$ 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 $\Im_\infty^B$. We define the left-continuous filtration

$$\overline{\Im}_t := \sigma(\overline{X}) \vee \Im_t^B = \sigma(\overline{X}, B_s; 0 \leq s \leq t), \qquad 0 \leq t < \infty, \tag{2.2}$$

and the collection of null sets

$$\aleph := \{N \subseteq \Omega; \exists G \in \overline{\mathfrak{I}}_\infty \text{ with } N \subseteq G \text{ and } P(G) = 0\}. \tag{2.3}$$

We create the augmented filtration

$$\widetilde{\mathfrak{I}}_t^X := \sigma(\overline{\mathfrak{I}}_t \cup \aleph), \qquad 0 \leq t < \infty. \tag{2.4}$$

The following conditions will be used in the study of (2.1). In what follows, the symbol $\mathfrak{D}$ denotes the admissible range of the process X_t.

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 twice continuously differentiable. They satisfy local Lipschitz and growth conditions. Thus, for every compact subset J of the range $\mathfrak{D}$ of the process, there exist constants C_1^J and C_2^J such that, for all x and y in J,*

$$|\mu(x) - \mu(y)| + |\sigma(x) - \sigma(y)| \leq C_1^J |x - y|, \tag{2.5}$$

and

$$|\mu(x)| + |\sigma(x)| \leq C_2^J \{1 + |x|\}. \tag{2.6}$$

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

(iii) *We define $S(\alpha)$, the natural scale function, as*

$$S(\alpha) = \int_c^\alpha \exp\left\{\int_c^y \left[-\frac{2\mu(x)}{\sigma^2(x)}\right] dx\right\} dy, \tag{2.7}$$

where c is a generic fixed number belonging to $\mathfrak{D}$. We require $S(\alpha)$ to satisfy

$$\lim_{\alpha \to l} S(\alpha) = -\infty, \tag{2.8}$$

and

$$\lim_{\alpha \to u} S(\alpha) = \infty. \tag{2.9}$$

Condition (i) is sufficient for pathwise uniqueness of the solution to (2.1) (Karatzas and Shreve (1991, Theorem 5.2.5, p. 287)). Conditions (i) and (ii) assure the existence of a unique strong solution up to an explosion time (Karatzas and Shreve (1991, Theorem 5.5.15, p. 341, and Corollary 5.3.23, p. 310)). Condition (iii) guarantees that the exit time from $\mathfrak{D} = (l, u)$ is infinite (Karatzas and Shreve (1991, Proposition 5.5.22, p. 345)), but does not imply existence of a stationary probability measure for X_t. The same condition is necessary and sufficient for recurrence, meaning that, for each $c \in \mathfrak{D}$, there exists 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) (see Karatzas and Shreve (1991, Theorem 5.2.9, p. 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.

Thus, under Assumption 1, the stochastic differential equation has a unique strong solution X_t that is adapted to $\widetilde{\Im}_t^X$ and recurrent. Further, X_t satisfies

$$X_t = \overline{X} + \int_0^t \mu(X_s)\,ds + \int_0^t \sigma(X_s)\,dB_s \qquad \text{a.s.,} \tag{2.10}$$

with $\int_0^t E[X_s^2]\,ds < \infty$.

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

$$E^x[X_t - x] = t\mu(x) + o(t), \tag{2.11}$$

$$E^x[(X_t - x)^2] = t\sigma^2(x) + o(t), \tag{2.12}$$

where x is a generic initial condition and E^x is the expectation operator associated with the process started at x. Loosely speaking, (2.11) and (2.12) can be interpreted as representing the instantaneous conditional mean and the instantaneous conditional variance of the process when $X_t = x$. More precisely, (2.11) describes the conditional expected rate of change of the process for infinitesimal time changes, whereas (2.12) 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 (Revuz and Yor (1998) is a standard reference).

DEFINITION 1 (Continuous SMG): A *continuous SMG* is a continuous process M_t that can be written as $M_t = LM_t + A_t$, where LM_t is a continuous local martingale and A_t is a continuous adapted process of finite variation.

Stochastic differential equations like (2.1) are known to have solutions that are SMGs since $\overline{X} + \int_0^t \mu(X_s)\,ds$ 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_t is defined as follows:

LEMMA 1 (The Tanaka Formula): *For any real number a, there exists a nondecreasing continuous process $L_M(.,a)$ called the local time of M_t at a, such that*

$$|M_t - a| = |M_0 - a| + \int_0^t sgn(M_s - a)\,dM_s + L_M(t,a), \tag{3.1}$$

(3.2) $$(M_t - a)^+ = (M_0 - a)^+ + \int_0^t \mathbf{1}_{\{M_s > a\}}\, dM_s + \frac{1}{2} L_M(t, a),$$

(3.3) $$(M_t - a)^- = (M_0 - a)^- - \int_0^t \mathbf{1}_{\{M_s \leq a\}}\, dM_s + \frac{1}{2} L_M(t, a),$$

where $\mathbf{1}$ *is the indicator function. In particular,* $|M_t - a|$, $(M_t - a)^+$, *and* $(M_t - a)^-$ *are SMGs.*

LEMMA 2 (Continuity of SMG Local Time): *For any continuous SMG* M_t, *there exists a version of the local time such that the map* $(t, a) \mapsto L_M(t, a)$ *is a.s. continuous in* t *and càdlàg in* a.

LEMMA 3 (The Occupation Time Formula): *Let* M_t *be a continuous SMG with quadratic variation process* $[M]_t$ *and let* L^a *be the local time at* a. *Then,*

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

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

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

LEMMA 4: *If* M_t *is a continuous SMG then, almost surely,*

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

If M_t *is a continuous local martingale then, almost surely,*

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

The process $L_M(t, a)$ is called the local time of M_t 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 (3.7), the local time of a standard Brownian motion W_t is

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

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

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

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

$$(3.10)\qquad \overline{L}_B(t,a)=\lim_{\varepsilon\to 0}\frac{1}{2\varepsilon}\int_0^t \mathbf{1}_{(|B_s-a|<\varepsilon)}\,ds=\sigma^{-2}L_B(t,a)\quad \text{a.s.}\quad \forall\, a,t.$$

Equation (3.10) 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 is by

$$(3.11)\qquad \overline{L}_X(t,a)=\frac{1}{\sigma^2(a)}\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_0^t \mathbf{1}_{[a,\,a+\varepsilon[}(X_s)\sigma^2(X_s)\,ds$$
$$=\frac{1}{\sigma^2(a)}L_X(t,a)\quad \text{a.s.}\quad \forall\, a,t.$$

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 (see, e.g., Bosq (1998, p. 146) and Florens-Zmirou (1993)), where it is sometimes referred to simply as "local time."

Lemma 5 and 6 below contain additional results that will be used in the development of our limit theory. Lemma 5 generalizes to diffusion processes the limit theory for Brownian local time (see Revuz and Yor (1998, Exercises 2.11 and 2.12, Chapter 13)).

LEMMA 5 (Limit Theory for the Local Time of a Diffusion): *Let* X_t *satisfy the properties in Section* 2.[2] *Let* r *and* $a>0$ *be fixed real numbers and treat* $\{L_X(t,r+(a/\lambda))-L_X(t,r)\}$ *as a double indexed stochastic process in* (t,a). *Then, as* $\lambda\to\infty$

$$(3.12)\qquad \frac{1}{2}\sqrt{\lambda}\left\{L_X\left(t,r+\frac{a}{\lambda}\right)-L_X(t,r)\right\}\Rightarrow \mathfrak{B}(L_X(t,r),a),$$

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

$$(3.13)\qquad \frac{1}{2}\sqrt{\lambda}\left\{L_X\left(t,r+\frac{a}{\lambda}\right)-L_X(t,r)\right\}\Rightarrow \mathfrak{B}(L_X(t,r),-a).$$

Finally, Lemma 6 specializes to scalar diffusion processes a result that has wider applicability in the theory of occupation times for recurrent Markov processes (see, e.g., Azèma, Kaplan-Duflo, and Revuz (1967) and Revuz and Yor (1998, Theorem 3.12, Chapter 10)).

[2] Lemma 5 also applies to transient solutions to (2.1).

LEMMA 6: *Let X_t satisfy the properties in Section 2. Then, for any Borel measurable pair of functions $f(.)$ and $g(.)$ that are integrable with respect to the speed measure $s(dx) = 2\,dx/S'(x)\sigma^2(x)$ of X_t, where $S'(x)$ is the first derivative of the natural scale function, the ratio of the functionals $\int_0^T f(X_s)\,ds$ and $\int_0^T g(X_s)\,ds$ is such that*

$$P\left[\lim_{T\to\infty}\frac{\int_0^T f(X_s)\,ds}{\int_0^T g(X_s)\,ds} = \frac{\int_{-\infty}^{\infty} f(x)s(dx)}{\int_{-\infty}^{\infty} g(x)s(dx)}\right] = 1, \tag{3.14}$$

provided $\int_{-\infty}^{\infty} g(x)s(dx) > 0$.

We now turn to the estimation method.

4. ECONOMETRIC ESTIMATION

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

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 (2.11) and (2.12):

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

$$:= \frac{\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\tilde{\mu}_{n,T}(X_{i\Delta_{n,T}})}{\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} \tag{4.2}$$

$$\hat{\sigma}^2_{(n,T)}(x) := \frac{\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\left(\frac{1}{m_{n,T}(i\Delta_{n,T})\Delta_{n,T}}\sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1}[X_{t(i\Delta_{n,T})_j+\Delta_{n,T}} - X_{t(i\Delta_{n,T})_j}]^2\right)}{\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} \tag{4.3}$$

$$:= \frac{\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\tilde{\sigma}^2_{n,T}(X_{i\Delta_{n,T}})}{\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}, \tag{4.4}$$

where $K(.)$ 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:

$$t(i\Delta_{n,T})_0 = \inf\{t \geq 0 : |X_t - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}\}, \tag{4.5}$$

and

$$t(i\Delta_{n,T})_{j+1} = \inf\big\{t \geq t(i\Delta_{n,T})_j + \Delta_{n,T} : |X_t - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}\big\}. \tag{4.6}$$

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

$$m_{n,T}(i\Delta_{n,T}) = \sum_{j=1}^{n} \mathbf{1}_{[|X_{j\Delta_{n,T}} - X_{i\Delta_{n,T}}| \leq \varepsilon_{n,T}]}, \qquad \forall i \leq n, \tag{4.7}$$

where $\mathbf{1}_A$ denotes the indicator of the set 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.

The function $K(\cdot)$ that appears in (4.2) and (4.4) is assumed to satisfy the following condition.

ASSUMPTION 2: *The kernel $K(.)$ is a continuously differentiable, symmetric and nonnegative function whose derivative K' is absolutely integrable and for which*

$$\int_{-\infty}^{\infty} K(s)\,ds = 1, \qquad \int_{-\infty}^{\infty} K^2(s)\,ds < \infty, \qquad \sup_s K(s) < C_3, \tag{4.8}$$

and

$$\int_{-\infty}^{\infty} s^2 K(s)\,ds < \infty. \tag{4.9}$$

The method hinges on the simultaneous operation of infill and long span asymptotics. The intuition underlying the construction of (4.2) and (4.4) 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, namely the drift and diffusion function, 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 $\tilde{\sigma}^2_{n,T}(X_{i\Delta_{n,T}})$ and $\tilde{\mu}_{n,T}(X_{i\Delta_{n,T}})$ in (4.2) and (4.4) are defined as empirical analogs to the true functions for all i's. Provided suitable conditions on the spatial bandwidth $\varepsilon_{n,T}$ are satisfied, such components are expected to be 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. Given appropriate choice of the smoothing sequence, divergence of $m_{n,T}(i\Delta_{n,T})$ to infinity occurs with probability one when the process X_t is recurrent, as it is under Condition (iii) in Assumption 1. In this case, the process almost surely visits any point in its admissible range an infinite number of times over time, i.e., $P_x\{X_t$ hits z at a sequence of times increasing to $\infty\} = 1, \forall x, z$ (here, as before, x represents a possible initialization of X_t).

Second, we apply standard nonparametric smoothing to recover the two functions of interest from the crude estimates $\tilde{\sigma}^2_{n,T}(X_{i\Delta_{n,T}})$ and $\tilde{\mu}_{n,T}(X_{i\Delta_{n,T}})$ 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 $(1/h_{n,\overline{T}})(\Delta_{n,\overline{T}} \log(1/\Delta_{n,\overline{T}}))^{1/2} = o(1)$, the quantity*

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

converges to $\overline{L}_X(\overline{T}, x)$ with probability one.

REMARK 2: 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 $T/n = \Delta_{n,T} \to 0$ and $h_{n,T} \to 0$ (as $n, T \to \infty$) in such a way that*

$$\frac{\overline{L}_X(T,x)}{h_{n,T}} (\Delta_{n,T} \log(1/\Delta_{n,T}))^{1/2} = o_{a.s.}(1),$$

then

$$\frac{\Delta_{n,T}}{h_{n,T}} \sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}} - x}{h_{n,T}}\right) \xrightarrow{a.s.} \overline{L}_X(\sup\{t : X_t = x\}, x). \tag{5.1}$$

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

5.2. *Functional Estimation of the Drift*

We now develop the asymptotic theory for the drift estimator (4.2).

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*

$$\frac{\overline{L}_X(T,x)}{h_{n,T}} (\Delta_{n,T} \log(1/\Delta_{n,T}))^{1/2} = o_{a.s.}(1),$$

and provided $\varepsilon_{n,T} \to 0$ *(as* $n, T \to \infty$*) such that*

$$\frac{\overline{L}_X(T,x)}{\varepsilon_{n,T}}(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o_{a.s.}(1)$$

and $\varepsilon_{n,T}\overline{L}_X(T,x) \xrightarrow{a.s.} \infty$, *the estimator* (4.2) *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*

$$\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o_{a.s.}(1),$$

and provided $\varepsilon_{n,T} \to 0$ *(as* $n, T \to \infty$*) such that*

$$\frac{\overline{L}_X(T,x)}{\varepsilon_{n,T}}(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o_{a.s.}(1),$$

$\varepsilon_{n,T}\overline{L}_X(T,x) \xrightarrow{a.s.} \infty$ *and* $\varepsilon^5_{n,T}\overline{L}_X(T,x) = o_{a.s.}(1)$, *then the asymptotic distribution of the drift function estimator is of the form*

$$\text{(5.2)} \qquad \sqrt{\varepsilon_{n,T}\widehat{\overline{L}}_X(T,x)}\{\hat{\mu}_{(n,T)}(x) - \mu(x)\} \Rightarrow N\big(0, K_2^{ind}\sigma^2(x)\big),$$

where $K_2^{ind} = \frac{1}{4}\int_{-\infty}^{\infty}\mathbf{1}^2_{\{|a|\le 1\}}\,da = \frac{1}{2}$ *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*

$$\text{(5.3)} \qquad \sqrt{\varepsilon_{n,T}\widehat{\overline{L}}_X(T,x)}\{\hat{\mu}_{(n,T)}(x) - \mu(x)\} \Rightarrow N\left(0, \frac{1}{2}\theta_\phi\sigma^2(x)\right),$$

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

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

$$\text{(5.4)} \qquad \Gamma_\mu(x) = \varepsilon^2_{n,T}K_1^{ind}\left[\mu'(x)\frac{s'(x)}{s(x)} + \frac{1}{2}\mu''(x)\right],$$

where $K_1^{ind} = \frac{1}{2}\int_{-\infty}^{\infty} a^2\mathbf{1}_{\{|a|\le 1\}}\,da = \frac{1}{3}$, *if* $h_{n,T} = o(\varepsilon_{n,T})$, *and*

$$\text{(5.5)} \qquad \Gamma_\mu^{(\phi)}(x) = \varepsilon^2_{n,T}\big(K_1\phi^2 + K_1^{ind}\big)\left[\mu'(x)\frac{s'(x)}{s(x)} + \frac{1}{2}\mu''(x)\right],$$

with $K_1 = \int_{-\infty}^{\infty} a^2K(a)\,da$, *if* $h_{n,T} = O(\varepsilon_{n,T})$ *with* $h_{n,T}/\varepsilon_{n,T} \to \phi > 0$. *The function* $s(x)$ *in* (5.4) *and* (5.5) *is the speed function of the process* X_t, *namely* $s(x) = 2/S'(x)\sigma^2(x)$.

REMARK 3 (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 $\sqrt{\varepsilon_{n,T}}$ (see Theorem 2.1 in Bandi (2002) for an alternative treatment). 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 (see, also, Geman (1979), among others).

REMARK 4 (The Rate of Convergence): The normalizations in (5.2) and (5.3) are random because of the presence of the local time factor $\widehat{\overline{L}}_X(T,x)$. In general, therefore, the rate of convergence will be path-dependent. The precise rate of convergence in (5.2) and (5.3) will depend on the asymptotic divergence characteristics of the chronological local time of the underlying diffusion process. We consider the two cases for which closed-form expressions for the rates of convergence exist, namely Brownian motion and the wide class of stationary processes. First, assume X_t is a Brownian motion ($\mu(.)=0$ and $\sigma(.)=\sigma$, that is). Then,

$$\overline{L}_X(T,x)=\overline{L}_B(T,x)=T^{1/2}\frac{1}{\sigma}L_W\left(1,\frac{a}{T^{1/2}\sigma}\right)=O_{a.s.}(T^{1/2}). \tag{5.6}$$

In this case, the convergence rate (to zero) of $\hat{\mu}_{(n,T)}(x)$ 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, it is possible to show that

$$\frac{\overline{L}_X(T,x)}{T}\overset{a.s.}{\longrightarrow} f(x), \tag{5.7}$$

where $f(x)$ is the time-invariant stationary distribution function of the process at x (see, e.g., Bosq (1998, Theorem 6.3, p. 150)). 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 asymptotic distribution is normal, and the (nonrandom) limiting variance depends inversely on the marginal distribution function of X_t at x.

REMARK 5 (Single Smoothing): We can simplify (4.2) above and write the estimator as a weighted average of differences with weights based on simple kernels. Consider

$$\bar{\mu}_{(n,T)}(x)=\frac{1}{\Delta_{n,T}}\frac{\sum_{i=1}^{n-1}K\left(\frac{X_{i\Delta_{n,T}}-x}{g_{n,T}}\right)(X_{(i+1)\Delta_{n,T}}-X_{i\Delta_{n,T}})}{\sum_{i=1}^{n}K\left(\frac{X_{i\Delta_{n,T}}-x}{g_{n,T}}\right)}. \tag{5.8}$$

The limit theory in this paper allows us to show that $\bar{\mu}_{(n,T)}(x)$ is consistent almost surely for the unknown drift function provided the window width $g_{n,T}$ is such that

$$\frac{\overline{L}_X(T,x)}{g_{n,T}}(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o_{a.s.}(1)$$

and $g_{n,T}\overline{L}_X(T,x) \xrightarrow{a.s.} \infty$ as $n, T \to \infty$ with $T/n \to 0$. Furthermore, if $g^5_{n,T}\overline{L}_X(T,x) = o_{a.s.}(1)$, then

$$\sqrt{g_{n,T}\widehat{\overline{L}}_X(T,x)}\{\bar{\mu}_{(n,T)}(x) - \mu(x)\} \Rightarrow N(0, K_2\sigma^2(x)), \tag{5.9}$$

where $K_2 = \int_{-\infty}^{\infty} K^2(s)\,ds$. Additionally, if $g^5_{n,T}\overline{L}_X(T,x) = O_{a.s.}(1)$, then

$$\sqrt{g_{n,T}\widehat{\overline{L}}_X(T,x)}\{\bar{\mu}_{(n,T)}(x) - \mu(x) - \Gamma_\mu(x)\} \Rightarrow N(0, K_2\sigma^2(x)), \tag{5.10}$$

where

$$\Gamma_\mu(x) = g^2_{n,T}K_1\left[\mu'(x)\frac{s'(x)}{s(x)} + \frac{1}{2}\mu''(x)\right], \tag{5.11}$$

$s(x)$ is the speed function of the process at the generic level x, and $K_1 = \int_{-\infty}^{\infty} s^2K(s)\,ds$.

It is noted that (5.8) behaves asymptotically like (4.2) in the case where $h_{n,T} = o(\varepsilon_{n,T})$ and (4.2) 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 $K_\phi = K_1\phi^2 + K_1^{ind}$ (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 $\varepsilon^5_{n,T}\overline{L}_X(T,x) = o_{a.s.}(1)$ (which implies undersmoothing with respect to the optimal bandwidth, i.e., $\varepsilon^5_{n,T}\overline{L}_X(T,x) = O_{a.s.}(1)$), 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 $\varepsilon^5_{n,T}\overline{L}_X(T,x) = o_{a.s.}(1)$, then the limiting bias remains zero but the limiting variance becomes $\frac{1}{2}\theta_\phi\sigma^2(x)$ which is strictly smaller than $\frac{1}{2}\sigma^2(x)$. 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. *Functional Estimation of the Diffusion*

We now turn to the asymptotic theory for the diffusion estimator (4.4).

THEOREM 4 (Almost Sure Convergence to the Diffusion 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*

$$\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o_{a.s.}(1),$$

and provided $\varepsilon_{nT} \to 0$ (*as* $n, T \to \infty$) *such that*

$$\frac{\overline{L}_X(T,x)}{\varepsilon_{n,T}}(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o_{a.s.}(1),$$

the estimator (4.4) *converges to the true function with probability one*.

THEOREM 5 (The Limiting Distribution of the Diffusion Estimator): *Assume* $n \to \infty$, $T \to \infty$, $\Delta_{n,T} \to 0$, *and* $h_{n,T} \to 0$ (*as* $n, T \to \infty$) *such that*

$$\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o_{a.s.}(1).$$

Also, assume $\varepsilon_{n,T} \to 0$ (*as* $n, T \to \infty$) *such that*

$$\frac{\overline{L}_X(T,x)}{\varepsilon_{n,T}}(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o_{a.s.}(1) \qquad and$$

$$\frac{\varepsilon_{n,T}^5\overline{L}_X(T,x)}{\Delta_{n,T}} = o_{a.s.}(1).$$

Then, the asymptotic distribution of the diffusion function estimator is of the form

$$\sqrt{\frac{\varepsilon_{n,T}\widehat{\overline{L}}_X(T,x)}{\Delta_{n,T}}}\left\{\hat{\sigma}^2_{(n,T)}(x) - \sigma^2(x)\right\} \Rightarrow N\left(0, 4K_2^{ind}\sigma^4(x)\right), \tag{5.12}$$

where $K_2^{ind} = \frac{1}{4}\int_{-\infty}^{\infty}\mathbf{1}^2_{\{|a|\le 1\}}da = \frac{1}{2}$ *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*

$$\sqrt{\frac{\varepsilon_{n,T}\widehat{\overline{L}}_X(T,x)}{\Delta_{n,T}}}\left\{\hat{\sigma}^2_{(n,T)}(x) - \sigma^2(x)\right\} \Rightarrow N(0, 2\theta_\phi\sigma^4(x)), \tag{5.13}$$

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

Under the same conditions, but provided

$$\frac{\varepsilon_{n,T}^5 \overline{L}_X(T,x)}{\Delta_{n,T}} = O_{a.s.}(1),$$

the limiting distribution of the diffusion estimator displays an asymptotic bias term whose form is

$$\Gamma_{\sigma^2}(x) = \varepsilon_{n,T}^2 K_1^{ind}\left[(\sigma^2(x))'\frac{s'(x)}{s(x)} + \frac{1}{2}(\sigma^2(x))''\right], \tag{5.14}$$

with $K_1^{ind} = \frac{1}{2}\int_{-\infty}^{\infty} a^2 \mathbf{1}_{\{|a|\le 1\}}\, da = \frac{1}{3}$, *if* $h_{n,T} = o(\varepsilon_{n,T})$, *and*

$$\Gamma_{\sigma^2}^{(\phi)}(x) = \varepsilon_{n,T}^2 \big(K_1\phi^2 + K_1^{ind}\big)\left[(\sigma^2(x))'\frac{s'(x)}{s(x)} + \frac{1}{2}(\sigma^2(x))''\right], \tag{5.15}$$

with $K_1 = \int_{-\infty}^{\infty} a^2 K(a)\, da$, *if* $h_{n,T} = O(\varepsilon_{n,T})$ *with* $h_{n,T}/\varepsilon_{n,T} \to \phi > 0$. *The function* $s(x)$ *in* (5.14) *and* (5.15) *is the speed function of the process* X_t, *namely* $s(x) = 2/S'(x)\sigma^2(x)$.

Next, we consider the fixed $T(=\overline{T})$ case.

THEOREM 6 (The Limiting Distribution of the Diffusion Estimator for a Fixed Time Span): *Given* $n \to \infty$, $T = \overline{T}$, *and* $h_{n,\overline{T}} \to 0$ (*as* $n \to \infty$) *such that*

$$\frac{1}{h_{n,\overline{T}}}(\Delta_{n,\overline{T}} \log(1/\Delta_{n,\overline{T}}))^{1/2} = o(1),$$

and provided $\varepsilon_{n,\overline{T}} \to 0$ (*as* $n \to \infty$) *such that*

$$\frac{1}{\varepsilon_{n,\overline{T}}}(\Delta_{n,\overline{T}} \log(1/\Delta_{n,\overline{T}}))^{1/2} = o(1),$$

the estimator (4.4) *converges to the true function with probability one*.

If $h_{n,\overline{T}} = o(\varepsilon_{n,\overline{T}})$ *and* $n\varepsilon_{n,\overline{T}}^4 = o(1)$, *then the asymptotic distribution of the diffusion function estimator is driven by a "martingale effect" and has the form*

$$\sqrt{\varepsilon_{n,\overline{T}}\, n}\left\{\hat{\sigma}^2_{(n,\overline{T})}(x) - \sigma^2(x)\right\} \Rightarrow MN\left(0, \frac{2\sigma^4(x)}{\overline{L}_X(\overline{T},x)/\overline{T}}\right), \tag{5.16}$$

where MN *indicates a mixed normal distribution*.

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

$$\frac{1}{\varepsilon_{n,\overline{T}}^{3/2}}\left\{\hat{\sigma}^2_{(n,\overline{T})}(x) - \sigma^2(x)\right\} \Rightarrow MN\left(0, 16\varphi^{ind}\frac{(\sigma'(x))^2}{\overline{L}_X(\overline{T},x)}\right), \tag{5.17}$$

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

If $h_{n,\overline{T}} = O(\varepsilon_{n,\overline{T}})$ with $h_{n,\overline{T}}/\varepsilon_{n,\overline{T}} \to \phi > 0$ and $n\varepsilon^4_{n,\overline{T}} = o(1)$, then the asymptotic distribution of the diffusion function estimator is driven by a "martingale effect" and is of the form

$$\sqrt{\varepsilon_{n,\overline{T}}n}\{\hat{\sigma}^2_{(n,\overline{T})}(x) - \sigma^2(x)\} \Rightarrow MN\left(0, \frac{2\theta_\phi\sigma^4(x)}{\overline{L}_X(\overline{T},x)/\overline{T}}\right), \tag{5.18}$$

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

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

$$\frac{1}{\varepsilon^{3/2}_{n,\overline{T}}}\{\hat{\sigma}^2_{(n,\overline{T})}(x) - \sigma^2(x)\} \Rightarrow MN\left(0, 16(\varphi^{ind,K}(\phi))\frac{(\sigma'(x))^2}{\overline{L}_X(\overline{T},x)}\right), \tag{5.19}$$

where $\varphi^{ind,K}(\phi)$ is a positive function of ϕ such that $\varphi^{ind,K}(\phi) \to \varphi^{ind}$ as $\phi \to 0$.[3]

REMARK 6: 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 determining the magnitude of the two effects is the relation of the observation rate of the process, $\Delta_{n,\overline{T}}$ that is, to the spatial bandwidth parameter, namely $\varepsilon_{n,\overline{T}}$. 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. If the spatial bandwidth $\varepsilon_{n,\overline{T}}$ is small relative to the observation interval and $n\varepsilon^4_{n,\overline{T}} = o(1)$, then the bias effect is eliminated asymptotically and the martingale effect governs the limit theory.

REMARK 7: When T is fixed as in Theorem 6 above, the admissible bandwidth conditions can be written as a function of the number of observations. The variance term dominates if (approximately)

$$\varepsilon_{n,\overline{T}} \propto n^{-k_1} \quad \text{with} \quad k_1 \in \left(\frac{1}{4}, \frac{1}{2}\right) \tag{5.20}$$

and

$$h_{n,\overline{T}} \propto n^{-k_2} \quad \text{with} \quad k_2 \in \left(0, \frac{1}{2}\right). \tag{5.21}$$

On the other hand, if (approximately)

$$\varepsilon_{n,\overline{T}} \propto n^{-k_1} \quad \text{with} \quad k_1 \in \left(0, \frac{1}{4}\right) \tag{5.22}$$

and

$$h_{n,\overline{T}} \propto n^{-k_2} \quad \text{with} \quad k_2 \in \left(0, \frac{1}{2}\right), \tag{5.23}$$

then the bias term drives the limiting distribution.

[3] See the proof of Theorem 6.

REMARK 8 (The Rate of Convergence): The diffusion function estimator converges at a faster rate than the drift estimator, namely

$$\sqrt{\frac{\varepsilon_{n,T}\widehat{\overline{L}}_X(T,x)}{\Delta_{n,T}}} \quad \text{versus} \quad \sqrt{\varepsilon_{n,T}\widehat{\overline{L}}_X(T,x)},$$

when $T \to \infty$. Using the results in Remark 4 above, in the Brownian motion and stationary case the normalizations in (5.12) and (5.13) are

$$\sqrt{\frac{\varepsilon_{n,T}T^{\frac{1}{2}}}{\Delta_{n,T}}} = \sqrt{\frac{n\varepsilon_{n,T}}{T^{1/2}}} \qquad \text{and} \qquad \sqrt{\frac{\varepsilon_{n,T}T}{\Delta_{n,T}}} = \sqrt{n\varepsilon_{n,T}},$$

respectively.

REMARK 9 (Single Smoothing): Coherently with the drift case, we can consider a simpler version of our infinitesimal volatility estimator based on single smoothing. Define

$$\bar{\sigma}^2_{(n,T)}(x) = \frac{1}{\Delta_{n,T}} \frac{\sum_{i=1}^{n-1} K\left(\frac{X_{i\Delta_{n,T}}-x}{g_{n,T}}\right)(X_{(i+1)\Delta_{n,T}} - X_{i\Delta_{n,T}})^2}{\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{g_{n,T}}\right)}. \tag{5.24}$$

Following our derivations in the convoluted case, we can prove that (5.24) is consistent almost surely for the unknown function $\sigma^2(.)$ provided the window width $g_{n,T}$ is such that

$$\frac{\overline{L}_X(T,x)}{g_{n,T}}(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o_{a.s.}(1)$$

$$\text{as} \quad n, T \to \infty \quad \text{with} \quad \frac{T}{n} \to 0.$$

Furthermore, if

$$\frac{g^5_{n,T}\overline{L}_X(T,x)}{\Delta_{n,T}} = o_{a.s.}(1),$$

then

$$\sqrt{\frac{g_{n,T}\widehat{\overline{L}}_X(T,x)}{\Delta_{n,T}}}\{\bar{\sigma}^2_{(n,T)}(x) - \sigma^2(x)\} \Rightarrow N(0, 4K_2\sigma^4(x)), \tag{5.25}$$

where $K_2 = \int_{-\infty}^{\infty} K^2(s)\,ds$. Additionally, if

$$\frac{g^5_{n,T}\overline{L}_X(T,x)}{\Delta_{n,T}} = O_{a.s.}(1),$$

then

$$(5.26)\qquad \sqrt{\frac{g_{n,T}\widehat{\overline{L}}_X(T,x)}{\Delta_{n,T}}}\{\bar{\sigma}^2_{(n,T)}(x)-\sigma^2(x)-\Gamma_{\sigma^2}(x)\}\Rightarrow N(0,4K_2\sigma^4(x)),$$

where

$$(5.27)\qquad \Gamma_{\sigma^2}(x)=g^2_{n,T}K_1\left[(\sigma^2(x))'\frac{s'(x)}{s(x)}+\frac{1}{2}(\sigma^2(x))''\right],$$

$s(x)$ is the speed function of the process at the generic level x, and $K_1 = \int_{-\infty}^{\infty} s^2K(s)\,ds$.

As in the case of drift estimation (see Remark 5 above), double-smoothing can reduce the asymptotic mean-squared error of the diffusion estimator for some processes and some levels x, thus offering increased flexibility over its simple counterpart. Contrary to drift estimation, the finite sample performance of alternative diffusion estimators based on simple and convoluted kernels is quite similar (see Bandi and Nguyen (2000)).

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

There is an important similarity between (5.16) 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 $\{t = t_1, t_2, \dots, t_n\}$ in the time interval $[0, \overline{T}]$, where $\overline{T}$ can be normalized to 1. Also, the data is equispaced. Consequently, $\{X_t = X_{\Delta_n}, X_{2\Delta_n}, X_{3\Delta_n}, \dots, X_{n\Delta_n}\}$ are n observations at points $\{t_1 = \Delta_n, t_2 = 2\Delta_n, \dots, t_n = \Delta_n\}$, where $\Delta_n = 1/n$. The estimator*

$$(5.28)\qquad \hat{\sigma}^2_{(n,1)}(x)=\frac{1}{\Delta_n}\frac{\sum_{i=1}^{n-1}\mathbf{1}_{\{|X_{i/n}-x|\le h_n\}}[X_{(i+1)/n}-X_{i/n}]^2}{\sum_{i=1}^{n}\mathbf{1}_{\{|X_{i/n}-x|\le h_n\}}}\overset{L_2}{\to}\sigma^2(x),$$

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

$$(5.29)\qquad \sqrt{h_n n}\{\hat{\sigma}^2_{(n,1)}(x)-\sigma^2(x)\}\Rightarrow MN\left(0,\frac{2\sigma^4(x)}{\overline{L}_X(1,x)}\right).$$

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, $\tilde{\sigma}^2_{n,\overline{T}}(X_{i\Delta_{n,\overline{T}}})$ can be rearranged as follows $\forall\, i \le n$:

$$\tilde{\sigma}^2_{n,\overline{T}}(X_{i\Delta_{n,\overline{T}}}) = \frac{1}{m_{n,\overline{T}}(i\Delta_{n,\overline{T}})\Delta_{n,\overline{T}}} \sum_{j=0}^{m_{n,\overline{T}}(i\Delta_{n,\overline{T}})-1} [X_{t(i\Delta_{n,\overline{T}})j+\Delta_{n,\overline{T}}} - X_{t(i\Delta_{n,\overline{T}})j}]^2 \tag{5.30}$$

$$= \frac{1}{\Delta_{n,\overline{T}}} \frac{\sum_{j=1}^{n-1} \mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}} - X_{i\Delta_{n,\overline{T}}}| \le \varepsilon_{n,\overline{T}}\}} [X_{(j+1)\Delta_{n,\overline{T}}} - X_{j\Delta_{n,\overline{T}}}]^2}{\sum_{j=1}^{n} \mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}} - X_{i\Delta_{n,\overline{T}}}| \le \varepsilon_{n,\overline{T}}\}}}. \tag{5.31}$$

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

$$\frac{1}{h_n^{3/2}} \{\hat{\sigma}^2_{(n,1)}(x) - \sigma^2(x)\} \Rightarrow MN\left(0, 16\varphi^{ind} \frac{(\sigma'(x))^2}{(\overline{L}_X(1,x))}\right), \tag{5.32}$$

where $\varphi^{ind} = 2\int_0^\infty \int_0^\infty ab(\frac{1}{2}\mathbf{1}_{\{|a|\le 1\}})(\frac{1}{2}\mathbf{1}_{\{|b|\le 1\}}) \min(a,b)\, da\, db$.

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 9 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 (as in (5.24)). 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 (5.17) and variance (5.16) 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 (see, e.g., Härdle (1990) and Pagan and Ullah (1999) for a more recent discussion).

COROLLARY 2 (c.f. Theorem 3): *Assume X_t 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 $(T/h_{n,T})(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o(1)$, and $\varepsilon_{n,T} \to 0$ (as $n, T \to \infty$) such that $(T/\varepsilon_{n,T})(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o(1)$ and $\varepsilon_{n,T}T \to \infty$. Then, $\hat{\mu}_{(n,T)}(x) \stackrel{a.s.}{\longrightarrow} \mu(x)$. Additionally, the asymptotic distribution of the drift function estimator is of the form*

$$\sqrt{\varepsilon_{n,T}T}\{\hat{\mu}_{(n,T)}(x) - \mu(x) - \Gamma_\mu(x)\} \Rightarrow N\left(0, \frac{1}{2}\frac{\sigma^2(x)}{f(x)}\right), \tag{5.33}$$

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

$$\Gamma_\mu(x) = \varepsilon^2_{n,T}\frac{1}{3}\left[\mu'(x)\frac{f'(x)}{f(x)} + \frac{1}{2}\mu''(x)\right], \tag{5.34}$$

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

Equivalently, in the diffusion case we obtain the following result.

COROLLARY 3 (c.f. Theorem 5): *Assume X_t 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 $(T/h_{n,T})(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o(1)$ and $\varepsilon_{n,T} \to 0$ (as $n, T \to \infty$) such that $(T/\varepsilon_{n,T})(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o(1)$. Then, $\hat{\sigma}^2_{(n,T)}(x) \xrightarrow{a.s.} \sigma^2(x)$. Additionally, the asymptotic distribution of the diffusion function estimator is of the form*

$$\text{(5.35)} \qquad \sqrt{\varepsilon_{n,T}n}\{\hat{\sigma}^2_{(n,T)}(x) - \sigma^2(x) - \Gamma_{\sigma^2}(x)\} \Rightarrow N\left(0, \frac{2\sigma^4(x)}{f(x)}\right),$$

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

$$\text{(5.36)} \qquad \Gamma_{\sigma^2}(x) = \varepsilon^2_{n,T}\frac{1}{3}\left[(\sigma^2(x))'\frac{f'(x)}{f(x)} + \frac{1}{2}(\sigma^2(x))''\right],$$

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

Interestingly, Corollaries 2 and 3 apply to the strictly 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 distribution (to which the process converges, eventually). The latter situation is known as positive-recurrence and is such that the speed measure of the process (as defined in Lemma 6) is finite, i.e., $s(\mathfrak{D}) < \infty$. In this case the normalized speed measure coincides with the stationary distribution of X_t (see Pollack and Siegmund (1985), for instance). Specifically,

$$\text{(5.37)} \qquad \lim_{t\to\infty} P^x(X_t < z) = \frac{s((l, z))}{s(\mathfrak{D})} \qquad \forall x, z \in \mathfrak{D}.$$

5.6. *Some Observations on the 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 5 and 9 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 ($\widehat{\overline{L}}_X(T, x)$, 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 $h_{n,\overline{T}}$ is (approximately) proportional to n^{-k} with $k \in (0, \frac{1}{2})$. Since the rate of convergence of the estimated local time to the true process is $1/\sqrt{h_{n,\overline{T}}}$ (see Bandi (2002)), it is sometimes convenient to set $h^{ltime}_{n,\overline{T}}$ equal to $c_{ltime}n^{-\frac{1}{2}}$, where c_{ltime} is a constant of proportionality that might be chosen using automated methods for bandwidth selection in density estimation (see Härdle (1990) and Pagan and Ullah (1999), among others). From a practical standpoint, functional estimation of the local time factor is 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 (see Bandi (2002) for additional discussions in the diffusion case). 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 $h_{n,T} = o(\varepsilon_{n,T})$) bandwidth $\varepsilon_{n,T}$ is more unusual. Consider the diffusion case and normalize T to 1. Remark 6 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 7 it is convenient to set $\varepsilon^{diff}_{n,\overline{T}}$ equal to $c_{diff}(1/\log(n))n^{-\frac{1}{4}}$. We undersmooth slightly with respect to the optimal rate 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 generally vanishes at a slower pace than the feasible diffusion bandwidth, a simple rule-of-thumb can be applied: we can set $\varepsilon^{drift}_{n,\overline{T}} = c_{drift}(1/\log(n))n^{-\frac{1}{4}}$ and choose c_{drift} using automated methods under the constraint that $\varepsilon^{drift}_{n,\overline{T}} > \varepsilon^{diff}_{n,\overline{T}}$. More rigorously, one could recognize the role played by local time in the functional estimation of the drift (Theorem 3) and set $\varepsilon^{drift}_{n,T}(x) = c_{drift}(1/\log(n))\widehat{\overline{L}}_X(T,x)^{-\frac{1}{5}}$. Again, we undersmooth slightly with respect to the optimal case ($\varepsilon^{drift}_{n,T}(x) \propto \widehat{\overline{L}}_x(T,x)^{-\frac{1}{5}}$) 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, in our set-up 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. Since the diffusion function is estimable over a fixed span of time, the need for level-dependent bandwidth choices appears to be less compelling than in the drift case. Nonetheless, standard arguments in favor of level-specific choices leading to bias reduction (see Pagan and Ullah (1999), among others) can still be made in our framework, even in the diffusion case.

In light of the limiting results in Remarks 5 and 9 above, it is noted that bandwidth choice in the simple case entails the same procedures as in the convoluted

that might not possess a time-invariant density. Such development is to be considered of primary importance mainly in light of the relevance of multifactor continuous-time models of the diffusion type in the pricing and hedging of derivative securities.

Graduate School of Business, University of Chicago, 1101 East 58th St., Chicago, IL 60637, U.S.A.; federico.bandi@gsb.uchicago.edu; http://gsb.uchicago.edu/fac/federico.bandi

and

Cowles Foundation for Research in Economics, Yale University, Box 208281, New Haven, CT 06520-8281, U.S.A.; University of Auckland, Auckland, New Zealand; University of York, York, UK; peter.phillips@yale.edu; http://korora.econ.yale.edu

Manuscript received August, 1999; final revision received September, 2001.

APPENDIX: PROOFS

PROOF OF LEMMA 1: See Revuz and Yor (1998, Theorem 1.2, Chapter 6, p. 213).

PROOF OF LEMMA 2: See Revuz and Yor (1998, Theorem 1.7, Chapter 6, p. 215).

PROOF OF LEMMA 3: See Revuz and Yor (1998, Exercise 1.15, Chapter 6, p. 222) and Revuz and Yor (1998, Corollary 1.6, Chapter 6, p. 215).

PROOF OF LEMMA 4: See Revuz and Yor (1998, Corollary 1.9, Chapter 6, p. 218).

PROOF OF LEMMA 5: The first part of the result is stated in Yor (1978). We prove the result in the second case ($a < 0$, that is). Start by considering a simple application of the Tanaka formula (Lemma 1), namely

$$X_t^+ = X_0^+ + \int_0^t \mathbf{1}_{(X_s>0)}\, dX_s + \frac{1}{2} L_X(t,0),$$

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

Subtract the second expression from the first expression, giving

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

Equivalently, we can write

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

where λ is a positive number. Now, multiply through by $\sqrt{\lambda}$. This gives

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

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

$$\sqrt{\lambda}\int_0^t \mathbf{1}_{(a/\lambda<X_s\leq 0)}\,dX_s,$$

as $\lambda\to\infty$. Further,

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

Now, notice that $\sqrt{\lambda}\int_0^t \mathbf{1}_{(a/\lambda<X_s\leq 0)}\mu(X_s)\,ds\overset{a.s.}{\to}0$ as $\lambda\to\infty$. In fact, by the occupation time formula (Lemma 3), we can write

$$\sqrt{\lambda}\int_0^t \mathbf{1}_{(a/\lambda<X_s\leq 0)}\mu(X_s)\,ds=\sqrt{\lambda}\int_{-\infty}^{\infty}\mathbf{1}_{(a/\lambda<b\leq 0)}\frac{\mu(b)}{\sigma^2(b)}L_X(t,b)\,db,$$

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

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

Next, note that the map $(t,a)\to L_X(t,a)$ is continuous in t and càdlàg in a with probability one (from Lemma 2) since the solution to (2.1) is a continuous SMG. In addition,

$$L_t^a-L_t^{a-}=2\int_0^t \mathbf{1}_{(X_s=a)}\,dV_s=0,$$

where V_t is the finite variation component of X_t. Hence, there exists a bicontinuous modification of the family of local times and the map $(t,a)\to L_X(t,a)$ can be taken to be continuous in both time and space. Thus, using continuity and dominated convergence, we obtain

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

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

$$\sqrt{\lambda}\int_0^t \mathbf{1}_{(a/\lambda<X_s\leq 0)}\mu(X_s)\,ds\xrightarrow[\lambda\to\infty]{a.s.}0,$$

as stated earlier. This, in turn, implies that the asymptotic behavior of (7.1) is determined by

$$\sqrt{\lambda}\int_0^t \mathbf{1}_{(a/\lambda<X_s\leq 0)}\sigma(X_s)\,dB_s.$$

Now define

$$M_t^\lambda:=\sqrt{\lambda}\int_0^t \mathbf{1}_{(a/\lambda<X_s\leq 0)}\sigma(X_s)\,dB_s.$$

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

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

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

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

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 (see, e.g., Revuz and Yor (1998, Theorem 1.6, Chapter 5, p. 173, and, for an asymptotic version, Theorem 2.3, Chapter 13, p. 496)). It follows that

$$\begin{aligned} M_t^\lambda &:= \sqrt{\lambda}\int_0^t \mathbf{1}_{(a<\lambda X_s\leq 0)}\sigma(X_s)\,dB_s \\ &\xrightarrow[\lambda\to\infty]{d} \widetilde{B}_{-aL_X(t,0)} \\ &\overset{d}{=} \sqrt{-a}\widetilde{B}_{L_X(t,0)} \\ &\overset{d}{=} \mathfrak{B}(L_X(t,0),-a), \end{aligned}$$

where $L_X(t,0) = \lim_{\varepsilon\to 0}(1/\varepsilon)\int_0^t \mathbf{1}_{[0,\varepsilon[}\sigma^2(X_s)\,ds$ a.s. and $\mathfrak{B}(\cdot,\cdot)$ 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 $(t,a)\in\Re_+\times\Re_-$)

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

converges weakly to

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

where $(\mathfrak{B}(s,c);\ (s,c)\in\Re_+^2)$ is a standard Brownian sheet independent of X_t. (For the independence property, see Revuz and Yor (1998, Exercises 2.11 and 2.12, Chapter 13, p. 501).) Then, a simple generalization of the previous finding to the spatial location $x\neq 0$ gives

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

as $\lambda\to\infty$. *Q.E.D.*

Proof of Lemma 6: Immediate given the Ratio-limit Theorem (see Azema, Kaplan-Duflo, and Revuz (1967) for the original treatment and Revuz and Yor (1998, Theorem 3.12, Chapter 10, p. 408) for additional comments) and the observation that the unique (up to multiplication by a constant) invariant measure of a recurrent diffusion is equal to the speed measure (Karatzas and Shreve (1991, Exercise 5.40, Chapter 5, p. 352, and Remark 6.19, Chapter 5, p. 362), for instance). *Q.E.D.*

Proof of Theorem 1: See Florens-Zmirou (1993) for the case involving a discontinuous kernel function. The derivation in the case of a continuous kernel (as implied by Assumption 2), follows the line of the proof of Theorem 2 below and is omitted here for brevity. *Q.E.D.*

Proof of Corollary 1: If $n, T\to\infty$ and $T/n = \Delta_{n,T}\to 0$, then

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

converges to $\overline{L}_X(\infty, x)$ provided $h_{n,T} \to 0$ (as $n, T \to \infty$) in such a way that

$$\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o_{a.s.}(1).$$

But $\overline{L}_X(\infty, x) = \overline{L}_X(\sup\{t : X_t = x\}, x)$ a.s. (Revuz and Yor (1998, Proposition 1.3, Remark 2, Chapter 6, p. 214)). If the process is recurrent, then $\overline{L}_X(\sup\{t : X_t = x\}, x) = \infty$ with probability one. *Q.E.D.*

Proof of Theorem 2: We start by considering the expression

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

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

First, we examine (7.3). We want to prove that

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

We begin with the numerator and look at the quantity

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

Given the properties of $K(.)$ (from Assumption 2) and the properties of $\mu(.)$ (from Assumption 1), (7.4) is seen to be bounded as follows:

$$\begin{aligned}
&\frac{1}{h_{n,T}}\left|\sum_{i=0}^{n-1}\int_{iT/n}^{(i+1)T/n}\left[K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\mu(X_{i\Delta_{n,T}}) - K\left(\frac{X_s-x}{h_{n,T}}\right)\mu(X_s)\right]ds\right| \\
&\quad + \left|\frac{\Delta_{n,T}}{h_{n,T}}K\left(\frac{X_0-x}{h_{n,T}}\right)\mu(X_0)\right| + \left|\frac{\Delta_{n,T}}{h_{n,T}}K\left(\frac{X_{n\Delta_{n,T}}-x}{h_{n,T}}\right)\mu(X_{n\Delta_{n,T}})\right| \\
&\le \frac{1}{h_{n,T}}\left|\sum_{i=0}^{n-1}\int_{i\Delta_{n,T}}^{(i+1)\Delta_{n,T}}\left[K\left(\frac{X_s-x}{h_{n,T}}\right)\mu(X_{i\Delta_{n,T}}) - K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\mu(X_{i\Delta_{n,T}})\right]ds\right| \\
&\quad + \frac{1}{h_{n,T}}\left|\sum_{i=0}^{n-1}\int_{i\Delta_{n,T}}^{(i+1)\Delta_{n,T}}\left[K\left(\frac{X_s-x}{h_{n,T}}\right)\mu(X_s) - K\left(\frac{X_s-x}{h_{n,T}}\right)\mu(X_{i\Delta_{n,T}})\right]ds\right| \\
&\quad + C_3 O_{a.s.}\left(\frac{\Delta_{n,T}}{h_{n,T}}\right)
\end{aligned}$$

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

$$(7.6)\qquad +\frac{1}{h_{n,T}}\left|\sum_{i=0}^{n-1}\int_{i\Delta_{n,T}}^{(i+1)\Delta_{n,T}} K\left(\frac{X_s-x}{h_{n,T}}\right)(\mu(X_s)-\mu(X_{i\Delta_{n,T}}))\,ds\right|$$
$$+C_3 O_{a.s.}\left(\frac{\Delta_{n,T}}{h_{n,T}}\right),$$

where C_3 is a suitable constant (from Assumption 2) and $\widetilde{X}_{is}$ in (7.5) is a value on the line segment connecting X_s to $X_{i\Delta_{n,T}}$. Now define

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

By the Levy's modulus of continuity of diffusions (see, e.g., Karatzas and Shreve (1991, Theorem 9.25, Chapter 2, p. 114)),

$$(7.8)\qquad P\left(\left[\lim\sup_{\Delta_{n,T}\to 0}\frac{\kappa_{n,T}}{(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2}}=C_4\right]\right)=1,$$

where C_4 is a suitable constant. In turn, (7.8) implies that

$$\kappa_{n,T}=O_{a.s.}\big((\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2}\big).$$

Hence, if $h_{n,T}$ is such that $\frac{1}{h_{n,T}}(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2}=o(1)$, then

$$(7.9)\qquad \frac{\kappa_{n,T}}{h_{n,T}}=o_{a.s.}(1)$$

as $n,T\to\infty$. In view of (7.9) we have

$$(7.10)\qquad K'\left(\frac{\widetilde{X}_{is}-x}{h_{n,T}}\right)=K'\left(\frac{X_s-x}{h_{n,T}}+o_{a.s.}(1)\right),$$

uniformly over $i=1,\ldots,n$. It follows from (7.7) and (7.10) that (7.5) is bounded by

$$\begin{aligned}&\left(\frac{\kappa_{n,T}}{h_{n,T}}\right)\frac{1}{h_{n,T}}\int_0^T\left|K'\left(\frac{X_s-x}{h_{n,T}}+o_{a.s.}(1)\right)\right||\mu(X_s+o_{a.s.}(1))|\,ds\\&=\left(\frac{\kappa_{n,T}}{h_{n,T}}\right)\frac{1}{h_{n,T}}\int_{-\infty}^{\infty}\left|K'\left(\frac{p-x}{h_{n,T}}+o_{a.s.}(1)\right)\right||\mu(p)|\overline{L}_X(T,p)\,dp\\&=\left(\frac{\kappa_{n,T}}{h_{n,T}}\right)\int_{-\infty}^{\infty}|K'(q+o_{a.s.}(1))||\mu(qh_{n,T}+x)|\overline{L}_X(T,qh_{n,T}+x)\,dq\\&\leq C_5\left(\frac{\kappa_{n,T}}{h_{n,T}}\right)O_{a.s.}(\overline{L}_X(T,x)),\end{aligned}$$

for some suitable constant C_5, by virtue of the absolute integrability of K', the continuity of $\overline{L}_X(t,\cdot)$ (see the proof of Lemma 5) and $\mu(\cdot)$ (from Assumption 1), and the occupation time formula. Employing similar methods we can prove that (7.6) is bounded by

$$C_6(\kappa_{n,T})O_{a.s.}(\overline{L}_X(T,x)).$$

In consequence, the formula for the numerator of (7.3) holds. As for the denominator of (7.3), we can show the stated result using the same steps as for (7.5) above. Next, it is easy to prove that

$$\begin{aligned}&\frac{\int_0^T\frac{1}{h_{n,T}}K\left(\frac{X_s-x}{h_{n,T}}\right)\mu(X_s)\,ds+O_{a.s.}\left(\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2}\right)}{\int_0^T\frac{1}{h_{n,T}}K\left(\frac{X_s-x}{h_{n,T}}\right)ds+O_{a.s.}\left(\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2}\right)}\\&=\frac{\mu(x)s(x)+o_{a.s.}(1)}{s(x)+o_{a.s.}(1)}+o_{a.s.}(1)\overset{a.s.}{\to}\mu(x),\end{aligned}$$

where $s(\cdot)$ is the speed function of the process, by virtue of Lemma 6, the continuity of $so\mu(.)$ and dominated convergence as $h_{n,T} \to 0$, with $n, T \to \infty$, so that

$$\frac{\overline{L}_X(T,x)}{h_{n,T}}(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o_{a.s.}(1).$$

We now turn to the analysis of (7.2). It is sufficient to show that

$$\tilde{\mu}_{n,T}(X_{i\Delta_{n,T}}) = \mu(X_{i\Delta_{n,T}}) + o_{a.s.}(1) \tag{7.11}$$

for a fixed $X_{i\Delta_{n,T}}$, in order to verify the stated result. To do so, using the Lipschitz property of $\mu(.)$ from Assumption 1, we bound

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

as follows:

$$\begin{aligned}
&\frac{1}{m_{n,T}(i\Delta_{n,T})}\sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1}\frac{\left[X_{t(i\Delta_{n,T})j+\Delta_{n,T}} - X_{t(i\Delta_{n,T})j}\right]}{\Delta_{n,T}} - \mu(X_{i\Delta_{n,T}})\\
&= \frac{1}{m_{n,T}(i\Delta_{n,T})\Delta_{n,T}}\sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1}\int_{t(i\Delta_{n,T})j}^{t(i\Delta_{n,T})j+\Delta_{n,T}}\left(\mu(X_s) - \mu(X_{i\Delta_{n,T}})\right)ds\\
&\quad + \frac{1}{m_{n,T}(i\Delta_{n,T})\Delta_{n,T}}\sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1}\int_{t(i\Delta_{n,T})j}^{t(i\Delta_{n,T})j+\Delta_{n,T}}\sigma(X_s)\,dB_s\\
&= C_7 O_{a.s.}(\kappa_{n,T}) + \frac{1}{m_{n,T}(i\Delta_{n,T})\Delta_{n,T}}\sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1}\int_{t(i\Delta_{n,T})j}^{t(i\Delta_{n,T})j+\Delta_{n,T}}\sigma(X_s)\,dB_s, \qquad (7.12)
\end{aligned}$$

where $\kappa_{n,T}$ has its usual meaning. Now, define the stochastic integral $y_{t(i\Delta_{n,T})j+\Delta_{n,T}} = \int_{t(i\Delta_{n,T})j}^{t(i\Delta_{n,T})j+\Delta_{n,T}}\sigma(X_s)\,dB_s$, which is measurable with respect to $\widetilde{\Im}^X_{t(i\Delta_{n,T})j+\Delta_{n,T}}$. Further, notice that

$$E(y_{t(i\Delta_{n,T})j+\Delta_{n,T}}) = 0,$$

and, by the Ito isometry (see, e.g., Øksendal (1995)),

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

for all $j \le m_{n,T}$. Hence, $(y_{t(i\Delta_{n,T})j+\Delta_{n,T}}, \widetilde{\Im}^X_{t(i\Delta_{n,T})j+\Delta_{n,T}})$ is a martingale difference array with zero mean and finite variance $\theta_{t(i\Delta_{n,T})j+\Delta_{n,T}}$. Invoking a strong law of large numbers for martingale differences (see Hall and Heyde (1980, Theorem 2.19, p. 36), for instance), we have

$$\frac{1}{m_{n,T}(i\Delta_{n,T})}\sum_{j=0}^{m_{n,T}(i\Delta_{n,T})-1} y_{t(i\Delta_{n,T})j+\Delta_{n,T}} \xrightarrow{a.s.} 0 \quad \text{with} \quad n, T \to \infty,$$

as $m_{n,T} \xrightarrow{a.s.} \infty$, by recurrence (as implied by Assumption 1, Condition (iii)). We now explore the rate of convergence. Write

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

First, analyze the numerator of this expression. Consider

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

For any fixed $X_{i\Delta_{n,T}}$, $U_{n,T}^{X_{i\Delta_{n,T}}}(r)$ can be embedded in a continuous martingale whose quadratic variation process $[U_{n,T}^{X_{i\Delta_{n,T}}}]_r$ is

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

using standard methods (see the proof of Lemma 5), by virtue of (3.11). Now, as in Theorem 3.4 in Phillips and Ploberger (1996), expanding the probability space as needed, we have

$$\left(U_{n,T}^{X_{i\Delta_{n,T}}}(1)\right)^2 \Big/ \left[U_{n,T}^{X_{i\Delta_{n,T}}}\right]_1 = O_{a.s.}(1),$$

and then it follows that

$$\begin{aligned}
&\sqrt{\overline{L}_X(T, X_{i\Delta_{n,T}})\varepsilon_{n,T}}\left(\frac{\frac{1}{\Delta_{n,T}} \frac{\Delta_{n,T}}{2\varepsilon_{n,T}} \sum_{j=1}^{n-1} \mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}} \int_{j\Delta_{n,T}}^{(j+1)\Delta_{n,T}} \sigma(X_s)\,dB_s}{\frac{\Delta_{n,T}}{2\varepsilon_{n,T}} \sum_{j=1}^{n} \mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}}}\right) \\
&\quad= O_{a.s.}(1).
\end{aligned}$$

This result implies that the bound (7.12) becomes

$$C_7 O_{a.s.}(\kappa_{n,T}) + O_{a.s.}\left(\sqrt{\frac{1}{\overline{L}_X(T, X_{i\Delta_{n,T}})\varepsilon_{n,T}}}\right) \xrightarrow{a.s.} 0.$$

In fact, $\varepsilon_{n,T}\overline{L}_X(T, X_{i\Delta_{n,T}}) \xrightarrow{a.s.} \infty$ as $n, T \to \infty$ since, by assumption, we control the spatial bandwidth $\varepsilon_{n,T}$ to ensure that this property holds. *Q.E.D.*

PROOF OF THEOREM 3: Write the estimation error in two components as follows:

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

$$= \underbrace{\frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\tilde{\mu}_{n,T}(X_{i\Delta_{n,T}})}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} - \frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\mu(X_{i\Delta_{n,T}})}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}}_{\text{term } V}$$

$$+ \underbrace{\frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\mu(X_{i\Delta_{n,T}})}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} - \frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\mu(x)}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}}_{\text{term } B}$$

$$= \text{term } V + \text{term } B.$$

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

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

As in the proof of Theorem 2, using Lemma 6, we find that

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

$$= \frac{\frac{1}{h_{n,T}}\int_{-\infty}^{\infty} K\left(\frac{a-x}{h_{n,T}}\right)(\mu(a)-\mu(x))s(a)\,da + o_{a.s.}(1)}{\frac{1}{h_{n,T}}\int_{-\infty}^{\infty} K\left(\frac{a-x}{h_{n,T}}\right)s(a)\,da + o_{a.s.}(1)} + o_{a.s.}(1),$$

for a fixed $h_{n,T}$. Neglecting the smaller order terms, we can write

$$\frac{\frac{1}{h_{n,T}}\int_{-\infty}^{\infty} K\left(\frac{a-x}{h_{n,T}}\right)(\mu(a)-\mu(x))s(a)\,da}{\frac{1}{h_{n,T}}\int_{-\infty}^{\infty} K\left(\frac{a-x}{h_{n,T}}\right)s(a)\,da}$$

$$= \frac{\int_{-\infty}^{\infty} K(u)(\mu(x+uh_{n,T})-\mu(x))s(x+uh_{n,T})\,du}{\int_{-\infty}^{\infty} K(u)s(x+uh_{n,T})\,du}$$

$$= h_{n,T}^2 K_1\left[\frac{1}{2}\mu''(x)+\mu'(x)\frac{s'(x)}{s(x)}\right] + o(h_{n,T}^2),$$

where $K_1 = \int u^2 K(u)\,du < \infty$, by the second order properties of the kernel $K(.)$ (from Assumption 2) and the differentiability properties of $\sigma(.)$ and $\mu(.)$ (from Assumption 1). Now consider the term V, viz.,

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

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

The numerator can be written as

$$\begin{aligned}
&\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)(\tilde{\mu}_{n,T}(X_{i\Delta_{n,T}})-\mu(X_{i\Delta_{n,T}}))\\
&=\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\\
&\quad\times\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}}\frac{n}{T}\left[\left(X_{(j+1)\Delta_{n,T}}-X_{j\Delta_{n,T}}\right)-\mu\left(X_{i\Delta_{n,T}}\right)\Delta_{n,T}\right]}{\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}}}\\
&\quad-\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\frac{\mu\left(X_{i\Delta_{n,T}}\right)\frac{\Delta_{n,T}}{2\varepsilon_{n,T}}\mathbf{1}_{\{|X_{n\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}}}{\frac{\Delta_{n,T}}{2\varepsilon_{n,T}}\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}}}\\
&=V_1^{num}+V_2^{num}.
\end{aligned}$$

Using the occupation time formula, it is immediate to prove that $V_2^{num}=O_{a.s.}(\Delta_{n,T}/\varepsilon_{n,T})$. As for V_1^{num}, write

$$\begin{aligned}
V_1^{num}&=\underbrace{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}}\frac{n}{T}\left[\int_{j\Delta_{n,T}}^{(j+1)\Delta_{n,T}}(\mu(X_s)-\mu(X_{i\Delta_{n,T}}))\,ds\right]}{\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}}}}_{(A_{n,T})}\\
&\quad+\underbrace{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n}K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}}\frac{n}{T}\left[\int_{j\Delta_{n,T}}^{(j+1)\Delta_{n,T}}\sigma(X_s)\,dB_s\right]}{\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}}}}_{(B_{n,T}(1))}.
\end{aligned}$$

These two elements comprise an additional bias effect, $A_{n,T}$, and a variance effect, $B_{n,T}(1)$. First, examine $\sqrt{\varepsilon_{n,T}}B_{n,T}(r)$, viz.,

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

The term $B_{n,T}(r)$ can be embedded in a time changed Brownian motion with increasing process that is given by the limit of the discretized quadratic variation process $[B_{n,T}]_r$ defined as

$$\begin{aligned}
&[B_{n,T}]_r\\
&=\left(\frac{\Delta_{n,T}}{h_{n,T}}\right)^2\sum_{i=1}^{[nr]}\sum_{k=1}^{[nr]}K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)K\left(\frac{X_{k\Delta_{n,T}}-x}{h_{n,T}}\right)\\
&\quad\times\frac{\frac{1}{4}\left(\frac{1}{\sqrt{\varepsilon_{n,T}}}\right)^2\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{k\Delta_{n,T}}|\le\varepsilon_{n,T}\}}\left[\int_{j\Delta_{n,T}}^{(j+1)\Delta_{n,T}}\sigma^2(X_s)\,ds\right]}{\left(\frac{\Delta_{n,T}}{2\varepsilon_{n,T}}\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}}\right)\left(\frac{\Delta_{n,T}}{2\varepsilon_{n,T}}\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{k\Delta_{n,T}}|\le\varepsilon_{n,T}\}}\right)}.
\end{aligned}$$

(See, Knight (1971), for instance.) Now, notice that

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

provided $h_{n,T}$ and $\varepsilon_{n,T}$ are such that $(\overline{L}_X(T,x)/h_{n,T})(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o_{a.s.}(1)$ and $(\overline{L}_X(T,x)/\varepsilon_{n,T})(\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2} = o_{a.s.}(1)$ as $n, T \to \infty$. Let

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

Then, we obtain

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

Setting

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

this last expression becomes

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

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

$$[B_{n,T}]_r \xrightarrow{a.s.} \frac{1}{2}\sigma^2(x)\frac{(\overline{L}_X(rT,x))^2}{\overline{L}_X(T,x)},$$

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

$$[B_{n,T}]_r \xrightarrow{a.s.} \frac{1}{2}\theta_\phi\sigma^2(x)\frac{(\overline{L}_X(rT,x))^2}{\overline{L}_X(T,x)},$$

where

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

By earlier arguments (see, e.g., the proof of Lemma 5), the above results imply that

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

and

$$\sqrt{\varepsilon_{n,T}\overline{L}_X(T,x)}\left(\frac{B_{n,T}(1)}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}\right) \Rightarrow N\left(0, \frac{1}{2}\sigma^2(x)\right), \tag{7.13}$$

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

$$\sqrt{\varepsilon_{n,T}\overline{L}_X(T,x)}\left(\frac{B_{n,T}(1)}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}\right) \Rightarrow N\left(0, \frac{1}{2}\theta_\phi\sigma^2(x)\right). \tag{7.14}$$

Next, examine the additional bias term, $A_{n,T}$ that is. We have

$$\begin{aligned}
&\frac{A_{n,T}}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}\\
&\quad= \frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}}\left(\mu\left(X_{j\Delta_{n,T}}\right)-\mu\left(X_{i\Delta_{n,T}}\right)\right)}{\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}}}}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}\\
&\qquad+ \frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}}\frac{n}{T}\left[\int_{j\Delta_{n,T}}^{(j+1)\Delta_{n,T}}(\mu(X_s)-\mu(X_{jT/n}))\,ds\right]}{\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,T}}-X_{i\Delta_{n,T}}|\le\varepsilon_{n,T}\}}}}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)}\\
&\quad= \frac{\frac{1}{h_{n,T}}\int_{-\infty}^{\infty}K\left(\frac{a-x}{h_{n,T}}\right)\frac{\int_{-\infty}^{\infty}\mathbf{1}_{\{|b-a|\le\varepsilon_{n,T}\}}(\mu(b)-\mu(a))s(b)\,db}{\int_{-\infty}^{\infty}\mathbf{1}_{\{|b-a|\le\varepsilon_{n,T}\}}s(b)\,db}s(a)\,da}{\frac{1}{h_{n,T}}\int_{-\infty}^{\infty}K\left(\frac{a-x}{h_{n,T}}\right)s(a)\,da} + O_{a.s.}\left((\Delta_{n,T}\log(1/\Delta_{n,T}))^{1/2}\right),
\end{aligned}$$

for a fixed $h_{n,T}$. Neglecting the order term and letting $(a-x)/h_{n,T}=c$ we can write

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

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

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

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

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

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

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

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

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

$$= \varepsilon_{n,T}^2 K_1^{ind}\left[\frac{1}{2}\mu''(x)+\mu'(x)\frac{s'(x)}{s(x)}\right]+o(\varepsilon_{n,T}^2),$$

where $K_1^{ind}=\frac{1}{2}\int_{-\infty}^{\infty} a^2\mathbf{1}_{\{|a|\le 1\}}\,da=\frac{1}{3}$. Now suppose $h_{n,T}=O(\varepsilon_{n,T})$ with $(h_{n,T}/\varepsilon_{n,T})\to\phi$. Then,

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

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

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

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

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

which implies

$$\frac{\int_{-\infty}^{\infty} K(c) \frac{\int_{-\infty}^{\infty} \mathbf{1}_{\{|g|\le 1\}} \left(\mu'(x)\varepsilon_{n,T} g+\frac{1}{2}\mu''(x)\varepsilon_{n,T}^2 g^2+\mu''(x)\varepsilon_{n,T}^2 \phi g c+o\right) s(x+\varepsilon_{n,T}(g+\phi c))\, dg}{\int_{-\infty}^{\infty} \mathbf{1}_{\{|g|\le 1\}} s(x+\varepsilon_{n,T}(g+\phi c))\, dg} s(x+h_{n,T}c)\, dc}{\int_{-\infty}^{\infty} K(c) s(x+h_{n,T}c)\, dc} \tag{7.17}$$

$$+ h_{n,T}^2 K_1 \left[\frac{1}{2}\mu''(x) + \mu'(x)\frac{s'(x)}{s(x)}\right] + o\left(h_{n,T}^2\right).$$

Now, notice that (7.17) can be represented as

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

$$= \varepsilon_{n,T}^2 K_1^{ind}\left[\frac{1}{2}\mu''(x) + \mu'(x)\frac{s'(x)}{s(x)}\right] + o\left(\varepsilon_{n,T}^2\right).$$

Next, write (7.16) as

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

$$= h_{n,T}^2 K_1 \left[\frac{1}{2}\mu''(x) + \mu'(x)\frac{s'(x)}{s(x)}\right] + o\left(h_{n,T}^2\right).$$

To conclude, when $h_{n,T} = O(\varepsilon_{n,T})$ with $(h_{n,T}/\varepsilon_{n,T}) \to \phi \ge 0$, then

$$\frac{A_{n,T}}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} + B = \varepsilon_{n,T}^2\left(K_1\phi^2 + K_1^{ind}\right)\left[\frac{1}{2}\mu''(x) + \mu'(x)\frac{s'(x)}{s(x)}\right]. \tag{7.18}$$

In consequence, defining the estimation error decomposition E as

$$E \overset{d}{=} B + \frac{A_{n,T}}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} + \frac{B_{n,T}(1)}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)},$$

we can write

$$\sqrt{\varepsilon_{n,T}\widehat{\overline{L}}_X(T,x)}\left\{\frac{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)\tilde{\mu}_{n,T}\left(X_{i\Delta_{n,T}}\right)}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} - \mu(x)\right\}$$

$$\overset{d}{=} \sqrt{\varepsilon_{n,T}\widehat{\overline{L}}_X(T,x)}\left\{\frac{B_{n,T}(1)}{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,T}}-x}{h_{n,T}}\right)} + O\left(\varepsilon_{n,T}^2\right)\right\}$$

$$\Rightarrow N\left(0, \frac{1}{2}\sigma^2(x)\right),$$

from (7.13), if $h_{n,T} = o(\varepsilon_{n,T})$, $\varepsilon^5_{n,T}\overline{L}_X(T,x) = o_{a.s.}(1)$, and $\varepsilon_{n,T}\overline{L}_X(T,x) \xrightarrow{a.s.} \infty$. If $h_{n,T} = o(\varepsilon_{n,T})$, $\varepsilon_{n,T}\overline{L}_X(T,x) \xrightarrow{a.s.} \infty$, and $\varepsilon^5_{n,T}\overline{L}_X(T,x) = O_{a.s.}(1)$, then

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

from (7.13) and (7.18), where

$$\Gamma_\mu(x) = \varepsilon^2_{n,T}K_1^{ind}\left[\frac{1}{2}\mu''(x) + \mu'(x)\frac{s'(x)}{s(x)}\right],$$

with $K_1^{ind} = \frac{1}{2}\int_{-\infty}^{\infty} a^2 \mathbf{1}_{\{|a|\le 1\}}\,da = \frac{1}{3}$. If $h_{n,T} = O(\varepsilon_{n,T})$ with $h_{n,T}/\varepsilon_{n,T} \to \phi > 0$, $\varepsilon^5_{n,T}\overline{L}_X(T,x) = o_{a.s.}(1)$, and $\varepsilon_{n,T}\overline{L}_X(T,x) \xrightarrow{a.s.} \infty$, then

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

from (7.14), where $\theta_\phi = \frac{1}{2}\int_{-\infty}^{\infty}\int_{(z-1)/\phi}^{(z+1)/\phi}\int_{(z-1)/\phi}^{(z+1)/\phi} K(a)K(e)\,dz\,da\,de$. If $h_{n,T} = O(\varepsilon_{n,T})$ with $h_{n,T}/\varepsilon_{n,T} \to \phi > 0$, $\varepsilon_{n,T}\overline{L}_X(T,x) \xrightarrow{a.s.} \infty$, and $\varepsilon^5_{n,T}\overline{L}_X(T,x) = O_{a.s.}(1)$, then

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

from (7.14) and (7.18), where

$$\Gamma_\mu^{(\phi)}(x) = \varepsilon^2_{n,T}\left(K_1\phi^2 + K_1^{ind}\right)\left[\frac{1}{2}\mu''(x) + \mu'(x)\frac{s'(x)}{s(x)}\right]. \qquad \textit{Q.E.D.}$$

PROOF OF THEOREM 4: The proof follows that of Theorem 2 and is omitted here for brevity. *Q.E.D.*

PROOF OF THEOREM 5: See the proof of Theorem 3. *Q.E.D.*

PROOF OF THEOREM 6: Fix $T=\overline{T}$. We write the estimation error as follows:

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

As earlier in the drift case (see the proof of Theorem 3), we have a bias term, B, and a second effect, V. We start with the bias term B. Using Lemma 3 and the symmetry of $K(\cdot)$ from Assumption 2, the numerator of B, B^{num} say, is seen to be distributed as

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

By Lemma 5 the first term has the following limiting form as a functional of a Brownian sheet $\mathfrak{B}(\cdot,\cdot)$:

$$
\begin{aligned}
B^1_{num}&=2h^{3/2}_{n,\overline{T}}\frac{(\sigma^2(x))'}{\sigma^2(x)}\int_{-\infty}^{\infty}cK(c)\frac{1}{2\sqrt{h_{n,\overline{T}}}}(L_X(\overline{T},x+h_{n,\overline{T}}c)-L_X(\overline{T},x))\,dc\\
&=2h^{3/2}_{n,\overline{T}}\frac{(\sigma^2(x))'}{\sigma^2(x)}\int_{0}^{\infty}cK(c)\frac{1}{2\sqrt{h_{n,\overline{T}}}}(L_X(\overline{T},x+h_{n,\overline{T}}c)-L_X(\overline{T},x))\,dc\\
&\quad+2h^{3/2}_{n,\overline{T}}\frac{(\sigma^2(x))'}{\sigma^2(x)}\int_{-\infty}^{0}cK(c)\frac{1}{2\sqrt{h_{n,\overline{T}}}}(L_X(\overline{T},x+h_{n,\overline{T}}c)-L_X(\overline{T},x))\,dc\\
&\Rightarrow 2h^{3/2}_{n,\overline{T}}\frac{(\sigma^2(x))'}{\sigma^2(x)}\int_{0}^{\infty}cK(c)\mathfrak{B}^{\oplus}(L_X(\overline{T},x),c)\,dc\\
&\quad+2h^{3/2}_{n,\overline{T}}\frac{(\sigma^2(x))'}{\sigma^2(x)}\int_{-\infty}^{0}cK(c)\mathfrak{B}^{\otimes}(L_X(\overline{T},x),-c)\,dc\\
&\stackrel{d}{=}2h^{3/2}_{n,\overline{T}}\frac{(\sigma^2(x))'}{\sigma(x)}\sqrt{\overline{L}_X(\overline{T},x)}\left(\int_{0}^{\infty}cK(c)\mathfrak{B}^{\oplus}(1,c)\,dc+\int_{-\infty}^{0}cK(c)\mathfrak{B}^{\otimes}(1,-c)\,dc\right),
\end{aligned}
\tag{7.19}
$$

where $\mathfrak{B}^{\oplus}$ and $\mathfrak{B}^{\otimes}$ are independent Brownian sheets (see Revuz and Yor (1998, Theorem 2.3, Chapter 13, p. 496), for instance). It follows that,

$$\int_0^\infty cK(c)\mathfrak{B}^{\oplus}(1,c)\,dc \stackrel{d}{=} \int_0^\infty cK(c)B^{\oplus}(c)\,dc \stackrel{d}{=} N\left(0, \int_0^\infty\int_0^\infty usK(u)K(s)\min(u,s)\,du\,ds\right).$$

Analogously,

$$\int_{-\infty}^0 cK(c)\mathfrak{B}^{\otimes}(1,-c)\,dc \stackrel{d}{=} \int_{-\infty}^0 cK(c)B^{\otimes}(-c)\,dc \stackrel{d}{=} N\left(0, -\int_{-\infty}^0\int_{-\infty}^0 usK(u)K(s)\max(u,s)\,du\,ds\right).$$

In consequence,

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

where $\varphi = 2\int_0^\infty\int_0^\infty ueK(u)K(e)\min(u,e)\,du\,de$. Then,

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

As for

$$\frac{B_{num}^2}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)} \quad\text{and}\quad \frac{B_{num}^3}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)},$$

by the Lipschitz property of $\sigma^2(.)$, we can write

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

Then,

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

Next, consider the numerator of the term V which, by an application of Itô's Lemma, can be expressed as

$$V = V_1^{num} + V_2^{num},$$

where $V_2^{num} = O_{a.s.}(\Delta_{n,\overline{T}}/\varepsilon_{n,\overline{T}})$ and

$$
\begin{aligned}
(7.20)\qquad & V_1^{num} \\
&= \underbrace{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}}-X_{i\Delta_{n,\overline{T}}}|\le\varepsilon_{n,\overline{T}}\}}\frac{n}{\overline{T}}\left[\int_{j\Delta_{n,\overline{T}}}^{(j+1)\Delta_{n,\overline{T}}}(\sigma^2(X_s)-\sigma^2(X_{i\Delta_{n,\overline{T}}}))\,ds\right]}{\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}}-X_{i\Delta_{n,\overline{T}}}|\le\varepsilon_{n,\overline{T}}\}}}}_{(A_{n,\overline{T}})} \\
&\quad + \underbrace{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}}-X_{i\Delta_{n,\overline{T}}}|\le\varepsilon_{n,\overline{T}}\}}\frac{n}{\overline{T}}\left[\int_{j\Delta_{n,\overline{T}}}^{(j+1)\Delta_{n,\overline{T}}}2(X_s-X_{j\Delta_{n,\overline{T}}})\sigma(X_s)\,dB_s\right]}{\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}}-X_{i\Delta_{n,\overline{T}}}|\le\varepsilon_{n,\overline{T}}\}}}}_{(B_{n,\overline{T}}(1))} \\
&\quad + \underbrace{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)\frac{\sum_{j=1}^{n-1}\mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}}-X_{i\Delta_{n,\overline{T}}}|\le\varepsilon_{n,\overline{T}}\}}\frac{n}{\overline{T}}\left[\int_{j\Delta_{n,\overline{T}}}^{(j+1)\Delta_{n,\overline{T}}}2(X_s-X_{j\Delta_{n,\overline{T}}})\mu(X_s)\,ds\right]}{\sum_{j=1}^{n}\mathbf{1}_{\{|X_{j\Delta_{n,\overline{T}}}-X_{i\Delta_{n,\overline{T}}}|\le\varepsilon_{n,\overline{T}}\}}}}_{(C_{n,\overline{T}})} \\
&= A_{n,\overline{T}} + B_{n,\overline{T}}(1) + C_{n,\overline{T}}.
\end{aligned}
$$

These three terms comprise an additional bias effect, $A_{n,\overline{T}}$, a martingale effect, $B_{n,\overline{T}}(1)$, and a residual effect, $C_{n,\overline{T}}$. As we shall see, depending on the bandwidth choices, either $A_{n,\overline{T}}$ (eventually in conjunction with B) or $B_{n,\overline{T}}(1)$ may dominate the asymptotic distribution. Using embedding arguments as in the proof of Theorem 3, we can show that

$$(7.21)\qquad \sqrt{\frac{\varepsilon_{n,\overline{T}}}{\Delta_{n,\overline{T}}}}\left(\frac{B_{n,\overline{T}}(1)}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}\right) \Rightarrow MN\left(0, \frac{2\sigma^4(x)}{\overline{L}_X(\overline{T},x)}\right)$$

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

$$(7.22)\qquad \sqrt{\frac{\varepsilon_{n,\overline{T}}}{\Delta_{n,\overline{T}}}}\left(\frac{B_{n,\overline{T}}(1)}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}\right) \Rightarrow MN\left(0, \frac{2\theta_\phi\sigma^4(x)}{\overline{L}_X(\overline{T},x)}\right)$$

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

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

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

By virtue of Lemma 5, and proceeding as earlier, we find that

$$\frac{1}{\varepsilon_{n,\overline{T}}^{3/2}}\left(\frac{A_{n,\overline{T}}}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^{n}K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}\right)\Rightarrow MN\left(0,16\varphi^{ind}\frac{(\sigma'(x))^2}{\overline{L}_X(\overline{T},x)}\right), \tag{7.23}$$

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

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

Only the first and the third term can affect the asymptotic distribution of $A_{n,\overline{T}}$. Write the first term as follows:

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

As for the third term, write

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

Then,

(7.24)
$$A_{n,\overline{T}} + B^{num} \stackrel{d}{=} \phi\varepsilon_{n,\overline{T}}^{3/2} \frac{(\sigma^2(x))'}{\sigma^2(x)} \frac{1}{\sqrt{\varepsilon_{n,\overline{T}}}} \int_{-\infty}^{\infty} cK(c)(L_X(\overline{T}, x+\phi\varepsilon_{n,\overline{T}}c) - L_X(\overline{T}, x))\,dc + \varepsilon_{n,\overline{T}}^{3/2} \frac{(\sigma^2(x))'}{\sigma^2(x)} \frac{1}{\sqrt{\varepsilon_{n,\overline{T}}}} \int_{-\infty}^{\infty} \Lambda(\phi, u)(L_X(\overline{T}, x+\varepsilon_{n,\overline{T}}u) - L_X(\overline{T}, x))\,du,$$

where $\Lambda(\phi, u) = \int_{-\infty}^{\infty} K(c)\frac{1}{2}\mathbf{1}_{\{|u-\phi c|\le 1\}}(u-\phi c)\,dc$. Now, we note that

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

is

$$o_{a.s.}\left(\frac{B_{n,\overline{T}}(1)}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n\overline{T}}}\right)}\right).$$

Then, defining the overall estimation error E as

$$E \stackrel{d}{=} B + \frac{A_{n,\overline{T}}}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)} + \frac{C_{n,\overline{T}}}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)} + \frac{B_{n,\overline{T}}(1)}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)}$$

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

$$\sqrt{\frac{\varepsilon_{n,\overline{T}}}{\Delta_{n,\overline{T}}}}\left\{\frac{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)\tilde{\sigma}^2_{n,\overline{T}}(X_{i\Delta_{n,\overline{T}}})}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)} - \sigma^2(x)\right\}$$
$$\stackrel{d}{=} \sqrt{\frac{\varepsilon_{n,\overline{T}}}{\Delta_{n,\overline{T}}}}\left\{\frac{B_{n,\overline{T}}(1)}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)} + O_p(\varepsilon_{n,\overline{T}}^{3/2})\right\}$$
$$\Rightarrow MN\left(0, \frac{2\sigma^4(x)}{\overline{L}_X(\overline{T}, x)}\right),$$

from (7.21), for choices of $\varepsilon_{n,\overline{T}}$ such that $(\varepsilon^4_{n,\overline{T}}/\Delta_{n,\overline{T}}) \to 0$ and $h_{n,\overline{T}} = o(\varepsilon_{n,\overline{T}})$. If $(\varepsilon^4_{n,\overline{T}}/\Delta_{n,\overline{T}}) \to 0$ and $h_{n,\overline{T}} = O(\varepsilon_{n,\overline{T}})$ with $h_{n,\overline{T}}/\varepsilon_{n,\overline{T}} \to \phi > 0$, then

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

from (7.22), where $\theta_\phi = \frac{1}{2}\int_{-\infty}^{\infty}\int_{(z-1)\phi}^{(z+1)/\phi}\int_{(z-1)/\phi}^{(z+1)/\phi} K(a)K(e)\,dz\,da\,de$. Finally, provided $(\varepsilon^4_{n,\overline{T}}/\Delta_{n,\overline{T}}) \to \infty$ and $h_{n,\overline{T}} = o(\varepsilon_{n,\overline{T}})$, then the bias term dominates, leading to

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

from (7.23), where $\varphi^{ind} = 2\int_0^\infty \int_0^\infty (\frac{1}{2}\mathbf{1}_{\{|a|\le 1\}}a)(\frac{1}{2}\mathbf{1}_{\{|b|\le 1\}}b)\min(a,b)\,da\,db$. Under the same conditions, but when $h_{n,\overline{T}} = O(\varepsilon_{n,\overline{T}})$ with $h_{n,\overline{T}}/\varepsilon_{n,\overline{T}} \to \phi > 0$, we have

$$\frac{1}{\varepsilon_{n,\overline{T}}^{3/2}}\left\{\frac{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)\tilde{\sigma}^2_{n,\overline{T}}(X_{i\Delta_{n,\overline{T}}})}{\frac{\Delta_{n,\overline{T}}}{h_{n,\overline{T}}}\sum_{i=1}^n K\left(\frac{X_{i\Delta_{n,\overline{T}}}-x}{h_{n,\overline{T}}}\right)} - \sigma^2(x)\right\}$$
$$\Rightarrow MN\left(0, 16(\phi^2\varphi_1 + 2\phi\varphi_3 + \varphi_2)\frac{(\sigma'(x))^2}{\overline{L}_X(\overline{T},x)}\right)$$
$$\stackrel{d}{=} MN\left(0, 16\varphi^{ind,K}(\phi)\frac{(\sigma'(x))^2}{\overline{L}_X(\overline{T},x)}\right),$$

from (7.24), where

$$\varphi_1 = \int_0^\infty\int_0^\infty abK(a)K(b)\min(\phi a,\phi b)\,da\,db - \int_{-\infty}^0\int_{-\infty}^0 abK(a)K(b)\max(\phi a,\phi b)\,da\,db,$$
$$\varphi_2 = \int_0^\infty\int_0^\infty \Lambda(\phi,a)\Lambda(\phi,b)\min(a,b)\,da\,db - \int_{-\infty}^0\int_{-\infty}^0 \Lambda(\phi,a)\Lambda(\phi,b)\max(a,b)\,da\,db,$$

and

$$\varphi_3 = \int_0^\infty\int_0^\infty aK(a)\Lambda(\phi,b)\min(\phi a,b)\,da\,db - \int_{-\infty}^0\int_{-\infty}^0 aK(a)\Lambda(\phi,b)\max(\phi a,b)\,da\,db.$$

Q.E.D.

PROOF OF COROLLARY 2: Immediate after noticing that under strict stationarity (or positive recurrence)

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

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 distribution function of the process at x. *Q.E.D.*

PROOF OF COROLLARY 3: See the proof of Corollary 2. *Q.E.D.*

REFERENCES

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

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

AZEMA, J., M. KAPLAN-DUFLO, AND D. REVUZ (1967): "Mesure Invariante sur les Classes Récurrentes des Processus de Markov," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 8, 157–181.

BANDI, F. M. (2002): "Short-Term Interest Rate Dynamics: A Spatial Approach," *Journal of Financial Economics*, 65, 73–110.

BANDI, F. M., AND T. NGUYEN (2000): "Fully Nonparametric Estimators for Diffusions: a Small Sample Analysis," Unpublished Manuscript, University of Chicago.

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

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

BILLINGSLEY, P. (1968): *Convergence of Probability Measures*. New York: Wiley.

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

BRUGIÈRE, P. (1993): "Théorèm de Limite Centrale pour un Estimateur Non Paramétrique de la Variance d'un Processus de Diffusion Multidimensionnelle," *Annales de l'Institute Henri Poincaré*, 29, 357–389.

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," Unpublished Manuscript, Brown University.

HALL, P., AND C. C. HEYDE (1980): *Martingale Limit Theory and its Application*. New York: Academic Press.

HANSEN, L. P., AND T. J. SARGENT (1983): "The Dimensionality of the Aliasing Problem in Models with Rational Spectral Densities," *Econometrica*, 50, 377–387.

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

JACOD, J. (1997): "Nonparametric Kernel Estimation of the Diffusion Coefficient of a Diffusion," Prépublication N. 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 (1991): *Brownian Motion and Stochastic Calculus*. New York: Springer-Verlag.

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

KNIGHT, F. B. (1971): *A Reduction of Continuous Square Integrable Martingales to Brownian Motion*. Berlin: Springer-Verlag.

MERTON, R. C. (1980): "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics*, 8, 323–361.

ØKSENDAL, B. (1995): *Stochastic Differential Equations*. Berlin: Springer-Verlag.

PAGAN, A., AND A. ULLAH (1999): *Nonparametric Statistics*. Cambridge, U.K.: Cambridge University Press.

PHILLIPS, P. C. B. (1973): "The Problem of Identification in Finite Parameter Continuous-Time Models," *Journal of Econometrics*, 4, 351–362.

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," *Biometrika*, 72, 267–280.

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

YOR, M. (1978): "Temps Locaux," *Astérisque*, 52–53, 17–35.