Probability, Uncertainty and Quantitative Risk
Vol. 10, No. 2, 2025, 213–240

https://www.aimsciences.org/PUQR
DOI:10.3934/puqr.2025010

Nonparametric estimation of forward-backward stochastic differential equations with random terminal time

Shaolin Ji, Chenyao Yu, Linlin Zhu*

Zhongtai Securities Institute for Financial Studies, Shandong University, Jinan 250100, *China*

Email: jsl@sdu.edu.cn, yucy@mail.sdu.edu.cn, 201611343@mail.sdu.edu.cn

Abstract This paper investigates the nonparametric estimation of the functional coefficients of the forward-backward stochastic differential equations with random terminal time, focusing on both local constant and local linear estimators. We establish the asymptotic properties of these estimators under both long observation time spans and short sampling intervals, providing precise expressions for the bias and variance terms. Moreover, we propose an empirical likelihood method to construct data-driven confidence intervals for these functional coefficients. We conduct numerical simulations to examine the finite-sample properties of the estimators and to compare the performance of the empirical likelihood method with the conventional approach for constructing confidence intervals based on asymptotic normality.

Keywords Backward stochastic differential equations, Nonparametric estimation, Asymptotic normality, Empirical likelihood

2020 Mathematics Subject Classification 60J60, 62G20, 62M05

1. Introduction

The nonlinear backward stochastic differential equation (BSDE) was introduced by Pardoux and Peng [12]. They showed that when BSDEs are coupled with forward stochastic differential equations (SDEs), these forward-backward stochastic differential equations (FBSDEs) provide a probabilistic interpretation for systems of quasilinear parabolic and elliptic partial differential equations (PDEs). This generalized the classical Feynman–Kac formula for linear parabolic and elliptic PDEs, as discussed in Peng [14]. Subsequently, many researchers developed theories for FBSDEs and their applications. In mathematical finance, FBSDEs are used in the theory of stochastic differential utility and the evaluation of contingent claims for large investors. The existence and uniqueness of solutions to certain types of FBSDEs are closely linked to optimal stochastic control problems. In many applications, there is a lack of prior information about model structure. Therefore, it becomes important to identify and estimate, using discretely observed data, the parameters and functionals of the process. Over the past few decades, there

Received 22 April 2024; Accepted 21 January 2025; Early access 14 April 2025
*Corresponding author

have been extensive studies of the nonparametric estimation of the (jump-) diffusion processes and stochastic volatility models (see [1–4, 7–9, 16, 17, 19–21]). However, although it is an important tool for describing stochastic processes in various fields, research pertaining to the statistical inference of BSDEs is relatively scarce. In this study, we consider the nonparametric estimation of FBSDEs with random terminal time.

Consider the following FBSDE:

$$\begin{cases} X_t = x + \int_0^t b(X_s)\mathrm{d}s + \int_0^t \sigma(X_s)\mathrm{d}W_s, \\ Y_t = g(X_\tau) + \int_{t\wedge\tau}^{\tau} f(X_s, Y_s, Z_s)\mathrm{d}s - \int_{t\wedge\tau}^{\tau} Z_s \mathrm{d}W_s, \end{cases} \tag{1.1}$$

where $\{W_t\}_{t\geqslant 0}$ is a d-dimensional standard Brownian motion defined on a complete filtered probability space $(\Omega, \mathscr{F}, \mathbb{P})$, and $\{\mathscr{F}_t\}_{t\geqslant 0}$ is the natural filtration of this Brownian motion, where $\mathscr{F}_0$ contains all the $\mathbb{P}$-null elements of $\mathscr{F}$. The function g is continuous, and τ is the $\mathscr{F}_t$ stopping time with values in $[0,\infty]$. The processes X, Y, and Z take values in $\mathbb{R}^d$, $\mathbb{R}^n$, and $\mathbb{R}^{n\times d}$, respectively. The functional coefficients b, σ, and f are assumed to be unknown, and we observe the processes $\{X_t\}_{t\geqslant 0}$ and $\{Y_t\}_{t\geqslant 0}$ at discrete times. In this study, we are interested in developing nonparametric estimators for the generator f and the diffusion term Z of the FBSDE.

Studies on statistical inference of BSDEs have been rarely reported. Kutoyants [10, 11] studied the approximation of the solutions to several types of FBSDEs with unknown parameters in forward SDEs. These unknown parameters could be estimated from continuous or discrete observations of the SDE. By substituting these estimators into the solution to the corresponding PDE, an estimate of the solution to the BSDE can be obtained. For the nonparametric estimation of BSDEs, Su and Lin [18] proposed a local constant estimator for the diffusion term in a linear FBSDE with a fixed terminal time. They also established a least squares estimator for the unknown parameter in the generator. Chen and Lin [6] considered a coupled Markovian FBSDE with fixed terminal time, and they constructed local linear estimators for the generator and diffusion term and derived the asymptotic properties of these estimators.

Some of the assumptions in these studies are quite strict, such as the stationary assumptions on the process X, Y and their increments. Moreover, their estimators are based on the relationship between FBSDEs with deterministic terminal time and quasilinear parabolic PDEs, where the solutions (Y_t, Z_t) and the functional coefficients of the FBSDEs are deterministic functions of X_t and t. However, this framework provides insufficient information for estimating bivariate functions when only a single trajectory of the process is observed. BSDEs that have random terminal time are reportedly connected with quasilinear elliptic PDEs, where the solutions to these BSDEs are deterministic functions of X_t. That is, the solutions depend on the current state of X_t, rather than on time. Applying this property, we construct the local constant and local linear estimators for the generator and diffusion term in FBSDEs with random terminal time. The asymptotic properties of these estimators are derived under mild conditions (the process X and Y need not be stationary). Furthermore, the local polynomial estimator and the double-smoothing estimator can be easily generalized.

Common confidence intervals for these estimators can be obtained based on their asymptotic normality, provided that consistent estimators of the asymptotic variances are available. However, this type of confidence interval is always symmetric, and the estimated variances usually have large biases. We apply the empirical likelihood method in conjunction with the local constant estimators to construct data-driven confidence intervals that can account for possible skewness of the estimators and prevent imprecise variance estimation. We focus on local constant smoothing, and the extensions to the local linear or local polynomial estimators are straightforward.

This paper is organized as follows. The models and some preliminaries are introduced in Section 2. The estimators and their asymptotic properties are presented in Section 3. Section 4 introduces the empirical likelihood method for constructing point-wise confidence intervals. The proofs are given in Section 5, and some numerical examples are presented in Section 6.

2. Model setup and preliminaries

Denote $L^{2,k}(0,\tau;\mathbb{R}^n)$, where $k\in\mathbb{R}$, as the space of $\mathbb{R}^n$-valued progressively measurable processes f such that

$$\|f(\cdot)\|_{L^{2,k}} := \left(\mathbb{E}\int_0^\tau |f(\cdot)|^2 e^{kt}\mathrm{d}t\right)^{1/2} < \infty,$$

and $L^2(0,\tau;\mathbb{R}^n) := L^{2,0}(0,\tau;\mathbb{R}^n)$. Further, define

$$L^{2,k}(0,\tau) := L^{2,k}(0,\tau;\mathbb{R}^n)\times L^{2,k}(0,\tau;\mathbb{R}^{n\times d}),\quad \mathscr{R} := \mathbb{R}^d\times\mathbb{R}^n\times\mathbb{R}^{n\times d}.$$

Let $\{X_t^x; t\geqslant 0\}$ denote the solution to the following forward SDE:

$$X_t^x = x + \int_0^t b\left(X_s^x\right)\mathrm{d}s + \int_0^t \sigma\left(X_s^x\right)\mathrm{d}W_s,\quad t\geqslant 0, \tag{2.1}$$

where $x\in\mathbb{R}^d$, $b:\mathbb{R}^d\to\mathbb{R}^d$, and $\sigma:\mathbb{R}^d\to\mathbb{R}^{d\times d}$ are globally Lipschitz and twice continuously differentiable functions.

2.1 BSDEs with infinite horizon

Consider the infinite-horizon BSDE, which is defined as follows:

$$Y_t^x = Y_T^x + \int_t^T f\left(X_s^x, Y_s^x, Z_s^x\right)\mathrm{d}s - \int_t^T Z_s^x\mathrm{d}W_s,\quad \forall\, t,T\ s.t.\ 0\leqslant t\leqslant T, \tag{2.2}$$

where $f:\mathscr{R}\to\mathbb{R}^n$ is a continuous function, and the process $\{X_t^x; t\geqslant 0\}$ is defined as (2.1).

Assumption 2.1 *There exists some constants $C_1>0$, $C_2>0$, $\mu<0$, $p>0$ such that*

$$|f(x,y,z)|\leqslant C_1\left(1+|x|^p+|y|+\|z\|\right),$$

$$\langle f(x,y_1,z)-f(x,y_2,z), y_1-y_2\rangle \leqslant \mu|y_1-y_2|^2,$$

$$|f(x,y,z_1)-f(x,y,z_2)|\leqslant C_2\|z_1-z_2\|.$$

Moreover, for some $\lambda>2\mu+C_2^2$ and for all $x\in\mathbb{R}^d$,

$$\mathbb{E}\int_0^\infty |f(X_t^x,0,0)|^2 e^{\lambda t}\mathrm{d}t<\infty.$$

According to Theorem 3.1 in Pardoux [13], under Assumption 2.1, BSDE (2.2) has a unique solution $\{(Y_t^x, Z_t^x); t\geqslant 0\}$ that belongs to $L^{2,\lambda}(0,\infty)$. Consider a semilinear elliptic PDE in $\mathbb{R}^d$ with the form

$$\mathscr{L}u_i(x) + f_i\left(x, u(x), (\nabla u\sigma)(x)\right) = 0,\quad x\in\mathbb{R}^d,\quad 0\leqslant i\leqslant n, \tag{2.3}$$

where

$$\mathscr{L} = \sum_i^d b_i(x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j}^d (\sigma\sigma^*)_{ij}(x)\frac{\partial^2}{\partial x_i\partial x_j}$$

is the infinitesimal generator of the Markov process $\{X_t^x; t \geqslant 0\}$.

Lemma 2.2 *([13], Theorem 4.1) Let Assumption 2.1 hold. If $u \in C^2(\mathbb{R}^d;\mathbb{R}^n)$ is a classical solution to PDE (2.3) such that*

$$\mathbb{E}\left[\int_0^\infty e^{\lambda t}\|(\nabla u\sigma)(X_t^x)\|^2 \mathrm{d}t\right] < \infty, \quad x \in \mathbb{R}^d,$$

then for each $x \in \mathbb{R}^d$, $\{(u(X_t^x), (\nabla u\sigma)(X_t^x)); t > 0\}$ is the unique solution to BSDE (2.2).

2.2 BSDEs with random terminal time

Let G be an open bounded subset of $\mathbb{R}^d$ with a boundary of class C^1. For each $x \in \bar{G}$, define the stopping time $\tau_x = \inf\{t \geqslant 0 : X_t^x \notin \bar{G}\}$. Assume that $\mathbb{P}(\tau_x < \infty) = 1$, and for all $x \in \bar{G}$, the set $\Gamma = \{x \in \partial G;\ \mathbb{P}(\tau_x > 0) = 0\}$ is closed. Consider the following BSDE:

$$Y_t^x = g\left(X_{\tau_x}^x\right) + \int_{t\wedge\tau_x}^{\tau_x} f\left(X_s^x, Y_s^x, Z_s^x\right) \mathrm{d}s - \int_{t\wedge\tau_x}^{\tau_x} Z_s^x \mathrm{d}W_s, \tag{2.4}$$

where $g \in C(\mathbb{R}^d)$, f is continuous and satisfies Assumption 2.1. Additionally, assume that for some $\lambda > 2\mu + C_2^2$, $\mathbb{E}\left[\exp\left(\lambda\tau_x\right)\right] < \infty$ for all $x \in \bar{G}$. Then, BSDE (2.4) has a unique solution $\{(Y_t^x, Z_t^x); t \geqslant 0\}$ in $L^{2,\lambda}(0, \tau_x)$. Next, consider the PDE with the Dirichlet boundary condition

$$\begin{cases} \mathscr{L}u_i(x) + f_i(x, u(x), (\nabla u\sigma)(x)) = 0, \quad 0 \leqslant i \leqslant n, \quad x \in G, \\ u_i(x) = g_i(x), \quad 0 \leqslant i \leqslant n, \quad x \in \partial G. \end{cases} \tag{2.5}$$

Lemma 2.3 *([13]) Let Assumption 2.1 hold. If PDE (2.5) has a classical solution $u \in C^2(\bar{G};\mathbb{R}^n)$, then $\{(u(X_t^x), (\nabla u\sigma)(X_t^x)); 0 \leqslant t \leqslant \tau_x\}$ is the unique solution to BSDE (2.4).*

Remark 2.4 *The conditions that guarantee a unique solution $u \in C^2$ to elliptic PDE (2.5) can be found in Theorem 4.1 in Peng [14].*

3. Nonparametric estimation of the coefficients

In this section, we construct the nonparametric estimators for the generator and the diffusion term in models (2.2) and (2.4). We consider the one-dimensional case for simplicity; similar results can be generalized to the multi-dimensional case under a more complex proof. Let $\mathscr{D} = (\bar{l}, \bar{r})$ denote the range of the process X and Y in the two models, where $-\infty \leqslant \bar{l} \leqslant \bar{r} \leqslant \infty$. Assume the processes X and Y in (2.2) and (2.4) are observed at $\{t = t_0, t_1, t_2, \ldots, t_n\}$ within the time interval $[0, T]$ with $t_0 = 0$ and $t_n = T$, where the observations are equispaced. The observations are given by $\left\{X_0, X_\Delta, \ldots, X_{(n-1)\Delta}, X_{n\Delta}\right\}$ and $\left\{Y_0, Y_\Delta, \ldots, Y_{(n-1)\Delta}, Y_{n\Delta}\right\}$, where $\Delta = T/n$ is the observation interval.

In the following, assume that PDEs (2.3) and (2.5) have unique solutions and that Assumption 2.1 holds. Under these assumptions, the solutions to the above FBSDEs (Y_t^x, Z_t^x) could be represented as deterministic functions of X_t^x. Accordingly, we denote $-f(X_t^x, Y_t^x, Z_t^x)$ and Z_t^x as $\tilde{f}(X_t)$ and $\tilde{Z}(X_t)$, respectively. Finally, for the diffusion process X that satisfies SDE (2.1) and a smooth function $\tilde{g} \in C^4$, the conditional expected increment can be expressed as

$$\mathbb{E}\left[\tilde{g}(X_{(i+1)\Delta} - \tilde{g}(X_{i\Delta}) \mid X_{i\Delta}\right] = \mathscr{L}\tilde{g}(X_{i\Delta})\Delta + \widetilde{R}, \tag{3.1}$$

where

$$\widetilde{R} = \int_{i\Delta}^{(i+1)\Delta} \int_{i\Delta}^{t} \mathbb{E}\left[\mathscr{L}^2\tilde{g}(X_s) \mid X_{i\Delta}\right] \mathrm{d}s\mathrm{d}t.$$

If $\mathbb{E}\left[|\mathscr{L}^2\tilde{g}(X_t)|\right]<\infty$, where $t\geqslant 0$, then $\widetilde{R}=O_{\mathbb{P}}(\Delta^2)$. Letting

$$\tilde{g}(x)=u(x)-u(X_{i\Delta})\quad\text{and}\quad (u(x)-u(X_{i\Delta}))^2,$$

we have

$$\mathscr{L}\tilde{g}(X_{i\Delta})=(u'b+u''\sigma^2/2)(X_{i\Delta})\quad\text{and}\quad (u'\sigma)^2(X_{i\Delta}).$$

If $\mathbb{E}\left[\left|\mathscr{L}^2u(X_t)\right|\right]<\infty, \mathbb{E}\left[\left|\mathscr{L}^2u^2(X_t)\right|\right]<\infty, \mathbb{E}\left[|u(X_{t-s})\mathscr{L}^2u(X_t)|\right]<\infty$, where $t>0,\, s>0$, then

$$\tilde{f}(X_{i\Delta})=\lim_{\Delta\to 0}\frac{1}{\Delta}\mathbb{E}\left[(Y_{(i+1)\Delta}-Y_{i\Delta})|X_{i\Delta}\right] \tag{3.2}$$

and

$$\tilde{Z}^2(X_{i\Delta})=\lim_{\Delta\to 0}\frac{1}{\Delta}\mathbb{E}\left[(Y_{(i+1)\Delta}-Y_{i\Delta})^2|X_{i\Delta}\right]. \tag{3.3}$$

According to the infinitesimal moment conditions (3.2) and (3.3), common nonparametric regression methods, such as local constant regression and local polynomial regression, can be employed to estimate the generator $\tilde{f}$ and diffusion term $\tilde{Z}^2$ at each spatial point x. In this section, we focus on local constant and local linear estimators, as the statistical properties of other estimators can be obtained by simple generalization of these two methods.

3.1 Nonparametric kernel estimation of the generator

3.1.1 Local constant estimator

The local constant estimator $\tilde{f}(x)$ for the BSDEs (2.2) and (2.4) is defined as follows:

$$\widehat{f}_{\mathrm{NW}}(x):=\frac{\frac{1}{h}\sum_{i=0}^{n-1}K\left(\frac{X_{i\Delta}-x}{h}\right)\left(Y_{(i+1)\Delta}-Y_{i\Delta}\right)}{\frac{\Delta}{h}\sum_{i=0}^{n-1}K\left(\frac{X_{i\Delta}-x}{h}\right)}:=\frac{N(K,\tilde{f})}{D(K)}, \tag{3.4}$$

where K is the kernel function whose property is specified below, and h is the bandwidth. To evaluate the property of the estimator, we decompose $\widehat{f}_{\mathrm{NW}}(x)$ as follows:

$$\widehat{f}_{\mathrm{NW}}(x)=\hat{g}_{p,\mathrm{NW}}(x)+\hat{g}_{q,\mathrm{NW}}(x)+\hat{g}_{r,\mathrm{NW}}(x),$$

where

$$\hat{g}_{p,\mathrm{NW}}(x)=\tilde{f}(x)+\frac{B(K,\tilde{f})}{D(K)},\quad \hat{g}_{q,\mathrm{NW}}(x)=\frac{M(K,\tilde{f})}{D(K)},\quad \hat{g}_{r,\mathrm{NW}}(x)=\frac{R(K,\tilde{f})}{D(K)},$$

with

$$B(K,\tilde{f})=\frac{\Delta}{h}\sum_{i=0}^{n-1}K\left(\frac{X_{i\Delta}-x}{h}\right)\left(\tilde{f}\left(X_{i\Delta}\right)-\tilde{f}(x)\right),$$
$$M(K,\tilde{f})=\frac{1}{h}\sum_{i=0}^{n-1}K\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}Z_s\mathrm{d}W_s,$$
$$R(K,\tilde{f})=\frac{1}{h}\sum_{i=0}^{n-1}K\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}\left(\tilde{f}(X_s)-\tilde{f}(X_{i\Delta})\right)\mathrm{d}s.$$

For the asymptotic property of the estimator, we make the following assumptions:

Assumption 3.1 *The kernel function $K(\cdot)$ is bounded, twice continuously differentiable, has support $[-1,1]$, and satisfies $\int_{-\infty}^{\infty}K(u)\mathrm{d}u=1$, $\int_{-\infty}^{\infty}uK(u)\mathrm{d}u=0$. Moreover, define $l(K_1)=\int_{-\infty}^{\infty}u^2K(u)\mathrm{d}u$, $l(K_2)=\int_{-\infty}^{\infty}K^2(u)\mathrm{d}u$.*

Remark 3.2 *These conditions can be satisfied by many kernel functions, including the Epanechnikov kernel* $K(u) = \frac{3}{4}(1-u^2)\mathbb{I}_{\{|u|\leqslant 1\}}$ *and the quartic kernel* $K(u) = \frac{15}{16}(1-u^2)^2\mathbb{I}_{\{|u|\leqslant 1\}}$, *where* $\mathbb{I}$ *denotes the indicator function.*

Assumption 3.3 *The process* X *is recurrent, i.e., the scale function of* X, *which is*

$$S(\alpha) = \int_c^\alpha exp\left(\int_c^y \frac{-2b(x)}{\sigma^2(x)}\mathrm{d}x\right)\mathrm{d}y,$$

where c *is a generic fixed number belonging to* $\mathscr{D}$, *satisfies*

$$\lim_{\alpha\to\bar{l}} S(\alpha) = -\infty, \quad \lim_{\alpha\to\bar{r}} S(\alpha) = \infty.$$

Remark 3.4 *Intuitively, recurrence ensures that the Markov process* X *can visit every spatial point* $x \in \mathscr{D}$ *infinite times with probability one when* $T \to \infty$. *The recurrence-related concepts and their applications for statistical inference of stochastic processes can be found in [1, 2]. This condition does not imply the existence of a time-invariant distribution for* X, *therefore, nonstationarity is allowed.*

To illustrate the asymptotic properties of the estimators, we use the local time $\ell(T,x)$ of the Markov process X defined in Aït-Sahalia and Park [1]. It measures the amount of calendar time spent by the process in the neighborhood of x. For a recurrent diffusion process, $\ell(T,x)$ diverges as $T \to \infty$ at every spatial point x. In this paper, the local time $\ell(T,x)$ of the Markov process X in FBSDEs (2.2) and (2.4) must satisfy the following assumptions.

Assumption 3.5 (i) *There exists* $\varepsilon > 0$ *such that* $\bar{\ell}_\varepsilon(T,x) = O_{\mathbb{P}}\left(\ell(T,x)^2\right)$, *where* $\bar{\ell}_\varepsilon(T,x) = \sup_{|x-y|\leqslant\varepsilon} \ell(T,y)$.

(ii) $\ell(T,x) = O_{\mathbb{P}}(\alpha_T)$ *for some nonrandom sequence* (α_T).

Remark 3.6 *In Assumption 3.5, condition (i) primarily controls the rate of the local time in the neighborhood of a spatial point. It is trivially satisfied when* T *is fixed. As* $T \to \infty$, *it is satisfied for Brownian motion, for which we have* $\frac{\sup_{x\in\mathbb{R}}\ell(T,x)}{T^{1/2}\log\log T} = O_{\mathbb{P}}(1)$. *Condition (ii) applies more generally to diffusion processes and is automatically satisfied for fixed* T. *For* $T \to \infty$, *the asymptotic behavior of* $\ell(T,x)$ *depends on the recurrence properties of the process* X. *For example, in the case of a positive recurrent process,* $\ell(T,x) = O_{\mathbb{P}}(T)$ *as* $T \to \infty$. *Specifically, for Brownian motion,* $\ell(T,x) = O_{\mathbb{P}}(\sqrt{T})$.

Assumption 3.7 *Given* $\Delta \to 0$ *and* $h \to 0$ *such that*

(i) $h^{-4}\Delta \to 0$,

(ii) $\sqrt{\Delta}\mathscr{M}(\mathscr{L}\tilde{Z}^2) = o_{\mathbb{P}}(1)$, $\sqrt{\Delta}\mathscr{M}((\tilde{Z}^2)'\sigma) = o_{\mathbb{P}}(\sqrt{h\ell(T,x)})$,

(iii) $\sqrt{\Delta}\mathscr{M}(\mathscr{L}\tilde{f}) = o_{\mathbb{P}}(1)$, $\sqrt{\Delta}\mathscr{M}(\tilde{f}'\sigma) = o_{\mathbb{P}}(\sqrt{h\ell(T,x)})$,

where $\mathscr{M}(H) := \sup_{s\leqslant T}|H(X_s)|$, *and* $H(X_s)$ *is a function of the process* X.

Remark 3.8 *For several processes that are commonly used in economics, the exact order of the extremal process is known. For example, the extremal processes of the Ornstein–Uhlenbeck process and the Cox–Ingersoll–Ross process are of orders* $O_{\mathbb{P}}(\sqrt{\log T})$ *and* $O_{\mathbb{P}}(\log T)$, *respectively. The divergence rate of the extremal process for a class of continuous-time diffusion processes is discussed in [4, 5].*

Theorem 3.9 *Let Assumptions 2.1, 3.1, 3.3, 3.5, and 3.7 hold. If PDEs (2.3) and (2.5) have a unique solution* $u \in C^4$, *and*

$$\mathbb{E}\left[|\mathscr{L}^2 u(X_t)|\right] < \infty \text{ for all } t \geqslant 0.$$

Then we have

$$\hat{g}_{p,\mathrm{NW}}(x) = \tilde{f}(x) + h^2 l(K_1)\left(\frac{\tilde{f}''(x)}{2} + \tilde{f}'(x)\frac{s'(x)}{s(x)}\right) + o_{\mathbb{P}}(h^2) + O_{\mathbb{P}}(h^{3/2}\ell(T,x)^{-1/2})$$

uniformly in T *as* $h \to 0$ *and* $\Delta \to 0$, *where* $s(x)$ *is the speed function of the process* X *at the generic lever* x, *given by* $s(x) = 2/S^{'}(x)\sigma^2(x)$, *and*

$$\sqrt{h\ell(T,x)}\hat{g}_{q,\mathrm{NW}}(x) \xrightarrow{d} \sqrt{l(K_2)}\tilde{Z}(x)\mathbf{N},$$

where $\mathbf{N}$ *is a standard normal random variable. Furthermore,* $\hat{g}_{r,\mathrm{NW}}(x) = o_{\mathbb{P}}(h^2)$.

Remark 3.10 *Note that* $\hat{g}_{p,\mathrm{NW}}(x)$ *and* $\hat{g}_{q,\mathrm{NW}}(x)$ *represent the bias and variance terms of the estimator* $\widehat{f}_{\mathrm{NW}}(x)$, *respectively.* $\hat{g}_{r,\mathrm{NW}}(x)$ *is a negligible bias. If* $h\ell(T,x) \xrightarrow{\mathbb{P}} \infty$, $h^5\ell(T,x) = o_{a.s.}(1)$, *we have*

$$\sqrt{h\ell(T,x)}\widehat{f}_{\mathrm{NW}}(x) \xrightarrow{d} \sqrt{l(K_2)}\tilde{Z}(x)\mathbf{N}.$$

3.1.2 Local linear estimator

The local linear estimator of $\tilde{f}(x)$ for (2.2) and (2.4) is given by

$$\widehat{f}_{\mathrm{LL}}(x) := \frac{N(K,\tilde{f})D(K_2) - N(K_1,\tilde{f})D(K_1)}{D(K)D(K_2) - (D(K_1))^2}, \tag{3.5}$$

where

$$N(K_1,\tilde{f}) := \frac{1}{h}\sum_{i=0}^{n-1}\left(\frac{X_{i\Delta}-x}{h}\right)K\left(\frac{X_{i\Delta}-x}{h}\right)\left(Y_{(i+1)\Delta} - Y_{i\Delta}\right),$$

$$D(K_1) := \frac{\Delta}{h}\sum_{i=0}^{n-1}\left(\frac{X_{i\Delta}-x}{h}\right)K\left(\frac{X_{i\Delta}-x}{h}\right),$$

$$D(K_2) := \frac{\Delta}{h}\sum_{i=0}^{n-1}\left(\frac{X_{i\Delta}-x}{h}\right)^2 K\left(\frac{X_{i\Delta}-x}{h}\right).$$

$\widehat{f}_{\mathrm{LL}}(x)$ can be decomposed as

$$\widehat{f}_{\mathrm{LL}}(x) = \hat{g}_{p,\mathrm{LL}}(x) + \hat{g}_{q,\mathrm{LL}}(x) + \hat{g}_{r,\mathrm{LL}}(x),$$

where

$$\hat{g}_{p,\mathrm{LL}}(x) = \tilde{f}(x) + \frac{B(K,\tilde{f})D(K_2) - B(K_1,\tilde{f})D(K_1)}{D(K)D(K_2) - (D(K_1))^2},$$

$$\hat{g}_{q,\mathrm{LL}}(x) = \frac{M(K,\tilde{f})D(K_2) - M(K_1,\tilde{f})D(K_1)}{D(K)D(K_2) - (D(K_1))^2},$$

$$\hat{g}_{r,\mathrm{LL}}(x) = \frac{R(K,\tilde{f})D(K_2) - R(K_1,\tilde{f})D(K_1)}{D(K)D(K_2) - (D(K_1))^2},$$

similar to the definition of $N(K_1,\tilde{f})$ and $N(K,\tilde{f})$, $B(K_1,\tilde{f})$ is defined as $B(K,\tilde{f})$, $M(K_1,\tilde{f})$ as $M(K,\tilde{f})$, $R(K_1,\tilde{f})$ as $R(K,\tilde{f})$, except that K is replaced by K_1, $K_1(x) = xK(x)$.

Theorem 3.11 *Let Assumptions 2.1, 3.1, 3.3, 3.5, and 3.7 hold. If PDEs (2.3) and (2.5) have a unique solution $u \in C^4$, and $\mathbb{E}\left[|\mathscr{L}^2 u(X_t)|\right] < \infty$ for all $t \geqslant 0$, Then, we have*

$$\hat{g}_{p,\mathrm{LL}}(x) = \tilde{f}(x) + h^2 l(K_1)\frac{\tilde{f}''(x)}{2} + o_{\mathbb{P}}(h^2) + O_{\mathbb{P}}(h^{3/2}\ell(T,x)^{-1/2})$$

uniformly in T as $h \to 0$ and $\Delta \to 0$, and

$$\sqrt{h\ell(T,x)}\hat{g}_{q,\mathrm{LL}}(x) \xrightarrow{d} \sqrt{l(K_2)}\tilde{Z}(x)\mathbf{N}.$$

Furthermore, $\hat{g}_{r,\mathrm{LL}}(x) = o_{\mathbb{P}}(h^2)$.

Remark 3.12 *Note that $\widehat{f}_{\mathrm{NW}}(x)$ and $\widehat{f}_{\mathrm{LL}}(x)$ are consistent only if $h\ell(T,x) \xrightarrow{\mathbb{P}} \infty$. For a recurrent process, we have $\ell(T,x) \xrightarrow{\mathbb{P}} \infty$ as $T \to \infty$ for each $x \in \mathscr{D}$, whereas $\ell(T,x)$ does not diverge for a transient process. Therefore, for recurrent processes X, the local constant and local linear estimators of the generator for the infinite-horizon BSDE (2.2) are consistent, but this does not hold for the BSDE with random terminal time (2.4).*

Similar to the local constant estimator, $\hat{g}_{p,\mathrm{LL}}(x)$ represents the main bias term of the local linear estimator. The local linear estimator can effectively reduce the bias, as its bias term depends on the second-order derivative of the true function, rather than its first-order derivative and the speed function of the process X. Additionally, $\hat{g}_{q,\mathrm{LL}}(x)$ corresponds to the variance term of the estimator.

For both the local constant and local linear estimators of the generator, the optimal bandwidths that minimize their mean squared errors are, respectively,

$$\begin{aligned} h^*_{f,\mathrm{NW}} &= c(K)\tilde{Z}^{2/5}(x)\left(\tilde{f}''(x) + 2\tilde{f}'(x)\frac{s'(x)}{s(x)}\right)^{-2/5}\ell(T,x)^{-1/5}, \\ h^*_{f,\mathrm{LL}} &= c(K)\tilde{Z}^{2/5}(x)\tilde{f}''(x)^{-2/5}\ell(T,x)^{-1/5}, \end{aligned}$$

where $c(K) = \frac{l(K_2)^{1/5}}{l(K_1)^{2/5}}$.

Multivariate regression analysis, machine learning techniques, and other statistical methods can be employed to analyze the functional form of the generator, based on the estimators of the generator and diffusion term, along with discrete observations of X and Y.

3.2 Nonparametric kernel estimation of the diffusion term

3.2.1 Local constant estimator

The local constant estimator for $\tilde{Z}^2(x)$ is given by

$$\begin{aligned} \widehat{Z}^2_{\mathrm{NW}}(x) &:= \frac{\frac{1}{h}\sum_{i=0}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right)\left(Y_{(i+1)\Delta} - Y_{i\Delta}\right)^2}{\frac{\Delta}{h}\sum_{i=0}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right)} := \frac{N(K,Z)}{D(K)}, \\ \widehat{Z}^2_{\mathrm{NW}}(x) &= \hat{Z}^2_{p,\mathrm{NW}}(x) + \hat{Z}^2_{q,\mathrm{NW}}(x) + \hat{Z}^2_{r,\mathrm{NW}}(x), \end{aligned} \tag{3.6}$$

where

$$\hat{Z}^2_{p,\mathrm{NW}}(x) = \tilde{Z}^2(x) + \frac{B(K,Z)}{D(K)}, \quad \hat{Z}^2_{q,\mathrm{NW}}(x) = \frac{M(K,Z)}{D(K)}, \quad \text{and } \hat{Z}^2_{r,\mathrm{NW}}(x) = \frac{R(K,Z)}{D(K)} + \frac{S(K)}{D(K)},$$

with

$$
\begin{aligned}
B(K,Z) &= \frac{\Delta}{h}\sum_{i=0}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right)\left(\tilde{Z}^2(X_{i\Delta})-\tilde{Z}^2(x)\right),\\
M(K,Z) &= \frac{1}{h}\sum_{i=0}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta} 2(Y_t-Y_{i\Delta})Z_t\mathrm{d}W_t,\\
R(K,Z) &= \frac{1}{h}\sum_{i=0}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta} \left(\tilde{Z}^2(X_t)-\tilde{Z}^2(X_{i\Delta})\right)\mathrm{d}t,\\
S(K) &= \frac{1}{h}\sum_{i=0}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta} 2(Y_t-Y_{i\Delta})\tilde{f}(X_t)\mathrm{d}t.
\end{aligned}
$$

Assumption 3.13 *Given $\Delta\to 0$ and $h\to 0$ such that*
(i) $h^{-4}\Delta\to 0$,
(ii) $\Delta\mathscr{M}(b)\mathscr{M}((\tilde{Z}^2)')=o_{\mathbb{P}}(1)$, $\sqrt{\Delta}\mathscr{M}(\sigma)\mathscr{M}((\tilde{Z}^2)')=o_{\mathbb{P}}(1)$,
(iii) $\sqrt{\Delta}\mathscr{M}(\mathscr{L}H_1)=o_{\mathbb{P}}(1)$, $\sqrt{\Delta}\mathscr{M}((H_1)'\sigma)=o_{\mathbb{P}}(\sqrt{h\ell(T,x)})$, $H_1=u\tilde{f},\ \tilde{f},\ u\tilde{Z},\ \tilde{Z}$.

Theorem 3.14 *Let Assumptions 2.1, 3.1, 3.3, 3.5, and 3.13 hold. If PDEs (2.3) and (2.5) have a unique solution $u\in C^4$, and $\mathbb{E}\left[|\mathscr{L}^2u^2(X_t)|\right]<\infty$, $\mathbb{E}\left[|u(X_{t-s})\mathscr{L}^2u(X_t)|\right]<\infty$, where $t>0$, $s>0$; then we have*

$$
\hat{Z}^2_{p,\mathrm{NW}}(x)=\tilde{Z}^2(x)+h^2l(K_1)\left(\frac{(\tilde{Z}^2)''(x)}{2}+(\tilde{Z}^2)'(x)\frac{s'(x)}{s(x)}\right)+o_{\mathbb{P}}(h^2)+O_{\mathbb{P}}(h^{3/2}\ell(T,x)^{-1/2}) \quad (3.7)
$$

uniformly in T as $\Delta\to 0$ and $h\to 0$, and

$$
\left(\frac{h\ell(T,x)}{\Delta}\right)^{1/2}\hat{Z}^2_{q,\mathrm{NW}}(x)\xrightarrow{d}\sqrt{2l(K_2)}\tilde{Z}^2(x)\mathbf{N}.
$$

Moreover, $\hat{Z}^2_{r,\mathrm{NW}}(x)=o_{\mathbb{P}}(h^2)$.

3.2.2 Local linear estimator

The local linear estimator of $\tilde{Z}(x)$ for (2.2) and (2.4) is defined as follows:

$$
\widehat{Z}^2_{\mathrm{LL}}(x):=\frac{N(K,Z)D(K_2)-N(K_1,Z)D(K_1)}{D(K)D(K_2)-D(K_1)^2}, \quad (3.8)
$$

where $N(K_1,Z)$ is defined as $N(K,Z)$ with K replaced by K_1.

$$
\widehat{Z}^2_{\mathrm{LL}}(x)=\hat{Z}^2_{p,\mathrm{LL}}(x)+\hat{Z}^2_{q,\mathrm{LL}}(x)+\hat{Z}^2_{r,\mathrm{LL}}(x),
$$

where

$$
\begin{aligned}
\hat{Z}^2_{p,\mathrm{LL}}(x) &= \tilde{Z}^2(x)+\frac{B(K,Z)D(K_2)-B(K_1,Z)D(K_1)}{D(K)D(K_2)-D(K_1)^2},\\
\hat{Z}^2_{q,\mathrm{LL}}(x) &= \frac{M(K,Z)D(K_2)-M(K_1,Z)D(K_1)}{D(K)D(K_2)-D(K_1)^2},\\
\hat{Z}^2_{r,\mathrm{LL}}(x) &= \frac{(R(K,Z)+S(K))D(K_2)-(R(K_1,Z)+S(K_1))D(K_1)}{D(K)D(K_2)-D(K_1)^2},
\end{aligned}
$$

where $B(K_1,Z)$ is defined as $B(K,Z)$, and similarly, $M(K_1,Z)$, $R(K_1,Z)$ and $S(K_1)$ are defined as $M(K,Z)$, $R(K,Z)$, and $S(K)$ respectively, with K replaced by K_1.

Theorem 3.15 *Let Assumptions 2.1, 3.1, 3.3, 3.5, and 3.13 hold. If PDEs (2.3) and (2.5) have a unique solution* $u \in C^4$, *and* $\mathbb{E}\left[|\mathscr{L}^2 u^2(X_t)|\right] < \infty$, $\mathbb{E}\left[|u(X_{t-s})\mathscr{L}^2 u(X_t)|\right] < \infty$, *where* $t > 0$, $s > 0$, *then we have*

$$\hat{Z}^2_{p,\mathrm{LL}}(x) = \tilde{Z}^2(x) + h^2 l(K_1)\frac{(\tilde{Z}^2)''(x)}{2} + o_{\mathbb{P}}(h^2) + O_{\mathbb{P}}(h^{3/2}\ell(T,x)^{-1/2}) \tag{3.9}$$

uniformly in T *as* $\Delta \to 0$ *and* $h \to 0$, *and*

$$\left(\frac{h\ell(T,x)}{\Delta}\right)^{1/2} \hat{Z}^2_{\mathrm{LL}}(x) \xrightarrow{d} \sqrt{2l(K_2)}\tilde{Z}^2(x)\mathbf{N}.$$

Moreover, $\hat{Z}^2_{r,\mathrm{LL}}(x) = o_{\mathbb{P}}(h^2)$.

Remark 3.16 *To ensure the consistency of* $\widehat{Z}^2_{\mathrm{NW}}(x)$ *and* $\widehat{Z}^2_{\mathrm{LL}}(x)$, *it is sufficient to have* $\Delta \to 0$ *and* $h \to 0$; *there is no need for* $\ell(T,x) \xrightarrow{\mathbb{P}} \infty$ *or* $T \to \infty$. *In other words, these estimators can be consistent as long as* Δ *tends to* 0 *sufficiently fast relative to* h, *and the process* X *need not be recurrent. Therefore, both estimators can be applied to the infinite-horizon BSDE (2.2) and the BSDE with random terminal time (2.4).*

If $\frac{h^5\ell(T,x)}{\Delta} = o_{a.s.}(1)$, then we have

$$\left(\frac{h\ell(T,x)}{\Delta}\right)^{1/2} \widehat{Z}^2(x) \xrightarrow{d} \sqrt{2l(K_2)}\tilde{Z}^2(x)\mathbf{N},$$

where $\widehat{Z}^2(x)$ represents both $\widehat{Z}^2_{\mathrm{NW}}(x)$ and $\widehat{Z}^2_{\mathrm{LL}}(x)$.

For the local constant and local linear estimators of the diffusion term, assuming $\ell(T,x)/\Delta \xrightarrow{\mathbb{P}} \infty$, the optimal bandwidths that minimize their mean squared errors are, respectively,

$$h^*_{Z,\mathrm{NW}} = 2^{1/5}c(K)\tilde{Z}^{4/5}(x)\left((\tilde{Z}^2)''(x) + 2(\tilde{Z}^2)'(x)\frac{s'(x)}{s(x)}\right)^{-2/5}\Delta^{1/5}\ell(T,x)^{-1/5},$$
$$h^*_{Z,\mathrm{LL}} = 2^{1/5}c(K)\tilde{Z}^{4/5}(x)((\tilde{Z}^2)''(x))^{-2/5}\Delta^{1/5}\ell(T,x)^{-1/5},$$

where $c(K) = \frac{l(K_2)^{1/5}}{l(K_1)^{2/5}}$.

Although the above theorems provide the asymptotic conditional variance of the estimators, calculating this variance can be quite challenging in practice due to lack of knowledge about the true function and the availability of only a single trajectory. In such cases, the bootstrap method can be used for statistical inference of the terms of interest.

4. Empirical likelihood confidence intervals

This section investigates the empirical likelihood method for constructing the data-driven pointwise confidence intervals for the generator and diffusion term of the two types of FBSDEs.

4.1 Confidence interval for the generator

Denote

$$g_{f_i}(x,h,\theta) = K\left(\frac{X_i - x}{h}\right)\left(\frac{Y_{(i+1)\Delta} - Y_{i\Delta}}{\Delta} - \theta\right),$$

where θ is the candidate value of the target quantity $\tilde{f}(x)$. Define

$$L_f(x,h,\theta) := \max_{(p_1,\cdots,p_n)}\left\{\prod_{i=1}^n np_i \mid \sum_{i=1}^n p_i g_{f_i}(x,h,\theta) = 0,\ p_i \geqslant 0,\ \sum_{i=1}^n p_i = 1\right\}, \tag{4.1}$$

and

$$l_f(x,h,\theta) := -2\log L_f(x,h,\theta). \tag{4.2}$$

The following theorem describes the asymptotic properties of $l_f(x,h,\theta)$, which helps to construct the confidence interval for $\tilde{f}(x)$.

Theorem 4.1 *Let the assumptions in Theorem 3.9 hold. Suppose that* $\Delta^2\mathscr{M}(\tilde{f}^3) = o_{\mathbb{P}}(h\ell(T,x))$, *and* $\sqrt{\Delta}\mathscr{M}(\tilde{Z}^3) = o_{\mathbb{P}}(h\ell(T,x))$ *uniformly in* T *as* $\Delta \to 0$ *and* $h \to 0$. *Furthermore, assume that*

$$h\ell(T,x) \xrightarrow{a.s.} \infty,$$

and

$$h^5\ell(T,x) \xrightarrow{a.s.} 0,$$

as $n, T \to \infty$. *Then,*

$$l_f(x,h,\tilde{f}(x)) \xrightarrow{d} \chi^2(1),$$

and

$$l_f\left(x,h,\tilde{f}(x) + \frac{\bar{\tau}(x)}{\sqrt{h\ell(T,x)}}\right) \xrightarrow{d} \chi^2\left(1, \frac{\bar{\tau}^2(x)}{l(K_2)\tilde{Z}^2(x)}\right),$$

as $n, T \to \infty$, *where* $\bar{\tau}(x)$ *is fixed and* $\chi^2(p_1,p_2)$ *is the chi-squared distribution with degree of freedom* p_1 *and non-central parameter* p_2. *The* $1-\alpha$ *empirical likelihood confidence interval for* $\tilde{f}(x)$ *is then defined as*

$$I^{EL}_{f,\alpha} = \{l_f(x,h,\theta) \leqslant \chi^2_{1-\alpha}(1)\},$$

where α *is the significance level and* $\chi^2_{1-\alpha}(1)$ *is the inverse cumulative distribution function of the* $\chi^2(1)$ *distribution evaluated at* $1-\alpha$.

Remark 4.2 (i) *Here, we focus on local constant smoothing, although the extension to local linear (or polynomial) smoothing is straightforward. In fact, let*

$$g^{\mathrm{LL}}_{f_i}(x,h,\theta) = \widetilde{K}(x,h)\left(\frac{Y_{(i+1)\Delta} - Y_{i\Delta}}{\Delta} - \theta\right),$$

where $\widetilde{K}(x,h) = K_h(X_{i\Delta}-x)s_{n,2} - (X_{i\Delta}-x)K_h(X_{i\Delta}-x)s_{n,1}$, *and*

$$s_{n,j} = \sum_{i=1}^n (X_{i\Delta}-x)^j K_h(X_{i\Delta}-x),\ j=1,2. \quad K_h(\cdot) = K(\cdot/h)/h.$$

Define

$$L^{\mathrm{LL}}_f(x,h,\theta) := \max_{(p_1,\cdots,p_n)}\left\{\prod_{i=1}^n np_i \mid \sum_{i=1}^n p_i g^{\mathrm{LL}}_{f_i}(x,h,\theta) = 0,\ p_i \geqslant 0,\ \sum_{i=1}^n p_i = 1\right\},$$

and

$$l^{\mathrm{LL}}_f(x,h,\theta) := -2\log L^{\mathrm{LL}}_f(x,h,\theta).$$

The conclusion, which is analogous to Theorem 4.1, can then be established, and the corresponding empirical likelihood confidence interval can similarly be constructed.

(ii) $l_f(x,h,\theta)$ *can easily be calculated. Using Lagrange multipliers, the constrained optimization problem (4.1) is solved by*

$$\widehat{p}_i = \frac{1}{n\left(1 + \lambda_f(x,h,\theta) g_{f_i}(x,h,\theta)\right)}, \tag{4.3}$$

where $\lambda_f(x,h,\theta)$ satisfies

$$\sum_{i=1}^{n} \frac{g_{f_i}(x,h,\theta)}{1 + \lambda_f(x,h,\theta) g_{f_i}(x,h,\theta)} = 0. \tag{4.4}$$

After θ is obtained, we can solve (4.4) numerically and compute

$$l_f(x,h,\theta) = 2\sum_{i=1}^{n-1} \log\left(1 + \lambda_f(x,h,\theta) g_{f_i}(x,h,\theta)\right).$$

(iii) *In the previous section, Theorem 3.9 established the asymptotic normality of the local constant estimator for $\tilde{f}(x)$. Hence, its confidence interval at the $1-\alpha$ level is expressed as*

$$I^{\mathrm{ASY}}_{f,\alpha} = \left[\widehat{f}_{\mathrm{NW}}(x) - z_{1-\alpha/2}(\hat{v}(\tilde{f}(x)))^{1/2}, \quad \widehat{f}_{\mathrm{NW}}(x) + z_{1-\alpha/2}(\hat{v}(\tilde{f}(x)))^{1/2}\right],$$

where $\hat{v}(\tilde{f}(x)) = (h\ell(T,x))^{-1} l(K_2) \tilde{Z}^2(x)$ is the asymptotic variance of the estimator, and $z_{1-\alpha/2}$ is the inverse cumulative distribution function of the standard normal distribution at $1-\alpha/2$. It is important to note that the confidence interval involves two unknown quantities: $\ell(T,x)$ and $\tilde{Z}^2(x)$. Estimating these quantities is crucial to constructing the normal confidence interval, and the symmetry of the interval can lead to inaccuracies in the estimation. In practice, $\ell(T,x)$ can be estimated by $\bar{\ell}(T,x) = \frac{\Delta}{h}\sum_{i=1}^{n} K\left(\frac{X_{i\Delta}-x}{h}\right)$, and $\tilde{Z}^2(x)$ can be estimated using the local constant or local linear estimators described earlier.

4.2 Confidence interval for the diffusion term

The empirical likelihood confidence interval for $\tilde{Z}^2(x)$ can be constructed in a similar way. Denote

$$g_{Z_i}(x,h,\theta) := K\left(\frac{X_i - x}{h}\right)\left(\frac{(Y_{(i+1)\Delta} - Y_{i\Delta})^2}{\Delta} - \theta\right),$$

and

$$L_Z(x,h,\theta) := \max_{(p_1,\cdots,p_n)} \left\{\prod_{i=1}^{n} np_i \mid \sum_{i=1}^{n} p_i g_{Z_i}(x,h,\theta) = 0,\ p_i \geqslant 0,\ \sum_{i=1}^{n} p_i = 1\right\}, \tag{4.5}$$

$$l_Z(x,h,\theta) := -2\log L_Z(x,h,\theta). \tag{4.6}$$

The following theorem describes the asymptotic properties of $l_Z(x,h,\theta)$, which helps to construct the confidence interval for $\tilde{Z}^2(x)$.

Theorem 4.3 *Let the assumptions in Theorem 3.14 hold, and let $\Delta^2\mathscr{M}(\tilde{f}^3) = o_{\mathbb{P}}(h\ell(T,x))$, $\sqrt{\Delta}\mathscr{M}(\tilde{Z}^3) = o_{\mathbb{P}}(h\ell(T,x))$ uniformly in T as $\Delta \to 0,\ h \to 0$. Furthermore,*

$$\frac{h\ell(T,x)}{\Delta} \xrightarrow{a.s.} \infty,$$

and

$$\frac{h^5\ell(T,x)}{\Delta} \xrightarrow{a.s.} 0,$$

as $n, T \to \infty$. Then,

$$l_Z(x,h,\tilde{Z}^2(x)) \xrightarrow{d} \chi^2(1),$$

and

$$l_Z\left(x,h,\tilde{Z}^2(x)+\frac{\bar{\tau}(x)}{\sqrt{h\ell(T,x)/\Delta}}\right)\xrightarrow{d}\chi^2\left(1,\frac{\bar{\tau}^2(x)}{2l(K_2)\tilde{Z}^4(x)}\right),$$

as $n,T\to\infty$. Then, the empirical likelihood confidence interval at the $1-\alpha$ confidence level for $\tilde{Z}^2(x)$ is defined as

$$I_{Z,\alpha}^{EL}=\{l_Z(x,h,\theta)\leqslant\chi^2_{1-\alpha}(1)\}.$$

Remark 4.4 *Based on the results from Theorem 3.14 in the previous section, the normal confidence interval for $\tilde{Z}^2(x)$ at the $1-\alpha$ confidence level is*

$$I_{Z,\alpha}^{\mathrm{ASY}}=\left[\widehat{Z}^2(x)-Z_{1-\alpha/2}(\hat{v}(\tilde{Z}^2(x)))^{1/2},\quad \widehat{Z}^2(x)+Z_{1-\alpha/2}(\hat{v}(\tilde{Z}^2(x)))^{1/2}\right],$$

where $\hat{v}(\tilde{Z}^2(x))=2(h\bar{\ell}(T,x)/\Delta)^{-1}l(K_2)\widehat{Z}^4(x)$. Here, $\bar{\ell}(T,x)$ is the estimate of $\ell(T,x)$, and $\widehat{Z}(x)$ is the local constant or local linear estimator of $\tilde{Z}(x)$. The method for constructing confidence intervals using weights from local linear estimation can be found in Remark 4.2.

5. Proofs

To give proofs for the main theorems, we first introduce some useful lemmas. In these lemmas, we assume that $\bar{w}$ is a nonnegative, bounded, and twice continuously differentiable function on $\mathbb{R}$ with a support $[-1,1]$, as assumed for the kernel function in Assumption 3.1, and that w is twice continuously differentiable on $\mathscr{D}$. The proofs for certain lemmas are not provided here, but can be found in Lemmas 6, 8, 9, 10, 11, and 12 in the "Proofs" section of Aït-Sahalia and Park [1].

Lemma 5.1

$$\begin{aligned}\frac{\Delta}{h}\sum_{i=0}^{n-1}\bar{w}\left(\frac{X_{i\Delta}-x}{h}\right)&=\frac{1}{h}\int_0^T\bar{w}\left(\frac{X_t-x}{h}\right)\mathrm{d}t+O_{\mathbb{P}}\left(h^{-2}\Delta\ell(T,x)\right)\\&=l(\bar{w})\ell(T,x)+o_{\mathbb{P}}(\ell(T,x))+O_{\mathbb{P}}\left(h^{-2}\Delta\ell(T,x)\right),\end{aligned}$$

uniformly in T as $h\to0$ and $\Delta\to0$, where $l(\bar{w})=\int\bar{w}(x)\mathrm{d}x$.

Lemma 5.2

$$\begin{aligned}&\frac{\Delta}{h}\sum_{i=0}^{n-1}\bar{w}\left(\frac{X_{i\Delta}-x}{h}\right)w\left(X_{i\Delta}\right)\\&=\frac{1}{h}\int_0^T\bar{w}\left(\frac{X_s-x}{h}\right)w\left(X_s\right)\mathrm{d}s+O_{\mathbb{P}}\left(h^{-2}\Delta\ell(T,x)\right)\\&=l(\bar{w})w(x)\ell(T,x)+o_{\mathbb{P}}(\ell(T,x))+O_{\mathbb{P}}\left(h^{-2}\Delta\ell(T,x)\right),\end{aligned}$$

uniformly in T as $h\to0$ and $\Delta\to0$.

Lemma 5.3 *Assume that $h\to0$ and $\Delta\to0$ such that*

(1) $h^{-4}\Delta\to0$,

(2) $\sqrt{\Delta}\mathscr{M}(\mathscr{L}w)=o_{\mathbb{P}}(1)$, $\sqrt{\Delta}\mathscr{M}(w'\sigma)=o_{\mathbb{P}}(\sqrt{h\ell(T,x)})$.

Then

$$\begin{aligned}
&\frac{1}{h}\sum_{i=0}^{n-1}\bar{w}\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}\left(w\left(X_s\right)-w\left(X_{i\Delta}\right)\right)\mathrm{d}s\\
&=O_{\mathbb{P}}\left(\Delta\mathscr{M}(\mathscr{L}w)\ell(T,x)\right)+O_{\mathbb{P}}\left(h^{-1/2}\Delta\mathscr{M}(w'\sigma)\ell(T,x)^{1/2}\right)=o_{\mathbb{P}}\left(h^2\ell(T,x)\right),
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{\sqrt{\Delta h}}\sum_{i=0}^{n-1}\bar{w}\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}\left(w\left(X_s\right)-w\left(X_{i\Delta}\right)\right)\mathrm{d}s\\
&=O_{\mathbb{P}}\left(\sqrt{\Delta}\mathscr{M}(\mathscr{L}w)\sqrt{h}\ell(T,x)\right)+O_{\mathbb{P}}\left(\sqrt{\Delta}\mathscr{M}(w'\sigma)\ell(T,x)^{1/2}\right)=o_{\mathbb{P}}(\sqrt{h}\ell(T,x))
\end{aligned}$$

uniformly in T as $h\to 0$ and $\Delta\to 0$.

Lemma 5.4 *Assume that $h\to 0$ and $\Delta\to 0$ such that*
(1) $h^{-4}\Delta\to 0$,
(2) $\sqrt{\Delta}\mathscr{M}(\mathscr{L}(u\tilde{f}))=o_{\mathbb{P}}(1)$, $\sqrt{\Delta}\mathscr{M}((u\tilde{f})'\sigma)=o_{\mathbb{P}}(\sqrt{h\ell(T,x)})$,
(3) $\sqrt{\Delta}\mathscr{M}(\mathscr{L}\tilde{f})=o_{\mathbb{P}}(1)$, $\sqrt{\Delta}\mathscr{M}(\tilde{f}'\sigma)=o_{\mathbb{P}}(\sqrt{h\ell(T,x)})$.
Then

$$\frac{1}{h}\sum_{i=1}^{n-1}\bar{w}\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}(Y_t-Y_{i\Delta})\tilde{f}(X_t)\mathrm{d}t=o_{\mathbb{P}}\left(h^2\ell(T,x)\right),$$

and

$$\frac{1}{\sqrt{h\Delta}}\sum_{i=1}^{n-1}\bar{w}\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}(Y_t-Y_{i\Delta})\tilde{f}(X_t)\mathrm{d}t=o_{\mathbb{P}}(\sqrt{h}\ell(T,x)),$$

uniformly in T as $h\to 0$ and $\Delta\to 0$.

Proof It is easy to obtain that

$$\begin{aligned}
\left(Y_t-Y_{i\Delta}\right)\tilde{f}(X_t)&=\left(u(X_t)-u(X_{i\Delta})\right)\tilde{f}(X_t)\\
&=\left((u\tilde{f})(X_t)-(u\tilde{f})(X_{i\Delta})\right)-u(X_{i\Delta})\left(\tilde{f}(X_t)-\tilde{f}(X_{i\Delta})\right).
\end{aligned}$$

Then,

$$\begin{aligned}
&\frac{1}{h}\sum_{i=1}^{n-1}\bar{w}\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}(Y_t-Y_{i\Delta})\tilde{f}(X_t)\mathrm{d}t\\
&=\frac{1}{h}\sum_{i=1}^{n-1}\bar{w}\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}\left((u\tilde{f})(X_t)-(u\tilde{f})(X_{i\Delta})\right)\mathrm{d}t\\
&\quad-\frac{1}{h}\sum_{i=1}^{n-1}\bar{w}\left(\frac{X_{i\Delta}-x}{h}\right)u(X_{i\Delta})\int_{i\Delta}^{(i+1)\Delta}\left(\tilde{f}(X_t)-\tilde{f}(X_{i\Delta})\right)\mathrm{d}t\\
&:=C_1+C_2.
\end{aligned}$$

Applying Lemma 5.3 under conditions (1) and (2),

$$\begin{aligned}
C_1&=O_{\mathbb{P}}\left(\Delta\mathscr{M}\left(\mathscr{L}(u\tilde{f})\right)\ell(T,x)\right)+O_{\mathbb{P}}\left(h^{-1/2}\Delta\mathscr{M}\left((u\tilde{f})'\sigma\right)\ell(T,x)^{1/2}\right)\\
&=o_{\mathbb{P}}(\sqrt{h\ell(T,x)}),
\end{aligned}$$

uniformly in T as $h\to 0$ and $\Delta\to 0$. With the Itô formula, C_2 can be decomposed as

$$
\begin{aligned}
&\frac{1}{h}\sum_{i=1}^{n-1}\bar{w}\left(\frac{X_{i\Delta}-x}{h}\right)u(X_{i\Delta})\int_{i\Delta}^{(i+1)\Delta}\left(\tilde{f}(X_t)-\tilde{f}(X_{i\Delta})\right)\mathrm{d}t\\
&=\frac{1}{h}\sum_{i=1}^{n-1}\bar{w}\left(\frac{X_{i\Delta}-x}{h}\right)u(X_{i\Delta})\int_{i\Delta}^{(i+1)\Delta}((i+1)\Delta-t)\mathscr{L}\tilde{f}(X_t)\mathrm{d}t\\
&\quad+\frac{1}{h}\sum_{i=1}^{n-1}\bar{w}\left(\frac{X_{i\Delta}-x}{h}\right)u(X_{i\Delta})\int_{i\Delta}^{(i+1)\Delta}((i+1)\Delta-t)(\tilde{f}'\sigma)(X_t)\mathrm{d}W_t\\
&:=C_{21}+C_{22}.
\end{aligned}
$$

According to Lemma 5.2,

$$C_{21}\leqslant\Delta\mathscr{M}(\mathscr{L}\tilde{f})\frac{\Delta}{h}\sum_{i=1}^{n-1}|\bar{w}|\left(\frac{X_{i\Delta}-x}{h}\right)|u|(X_{i\Delta})=O_{\mathbb{P}}(\Delta\mathscr{M}(\mathscr{L}\tilde{f})\ell(T,x)),$$

uniformly in T as $h\to 0$ and $\Delta\to 0$. Moreover, C_{22} can be regarded as a continuous martingale whose quadratic variation is of $O_{\mathbb{P}}(h^{-1}\Delta^2\mathscr{M}\left((\tilde{f}'\sigma)^2\right)\ell(T,x))$. Therefore, $C_{22}=O_{\mathbb{P}}(h^{-1/2}\Delta\mathscr{M}(\tilde{f}'\sigma)\ell(T,x)^{1/2})$. Under conditions (1) and (3),

$$C_2=O_{\mathbb{P}}\left(\Delta\mathscr{M}(\mathscr{L}\tilde{f})\ell(T,x)+h^{-1/2}\Delta\mathscr{M}(\tilde{f}'\sigma)\ell(T,x)^{1/2}\right)=o_{\mathbb{P}}(\sqrt{h\ell(T,x)}),$$

uniformly in T as $h\to 0$ and $\Delta\to 0$. The proof of the second part is similar to the first part; thus, we omit it here. □

Lemma 5.5

$$\frac{1}{h}\int_0^T\frac{X_s-x}{h}K\left(\frac{X_s-x}{h}\right)\mathrm{d}s=l(K_1)\frac{s'(x)}{s(x)}h\ell(T,x)+O_{\mathbb{P}}(\sqrt{h\ell(T,x)})+o_{\mathbb{P}}(h\ell(T,x)),$$

uniformly in T as $h\to 0$ and $\Delta\to 0$.

Lemma 5.6

$$
\begin{aligned}
&\frac{1}{h}\int_0^T K\left(\frac{X_s-x}{h}\right)(w(X_s)-w(x))\,\mathrm{d}s\\
&=l(K_1)\left(w'(x)\frac{s'(x)}{s(x)}+\frac{w''(x)}{2}\right)h^2\ell(T,x)+O_{\mathbb{P}}(h^{3/2}\ell(T,x)^{1/2})+o_{\mathbb{P}}(h^2\ell(T,x)),
\end{aligned}
$$

uniformly in T as $h\to 0$ and $\Delta\to 0$.

Lemma 5.7

$$\frac{1}{h}\int_0^T\frac{X_s-x}{h}K\left(\frac{X_s-x}{h}\right)(w(X_s)-w(x))\,\mathrm{d}s=l(K_1)w'(x)h\ell(T,x)+o_{\mathbb{P}}(h\ell(T,x)),$$

uniformly in T as $h\to 0$ and $\Delta\to 0$.

Next, we proceed with the proof of the main theorems in Section 3. The proof of Theorem 3.9 is similar to that of Theorem 3.14, and the proof of Theorem 3.15 is similar to that of Theorem 3.11. Therefore, we omit the proofs of Theorems 3.9 and 3.15 here.

Proof of Theorem 3.11 Let $\bar{w}$ in Lemma 5.1 be K and K_1, respectively; then, we obtain

$$
\begin{aligned}
D(K)&=\ell(T,x)+o_{\mathbb{P}}(\ell(T,x))+O_{\mathbb{P}}\left(h^{-2}\Delta\ell(T,x)\right),\\
D(K_2)&=l(K_1)\ell(T,x)+o_{\mathbb{P}}(\ell(T,x))+O_{\mathbb{P}}\left(h^{-2}\Delta\ell(T,x)\right),
\end{aligned}
\tag{5.1}
$$

uniformly in T as $h \to 0$ and $\Delta \to 0$. Applying Lemma 5.1 and 5.5, we obtain

$$D(K_1) = l(K_1)\frac{s'(x)}{s(x)}h\ell(T,x) + O_{\mathbb{P}}(\sqrt{h\ell(T,x)}) + o_{\mathbb{P}}(h\ell(T,x)) + O_{\mathbb{P}}\left(h^{-2}\Delta\ell(T,x)\right), \tag{5.2}$$

uniformly in T as $h \to 0$ and $\Delta \to 0$. Then,

$$D(K)D(K_2) - (D(K_1))^2 = l(K_1)(\ell(T,x))^2 + o_{\mathbb{P}}(\ell(T,x)^2),$$

uniformly in T as $h \to 0$ and $\Delta \to 0$.

We first analyze $\hat{g}_{p,\mathrm{LL}}(x)$. According to Lemma 5.2 and 5.6, let $\bar{w} = K$, $w = \tilde{f}$, we obtain

$$\begin{aligned}
&B(K,\tilde{f})\\
&= \frac{1}{h}\int_0^T K\left(\frac{X_s - x}{h}\right)\left(\tilde{f}(X_s) - \tilde{f}(x)\right)\mathrm{d}s + O_{\mathbb{P}}(h^{-2}\Delta\ell(T,x))\\
&= l(K_1)\left(\tilde{f}'(x)\frac{s'(x)}{s(x)} + \frac{\tilde{f}''(x)}{2}\right)h^2\ell(T,x) + O_{\mathbb{P}}(h^{-2}\Delta\ell(T,x))\\
&\quad + O_{\mathbb{P}}(h^{3/2}\ell(T,x)^{1/2}) + o_{\mathbb{P}}(h^2\ell(T,x)),
\end{aligned} \tag{5.3}$$

uniformly in T as $h \to 0$ and $\Delta \to 0$. Similarly, taking $\bar{w} = K_1$ and $w = \tilde{f}$ in Lemma 5.2 and 5.7, we obtain

$$\begin{aligned}
B(K_1,\tilde{f}) &= \frac{1}{h}\int_0^T K_1\left(\frac{X_s - x}{h}\right)\left(\tilde{f}(X_s) - \tilde{f}(x)\right)\mathrm{d}s + O_{\mathbb{P}}(h^{-2}\Delta\ell(T,x))\\
&= l(K_1)\tilde{f}'(x)h\ell(T,x) + o_{\mathbb{P}}(h\ell(T,x)),
\end{aligned} \tag{5.4}$$

uniformly in T as $h \to 0$ and $\Delta \to 0$. Combining (5.1), (5.2), (5.3), and (5.4), we obtain

$$\begin{aligned}
&B(K,\tilde{f})D(K_2) - B(K_1,\tilde{f})D(K_1)\\
= &\left[l(K_1)\left(\tilde{f}'(x)\frac{s'(x)}{s(x)} + \frac{\tilde{f}''(x)}{2}\right)h^2\ell(T,x) + O_{\mathbb{P}}(h^{3/2}\ell(T,x)^{1/2}) + o_{\mathbb{P}}(h^2\ell(T,x))\right]\\
&\times\left[l(K_1)\ell(T,x) + o_{\mathbb{P}}(\ell(T,x))\right]\\
&- \left[l(K_1)\tilde{f}'(x)h\ell(T,x) + o_{\mathbb{P}}(h\ell(T,x)\right]\left[l(K_1)\frac{s'(x)}{s(x)}h\ell(T,x) + o_{\mathbb{P}}(h\ell(T,x))\right]\\
= &\frac{h^2}{2}(l(K_1))^2\tilde{f}''(x)(\ell(T,x))^2 + O_{\mathbb{P}}((h\ell(T,x))^{3/2}) + o_{\mathbb{P}}(h^2\ell(T,x)^2),
\end{aligned}$$

uniformly in T as $h \to 0$ and $\Delta \to 0$. Furthermore,

$$\begin{aligned}
\hat{g}_{p,\mathrm{LL}}(x) - \tilde{f}(x) &= \frac{\frac{h^2}{2}(l(K_1))^2\tilde{f}''(x)(\ell(T,x))^2 + O_{\mathbb{P}}((h\ell(T,x))^{3/2}) + o_{\mathbb{P}}(h^2\ell(T,x)^2)}{l(K_1)(\ell(T,x))^2 + o_{\mathbb{P}}(\ell(T,x)^2)}\\
&= \frac{h^2}{2}l(K_1)\tilde{f}''(x) + O_{\mathbb{P}}(h^{3/2}\ell(T,x)^{-1/2}) + o_{\mathbb{P}}(h^2),
\end{aligned}$$

uniformly in T as $h \to 0$ and $\Delta \to 0$.

To deal with $\hat{g}_{q,\mathrm{LL}}(x)$, we define a continuous martingale M as

$$M_T = \frac{1}{\sqrt{h}}\sum_{i=1}^{n-1} K_1\left(\frac{X_{i\Delta} - x}{h}\right)\int_{i\Delta}^{(i+1)\Delta} Z_s\mathrm{d}W_s.$$

Note that $M(K_1,\tilde{f}) = h^{-1/2}M_T$. According to Lemma 5.2 and 5.3, if the conditions (1) and (2) in Assumption 3.7 hold, we obtain

$$\begin{aligned}\langle M\rangle_T &= \frac{1}{h}\sum_{i=1}^{n-1} K_1^2\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta} Z_s^2 \mathrm{d}s\\ &= \frac{\Delta}{h}\sum_{i=1}^{n-1} K_1^2\left(\frac{X_{i\Delta}-x}{h}\right)\tilde{Z}^2(X_{i\Delta}) + \frac{1}{h}\sum_{i=1}^{n-1} K_1^2\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}\left(\tilde{Z}^2(X_s)-\tilde{Z}^2(X_{i\Delta})\right)\mathrm{d}s\\ &= l(K^2)\tilde{Z}^2(x)\ell(T,x) + o_{\mathbb{P}}(\ell(T,x)) + o_{\mathbb{P}}(h^2\ell(T,x)),\end{aligned}$$

uniformly in T as $h \to 0$ and $\Delta \to 0$. Therefore,

$$M(K_1,\tilde{f}) = O_{\mathbb{P}}(h^{-1/2}\ell(T,x)^{1/2}),$$

and

$$M(K_1,\tilde{f})D(K_1) = O_{\mathbb{P}}(h^{-1/2}\ell(T,x)^{1/2}) \times O_{\mathbb{P}}(h\ell(T,x)) = O_{\mathbb{P}}(\sqrt{h}\ell(T,x)^{3/2}).$$

Moreover,

$$\begin{aligned}&\sqrt{h\ell(T,x)}\hat{g}_{q,\mathrm{LL}}(x)\\ &= \sqrt{h\ell(T,x)}\frac{M(K,\tilde{f})D(K_2)-M(K_1,\tilde{f})D(K_1)}{D(K)D(K_2)-(D(K_1))^2}\\ &= \sqrt{h\ell(T,x)}\frac{l(K_1)\ell(T,x)(1+o_{\mathbb{P}}(1))}{l(K_1)\ell(T,x)^2(1+o_{\mathbb{P}}(1))}M(K,\tilde{f}) + \frac{O_{\mathbb{P}}(\sqrt{h}(\ell(T,x))^{3/2})}{O_{\mathbb{P}}(\ell(T,x)^2)}\\ &= \frac{O_{\mathbb{P}}(\sqrt{h}M(K,\tilde{f}))}{\sqrt{\ell(T,x)}} + o_{\mathbb{P}}(1),\end{aligned}$$

uniformly in T as $h \to 0$ and $\Delta \to 0$. Define another continuous martingale $\bar{M}$ as

$$\bar{M}_T = \frac{1}{\sqrt{h}}\sum_{i=1}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta} Z_s \mathrm{d}W_s.$$

In fact, $M(K,\tilde{f}) = h^{-1/2}\bar{M}_T$. With Lemma 5.2 and 5.3, we obtain

$$\begin{aligned}\langle\bar{M}\rangle_T &= \frac{1}{h}\sum_{i=1}^{n-1} K^2\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta} Z_s^2 \mathrm{d}s\\ &= l(K_2)\tilde{Z}^2(x)\ell(T,x) + o_{\mathbb{P}}(\ell(T,x)) + o_{\mathbb{P}}(h^2\ell(T,x)),\end{aligned}$$

and

$$\begin{aligned}\langle\bar{M},W\rangle_T &= \frac{1}{\sqrt{h}}\sum_{i=1}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta} Z_s \mathrm{d}s\\ &= \sqrt{h}\tilde{Z}(x)\ell(T,x)(1+o_{\mathbb{P}}(1)),\end{aligned}$$

uniformly in T as $h \to 0$ and $\Delta \to 0$. According to Theorem V.1.6 in Revuz and Yor [15], let $\bar{M}_T = V(\langle\bar{M}\rangle_T)$, then V is the DDS Brownian motion of $\bar{M}$. Define

$$V_T(\cdot) = \frac{V(a_T \times \cdot)}{\sqrt{a_T}},$$

where $a_T = l(K_2)\tilde{Z}^2(x)\alpha_T$ is a sequence of nonrandom vectors. Then,

$$\frac{\bar{M}_T}{\sqrt{a_T}} = V_T\left(\frac{\langle\bar{M}\rangle_T}{a_T}\right).$$

In fact, for any T, V_T is a standard Brownian motion. Based on the above results, as $h \to 0$ and

$\Delta \to 0$, we obtain

$$\frac{\langle \bar{M}, W\rangle_T}{\langle \bar{M}\rangle_T} \xrightarrow{\mathbb{P}} 0, \quad \frac{\langle \bar{M}, W\rangle_T}{\langle W\rangle_T} = \frac{\langle \bar{M}, W\rangle_T}{T} \xrightarrow{\mathbb{P}} 0.$$

Then, by applying Corollary XIII. 2.4 in Revuz and Yor [15], we obtain

$$(V_T, W) \xrightarrow{d} (V_0, W),$$

where V_0 is a standard Brownian motion independent of W. Note that

$$\frac{\langle \bar{M}\rangle_T}{a_T} = \frac{l(K_2)\tilde{Z}^2(x)\ell(T,x) + o_{\mathbb{P}}(\ell(T,x))}{l(K_2)\tilde{Z}^2(x)\alpha_T} = \frac{\ell(T,x)}{\alpha_T} + o_{\mathbb{P}}(1).$$

Then, as $h, \ \Delta \to 0$,

$$\frac{\bar{M}_T}{\sqrt{a_T}} \xrightarrow{d} V_0\left(\frac{\ell(T,x)}{\alpha_T}\right).$$

Furthermore,

$$\frac{\sqrt{h}M(K,\tilde{f})}{\sqrt{\ell(T,x)}} \xrightarrow{d} \sqrt{l(K_2)}\tilde{Z}(x)\mathbf{N},$$

uniformly in T as $h \to 0$ and $\Delta \to 0$, where $\mathbf{N}$ is a standard normal random variate independent of $\ell(T,x)$.

For $\hat{g}_{r,\mathrm{LL}}(x)$, we set $\bar{w} = K, K_1$ and $w = \tilde{f}$ in Lemma 5.3. If (1) and (3) in Assumption 3.7 hold, we have

$$R(K,\tilde{f}) = o_{\mathbb{P}}(h^2\ell(T,x)), \quad R(K_1,\tilde{f}) = o_{\mathbb{P}}(h^2\ell(T,x)),$$

and

$$R(K,\tilde{f})D(K_2) - R(K_1,\tilde{f})D(K_1) = o_{\mathbb{P}}(h^2(\ell(T,x))^2),$$

uniformly in T as $h \to 0$ and $\Delta \to 0$. Furthermore,

$$\hat{g}_{r,\mathrm{LL}}(x) = o_{\mathbb{P}}(h^2),$$

uniformly in T as $h \to 0$ and $\Delta \to 0$. In conclusion, the theorem is proved. □

Proof of Theorem 3.14 Applying Lemma 5.2 and 5.6, we have

$$\begin{aligned}
&\frac{\Delta}{h}\sum_{i=0}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right)\left(\tilde{Z}^2(X_{i\Delta}) - \tilde{Z}^2(x)\right)\\
&= l(K_1)\left((\tilde{Z}^2)'(x)\frac{s'(x)}{s(x)} + \frac{(\tilde{Z}^2)''(x)}{2}\right)h^2\ell(T,x) + o_{\mathbb{P}}(h^{3/2}\ell(T,x)^{1/2}) + O_{\mathbb{P}}(h^2\ell(T,x)),
\end{aligned}$$

and

$$D(K) = \ell(T,x) + o_{\mathbb{P}}(\ell(T,x)) + O_{\mathbb{P}}\left(h^{-2}\Delta\ell(T,x)\right),$$

uniformly in T as $h \to 0$ and $\Delta \to 0$. Then, (3.7) can be obtained. For $\hat{Z}^2_{q,\mathrm{NW}}(x)$, we first analyze the numerator $\big($ denoted as $\bar{A}(x)\big)$ Define a continuous martingale $\tilde{M}$ as

$$\tilde{M}_T = \sqrt{\frac{4}{h\Delta}}\sum_{i=0}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}(Y_t - Y_{i\Delta})Z_t \mathrm{d}W_t. \tag{5.5}$$

Note that $\tilde{M}_T = \sqrt{h/\Delta}\bar{A}(x)$. With the Itô formula, we have

$$\begin{aligned}\langle \tilde{M}\rangle_T &= \frac{4}{h\Delta}\sum_{i=0}^{n-1} K^2\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}(Y_t-Y_{i\Delta})^2\tilde{Z}^2(X_t)\mathrm{d}t\\
&= \frac{4}{h\Delta}\sum_{i=0}^{n-1} K^2\left(\frac{X_{i\Delta}-x}{h}\right)\left[\tilde{Z}^2\left(X_{i\Delta}\right)\int_{i\Delta}^{(i+1)\Delta}\left(Y_t-Y_{i\Delta}\right)^2\mathrm{d}t\right.\\
&\quad \left.+\int_{i\Delta}^{(i+1)\Delta}\left(Y_t-Y_{i\Delta}\right)^2(\tilde{Z}^2(X_t)-\tilde{Z}^2(X_{i\Delta}))\mathrm{d}t\right]\\
&=: A_1+A_2+A_3+A_4+A_5,\end{aligned}$$

where

$$\begin{aligned}A_1 &= \frac{2\Delta}{h}\sum_{i=0}^{n-1} K^2\left(\frac{X_{i\Delta}-x}{h}\right)\tilde{Z}^4(X_{i\Delta}),\\
A_2 &= \frac{8}{h\Delta}\sum_{i=0}^{n-1} K^2\left(\frac{X_{i\Delta}-x}{h}\right)\tilde{Z}^2(X_{i\Delta})\int_{i\Delta}^{(i+1)\Delta}\int_{i\Delta}^{t}\left(Y_s-Y_{i\Delta}\right)\tilde{f}(X_s)\mathrm{d}s\mathrm{d}t,\\
A_3 &= \frac{4}{h\Delta}\sum_{i=0}^{n-1} K^2\left(\frac{X_{i\Delta}-x}{h}\right)\tilde{Z}^2(X_{i\Delta})\int_{i\Delta}^{(i+1)\Delta}\int_{i\Delta}^{t}(\tilde{Z}^2(X_s)-\tilde{Z}^2(X_{i\Delta}))\mathrm{d}s\mathrm{d}t,\\
A_4 &= \frac{8}{h\Delta}\sum_{i=0}^{n-1} K^2\left(\frac{X_{i\Delta}-x}{h}\right)\tilde{Z}^2(X_{i\Delta})\int_{i\Delta}^{(i+1)\Delta}\int_{i\Delta}^{t}\left(Y_s-Y_{i\Delta}\right)\tilde{Z}(X_s)\mathrm{d}W_s\mathrm{d}t,\\
A_5 &= \frac{4}{h\Delta}\sum_{i=0}^{n-1} K^2\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}\left(Y_t-Y_{i\Delta}\right)^2(\tilde{Z}^2(X_t)-\tilde{Z}^2(X_{i\Delta}))\mathrm{d}t.\end{aligned}$$

We will analyze every part in the following. With Lemma 5.2, if (1) in Assumption 3.7 holds, we obtain

$$\begin{aligned}A_1 &= \frac{2}{h}\int_0^T K^2\left(\frac{X_t-x}{h}\right)\tilde{Z}^4(X_t)\mathrm{d}t + O_{\mathbb{P}}(h^{-2}\Delta\ell(T,x))\\
&= 2\tilde{Z}^4(x)l(K_2)\ell(T,x)(1+o_{\mathbb{P}}(1)),\end{aligned}$$

uniformly in T as $h\to 0$ and $\Delta\to 0$. According to Lemma 5.4, if

$$\sqrt{\Delta}\mathscr{M}(\mathscr{L}(u\tilde{f})) = o_{\mathbb{P}}(1),\quad \sqrt{\Delta}\mathscr{M}(\mathscr{L}\tilde{f}) = o_{\mathbb{P}}(1),$$

and

$$\sqrt{\Delta}\mathscr{M}((u\tilde{f})'\sigma) = o_{\mathbb{P}}(\sqrt{h\ell(T,x)}),\quad \sqrt{\Delta}\mathscr{M}(\tilde{f}'\sigma) = o_{\mathbb{P}}(\sqrt{h\ell(T,x)}),$$

we obtain

$$A_2 = o_{\mathbb{P}}\left(h^2\ell(T,x)\right),$$

uniformly in T as $h\to 0$ and $\Delta\to 0$. The specific steps of the proof can be found in A_4. For A_3, using Lemma 5.2 and 5.3, if

$$h^{-4}\Delta = o_{\mathbb{P}}(1),\quad \sqrt{\Delta}\mathscr{M}(\mathscr{L}(\tilde{Z}^2)) = o_{\mathbb{P}}(1),\quad \sqrt{\Delta}\mathscr{M}((\tilde{Z}^2)'\sigma) = o_{\mathbb{P}}(\sqrt{h\ell(T,x)}),$$

then

$$A_3 = O_{\mathbb{P}}\left(\Delta\mathscr{M}(\mathscr{L}(\tilde{Z}^2))\ell(T,x)\right) + O_{\mathbb{P}}\left(h^{-1/2}\Delta\mathscr{M}((\tilde{Z}^2)'\sigma)\ell(T,x)^{1/2}\right) = o_{\mathbb{P}}\left((h^2\ell(T,x))\right),$$

uniformly in T as $h\to 0$ and $\Delta\to 0$. Note that A_4 can be expressed as a continuous martingale whose quadratic variation can be bounded by

$$
\begin{aligned}
&\frac{64}{h^2}\sum_{i=0}^{n-1}K^4\left(\frac{X_{i\Delta}-x}{h}\right)\tilde{Z}^4(X_{i\Delta})\int_{i\Delta}^{(i+1)\Delta}\left(Y_t-Y_{i\Delta}\right)^2\tilde{Z}^2(X_t)\mathrm{d}t\\
&\leqslant O_{\mathbb{P}}\left((\Delta\mathscr{M}(\tilde{f})+\sqrt{\Delta}\mathscr{M}(\tilde{Z}))^2/h\right)\\
&\quad\times\left(\frac{1}{h}\sum_{i=0}^{n-1}K^4\left(\frac{X_{i\Delta}-x}{h}\right)\tilde{Z}^4(X_{i\Delta})\int_{i\Delta}^{(i+1)\Delta}(\tilde{Z}^2(X_t)-\tilde{Z}^2(X_{i\Delta}))\mathrm{d}t+\frac{\Delta}{h}\sum_{i=0}^{n-1}K^4\left(\frac{X_{i\Delta}-x}{h}\right)\tilde{Z}^6(X_{i\Delta})\right)\\
&\leqslant O_{\mathbb{P}}\left((\Delta\mathscr{M}(\tilde{f})+\sqrt{\Delta}\mathscr{M}(\tilde{Z}))^2/h\right)\\
&\quad\times O_{\mathbb{P}}\left(\Delta\mathscr{M}(\mathscr{L}\tilde{Z}^2)\ell(T,x)+h^{-1/2}\Delta\mathscr{M}((\tilde{Z}^2)'\sigma)\ell(T,x)^{1/2}+\int K^4(u)\mathrm{d}u\tilde{Z}^6(x)\ell(T,x)\right)\\
&\leqslant O_{\mathbb{P}}\left((\Delta\mathscr{M}(\tilde{f})+\sqrt{\Delta}\mathscr{M}(\tilde{Z}))^2/h\right)\times\left(o_{\mathbb{P}}(h^2\ell(T,x))+O_{\mathbb{P}}(\ell(T,x))\right),
\end{aligned}
$$

where the first inequality uses the following relationship:

$$
|Y_t-Y_{i\Delta}|\leqslant\left|\int_{i\Delta}^t\tilde{f}(X_s)\mathrm{d}s\right|+\left|\int_{i\Delta}^t\tilde{Z}(X_s)\mathrm{d}W_s\right|\leqslant O_{\mathbb{P}}\left(\Delta\mathscr{M}(\tilde{f})+\sqrt{\Delta\mathscr{M}(\tilde{Z}^2)}\right),\tag{5.6}
$$

where $t\in[i\Delta,(i+1)\Delta]$, $1\leqslant i\leqslant n-1$. Using Lemma 5.2 and 5.3, if

$$
h^{-4}\Delta\to 0,\quad \sqrt{\Delta}\mathscr{M}(\mathscr{L}\tilde{Z}^2)=o_{\mathbb{P}}(1),\quad \sqrt{\Delta}\mathscr{M}((\tilde{Z}^2)'\sigma)=o_{\mathbb{P}}(\sqrt{h\ell(T,x)}),
$$

the second inequality can be obtained. According to the last inequality, if

$$
\Delta\mathscr{M}(\tilde{f})=o_{\mathbb{P}}((h\ell(T,x))^{1/2}),\quad \sqrt{\Delta}\mathscr{M}(\tilde{Z})=o_{\mathbb{P}}((h\ell(T,x))^{1/2}),
$$

then

$$
A_4=o_{\mathbb{P}}(\ell(T,x)),
$$

uniformly in T as $h\to 0$ and $\Delta\to 0$. For A_5, according to

$$
\tilde{Z}^2(X_t)-\tilde{Z}^2(X_{i\Delta})\leqslant\mathscr{M}((\tilde{Z}^2)')|X_t-X_{i\Delta}|\leqslant O_{\mathbb{P}}\left(\Delta\mathscr{M}(b)\mathscr{M}((\tilde{Z}^2)')+(\Delta\mathscr{M}(\sigma^2))^{1/2}\mathscr{M}((\tilde{Z}^2)')\right),
$$

combining the Itô formula, we have

$$
\begin{aligned}
A_5&\leqslant O_{\mathbb{P}}\left(\Delta\mathscr{M}(b)\mathscr{M}((\tilde{Z}^2)')+\sqrt{\Delta}\mathscr{M}(\sigma)\mathscr{M}((\tilde{Z}^2)')\right)\\
&\quad\times\left(\frac{4}{h\Delta}\sum_{i=0}^{n-1}K^2\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}\left(Y_t-Y_{i\Delta}\right)^2\mathrm{d}t\right)\\
&=o_{\mathbb{P}}\left(\frac{1}{h\Delta}\sum_{i=0}^{n-1}K^2\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}\left(Y_t-Y_{i\Delta}\right)^2\mathrm{d}t\right)\\
&=o_{\mathbb{P}}\left(A_{51}+A_{52}+A_{53}\right),
\end{aligned}
$$

where

$$
\begin{aligned}
A_{51}&=\frac{2}{h\Delta}\sum_{i=0}^{n-1}K^2\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}\int_{i\Delta}^t\left(Y_s-Y_{i\Delta}\right)\tilde{f}(X_s)\mathrm{d}s\mathrm{d}t,\\
A_{52}&=\frac{1}{h\Delta}\sum_{i=0}^{n-1}K^2\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}\int_{i\Delta}^t\tilde{Z}^2(X_s)\mathrm{d}s\mathrm{d}t,\\
A_{53}&=\frac{2}{h\Delta}\sum_{i=0}^{n-1}K^2\left(\frac{X_{i\Delta}-x}{h}\right)\int_{i\Delta}^{(i+1)\Delta}\int_{i\Delta}^t\left(Y_s-Y_{i\Delta}\right)\tilde{Z}(X_s)\mathrm{d}W_s\mathrm{d}t.
\end{aligned}
$$

The proof of A_{51} is similar to that of A_2; the proof of A_{52} refers to A_1 and A_3; and the proof of A_{53} is the same as that of A_4. It can be shown that $A_5 = o_{\mathbb{P}}\left(\ell(T,x)\right)$ uniformly in T as $h \to 0$ and $\Delta \to 0$. Combining A_1, A_2, A_3, A_4, and A_5, we have

$$\langle \tilde{M} \rangle_T = 2\tilde{Z}^4(x) l(K_2) \ell(T,x)(1 + o_{\mathbb{P}}(1)).$$

On the other hand, if (1) in Assumption 3.13 holds, and $\sqrt{\Delta}\mathscr{M}(\mathscr{L}(u\tilde{Z}))$, $\sqrt{\Delta}\mathscr{M}(\mathscr{L}\tilde{Z}) = o_{\mathbb{P}}(1)$, then $\sqrt{\Delta}\mathscr{M}((u\tilde{Z})'\sigma)$, $\sqrt{\Delta}\mathscr{M}(\tilde{Z}'\sigma) = o_{\mathbb{P}}(\sqrt{h\ell(T,x)})$. Similar to the proof of Lemma 5.4, we obtain

$$\begin{aligned}\langle W, \tilde{M} \rangle_T &= \sqrt{\frac{4}{h\Delta}} \sum_{i=0}^{n-1} K\left(\frac{X_{i\Delta} - x}{h}\right) \int_{i\Delta}^{(i+1)\Delta} \left(Y_t - Y_{i\Delta}\right) \tilde{Z}(X_t) \mathrm{d}t \\ &= o_{\mathbb{P}}(\sqrt{h}\ell(T,x)),\end{aligned}$$

uniformly in T as $h \to 0$ and $\Delta \to 0$. Therefore,

$$\frac{\langle W, \tilde{M} \rangle_T}{\langle \tilde{M} \rangle_T} = o_{\mathbb{P}}(\sqrt{h}), \quad \text{and} \quad \frac{\langle W, \tilde{M} \rangle_T}{\langle W \rangle_T} = o_{\mathbb{P}}\left(\sqrt{h}\ell(T,x)T^{-1}\right).$$

Similar to the proof of the asymptotic normality of $\hat{g}_{q,\mathrm{LL}}(x)$ in Theorem 3.11, we have

$$\ell(T,x)^{-1/2} \tilde{M}_T \xrightarrow{d} \sqrt{2l(K_2)} \tilde{Z}^2(x) \mathbf{N}.$$

Then,

$$\sqrt{\frac{h\ell(T,x)}{\Delta}} \hat{Z}^2_{q,\mathrm{NW}}(x) = \frac{\sqrt{\ell(T,x)}}{\ell(T,x)(1 + o_{\mathbb{P}}(1))} \tilde{M}_T \xrightarrow{d} \sqrt{2l(K_2)} \tilde{Z}^2(x) \mathbf{N}.$$

For the bias term $\hat{Z}^2_{r,\mathrm{NW}}(x)$, using Lemma 5.3, if

$$h^{-4}\Delta \to 0, \quad \sqrt{\Delta}\mathscr{M}(\mathscr{L}\tilde{Z}^2) = o_{\mathbb{P}}(1), \quad \sqrt{\Delta}\mathscr{M}((\tilde{Z}^2)'\sigma) = o_{\mathbb{P}}(\sqrt{h\ell(T,x)}),$$

then

$$\frac{\frac{1}{h}\sum_{i=0}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right) \int_{i\Delta}^{(i+1)\Delta} \left(\tilde{Z}^2(X_t) - \tilde{Z}^2(X_{i\Delta})\right) \mathrm{d}t}{\frac{\Delta}{h}\sum_{i=0}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right)} = o_{\mathbb{P}}(h^2),$$

uniformly in T as $h \to 0$ and $\Delta \to 0$. For the second part of $\hat{Z}^2_{r,\mathrm{NW}}(x)$, according to Lemma 5.4, if

$$\sqrt{\Delta}\mathscr{M}(\mathscr{L}(u\tilde{f})), \quad \sqrt{\Delta}\mathscr{M}(\mathscr{L}\tilde{f}) = o_{\mathbb{P}}(1) \quad \sqrt{\Delta}\mathscr{M}((u\tilde{f})'\sigma), \quad \sqrt{\Delta}\mathscr{M}(\tilde{f}'\sigma) = o_{\mathbb{P}}(\sqrt{h\ell(T,x)}),$$

then we have

$$\frac{\frac{1}{h}\sum_{i=0}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right) \int_{i\Delta}^{(i+1)\Delta} 2(Y_t - Y_{i\Delta})\tilde{f}(X_t) \mathrm{d}t}{\frac{\Delta}{h}\sum_{i=0}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right)} = o_{\mathbb{P}}(h^2),$$

uniformly in T as $h \to 0$ and $\Delta \to 0$. In conclusion, the proof is complete. □

In Section 4, the proofs of Theorem 4.1 and Theorem 4.3 are analogous. Therefore, we will only prove Theorem 4.1, starting with a lemma.

Lemma 5.8 *Let*

$$A_f(x,h,\theta) = \frac{\sum_{i=1}^{n-1} g^2_{f_i}(x,h,\theta)}{\left(\sum_{i=1}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right)\right)^2}.$$

Then, under the conditions of Theorem 4.1, for all θ,

$$h\ell(T,x)A_f(x,h,\theta) \xrightarrow{\mathbb{P}} l(K_2)\tilde{Z}^2(x).$$

Proof According to the It ô formula, we have

$$\begin{aligned}
&h\ell(T,x)A_f(x,h,\theta)\\
=&\frac{h\ell(T,x)\sum_{i=1}^n K^2\left(\frac{X_{i\Delta}-x}{h}\right)\left(\frac{Y_{(i+1)\Delta}-Y_{i\Delta}}{\Delta}-\theta\right)^2}{(\sum_{i=1}^n K(\frac{X_{i\Delta}-x}{h}))^2}\\
=&\frac{\frac{h\ell(T,x)}{\Delta}}{\sum_{i=1}^n K(\frac{X_{i\Delta}-x}{h})}\times\frac{\sum_{i=1}^n K^2(\frac{X_{i\Delta}-x}{h})\frac{(Y_{(i+1)\Delta}-Y_{i\Delta})^2}{\Delta}}{\sum_{i=1}^n K(\frac{X_{i\Delta}-x}{h})}\\
&-2\Delta\frac{\theta\frac{h\ell(T,x)}{\Delta}}{\sum_{i=1}^n K(\frac{X_{i\Delta}-x}{h})}\times\frac{\sum_{i=1}^n K^2(\frac{X_{i\Delta}-x}{h})\frac{Y_{(i+1)\Delta}-Y_{i\Delta}}{\Delta}}{\sum_{i=1}^n K(\frac{X_{i\Delta}-x}{h})}\\
&+\Delta\frac{\theta^2\frac{h\ell(T,x)}{\Delta}}{\sum_{i=1}^n K(\frac{X_{i\Delta}-x}{h})}\times\frac{\sum_{i=1}^n K^2(\frac{X_{i\Delta}-x}{h})}{\sum_{i=1}^n K(\frac{X_{i\Delta}-x}{h})}\\
=:&A_{f_1}+A_{f_2}+A_{f_3}.
\end{aligned}$$

According to Lemma 5.1, we have $\sum_{i=1}^n K(\frac{X_{i\Delta}-x}{h}) = O_{\mathbb{P}}(h\ell(T,x)/\Delta)$. Similar to the proof of Theorem 3.14, we obtain

$$\frac{\sum_{i=1}^n K^2(\frac{X_{i\Delta}-x}{h})\frac{(Y_{(i+1)\Delta}-Y_{i\Delta})^2}{\Delta}}{\sum_{i=1}^n K(\frac{X_{i\Delta}-x}{h})} \xrightarrow{\mathbb{P}} l(K_2)\tilde{Z}^2(x).$$

Therefore, $A_{f_1} \xrightarrow{\mathbb{P}} l(K_2)\tilde{Z}^2(x)$. Similarly, $A_{f_2} \xrightarrow{\mathbb{P}} 0$, $A_{f_3} \xrightarrow{\mathbb{P}} 0$. Thus, the proof is complete. □

Proof of Theorem 4.1 For convenience, let $\lambda_f(x,h,\theta)$ in equation (4.3) be denoted as $\lambda_f(\theta)$. Similarly, denote $g_{f_i}(x,h,\theta)$ and $l_f\,(x,h,\theta)$ as $g_{f_i}(\theta)$ and $l_f(\theta)$, respectively.

We now aim to demonstrate that for $\theta = \tilde{f}(x) + \frac{\bar{\tau}(x)}{\sqrt{h\ell(T,x)}}$, where $\bar{\tau}(x)$ is given, the following holds,

$$l_f(\theta) = 2\sum_{i=1}^{n-1}\log\left(1+\lambda_f(\theta)g_{f_i}(\theta)\right) = \frac{\left(\sum_{i=1}^{n-1} g_{f_i}(\theta)\right)^2}{\sum_{i=1}^{n-1} g_{f_i}^2(\theta)} + o_{\mathbb{P}}(1),$$

which can be rewritten as

$$l_f(\theta) := \frac{B_f^2(x,h,\theta)}{A_f(x,h,\theta)} + o_{\mathbb{P}}(1), \tag{5.7}$$

where

$$B_f(x,h,\theta) = \frac{\sum_{i=1}^{n-1} g_{f_i}(\theta)}{\sum_{i=1}^{n-1} K\left(\frac{X_{i\Delta}-x}{h}\right)}. \tag{5.8}$$

According to Theorem 3.9, assuming that $h\ell(T,x) \xrightarrow{a.s.} \infty$, and $h^5\ell(T,x) \xrightarrow{a.s.} 0$, as $n,T\to\infty$, we have

$$\sqrt{h\ell(T,x)}B_f(x,h,\tilde{f}(x)) \xrightarrow{d} \sqrt{l(K_2)}\tilde{Z}(x)\mathbf{N}.$$

Finally, combining this with Lemma 5.8, we can deduce the conclusion of the Theorem. Hence, we now proceed with the proof of equation (5.7). Based on the boundedness of the kernel function K and equation (5.6), we obtain

$$\sup_{1\leqslant i\leqslant n}|g_{f_i}(\theta)|\leqslant O_{\mathbb{P}}\left(\mathscr{M}(\tilde{f})+\mathscr{M}(\tilde{Z})/\sqrt{\Delta}\right).$$

With the application of (4.4), we have

$$\begin{aligned}
0 &= \left|\sum_{i=1}^n \frac{g_{f_i}(\theta)}{1+\lambda_f(\theta)g_{f_i}(\theta)}\right| \\
&= \left|\sum_{i=1}^n \left(\frac{\lambda_f(\theta)g_{f_i}^2(\theta)}{1+\lambda_f(\theta)g_{f_i}(\theta)} - g_{f_i}(\theta)\right)\right| \\
&\geqslant \sum_{i=1}^n \frac{\lambda_f(\theta)g_{f_i}^2(\theta)}{1+\lambda_f(\theta)g_{f_i}(\theta)} - \left|\sum_{i=1}^n g_{f_i}(\theta)\right| \\
&\geqslant \frac{\lambda_f(\theta)\sum_{i=1}^n g_{f_i}^2(\theta)}{1+\sup_{1\leqslant i\leqslant n}|g_{f_i}(\theta)||\lambda_f(\theta)|} - \left|\sum_{i=1}^n g_{f_i}(\theta)\right| \\
&= \frac{\left(\sum_{i=1}^n g_{f_i}^2(\theta) - \sup_{1\leqslant i\leqslant n}|g_{f_i}(\theta)|\,|\sum_{i=1}^n g_{f_i}(\theta)|\right)\lambda_f(\theta) - |\sum_{i=1}^n g_{f_i}(\theta)|}{1+\sup_{1\leqslant i\leqslant n}|g_{f_i}(\theta)||\lambda_f(\theta)|}.
\end{aligned}$$

Therefore,

$$\left(1-\sup_{1\leqslant i\leqslant n}|g_{f_i}(\theta)|\frac{|\sum_{i=1}^n g_{f_i}(\theta)|}{\sum_{i=1}^n g_{f_i}^2(\theta)}\right)|\lambda_f(\theta)| \leqslant \frac{|\sum_{i=1}^n g_{f_i}(\theta)|}{\sum_{i=1}^n g_{f_i}^2(\theta)}. \tag{5.9}$$

With Lemma 5.1, $\frac{\Delta}{h}\sum_{i=1}^n K\left(\frac{X_{i\Delta}-x}{h}\right) = O_{\mathbb{P}}(\ell(T,x))$. Then,

$$\sum_{i=1}^n K\left(\frac{X_{i\Delta}-x}{h}\right) = O_{\mathbb{P}}\left(\frac{h\ell(T,x)}{\Delta}\right). \tag{5.10}$$

Moreover, according to Theorem 3.9, we have

$$\sqrt{h\ell(T,x)}\frac{\sum_{i=1}^n g_{f_i}(\theta)}{\sum_{i=1}^n K\left(\frac{X_{i\Delta}-x}{h}\right)} = O_{\mathbb{P}}(1).$$

Therefore,

$$\left|\sum_{i=1}^n g_{f_i}(\theta)\right| = O_{\mathbb{P}}\left(\frac{(h\ell(T,x))^{1/2}}{\Delta}\right).$$

Similarly, using (5.10) and Lemma 5.8, we obtain

$$\left|\sum_{i=1}^n g_{f_i}^2(\theta)\right| = O_{\mathbb{P}}\left(\frac{h\ell(T,x)}{\Delta^2}\right).$$

According to

$$\Delta^2\mathscr{M}(\tilde{f}^3) = o_{\mathbb{P}}(h\ell(T,x)),\ \sqrt{\Delta}\mathscr{M}(\tilde{Z}^3) = o_{\mathbb{P}}(h\ell(T,x)),$$

we have

$$\sup_{1\leqslant i\leqslant n}|g_{f_i}(\theta)|\leqslant o_{\mathbb{P}}\left(\left(\frac{h\ell(T,x)}{\Delta^2}\right)^{1/3}\right).$$

Substituting these results into (5.9), we obtain

$$|\lambda_f(\theta)| = O_{\mathbb{P}}\left(\frac{\Delta}{(h\ell(T,x))^{1/2}}\right), \quad \text{and} \quad |\lambda_f(\theta)|^2\cdot\sup_{1\leqslant i\leqslant n}|g_{f_i}(\theta)|^3 = o_{\mathbb{P}}(1).$$

Applying the Taylor expansion, we have

$$\begin{aligned}0 &= \frac{1}{n}\sum_{i=1}^{n}\frac{g_{f_i}(\theta)}{1+\lambda_f(\theta)g_{f_i}(\theta)} \\ &= \frac{1}{n}\sum_{i=1}^{n} g_{f_i}(\theta)\left[1-\lambda_f(\theta)g_{f_i}(\theta)+\frac{\lambda_f^2(\theta)g_{f_i}^2(\theta)}{1+\lambda_f(\theta)g_{f_i}(\theta)}\right] \\ &= \frac{1}{n}\sum_{i=1}^{n} g_{f_i}(\theta)-\frac{\lambda_f(\theta)}{n}\sum_{i=1}^{n} g_{f_i}^2(\theta)+\frac{1}{n}\sum_{i=1}^{n}\frac{\lambda_f^2(\theta)g_{f_i}^3(\theta)}{1+\lambda_f(\theta)g_{f_i}(\theta)},\end{aligned}$$

then we can obtain

$$\lambda_f(\theta)=\frac{\sum_{i=1}^{n} g_{f_i}(\theta)}{\sum_{i=1}^{n} g_{f_i}^2(\theta)}+o_{\mathbb{P}}(1). \tag{5.11}$$

Finally,

$$2\sum_{i=1}^{n-1}\log\left(1+\lambda_f(\theta)g_{f_i}(\theta)\right)=2\lambda_f(\theta)\sum_{i=1}^{n-1} g_{f_i}(\theta)-\lambda_f^2(\theta)\sum_{i=1}^{n-1} g_{f_i}^2(\theta)+\sum_{i=1}^{n-1}\eta_{f_i}(\theta), \tag{5.12}$$

where

$$\sum_{i=1}^{n-1}\eta_{f_i}(\theta)\leqslant\sum_{i=1}^{n-1}\left|\lambda_f^3(\theta)g_{f_i}^3(\theta)\right|\leqslant\lambda_f^3(\theta)\sup_{1\leqslant i\leqslant n}|g_{f_i}(\theta)|\sum_{i=1}^{n-1} g_{f_i}^2(\theta)=o_{\mathbb{P}}(1). \tag{5.13}$$

Combining (5.11), (5.12), and (5.13), we obtain (5.7). In summary, the theorem is proved. □

6. Simulation

In this section, we compare, under different sample sizes and observation time intervals, the finite-sample properties of the local constant and local linear estimators for the generator and diffusion term of the infinite horizon FBSDE. Furthermore, we compare the coverage rates and interval lengths of the confidence intervals based on the asymptotic normality of the estimators with those based on empirical likelihood. To evaluate the performance of the estimators, we employ the following measures:

$$\begin{aligned}\mathrm{MAE}_v &= \frac{1}{m}\sum_{i=1}^{m}\frac{1}{L}\sum_{l=1}^{L}\left|\hat{v}^l(x_i)-v(x_i)\right|, \\ \mathrm{MSE}_v &= \frac{1}{m}\sum_{i=1}^{m}\frac{1}{L}\sum_{l=1}^{L}\left|\hat{v}^l(x_i)-v(x_i)\right|^2,\end{aligned}$$

where v denotes the function to be estimated, which represents the generator and diffusion term of the FBSDE. l represents the l-th Monte Carlo replication, and $L=1000$. $\{x_i\}_{i=1}^m$ are chosen uniformly to cover the range of sample path of X. We use the Epanechnikov kernel $K(u)=\frac{3}{4}(1-u^2)I_{\{|u|\leqslant 1\}}$, and the bandwidth h is selected using the cross-validation rule.

Consider an infinite horizon FBSDE,

$$\begin{cases} X_t = x_0+\int_0^t \sigma \mathrm{d}W_s, \\ Y_t = Y_T+\int_t^T \left((2+\sigma^2)\sin X_s-2\sigma^2\cos X_s-Y_s+\sigma Z_s\right)\mathrm{d}s-\int_t^T Z_s\mathrm{d}W_s,\ \forall\ t,T>0. \end{cases} \tag{6.1}$$

The solution to (6.1) is $Y_t = 2\sin X_t$, $Z_t = 2\sigma\cos X_t$. The parameters we take are $x_0 = 0.05$, $\sigma = 0.5$.

Tables 1 and 2 report the MAE and MSE of the local constant (denoted as "NW") and local linear (denoted as "LL") estimators for the generator and diffusion term over the estimation interval $[0.45, 0.52]$, under different observation time spans T and sample sizes n. The results show that for the generator, as the observation time span T increases, the estimation error decreases significantly. For the diffusion term, an increase in sample size n leads to a reduction in the error, which is consistent with the theoretical results.

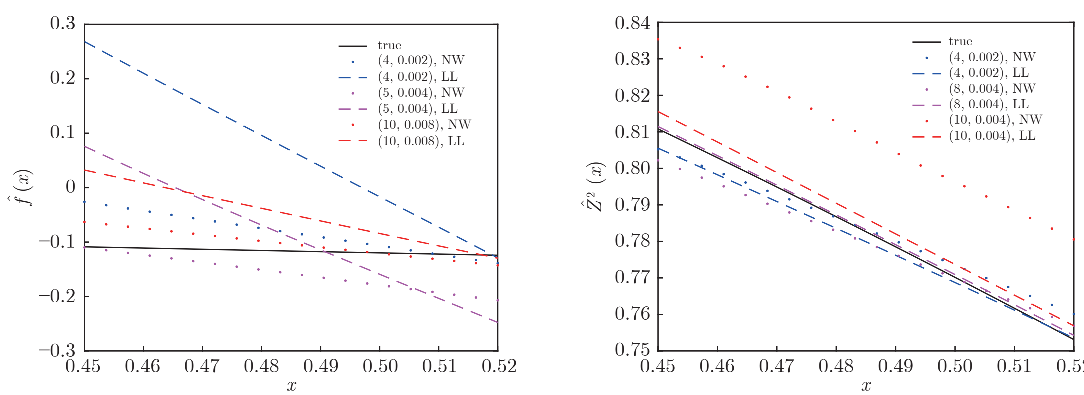
Table 1 MAE and MSE of the local constant and local linear estimators for the generator

(T, Δ)	MAE($\times 10^2$)		MSE($\times 10^4$)	
	NW	LL	NW	LL
$(4, 0.002)$	0.0525	0.0488	2.7578	2.3781
$(5, 0.004)$	0.0094	0.0105	0.0904	0.1151
$(5, 0.002)$	0.1083	0.1749	11.7613	30.6226
$(8, 0.004)$	0.0751	0.0779	5.6850	6.0629
$(10, 0.008)$	0.0617	0.0608	3.8141	3.6944

Table 2 MAE and MSE of the local constant and local linear estimator for the diffusion term

(T, n)	MAE($\times 10^3$)		MSE($\times 10^5$)	
	NW	LL	NW	LL
$(4, 2000)$	0.0435	0.0305	0.1993	0.0936
$(5, 1250)$	0.0858	0.0245	0.7936	0.0610
$(5, 2500)$	0.0359	0.0085	0.0085	0.0090
$(8, 2000)$	0.0375	0.0301	0.1423	0.0905
$(10, 2500)$	0.0476	0.0144	0.2522	0.0354

Figure 1 compares the means of the local constant and local linear estimators for the generator and diffusion term based on 1000 trajectories over the estimation interval [0.45, 0.52], with different values for T and n.



(a) Estimation of the generator (b) Estimation of the diffusion term

Figure 1 Estimation of the generator and diffusion term. $(\cdot, \cdot)$ represents (T, Δ); the "true" lines represent the true function

Tables 3 and 4 report the performance of the confidence intervals constructed using the asymptotic normality of the local constant estimator (denoted as "Am") and the empirical likelihood (denoted as "EL") for the generator and diffusion term at $x = 0.5216$. In this context, $\tilde{f}(x)$ and $\tilde{Z}^2(x)$ represent the true values of the generator and diffusion term at x, respectively, and $\widehat{f}_{\mathrm{NW}}(x)$ and $\widehat{Z}^2_{\mathrm{NW}}(x)$ denote the local constant estimators of the generator and diffusion term at x, respectively. We present the effects of different T, n, and h on the average interval length (denoted as "length") and coverage rate (denoted as "Cov"), based on 1000 sample paths. In Table 3, the variance estimator is given by $v\hat{a}r = (h\bar{\ell}(T,x))^{-1} l(K_2) \widehat{Z}^2_{\mathrm{NW}}(x)$, where $\bar{\ell}(T,x) = \frac{\Delta}{h} K(\frac{X_{i\Delta}-x}{h})$, and in Table 4, it is $v\hat{a}r = 2\Delta(h\bar{\ell}(T,x))^{-1} l(K_2) \widehat{Z}^4_{\mathrm{NW}}(x)$.

Table 3 Comparison of two types of confidence intervals for the generator

(T,Δ)	h	length		Cov(%)		$\tilde{f}(x)$	$\widehat{f}_{\mathrm{NW}}(x)$	$\sqrt{v\hat{a}r}$
		EL	Am	EL	Am			
(5, 0.004)	0.8	2.0489	2.0094	92.0	92.8	−0.1246	−0.1505	0.5126
	1.0	1.8415	1.7807	93.8	93.4		−0.1429	0.4543
	1.2	1.8415	1.6194	93.9	93.1		−0.1396	0.4131
(8, 0.004)	0.8	1.7890	1.7638	90.8	92.1	−0.1246	−0.1697	0.4499
	1.0	1.5912	1.5508	92.6	93.7		−0.1587	0.3956
	1.2	1.4505	1.3974	94.2	94.1		−0.1490	0.3565
(10, 0.004)	0.8	1.6649	1.6370	91.4	92.9	−0.1246	−0.1223	0.4176
	1.0	1.4756	1.4340	93.6	93.5		−0.1269	0.3658
	1.2	1.3357	1.2857	95.0	94.4		−0.1313	0.3280

Table 4 Comparison of two types of confidence intervals for the diffusion term

(T,Δ)	h	length		Cov(%)		$\tilde{Z}^2(x)$	$\widehat{Z}^2_{\mathrm{NW}}(x)$	$\sqrt{v\hat{a}r}$
		EL	Am	EL	Am			
(4, 0.002)	0.08	0.4010	0.4815	94.0	96.0	0.7517	0.7865	0.1228
	0.20	0.2455	0.2399	94.5	92.5		0.7547	0.0612
	0.40	0.1930	0.1684	67.5	57.5		0.7471	0.0430
(5, 0.004)	0.08	0.5007	0.6016	93.0	94.5	0.7517	0.7647	0.1535
	0.20	0.3603	0.3641	94.5	94.0		0.7588	0.0929
	0.40	0.2497	0.2322	78.5	75.5		0.7415	0.0592
(8, 0.004)	0.08	0.4380	0.5020	92.5	94.0	0.7517	0.7641	0.1281
	0.20	0.2909	0.2932	94.5	92.0		0.7522	0.0748
	0.40	0.2165	0.2030	77.0	73.0		0.7445	0.0518

From these results, it is evident that the confidence intervals based on the empirical likelihood method outperform those based on the traditional asymptotic normality. In particular, the empirical likelihood intervals generally exhibit shorter average lengths and higher coverage rates. This suggests that the empirical likelihood method is more efficient in utilizing sample information, and it leads to more accurate and reliable interval estimates in practical applications.

Acknowledgements

This work was supported by the National Key R&D Program of China (Grant No. 2023YFA1008701) and National Natural Science Foundation of China (Grant Nos. 12431017 and 12326332).

References

[1] Aït-Sahalia, Y. and Park, J. Y., Bandwidth selection and asymptotic properties of local nonparametric estimators in possibly nonstationary continuous-time models, Journal of Econometrics, 2016, 192(1): 119–138.

[2] Bandi, F. M. and Phillips, P. C., Fully nonparametric estimation of scalar diffusion models, Econometrica, 2003, 71(1): 241–283.

[3] Bandi, F. M. and Moloche, G., On the functional estimation of multivariate diffusion processes, Econometric Theory, 2018, 34(4): 896–946.

[4] Bandi, F. M. and Renò, R., Nonparametric stochastic volatility, Econometric Theory, 2018, 34(6): 1207–1255.

[5] Borkovec, M. and Klüppelberg, C., Extremal behavior of diffusion models in finance, Extremes, 1998, 1(1): 47–80.

[6] Chen, X. and Lin, L., Nonparametric estimation for FBSDEs models with applications in finance, Communications in Statistics-Theory and Methods, 2010, 39(14): 2492–2514.

[7] Fan, J. and Zhang, C., A reexamination of diffusion estimators with applications to financial model validation, Journal of the American Statistical Association, 2003, 98(461): 118–134.

[8] Florens-Zmirou, D., On estimating the diffusion coefficient from discrete observations, Journal of Applied Probability, 1993, 30(4): 790–804.

[9] Jacod, J., Non-parametric kernel estimation of the coefficient of a diffusion, Scandinavian Journal of Statistics, 2000, 27(1): 83–96.

[10] Kutoyants, Y. A., Approximation of the solution of the backward stochastic differential equation: Small noise, large sample and high frequency cases, Proceedings of the Steklov Institute of Mathematics, 2014, 287(1): 133–154.

[11] Kutoyants, Y. A., On approximation of BSDE and multi-step MLE-processes, Probability, Uncertainty and Quantitative Risk, 2016, 1: 4.

[12] Pardoux, E. and Peng, S., Adapted solution of a backward stochastic differential equation, Systems & Control Letters, 1990, 14(1): 55–61.

[13] Pardoux, E., Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order, In: Decreusefond, L., Øksendal, B., Gjerde, J. and Üstünel, A. S.(eds.), Stochastic Analysis and Related Topics VI. Progress in Probability, Vol. 42, Birkhäuser, Boston, MA, 1998.

[14] Peng, S., Probabilistic interpretation for systems of quasilinear parabolic partial differential equations, Stochastics and Stochastics Reports, 1991, 37(1–2): 61–74.

[15] Revuz, D. and Yor, M., Continuous Martingales and Brownian Motion, Springer, Berlin, Heidelberg, 1999.

[16] Stanton, R., A nonparametric model of term structure dynamics and the market price of interest rate risk, The Journal of Finance, 1997, 52(5): 1973–2002.

[17] Song, Y. and Wang, H., Central limit theorems of local polynomial threshold estimator for diffusion processes

with jumps, Scandinavian Journal of Statistics, 2018, 45(3): 644–681.

[18] Su, Y. and Lin, L., Semi-parametric estimation for forward-backward stochastic differential equations, Communications in Statistics-Theory and Methods, 2009, 38(11): 1759–1775.

[19] Xu, K. L., Empirical likelihood-based inference for nonparametric recurrent diffusions, Journal of Econometrics, 2009, 153(1): 65–82.

[20] Xu, K. L., Reweighted functional estimation of diffusion models, Econometric Theory, 2010, 26(2): 541–563.

[21] Zhang, D., Local linear estimators for reflected diffusions, Journal of Nonparametric Statistics, 2024, https://doi.org/10.1080/10485252.2024.2426099.