



MONASH University

Australia

Department of Econometrics and Business Statistics

<http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/>

**Expansion of Lévy Process Functionals and Its Application in
Statistical Estimation**

Chaohua Dong and Jiti Gao

January 2012

Working Paper 02/12

Expansion of Lévy Process Functionals and Its Application in Statistical Estimation

CHAOHUA DONG
School of Economics
The University of Adelaide

JITI GAO
Department of Econometrics and Business Statistics
Monash University

Abstract

In this paper, expansions of functionals of Lévy processes are established under some Hilbert spaces and their orthogonal bases. From practical standpoint, both time-homogeneous and time-inhomogeneous functionals of Lévy processes are considered. Several expansions and rates of convergence are established. In order to state asymptotic distributions for statistical estimators of unknown parameters involved in a general regression model, we develop a general asymptotic theory for partial sums of functionals of Lévy processes. The results show that these estimators of the unknown parameters in different situations converge to quite different random variables. In addition, the rates of convergence depend on various factors rather than just the sample size.

Key Words: Expansion, Lévy Process, Orthogonal Series, Statistical Estimation.

JEL Classifications: C13, C14, C22.

1 Introduction

Stochastic differential equations driven by a Lévy process under some conditions have solutions in the form of functionals of the underlying process. Such equations are used extensively in economics, finance and engineering disciplines to describe random phenomena in both theory and practice. Meanwhile, some empirical studies show that many

data sets admit nonlinearity and nonstationarity. Consequently, a number of nonparametric and semiparametric models and kernel-based methods have been proposed to deal with both nonlinearity and nonstationarity simultaneously. Existing studies mainly discuss the employment of nonparametric kernel estimation methods. Such studies include Phillips and Park (1998), Park and Phillips (1999), Park and Phillips (2001), Karlsen and Tjøstheim (2001), Karlsen et al. (2007), Cai et al. (2009), Phillips (2009), Wang and Phillips (2009a,b), Xiao (2009), and Gao and Phillips (2011).

However, such kernel-based estimation methods are not applicable to establish closed-form expansions of functionals of Lévy processes. In the stationary case, the literature already discusses how series approximations may be used in dealing with stationary time series models, such as, Ai and Chen (2003), Chapter 2 of Gao (2007) and Li and Racine (2007). In addition, although the celebrated Black-Scholes option pricing formula described the price of the financial product as a functional of Brownian motion, literature has pointed out that there are some significant drawbacks in this formula. For example, empirical evidence suggests that log returns do not behave according to a normal distribution (see Schoutens, 2003). Hence, the researcher realizes that one would need to include other stochastic processes (not just Brownian motion) when one needs to formulate a continuous-time stochastic model in order to depict some stochastic phenomenon or scientific data set.

Therefore, there is need to study functionals of Lévy process, $Z(t)$, in the both cases of time-homogeneity and time-inhomogeneity. Note that one powerful way of dealing with such problems is to decompose the process, say $f(Z(t))$ or $f(t, Z(t))$, where the functional form is unknown, into an orthogonal series in some Hilbert space, such that once one has obtained observed values of the process, the coefficients involved in the series can be estimated using an econometric method. Actually, there is long history that there exists a close connection between stochastic processes and orthogonal polynomials. For example, the so-called Karlin-McGregor representation expresses the transition probability of birth and death process by means of a spectral representation in terms of orthogonal polynomials. Some people clearly feel the potential importance of orthogonal polynomials in probability theory. Schoutens (2000), for instance, gives an extensive discussion about relations between stochastic processes and orthogonal polynomials.

In this paper, we establish some general theory and methodology for the expansion of a class of functionals of Lévy processes. As an application, we shall estimate an unknown

function of the form $m(t, z)$ involved in the following model:

$$Y(t) = m(t, Z(t)) + \varepsilon(t), \quad t \in [0, \infty), \quad (1.1)$$

where $Z(t)$ is a Lévy process that covers both the continuous (such as Brownian motion) and the discrete (such as Poisson process) cases, $\varepsilon(t)$ is an error process with zero mean and finite variance, and $m(t, z)$ is an unknown function of (t, z) .

As far as we are aware, there is no discussion about how to estimate $m(t, z)$ by a non- or semi-parametric method in the literature. Even in the discrete case where $t = 1, 2, \dots$, it is not clear whether a nonparametric kernel method can provide a consistent estimator for $m(t, z)$. Part of the contribution of this paper is to establish an asymptotically consistent estimator of $m(t, z)$ and the resulting asymptotic theory in each of the three sampling situations: a) the case where $Z_t = Z(t)$ at $t = 1, 2, \dots$, b) the case where $Z_{t,n} = Z\left(\frac{tT}{n}\right)$ at $t = 1, 2, \dots$, and c) the case where $Z_{t,n} = Z\left(\frac{tT_n}{n}\right)$ at $t = 1, 2, \dots$ and with $T_n \rightarrow \infty$.

The estimation methodology proposed in this paper is summarized as follows. We shall employ an appropriate polynomial sequence that is orthogonal with respect to either the probability density or the probability distribution of $Z(t)$ depending on whether $Z(\cdot)$ is continuous or discrete. We then expand the unknown function $m(t, Z(t))$ into an orthogonal series in some Hilbert space in terms of the polynomial sequence. We then propose using a semiparametric least squares (SLS) estimation method to estimate $m(t, z)$ by $\hat{m}(t, z)$. To establish an asymptotic theory for \hat{m} , we introduce a general asymptotic theory to deal with the sample mean and sample covariance of four classes of functionals of Lévy processes. It is noteworthy to point out that the established asymptotic theory considerably extends some existing results, such as Park and Phillips (1999, 2001), and Wang and Phillips (2009a).

With the advantage of expanding an unknown functional into an orthogonal series, the proposed method can be used to deal with some estimation problems in economics, finance and engineering. For example, there are a number of studies involving models with conditional moment restriction containing an unknown functional, such as Ai and Chen (2003, 2007), and Chen and Ludvigson (2009). Since existing theory for expansions of functionals of stationary processes is not directly applicable, the proposed expansion and estimation method in this paper is useful and significant in both theory and applications.

The organization of the rest of the paper is as follows. Section 2 is devoted to the expansions of functionals of Lévy processes for the cases of time-homogeneous function-

als and time-inhomogeneous on both finite and infinite horizons. Section 3 develops an asymptotic theory. Section 4 systematically discusses the estimation theory of model (1.1) for the cases of infinite-time, finite-time and infinite-time sequence horizons. Section 5 gives some conclusions and discussions. Appendix A summarizes some basic results for both the existence and explicit expression of an orthogonal polynomial system associated with an underlying Lévy process. Appendix B includes all remarks and justifications of some results and assumptions. Appendix C mainly shows several lemmas which are crucial for the proofs of the main results. Appendix D contains the proofs of the main theorems.

2 Orthogonal expansion of functionals of Lévy processes

This section is devoted to the expansion of both time-homogeneous functional $f(Z(t))$ and time-inhomogeneous functional $f(t, Z(t))$ for $t \in [0, T]$ with fixed T and $t \in (0, \infty)$ where $(Z(t), t \geq 0)$ is a Lévy process with $\rho(t, x)$ being either the probability density or the probability distribution function of $Z(t)$ depending on whether $Z(t)$ is continuous or discrete. Suppose that $Z(t)$ admits a classical orthonormal polynomial system $Q_i(t, x)$ with the weight function $\rho(t, x)$. Note that this requires $\rho(t, x)$ satisfy the so-called boundary condition and another one specified in Appendix A. $Q_i(t, x)$ is a normalized polynomial, viz., $Q_i(t, x) = \frac{1}{d_i(t)} y_i(t, x)$ so that $\|Q_i(t, x)\| = 1$ in the relevant space where $y_i(t, x)$ is a polynomial solution in both t and x of the so-called hypergeometric differential/difference equation. Readers consult Appendix A for detailed explanation. Many notations and assertions are related to Appendix A. Let I be the support of $\rho(t, x)$ and μ signify Lebesgue measure on line. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space on which $Z(t)$ is defined.

2.1 Expansion of homogeneous functionals of Lévy processes

Consider a function space for $t > 0$

$$L^2(I, d\Psi_t(x)) = \left\{ f(x) : \int_I f^2(x) d\Psi_t(x) < \infty \right\}, \quad (2.1)$$

where $\Psi_t(x)$ is the distribution function of Lévy process $Z(t)$.

According to Billingsley (1995, p249), $L^2(I, d\Psi_t(x))$ is a Hilbert space. Given that $\Psi_t(x)$ satisfies a sufficient condition, viz., there exists a constant $c > 0$, such that $\int e^{c|x|} d\Psi_t(x) < \infty$ for each fixed t , the system $Q_i(t, x)$ is not only orthonormal but also complete in $L^2(I, d\Psi_t(x))$ (see, for example, Nikiforov and Uvarov (1988, p. 57)). Indeed, there are

many Lévy processes satisfying this sufficient condition. For instance, the Laguerre polynomial system associated with the Gamma process satisfies it with $c < 1$, and both the Hermite polynomial system with the density of Brownian motion as its weight and the Charlier polynomial system with the probability distribution of Poisson process as its weight satisfy the condition with any $c > 0$.

Additionally, in the Hilbert space $L^2(I, d\Psi_t(x))$, the scalar product and the induced norm are defined as follows:

$$(f, g) = \int f(x)g(x)d\Psi_t(x) \quad \text{and} \quad \|f\| = \sqrt{(f, f)}.$$

Construct a mapping, for $f(x) \in L^2(I, d\Psi_t(x))$, $\mathcal{T} : f \mapsto f(Z(t))$. Since $E[f^2(Z(t))] = \int_I f^2(x)d\Psi_t(x) < \infty$, $f(Z(t))$ is an element of $L^2(\Omega)$, the collection of all random variables with finite second moments. Accordingly, the image of \mathcal{T} , denoted by Θ , is a subset of $L^2(\Omega)$. Hence, the elements in Θ admit the norms and scalar products, namely, $\langle f(Z(t)), g(Z(t)) \rangle_{\Theta} = E[f(Z(t))g(Z(t))]$ and $\|f(Z(t))\|_{\Theta} = \sqrt{\langle f(Z(t)), f(Z(t)) \rangle}$, the induced norm. The following lemmas give the properties of \mathcal{T} and Θ .

Lemma 2.1. *The mapping \mathcal{T} has the following properties:*

(1) \mathcal{T} is linear; (2) \mathcal{T} is an one-to-one mapping from $L^2(I, d\Psi_t(x))$ to Θ ; (3) \mathcal{T} is an isomorphism.

Lemma 2.2. Θ is a closed subspace of $L^2(\Omega)$, hence it is a Hilbert space.

Lemma 2.3. If $\{p_i(x)\}_{i=0}^{\infty}$ is any orthonormal basis in $L^2(I, d\Psi_t(x))$, then $\{\mathcal{T}(p_i)\}_{i=0}^{\infty}$ is an orthonormal basis in Θ . Particularly, $\{Q_i(t, Z(t))\}_{i=0}^{\infty}$, $t > 0$, is an orthonormal basis in Θ .

The following theorem is a consequence of the above lemmas.

Theorem 2.1. *Suppose that Lévy process $(Z(t), t > 0)$ admits a classical orthonormal polynomial system $Q_i(t, x)$ with weight $\rho(t, x)$. For any element $f(Z(t)) \in \Theta$, it has a Fourier series expansion*

$$f(Z(t)) = \sum_{i=0}^{\infty} c_i(t, f)Q_i(t, Z(t)), \quad (2.2)$$

where $c_i(t, f) = \langle f(Z(t)), Q_i(t, Z(t)) \rangle_{\Theta}$.

See some examples of expansion in Remark B.1.

Let k be a truncation parameter for i . The truncation series of (2.2) is defined as

$$f_k(Z(t)) = \sum_{i=0}^k c_i(t, f) Q_i(t, Z(t)).$$

With the notations D (differentiate or difference operator with respect to x only throughout the paper), $\rho_h(t, x)$ and $\psi(t) > 0$ defined in Appendix A, we have the following theorem.

Theorem 2.2. *Let $(Z(t), t > 0)$ be a Lévy process satisfying conditions in Theorem 2.1. Suppose further that $D^h f(x) \in L^2(I, \rho_h(t, x))$ for $h = 0, 1, \dots, r$. Then*

$$\|f(Z(t)) - f_k(Z(t))\|_{\Theta}^2 \leq \frac{1}{k^r} R_k^2(t, D^r f), \quad (2.3)$$

where $R_k^2(t, D^r f) = (1 + o(1))[\psi(t)]^r \sum_{i=k+1}^{\infty} c_i^2(t, D^r f) \rightarrow 0$ as $k \rightarrow \infty$ for every $t > 0$.

2.2 Expansion of time-inhomogeneous functionals of Lévy processes on finite time horizon

In this subsection we shall dwell on the expansion of $f(t, Z(t))$ for $t \in [0, T]$ with fixed T .

Consider function space

$$L^2([0, T] \times I, \nu) = \left\{ f(t, x) : \int_I f^2(t, x) d\Psi_t(x) < \infty, \text{ for each } t \in [0, T], \right. \\ \left. \text{and } \int_0^T \int_I f^2(t, x) d\nu < \infty \right\},$$

where ν is the product measure of $\Psi_t(x)$ and Lebesgue measure μ .

We abbreviate the notation of the space as $L^2([0, T] \times I)$. As a conventional L^2 space, $L^2([0, T] \times I)$ is a Hilbert space with scalar product

$$(f_1(t, x), f_2(t, x)) = \int_0^T \int_I f_1(t, x) f_2(t, x) d\Psi_t(x) dt.$$

Since $\{Q_i(t, x)\}$ is an orthonormal basis for $L^2(I, d\Psi_t(x))$ and $\{\varphi_{jT}(t)\}$, where $\varphi_{0T} = \sqrt{\frac{1}{T}}$, $\varphi_{jT} = \sqrt{\frac{2}{T}} \cos \frac{j\pi t}{T}$ for $j \geq 1$, is an orthonormal basis in $L^2([0, T], \mu)$, according to Problem 12 of Dudley (2003, p173), $\{Q_i(t, x) \varphi_{jT}(t)\}$ is an orthonormal basis in $L^2([0, T] \times I)$.

Construct a mapping \mathcal{T} from $L^2([0, T] \times I)$ to a set of stochastic processes,

$$\mathcal{T} : f(t, x) \mapsto f(t, Z(t)), \quad \text{for } f(t, x) \in L^2([0, T] \times I).$$

Denote the image of \mathcal{T} by Ξ . Define operation $\langle f_1(t, Z(t)), f_2(t, Z(t)) \rangle_{\Xi} = \int_0^T E[f_1(t, Z(t))f_2(t, Z(t))]dt$ on Ξ . Obviously, $\langle \cdot, \cdot \rangle_{\Xi}$ is an inner product on Ξ . Meanwhile, \mathcal{T} and Ξ enjoy the properties established in Lemma 2.1–2.3. We then assert that Ξ is a Hilbert space and $\{Q_i(t, Z(t))\varphi_{jT}(t)\}$ ($i, j = 0, 1, \dots$) is an orthonormal basis in Ξ . The following theorem is obtained from Hilbert space theory.

Theorem 2.3. *In Ξ , any element $f(t, Z(t))$ admits a Fourier series expansion*

$$f(t, Z(t)) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} c_{ij} \varphi_{jT}(t) Q_i(t, Z(t)), \quad (2.4)$$

where $c_{ij} = \langle f(t, Z(t)), Q_i(t, Z(t))\varphi_{jT}(t) \rangle_{\Xi}$.

Because $c_{ij} = \int_0^T E[f(t, Z(t))Q_i(t, Z(t))\varphi_{jT}(t)]dt := \int_0^T c_i(t, f)\varphi_{jT}(t)dt$, expansion (2.4) can be regarded as a two-step expansion, that is, expand $f(t, Z(t))$ first in terms of $\{Q_i(t, Z(t))\}$ obtaining coefficients $c_i(t, f) = E[f(t, Z(t))Q_i(t, Z(t))]$, then expand $c_i(t, f)$ in terms of $\{\varphi_{jT}(t)\}$ on $[0, T]$.

Notice that from Parseval equality it follows that

$$\|f(t, Z(t))\|_{\Xi}^2 = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} c_{ij}^2 = \sum_{i=0}^{\infty} \|c_i(t, f)\|_{L^2[0, T]}^2. \quad (2.5)$$

Given a bundle of truncation parameters k for i and p_i for j 's, we define the truncation series of (2.4) as follows

$$f_{k,p}(t, Z(t)) = \sum_{i=0}^k \sum_{j=1}^{p_i} c_{ij} Q_i(t, Z(t))\varphi_{jT}(t). \quad (2.6)$$

Denote $p_{\min} = \min\{p_1, \dots, p_k\}$ and $p_{\max} = \max\{p_1, \dots, p_k\}$ throughout the paper.

Theorem 2.4. *Suppose that functional $f(t, Z(t)) \in \Xi$ and that $D^h f(t, x)$, $h = 1, \dots, r$, are in the space $L^2(I, \rho_h(t, x))$ for each $t > 0$. Moreover, $\sqrt{\psi(t)}^r D^r f(t, x) \in L^2(I \times [0, T])$. In addition, for each $i \geq 0$, $c_i(t, f) \in C^2[0, T]$ and $\|c_i''(t, f)\|_{L^2[0, T]}$ is uniformly bounded in i . Then,*

$$\|f(t, Z(t)) - f_{k,p}(t, Z(t))\|_{\Xi}^2 \leq \frac{1}{k^r} R_k^2 + C(k, p) \frac{k}{p_{\min}^4}, \quad (2.7)$$

where $R_k^2 = (1 + o(1)) \sum_{i=k+1}^{\infty} \|c_i(t, \sqrt{\psi(t)}^r D^r f)\|_{L^2[0, T]}^2 \rightarrow 0$ as $k \rightarrow \infty$, $C(k, p) = \pi^{-4} \max_{0 \leq i \leq k} \sum_{j=p_{\min}+1}^{\infty} b_j^2(c_i'')$, in which $b_j(c_i'')$ stands for the j -th coefficient in the expansion of $c_i''(t, f)$. Here we assume that $\frac{k}{p_{\min}^4} \rightarrow 0$.

See Remark B.2 for some discussions.

2.3 Expansion of time-inhomogeneous functionals of Lévy processes on infinite time horizon

Let $t \in (0, \infty)$ and consider function space defined by

$$L^2(\mathbb{R}^+ \times I, \nu) = \left\{ f(t, x) : \int_I f^2(t, x) d\Psi_t(x) < \infty, \text{ for each } t \in (0, \infty), \right. \\ \left. \text{and } \int_0^\infty \int_I f^2(t, x) d\nu < \infty \right\},$$

where ν is the product of $\Psi_t(x)$ and Lebesgue measure μ .

We abbreviate the notation of the space as $L^2(\mathbb{R}^+ \times I)$. Apparently, it is a L^2 space so that it is a Hilbert space. The inner product is conventional $(f_1(t, x), f_2(t, x)) = \int_{\mathbb{R}^+} \int_I f_1(t, Z(t)) f_2(t, Z(t)) d\Psi_t(x) dt$, which can induce a norm $\|f(t, x)\|$.

As $\{Q_i(t, x)\}$ and $\{\mathcal{L}_j(t)\}$ ($\mathcal{L}_j(t) = e^{-t/2} L_j(t)$, with $L_j(t)$ being Leguerre polynomial sequence) are orthonormal bases in $L^2(I, d\Psi_t(x))$ and $L^2(\mathbb{R}^+, \mu)$ respectively, $\{Q_i(t, x)\mathcal{L}_j(t)\}$ is an orthonormal basis in $L^2(\mathbb{R}^+ \times I)$.

Similarly, construct a mapping from $L^2(\mathbb{R}^+ \times I)$ to a set of stochastic processes

$$\mathcal{T} : f(t, x) \mapsto f(t, Z(t)), \quad \text{for } f(t, x) \in L^2(\mathbb{R}^+ \times I).$$

It is clear that \mathcal{T} is linear, so that the image set, denoted by Λ , is a linear real vector space. Note that after defining $\langle f_1(t, Z(t)), f_2(t, Z(t)) \rangle_\Lambda = \int_0^\infty E[f_1(t, Z(t)) f_2(t, Z(t))] dt$, Λ becomes an inner product space equipped with induced norm. Analogously, \mathcal{T} and Λ enjoy the properties in Lemma 2.1–2.3. As a result, $\{Q_i(t, Z(t))\mathcal{L}_j(t)\}$ is an orthonormal basis in Λ .

Theorem 2.5. *In Λ , any element $f(t, Z(t))$ admits Fourier series expansion*

$$f(t, Z(t)) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} b_{ij} \mathcal{L}_j(t) Q_i(t, Z(t)), \quad (2.8)$$

where $b_{ij} = \langle f(t, Z(t)), Q_i(t, Z(t))\mathcal{L}_j(t) \rangle_\Lambda$.

Because $b_{ij} = \int_{\mathbb{R}^+} E[f(t, Z(t)) Q_i(t, Z(t))] \mathcal{L}_j(t) dt := \int_{\mathbb{R}^+} b_i(t, f) \mathcal{L}_j(t) dt$, expansion (2.8) can be regarded as a two-step expansion, that is, expand $f(t, Z(t))$ first in terms of $\{Q_i(t, Z(t))\}$ obtaining coefficients $b_i(t, f) = E[f(t, Z(t)) Q_i(t, Z(t))]$, then expand $b_i(t, f)$ in terms of $\{\mathcal{L}_j(t)\}$ on $(0, \infty)$.

Notice that from Parseval equality it follows that

$$\|f(t, Z(t))\|_{\Lambda}^2 = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} b_{ij}^2 = \sum_{i=0}^{\infty} \|b_i(t, f)\|_{L^2(\mathbb{R}^+)}^2. \quad (2.9)$$

Given a bundle of truncation parameters k for i and p_i for j 's, we define the truncation series of (2.8) as follows:

$$f_{k,p}(t, Z(t)) = \sum_{i=0}^k \sum_{j=0}^{p_i} b_{ij} Q_i(t, Z(t)) \mathcal{L}_j(t). \quad (2.10)$$

Theorem 2.6. *Suppose that functional $f(t, Z(t)) \in \Lambda$, that $D^h f(t, x)$, $h = 1, \dots, r_1$, are in the space $L^2(I, \rho_h(t, x))$ for all $t > 0$. Moreover, $\sqrt{\psi(t)}^{r_1} D^{r_1} f(t, x) \in L^2(I \times \mathbb{R}^+)$. In addition, for each $i \geq 0$, $b_i(t, f)$ is differentiable up to r_2 -th order and $b_i(t, f)$ and $\sqrt{v(t)}^{r_2} \frac{d^h}{dt^h} b_i(t, f)$ are all in $L^2(\mathbb{R}^+)$ for $h = 1, \dots, r_2$. Then,*

$$\|f(t, Z(t)) - f_{k,p}(t, Z(t))\|_{\Lambda}^2 \leq \frac{1}{k^{r_1}} R^2(k) + C(k, p) \frac{k}{p_{\min}^{r_2}}, \quad (2.11)$$

where $R^2(k) = (1 + o(1)) \sum_{i=k+1}^{\infty} \|b_i(t, \sqrt{\psi(t)}^{r_1} D^{r_1} f)\|_{L^2(\mathbb{R}^+)}^2 \rightarrow 0$, as $k \rightarrow \infty$, $C(k, p) = (1 + o(1)) \max_{0 \leq i \leq k} \sum_{j=p_{\min}+1}^{\infty} [a_{j-r_1}^{(r_1)}(\tilde{b}_i(t))]^2$, in which $\tilde{b}_i(t) = t^{r_2/2} e^{-t/2} [b_i(t, f) e^{t/2}]^{(r_2)}$ and $a_{j-r_2}^{(r_2)}(\tilde{b}_i(t))$ are the coefficients of the expansion of $\tilde{b}_i(t)$ in terms of $\mathcal{L}_j^{(r_2)}(t)$. Here we assume that $C(k, p) \frac{k}{p_{\min}^{r_2}} \rightarrow 0$.

We have a similar remark for Theorem 2.6 to that for Theorem 2.5.

Theorems 2.1–2.6 show that either a homogenous or in-homogenous functionals of $Z(t)$ can be expanded as an orthogonal series. In order to apply such expansions to establish an estimation theory for model (1.1) in Section 4 below, we first develop some asymptotic properties for partial sums of several classes of functionals of Lévy processes.

3 Asymptotic theory

In this section we shall establish some asymptotic theory for two basic classes of functionals $f(\cdot, \cdot)$ defined below. We then shall define two more general classes of functionals $F(t, x)$ for $t > 0$ and $x \in \mathbb{R}$, $\mathcal{T}(HI)$ and $\mathcal{T}(HH)$, and investigate the asymptotic theory for these functionals on both sample mean and sample covariance.

Given a triangular array $x_{s,n}$ ($x_{0,n} = 0$ by definition), $1 \leq s \leq n$, constructed from some underlying time series, we assume that $x_{[nr],n}$ ($0 \leq r \leq 1$) converges in distribution to

a stochastic process $W(r)$ on $D[0, 1]$ with respect to the Skorohod topology, where $D[0, 1]$ stands for the space of real-valued functions that are right continuous with left limits. It is known that there are many cases in which $\{x_{s,n}\}$ satisfies this condition, and in some suitable probability space it can be shown that $\sup_{0 \leq r \leq 1} |x_{[nr],n} - W(r)| = o_P(1)$. Readers consult Phillips (1987), Park and Phillips (1999, 2001), Wang and Phillips (2009a) and Gao and Phillips (2010) for detailed discussion.

We now state the following assumption on $x_{s,n}$.

Assumption A

- (a) Suppose that $x_{[nr],n}$ ($0 \leq r \leq 1$) converges in distribution to a stochastic process $W(r)$ on $D[0, 1]$ with respect to the Skorohod topology. Let $W(r)$ admit a continuous local-time $L_W(r, s)$.
- (b) In some suitable probability space there exists a stochastic process $W(r)$ that admits a continuous local-time $L_W(r, s)$ such that $\sup_{0 \leq r \leq 1} |x_{[nr],n} - W(r)| = o_P(1)$.
- (c) Denote for ϵ ($0 < \epsilon < 1$) that $\Omega_n(\epsilon) = \{(l, k) : \epsilon n \leq k \leq (1 - \epsilon)n, k + \epsilon n \leq l \leq n\}$. For all $0 \leq k < l \leq n$, there exist a sequence of constants $d_{l,k,n}$ and a sequence of σ -fields $\mathcal{F}_{n,k}$ where $\mathcal{F}_{n,0} = \{\emptyset, \Omega\}$, such that

- (i) for some $m_0 > 0$ and $C > 0$, $\inf_{(l,k) \in \Omega_n(\epsilon)} d_{l,k,n} \geq \epsilon^{m_0}/C$ as $n \rightarrow \infty$,

$$\lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=(1-\epsilon)n}^n \frac{1}{d_{l,0,n}} = 0, \quad (3.1)$$

$$\lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} \max_{0 \leq k \leq (1-\epsilon)n} \sum_{l=k+1}^{k+\epsilon n} \frac{1}{d_{l,k,n}} = 0, \quad (3.2)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \max_{0 \leq k \leq n-1} \sum_{l=k+1}^n \frac{1}{d_{l,k,n}} < \infty. \quad (3.3)$$

- (ii) Suppose that $x_{k,n}$ are adapted to $\mathcal{F}_{n,k}$. Moreover, if $x_{k,n}$ are continuous variables, conditional on $\mathcal{F}_{n,k}$, $(x_{l,n} - x_{k,n})/d_{l,k,n}$ has a density $h_{l,k,n}$ which is uniformly bounded by a constant K and

$$\lim_{\delta \rightarrow 0} \lim_{n \rightarrow \infty} \sup_{(l,k) \in \Omega_n(\delta^{1/(2m_0)})} \sup_{|u| < \delta} |h_{l,k,n}(u) - h_{l,k,n}(0)| = 0. \quad (3.4)$$

If $x_{k,n}$ are discrete variables, conditional on $\mathcal{F}_{n,k}$, $(x_{l,n} - x_{k,n})/d_{l,k,n}$ has a probability distribution $F_{l,k,n}(x)$ and its distribution function $F_{l,k,n}(x)$ satisfies

$$\lim_{\delta \rightarrow 0} \lim_{n \rightarrow \infty} \sup_{(l,k) \in \Omega_n(\delta^{1/(2m_0)})} \sup_{|u| < \delta} |F_{l,k,n}(u) - F_{l,k,n}(0)| = 0. \quad (3.5)$$

Remark 3.1. Assumption A is almost the same as the conditions in the univariate function case in Wang and Phillips (2009a) except that we concern both continuous and discrete variables in A (c). We shall discuss the condition (3.5) later. Note that Assumption A is quite weak which is discussed in the literature. As a consequence, the following theorems are generally applicable.

Also, we remark that this situation particularly accommodates any Lévy process. According to infinite divisibility, a Lévy process $Z(t)$ at point positive integer s can be rephrased as $Z(s) = \mu s + v_1 + \dots + v_s$ in distribution where $v_i = Z(i) - Z(i-1) - \mu$ ($i = 1, \dots, s$) form an i.i.d. sequence, and $\mu = E(Z(1))$. Whence, define $x_{s,n} = \frac{1}{\sqrt{n}\sigma_z} Z(s)$ for $s = 1, \dots, n$ and $n \geq 1$ where $\sigma_z^2 = Var(Z(1))$, then by virtue of functional central limit theorem $x_{s,n}$ converges in distribution to a Brownian motion $W(r)$ on $[0, 1]$ as $n \rightarrow \infty$. In addition, with $d_{l,k,n} = \sqrt{(l-k)/n}$, $x_{s,n}$ and $d_{l,k,n}$ satisfy Assumption A (a) and (c), and also A (b) can be achieved by the Skorohod representation theorem.

Take an example to verify the condition (3.5). Suppose now that $Z(t)$ is a Poisson process, viz., $Z(t) \sim Poi(\mu t)$. Because $\frac{1}{d_{l,k,n}}(x_{l,n} - x_{k,n}) =_D \frac{1}{\sqrt{l-k}\sigma_z}(Z(l-k) - (l-k)\mu)$,

$$F_{l,k,n}(0) = \sum_{i \leq (l-k)\mu} \frac{[(l-k)\mu]^i}{i!} e^{-(l-k)\mu}$$

$$F_{l,k,n}(u) = \sum_{i \leq (l-k)\mu + u\sqrt{l-k}\sigma_z} \frac{[(l-k)\mu]^i}{i!} e^{-(l-k)\mu}.$$

Thus, if $u > 0$

$$F_{l,k,n}(u) - F_{l,k,n}(0) = e^{-(l-k)\mu} \sum_{(l-k)\mu < i \leq (l-k)\mu + u\sqrt{l-k}\sigma_z} \frac{[(l-k)\mu]^i}{i!},$$

if $u < 0$,

$$|F_{l,k,n}(u) - F_{l,k,n}(0)| = e^{-(l-k)\mu} \sum_{(l-k)\mu + u\sqrt{l-k}\sigma_z < i \leq (l-k)\mu} \frac{[(l-k)\mu]^i}{i!}.$$

Because $e^{-(l-k)\mu} \rightarrow 0$ as $(l, k) \in \Omega_n(\epsilon)$, $n \rightarrow \infty$ and the sums are less than the tail of a convergent series, the condition (3.5) is fulfilled.

Notice also that in some situation, for continuous process the condition (3.5) implies the requirement (3.4), so that they merge as (3.5) which harbours both continuous and discrete cases.

Since we study the asymptotic theory not only for the sample mean but also for the sample covariance, the following assumption stipulates some necessary conditions for $x_{s,n}$ and error sequence e_s .

Assumption B

- (a) There is a martingale difference sequence $(e_s, \mathcal{F}_{n,s})$ with $E(e_s^2 | \mathcal{F}_{n,s-1}) = \sigma_e^2$ a.s. for all $s = 1, 2, \dots, n$ and $\sup_{1 \leq s \leq n} E(|e_s|^p | \mathcal{F}_{n,s-1}) < \infty$ a.s. for some $p > 2$.
- (b) $\{x_{s+1,n}\}$ is adapted to $\mathcal{F}_{n,s}$, $s \geq 0$.
- (c) Let, for $r \in [0, 1]$,

$$U_n(r) = \frac{1}{\sqrt{n}} \sum_{s=1}^{[nr]} e_s \quad \text{and} \quad W_n(r) = x_{[nr],n}.$$

Suppose that (U_n, W_n) converges in distribution to (U, W) on $D[0, 1]^2$ as $n \rightarrow \infty$, where (U, W) is a correlated Brownian motion vector.

Remark 3.2. As for Assumption A, Assumption B is also quite general and applicable in many situations. For example, Condition (b) holds when $\{e_s\}$ is a sequence of independent errors and $\mathcal{F}_{n,s} = \sigma(e_1, \dots, e_s, x_{s+1,n})$.

3.1 Time-normalized and integrable functionals

This subsection establishes an asymptotic theory that extends existing literature, such as Park and Phillips (1999, 2001) and Wang and Phillips (2009a), from the univariate case to the bivariate case. In what follows our asymptotic theory depends heavily on a local-time process of Brownian motion. A standard reference book for local-time process is Revuz and Yor (1999).

Let us now introduce some necessary conditions to establish important theorem. Such theorem is of general interest.

Assumption C

- (a) Suppose that $f(t, x)$ is defined on $[0, 1] \times (-\infty, \infty)$. Suppose further that both $|f(t, x)|$ and $f^2(t, x)$ are Lebesgue integrable with respect to x on $(-\infty, \infty)$.
- (b) There exists a function $c_f(x) : \mathbb{R} \rightarrow \mathbb{R}^+$ such that $|f(t, x)| \leq c_f(x)$ uniformly in $t \in [0, 1]$ and $c_f(x)$ is integrable on \mathbb{R} .
- (c) For each $x \in \mathbb{R}$, $f(t, x)$ is continuous in t and there are at most a finite number of points for t at which $\int f(t, x)dx = 0$.

Remark 3.3. We shall denote $G_1(t) = \int f(t, x)dx$, $G_2(t) = \int |f(t, x)|dx$ and $G_3(t) = \int f^2(t, x)dx$ for notational convenience. Notice that they are all continuous functions by the dominated convergence theorem.

Condition (a) is an extension of Assumption 2.1 in Wang and Phillips (2009a). Requirement on integrability of functions is a basic need to deal with such kind of problems. Note that if $f(t, x) = f(x)$ becomes time-homogeneous, Condition (a) reduces to Assumption 2.1 in Wang and Phillips (2009a).

Condition (b) requires that the function $f(t, x)$ be dominated uniformly in t over compact interval $[0, 1]$ by an integrable function $c_f(x)$. In the situations where $f(t, x)$ is the product of a continuous function of t and an integrable function of x or the superposition of such products, the condition is automatically fulfilled.

Condition (c) also excludes the situation where there are infinite many points $t_j \in [0, 1]$ such that $G_1(t_j) = 0$.

Theorem 3.1. *If Assumptions C and A (a) and (c) hold, we have for any $c_n \rightarrow \infty$, $n/c_n \rightarrow \infty$ and $r \in [0, 1]$,*

$$\frac{c_n}{n} \sum_{s=1}^{[nr]} f\left(\frac{s}{n}, c_n x_{s,n}\right) \rightarrow_D \int_0^r G_1(t) dL_W(t, 0), \quad (3.6)$$

where $G_1(\cdot) = \int f(\cdot, x)dx$ and $L_W(t, 0)$ is the local-time process of W at origin over time interval $[0, t]$.

If, in addition, Assumption A (a) is replaced by Assumption A(b), then for any $c_n \rightarrow \infty$, $n/c_n \rightarrow \infty$ and $r \in [0, 1]$,

$$\sup_{0 \leq r \leq 1} \left| \frac{c_n}{n} \sum_{s=1}^{[nr]} f\left(\frac{s}{n}, c_n x_{s,n}\right) - \int_0^r G_1(t) dL_W(t, 0) \right| = o_P(1), \quad (3.7)$$

under the same probability space as defined in Assumption A(b).

Moreover, suppose that $f^2(t, x)$ satisfies Assumption C, and that $\{e_s\}$ and $\{x_{s,n}\}$ satisfy Assumptions B and A (c). We have for $n \rightarrow \infty$, $c_n \rightarrow \infty$, $c_n/n \rightarrow 0$ and $r \in [0, 1]$,

$$\sqrt{\frac{c_n}{n}} \sum_{s=1}^{[nr]} f\left(\frac{s}{n}, c_n x_{s,n}\right) e_s \rightarrow_D \left(\int_0^r G_3(t) dL_W(t, 0) \right)^{\frac{1}{2}} N, \quad (3.8)$$

where $G_3(\cdot) = \int f^2(\cdot, x) dx$ and N is a standard normal random variable independent of W .

Remark 3.4. Note that if function $f(t, x)$ reduces to $f(x)$, equations (3.6) and (3.7) reduce to Theorem 2.1 of Wang and Phillips (2009a) and with $c_n = \sqrt{n}$ to Theorem 5.1 of Park and Phillips (1999), since $G_1(t) = \int f(x) dx$ becomes a constant and $\int_0^1 dL_W(r, 0) = L_W(1, 0)$. Also, these reduced cases of (3.7) and (3.8) can be viewed as a special case of Theorem 3.2 in Park and Phillips (2001) by taking parameter set Π as singleton since in the situation $G_3 = \int f^2(x) dx$ is a constant.

3.2 Asymptotic time-homogeneous and integrable functionals

Since in most cases the interested stochastic quantities are $L_n = \sum_{s=1}^n F(s, c_n x_{s,n})$ and $M_n = \sum_{s=1}^n F(s, c_n x_{s,n}) e_s$, the results in the last subsection could not be used directly. To tackle this issue, the key point is how we can normalize the time variable involved in the functionals. Noting that if s in function F is in the form of some polynomial, we would be able to deal with the normalization issue of time variable given that the F has some convenient form. Motivated by this idea, we introduce the following definition of asymptotic homogeneity with respect to t .

Definition 3.1. Let $F(t, x)$ be defined on $t \geq 0$ and $x \in \mathbb{R}$. Suppose for every $x \in \mathbb{R}$, $\forall \eta > 0$, and $t \in [0, 1]$,

$$F(\eta t, x) = v(\eta) f(t, x) + R_\eta(t, x),$$

where

- (a) $f(t, x)$ satisfies Assumption C.
- (b) $R_\eta(t, x)$ is chosen such that it satisfies either (i) or (ii) below:
 - (i) $|R_\eta(t, x)| \leq q_\eta(t) P(x)$, where both $P(x)$ and $P^2(x)$ are Lebesgue integrable, and $q_\eta(t)/v(\eta) \rightarrow 0$ uniformly in $t \in [0, 1]$ as $\eta \rightarrow \infty$.

- (ii) $|R_\eta(t, x)| \leq q_\eta(t)Q(\eta t)P(x)$ where $P(x)$ and $P^2(x)$ are Lebesgue integrable, $\lim_{\eta \rightarrow \infty} \frac{q_\eta(t)}{v(\eta)} = l(t)$ which is bounded on $[0,1]$ and $Q(y)$ is bounded on any compact interval and $\lim_{y \rightarrow +\infty} Q(y) = 0$.

Such functions $F(t, x)$ are asymptotic homogeneous with respect to t and integrable with respect to x , thus $F(t, x)$ is called homogeneously-integrable functions, said to be in Class (HI), denoted by $\mathcal{T}(HI)$. Functions v and f are called homogeneity power and normal function respectively. Function $F(t, x)$ with $R(t, x)$ satisfying (i) and (ii) is said to be in $\mathcal{T}(HI_1)$ and $\mathcal{T}(HI_2)$, respectively.

Theorem 3.2. *Suppose that $F(t, x)$ is in the class $\mathcal{T}(HI)$ with homogeneity power v and normal function f . Then, when Assumption A (a) and (c) hold, for any $c_n \rightarrow \infty$, $n/c_n \rightarrow \infty$ and $r \in [0, 1]$,*

$$\frac{c_n}{nv(n)} \sum_{s=1}^{[nr]} F(s, c_n x_{s,n}) \rightarrow_D \int_0^r G_1(t) dL_W(t, 0), \quad (3.9)$$

where $G_1(\cdot) = \int f(\cdot, x) dx$ and L_W is the local-time process of W .

If Assumption A (a) is replaced by A (b), for any $c_n \rightarrow \infty$, $n/c_n \rightarrow \infty$,

$$\frac{c_n}{nv(n)} \sum_{s=1}^{[nr]} F(s, c_n x_{s,n}) \rightarrow_P \int_0^r G_1(t) dL_W(t, 0), \quad (3.10)$$

uniformly in $r \in [0, 1]$ as $n \rightarrow \infty$ where $G_1(\cdot) = \int f(\cdot, x) dx$ and L_W is the local-time process of W .

Moreover, if $\{e_s\}$ and $\{x_{s,n}\}$ satisfy Assumption B, and $f^2(t, x)$ satisfies Assumption C, we have for $n \rightarrow \infty$, $c_n \rightarrow \infty$, $n/c_n \rightarrow \infty$ and $r \in [0, 1]$,

$$\sqrt{\frac{c_n}{n}} \frac{1}{v(n)} \sum_{s=1}^{[nr]} F(s, c_n x_{s,n}) e_s \rightarrow_D \left(\int_0^r G_3(t) dL_W(t, 0) \right)^{\frac{1}{2}} N, \quad (3.11)$$

where $G_3(\cdot) = \int f^2(\cdot, x) dx$ and N is a standard normal random variable independent of W .

3.3 Regular functionals

In this subsection, we establish an asymptotic theory for the sample mean and sample covariance of a regular functional $f(t, x)$ to be defined below. The idea here is to deal with

the convergence of the sample moment for functionals which have discontinuous point but can be sandwiched between two continuous functions that the integral of whose difference on some compact interval approaches to zero. Also, the definition of regularity extends that in Park and Phillips (1999, 2001).

Definition 3.2. Let $f(t, x)$ be defined on $[0, 1] \times \mathbb{R}$. We say that $f(t, x)$ is regular, if

(a) for each $x \in \mathbb{R}$, $f(t, x)$ is Lipschitz with respect to t , that is, there exists a constant $L(x)$ relative to x such that for any $t_1, t_2 \in [0, 1]$,

$$|f(t_1, x) - f(t_2, x)| \leq L(x)|t_1 - t_2|, \quad (3.12)$$

where $L(x)$, viewed as a function of x , satisfies condition (b) and (c) below (ignoring t);

(b) for each t , $f(t, x)$ is continuous in x in a neighborhood of infinity;

(c) on any compact interval J of \mathbb{R} , for any given $\epsilon > 0$, there exist functions $\underline{f}_\epsilon(t, x)$, $\bar{f}_\epsilon(t, x)$, which are continuous in t and x , and $\delta > 0$ such that whenever $|y - x| < \delta$ on J , for each $t \in [0, 1]$,

$$\underline{f}_\epsilon(t, x) \leq f(t, y) \leq \bar{f}_\epsilon(t, x), \quad (3.13)$$

and as $\epsilon \rightarrow 0$,

$$\int_J \sup_{t \in [0, 1]} (\bar{f}_\epsilon(t, x) - \underline{f}_\epsilon(t, x)) dx \rightarrow 0. \quad (3.14)$$

Remark 3.5. Note that if $f(t, x)$ reduces to $f(x)$, the conditions in (b) and (c) (ignoring t) can be viewed as the definition of regularity of $f(x)$, identical (with negligible difference) to that in Park and Phillips (1999, 2001). Hence, since t is in $[0, 1]$, any type of functions $f(t, x) = q(t)L(x)$ is regular where $q(t) \in C^1[0, 1]$ and $L(x)$ is regular. For detailed discussion consult the papers above.

Note also that the main difference between this definition for $f(t, x)$ and Definition 3.2 in Park and Phillips (2001) for function $F(x, \pi)$, $\pi \in \Pi$, is that π is a parameter in a compact set Π , while $t \in [0, 1]$ is not parameter, which is involved as a variable in our discussion below.

Theorem 3.3. Let $f(t, x)$ be regular. If Assumption B is verified for triangular array $x_{s,n}$, $1 \leq s \leq n$, $n = 1, 2, \dots$, and martingale difference $(e_s, \mathcal{F}_{n,s})$, then

$$\frac{1}{n} \sum_{s=1}^n f\left(\frac{s}{n}, x_{s,n}\right) \rightarrow_D \int_0^1 f(r, W(r)) dr, \quad (3.15)$$

$$\frac{1}{\sqrt{n}} \sum_{s=1}^n f\left(\frac{s}{n}, x_{s,n}\right) e_s \xrightarrow{D} \int_0^1 f(r, W(r)) dU(r), \quad (3.16)$$

as $n \rightarrow \infty$.

The following lemma gives the closure of the usual operations: addition, multiplication by a scalar and product for regular functions.

Lemma 3.1. *Suppose that both $f(t, x)$ and $g(t, x)$ are regular, then $f(t, x) + g(t, x)$, $cf(t, x)$ for any $c \in \mathbb{R}$ and $f(t, x)g(t, x)$ are regular.*

3.4 Asymptotic homogeneous regular functionals

We borrow some notations from Park and Phillips (2001) for notational brevity. Let \mathcal{T}_{LB} denote a class of locally bounded transformations on \mathbb{R} ; let \mathcal{T}_{LB}^0 be a subclass of \mathcal{T}_{LB} consisting only of locally bounded transformations which are exponential bounded, i.e. transformations P such that $P(x) = O(e^{c|x|})$ for some $c > 0$; a class of bounded transformations on \mathbb{R} is denoted by \mathcal{T}_B , and a subclass \mathcal{T}_B^0 of \mathcal{T}_B is the collection of transformations that are bounded and vanish at infinity, i.e. transformations P such that $P(x) \rightarrow 0$ as $|x| \rightarrow \infty$. Clearly, $\mathcal{T}_B^0 \subset \mathcal{T}_B \subset \mathcal{T}_{LB}^0 \subset \mathcal{T}_{LB}$.

Definition 3.3. We say that function $F(t, x)$ is asymptotically homogeneous with respect to both t and x , if for all $\xi, \eta > 0$ and $t \in [0, 1]$,

$$F(\xi t, \eta x) = v_1(\xi)v_2(\eta)f(t, x) + R(\xi, \eta; t, x), \quad (3.17)$$

where $f(t, x)$ is regular on $[0, 1] \times \mathbb{R}$, and $|R(\xi, \eta; t, x)| \leq A_\xi(t)a(\eta)P(x) + q(t)b(\xi)B_\eta(x)$ with positive functions A, a, P, q, b, B such that

- a) $P(x) \in \mathcal{T}_{LB}^0$, $\limsup_{\eta \rightarrow \infty} \frac{a(\eta)}{v_2(\eta)} < \infty$ and either $\limsup_{\xi \rightarrow \infty} \frac{A_\xi(t)}{v_1(\xi)} = 0$ uniformly in $t \in [0, 1]$; or $v_1(\xi) \rightarrow \infty$ as $\xi \rightarrow \infty$ and $A_\xi(t) = A(t)$ which is Riemann integrable on $[0, 1]$; or $A_\xi(t) = \bar{A}_\xi(t)Q(\xi t)$ with $\limsup_{\xi \rightarrow \infty} \frac{\bar{A}_\xi(t)}{v_1(\xi)} = l(t)$ which is bounded on $[0, 1]$ and $Q(\cdot) \in \mathcal{T}_B^0$. And,
- b) $q(t)$ is bounded on $[0, 1]$, $\limsup_{\xi \rightarrow \infty} \frac{b(\xi)}{v_1(\xi)} < \infty$ and either $B_\eta(x) = \bar{B}(\eta)V(x)$ with $\limsup_{\eta \rightarrow \infty} \frac{\bar{B}(\eta)}{v_2(\eta)} = 0$ and $V(x) \in \mathcal{T}_{LB}^0$, or $B_\eta(x) = \bar{B}(\eta)V(\eta x)$, where $V(\cdot) \in \mathcal{T}_B^0$ and $\limsup_{\eta \rightarrow \infty} \frac{\bar{B}(\eta)}{v_2(\eta)} < \infty$.

In the definition of asymptotic homogeneity, we denote $F(t, x) \in \mathcal{T}(HH)$ and call $f(t, x)$ the normal function of $F(t, x)$, and $v_1(\cdot)$ and $v_2(\cdot)$ the homogeneity powers with respect to t and x respectively. See Remark B.3 for discussion and examples.

Theorem 3.4. *Let $F(t, x)$ be in Class $\mathcal{T}(HH)$ with homogeneity powers $v_1(\cdot)$ and $v_2(\cdot)$ and normal function $f(t, x)$. Let martingale difference $(e_s, \mathcal{F}_{n,s})$ and $x_{s,n}$ satisfy Assumption B. We then have*

$$\frac{1}{nv_1(n)v_2(c_n)} \sum_{s=1}^n F(s, c_n x_{s,n}) \rightarrow_D \int_0^1 f(r, W(r)) dr, \quad (3.18)$$

$$\frac{1}{\sqrt{n}v_1(n)v_2(c_n)} \sum_{s=1}^n F(s, c_n x_{s,n}) e_s \rightarrow_D \int_0^1 f(r, W(r)) dU(r), \quad (3.19)$$

where $(U(r), W(r))$ is the limit of $(U_n(r), W_n(r))$ for $r \in [0, 1]$ stipulated in Assumption B.

Remark 3.6. Note that if $F(t, x)$ reduces to a univariate function $F(x)$, with $c_n = \sqrt{n}$, equation (3.18) becomes Theorem 5.3 of Park and Phillips (1999) and the first part of Theorem 3.3 with singleton Π of Park and Phillips (2001); equation (3.19) becomes the second part of Theorem 3.3 with singleton Π in Park and Phillips (2001).

4 Statistical estimation

We consider a general statistical regression model of the form

$$Y(t) = m(t, Z(t)) + \varepsilon(t), \quad (4.1)$$

where $m(\cdot, \cdot)$ is an unknown functional, $\varepsilon(t)$ is an error process with zero mean and finite variance, and $Z(t)$ is a Lévy process.

Suppose that $Z(t)$ admits a classical orthonormal polynomial system $Q_i(t, x)$ with weight $\rho(t, x)$, the density function or the probability distribution function of $Z(t)$. Let the support of $\rho(t, x)$ be denoted by I , which can be \mathbb{R} , \mathbb{R}^+ or \mathbb{N} . Note that, as before, the operator D signifies either differentiation or difference and it is conducted only with respect to x .

This section is devoted to the estimation of $m(\cdot, \cdot)$ given observations of $Y(t)$. We shall divide the section into three subsections according to the different types of time horizons, viz., on $(0, \infty)$, $[0, T]$ with fixed T and $[0, T_n]$ where T_n is increasing with sample size n .

4.1 Infinite time horizon

Suppose t is in the interval $(0, \infty)$. In this subsection we are going to work with the situation where $m(\cdot, \cdot)$ is defined on $[0, \infty) \times I$ and our sampling points are $t_s = s$, $s = 1, 2, \dots, n$. Given that we have observations (Z_s, Y_s) where $Y_s = Y(s)$ and $Z_s = Z(s)$ for $s = 1, 2, \dots, n$, our aim is to estimate $m(\tau, Z_\tau)$ at $t = \tau$.

At each point of observations, model (4.1) now becomes

$$Y_s = m(s, X_s) + e_s, \quad s = 1, \dots, n, \quad (4.2)$$

where $X_s = Z(s)$ denotes the Lévy process at point s , $e_s = \varepsilon(s)$ ($s = 1, \dots, n$) form an error sequence with mean zero and finite variance.

Observe that because $Z(t)$ is a Lévy process, $E[Z(t)] = t\mu$ where $\mu = E(Z(1))$ and $\text{Var}(Z(t)) = t\sigma_z^2$ where $\sigma_z^2 = \text{Var}(Z(1))$. Observe also that $X_s = s\mu + X_s - s\mu = s\mu + \sum_{i=1}^s (X_i - X_{i-1} - \mu) = s\mu + \sqrt{n}\sigma_z x_{s,n}$, where $x_{s,n} = \frac{1}{\sqrt{n}\sigma_z} \sum_{i=1}^s (X_i - X_{i-1} - \mu)$. Since $X_i - X_{i-1} - \mu$ form an i.i.d $(0, \sigma_z^2)$ sequence, it follows from the functional central limit theorem, $x_{s,n}$ converges in distribution to a Brownian motion on $[0, 1]$. In addition, $x_{s,n}$ satisfies Assumption A in the preceding section.

We firstly need to impose some conditions on $m(t, x)$.

Assumption 4.1

- (a) For every $t > 0$, $m(t, x)$ and $D^r m(t, x)$ are in $L^2(I, \rho_r(t, x))$, $r = 1, 2, 3$.
- (b) For each i , the coefficient function $c_i(t, m) = E[m(t, Z(t))Q_i(t, Z(t))]$, and its derivatives of up to third order all belong to $L^2(\mathbb{R}^+)$.
- (c) For i large enough, the coefficient functions $c_i(t, D^3 m)$ of $D^3 m(t, Z(t))$ expanded by the system $\{Q_{3i}(t, Z(t))\}$ are chosen such that $\psi(t)^3 c_i^2(t, D^3 m)$ are bounded on $(0, \infty)$ uniformly in i .

See Remark B.4 for discussion and examples.

Having expanded function m at sampling points, given truncation parameters k and p_i , model (4.2) can be written as for $s = 1, 2, \dots, n$,

$$Y_s = \left(\sum_{i=0}^k \sum_{j=0}^{p_i} + \sum_{i=0}^k \sum_{j=p_i+1}^{\infty} + \sum_{i=k+1}^{\infty} \sum_{j=0}^{\infty} \right) c_{ij} \mathcal{L}_j(s) Q_i(s, X_s) + e_s. \quad (4.3)$$

As we know from the preceding section, $\sum_{j=0}^{\infty} c_{ij} \mathcal{L}_j(s) = c_i(s, m)$ or more simply, $c_i(s)$ if there is no confusion occurred. We now may rewrite equations (4.3) in the following matrix form:

$$Y = X\theta + \delta + \gamma + \varepsilon, \quad (4.4)$$

where

$$\begin{aligned} Y' &= (Y_1, Y_2, \dots, Y_n); \quad \theta' = (c_{00}, c_{01}, \dots, c_{0p_0}, c_{10}, \dots, c_{1p_1}, \dots, c_{k0}, \dots, c_{kp_k}); \\ x_1 &= (\mathcal{L}_0(1)Q_0(1, X_1), \mathcal{L}_1(1)Q_0(1, X_1), \dots, \mathcal{L}_{p_0}(1)Q_0(1, X_1), \\ &\quad \mathcal{L}_0(1)Q_1(1, X_1), \mathcal{L}_1(1)Q_1(1, X_1), \dots, \mathcal{L}_{p_1}(1)Q_1(1, X_1), \\ &\quad \dots, \mathcal{L}_0(1)Q_k(1, X_1), \mathcal{L}_1(1)Q_k(1, X_1), \dots, \mathcal{L}_{p_k}(1)Q_k(1, X_1)), \\ &\quad \vdots \\ x_n &= (\mathcal{L}_0(n)Q_0(n, X_n), \mathcal{L}_1(n)Q_0(n, X_n), \dots, \mathcal{L}_{p_0}(n)Q_0(n, X_n), \\ &\quad \mathcal{L}_0(n)Q_1(n, X_n), \mathcal{L}_1(n)Q_1(n, X_n), \dots, \mathcal{L}_{p_1}(n)Q_1(n, X_n), \\ &\quad \dots, \mathcal{L}_0(n)Q_k(n, X_n), \mathcal{L}_1(n)Q_k(n, X_n), \dots, \mathcal{L}_{p_k}(n)Q_k(n, X_n)), \end{aligned}$$

and $X = (x'_1, x'_2, \dots, x'_n)'$, $\delta' = (\delta_1, \dots, \delta_n)$, $\gamma' = (\gamma_1, \gamma_2, \dots, \gamma_n)$ with $\delta_s = \sum_{i=1}^k \sum_{j=p_i+1}^{\infty} c_{ij} \mathcal{L}_j(s) Q_i(s, X_s)$ and $\gamma_s = \sum_{i=k+1}^{\infty} c_i(s) Q_i(s, X_s)$, $s = 1, 2, \dots, n$; $\varepsilon' = (e_1, e_2, \dots, e_n)$.

The Semiparametric Least Squares (SLS) estimator of θ is given by

$$\hat{\theta} = (X'X)^{-1} X'Y. \quad (4.5)$$

After obtaining the estimators of the coefficients in the expansion of functional $m(t, Z(t))$, we are able to estimate the function $m(\tau, x)$ at point (τ, x) , where $\forall \tau > 0$ and $x \in \mathbb{R}$ is any point on the trajectory of $X_\tau = Z(\tau)$, namely, we can have $\hat{m}(\tau, x)$ by superseding $\hat{\theta}$ in lieu of θ and getting rid of residues in the expansion of $m(\tau, x)$. More precisely, as $m(\tau, x) = A'(\tau, x)\theta + \delta(\tau, x) + \gamma(\tau, x)$, $\hat{m}(\tau, x) = A'(\tau, x)\hat{\theta}$, where θ is defined as before and

$$\begin{aligned} \delta(\tau, x) &= \sum_{i=0}^k \sum_{j=p_i+1}^{\infty} c_{ij} \mathcal{L}_j(\tau) Q_i(\tau, x), \quad \gamma(\tau, x) = \sum_{i=k+1}^{\infty} c_i(\tau, m) Q_i(\tau, x), \\ A'(\tau, x) &= (\mathcal{L}_0(\tau)Q_0(\tau, x), \dots, \mathcal{L}_{p_0}(\tau)Q_0(\tau, x), \\ &\quad \dots, \mathcal{L}_0(\tau)Q_k(\tau, x), \dots, \mathcal{L}_{p_k}(\tau)Q_k(\tau, x)). \end{aligned}$$

We shall investigate the limit of

$$\begin{aligned}\widehat{m}(\tau, x) - m(\tau, x) &= A'(\tau, x)(\widehat{\theta} - \theta) - \delta(\tau, x) - \gamma(\tau, x) \\ &= A'(\tau, x)(X'X)^{-1}X'(\delta + \gamma + \varepsilon) - \delta(\tau, x) - \gamma(\tau, x).\end{aligned}\tag{4.6}$$

For late use, denote $A_{p \times p}$ and $B_{p \times p}$ by

$$A = \frac{A(\tau, x)A'(\tau, x)}{\|A(\tau, x)\|^2} \quad \text{and} \quad B = (X'X)A(X'X)^{-1},$$

where $\|\cdot\|$ signifies Euclidean norm and dimension $p = p_0 + \dots + p_k + k + 1$.

Before proceeding further let us establish the following useful lemma.

Lemma 4.1. *Let v be an $1 \times p$ unit row vector. Define $p \times p$ matrix $V = v'v$. Then V has eigenvalues $\lambda_1 = 1$, $\lambda_i = 0$, $i = 2, \dots, p$.*

Because of their similarity, in view of Lemma (4.1), A and B share the same eigenvalues, $\lambda_1 = 1$, $\lambda_2 = \dots = \lambda_p = 0$.

Let α be the unit left eigenvector of B pertaining to eigenvalue 1, viz., $\alpha'B = \alpha'$ and $\|\alpha\| = 1$. Denote $\alpha' = (\alpha_{00}, \dots, \alpha_{0p_0}, \dots, \alpha_{k0}, \dots, \alpha_{kp_k})$, in accordance with $A(\tau, x)$.

The following assumption imposes some conditions on a double-index sequence we are working with.

Assumption 4.2

- (a) Let $\mathcal{S} = \{a_0, a_1, a_2, \dots\}$, where $a_i = \{a_{ij}\}_{j=0}^\infty$ is a sequence such that $\sum_{j=1}^\infty j|a_{i,j}| < \infty$ for $i = 0, 1, 2, \dots$.
- (b) Suppose further that $\sum_{i=1}^\infty i \left(\sum_{j=0}^\infty |a_{i,j}| \right)^2 < \infty$.

Remark 4.1. Note that Assumption 4.2(a) and Assumption 4.2(b) are both required. This is because the first condition is the requirement of the decay rate of $|a_{ij}|$ in terms of j , while the second one postulates that for each $i > 0$, $\varsigma_i = \sum_{j=0}^\infty |a_{i,j}|$ is approximately of $O\left(\frac{1}{i^{1+\eta}}\right)$ for some $\eta > 0$. Obviously, if there are some $\epsilon > 0$ and $\eta > 0$ such that $a_{ij} = O\left(\frac{1}{(1+j)^{2+\epsilon}(1+i)^{1+\eta}}\right)$ for $i, j \geq 0$, both conditions are fulfilled.

Using α and $\frac{1}{\|A'(\tau, x)\|}A'(\tau, x)$, let us reshuffle the set \mathcal{S} as $\widetilde{\mathcal{S}}$ and $\bar{\mathcal{S}}$ by defining

- 1) $\widetilde{\mathcal{S}} = \{\widetilde{a}_0, \dots, \widetilde{a}_i, \dots\}$, and $\bar{\mathcal{S}} = \{\bar{a}_0, \dots, \bar{a}_i, \dots\}$;
- 2) $\widetilde{a}_i = \{\widetilde{a}_{ij}\}$ where $\widetilde{a}_{ij} = \frac{1}{\sqrt{p_{\max}}}\alpha_{ij}$ for $0 \leq i \leq k$ and $0 \leq j \leq p_i$; otherwise, $\widetilde{a}_{ij} = a_{ij}$;

3) $\bar{a}_i = \{\bar{a}_{ij}\}$ where $\bar{a}_{ij} = \frac{1}{\sqrt{p_{\max}}\|A(\tau, x)\|} \mathcal{L}_j(\tau) Q_i(\tau, x)$ for $0 \leq i \leq k$ and $0 \leq j \leq p_i$; otherwise, $\bar{a}_{ij} = a_{ij}$.

Obviously, $\tilde{a}_{ij} = a_{ij} = \bar{a}_{ij}$ if $i > k$ or $j > p_i$. Meanwhile, since $\tilde{\mathcal{S}}$ and $\bar{\mathcal{S}}$ satisfy Riesz-Fischer theorem, there exist functions, denoted by $\tilde{F}(t, x)$ and $\tilde{G}(t, x)$, such that

$$\tilde{F}(t, Z(t)) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \tilde{a}_{ij} \mathcal{L}_j(t) Q_i(t, Z(t)), \quad (4.7)$$

$$\tilde{G}(t, Z(t)) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \bar{a}_{ij} \mathcal{L}_j(t) Q_i(t, Z(t)), \quad (4.8)$$

for any $t > 0$.

Therefore, in view of (4.7) and (4.8), we have

$$\frac{1}{\sqrt{p_{\max}}} \alpha' X' = \tilde{\mathbf{F}}' - \tilde{\delta}' - \tilde{\gamma}', \quad (4.9)$$

$$\frac{1}{\sqrt{p_{\max}}\|A'(\tau, x)\|} A'(\tau, x) X' = \tilde{\mathbf{G}}' - \tilde{\delta}' - \tilde{\gamma}', \quad (4.10)$$

where

$$\begin{aligned} \tilde{\mathbf{F}}' &= (\tilde{F}(1, X_1), \dots, \tilde{F}(n, X_n)), \quad \tilde{\mathbf{G}}' = (\tilde{G}(1, X_1), \dots, \tilde{G}(n, X_n)), \\ \tilde{\delta}' &= (\tilde{\delta}_1, \dots, \tilde{\delta}_n) \text{ with } \tilde{\delta}_s = \sum_{i=0}^k \sum_{j=p_i+1}^{\infty} a_{ij} \mathcal{L}_j(s) Q_i(s, X_s), \\ \tilde{\gamma}' &= (\tilde{\gamma}_1, \dots, \tilde{\gamma}_n) \text{ with } \tilde{\gamma}_s = \sum_{i=k+1}^{\infty} \sum_{j=0}^{\infty} a_{ij} \mathcal{L}_j(s) Q_i(s, X_s). \end{aligned}$$

We have the following proposition for the functions $\tilde{F}(t, x)$ and $\tilde{G}(t, x)$.

Proposition 4.1. *For any $t > 0$, (a) $E[\tilde{G}(t, Z(t))]^2 < \infty$, and (b) $E[\tilde{F}(t, Z(t))]^2 < \infty$.*

Notice that $E[Z(t)] = \mu t$. Denote $F(t, x - \mu t) = \tilde{F}(t, x)$ and $G(t, x - \mu t) = \tilde{G}(t, x)$. This is only a change in the form of functions since the process $Z(t)$ has to be centralized in order to acquire the limit distribution of \hat{m} .

The following assumption is stipulated for the truncation parameters, which is crucial for obtaining the limit distribution of the estimator.

Assumption 4.3

(a) $k = [n^{\kappa_1}]$ with $\frac{1}{2} < \kappa_1 < 1$;

(b) $p_{\min} = \lceil n^{\kappa_2} \rceil$ and $p_{\max} = \lceil n^{\bar{\kappa}_2} \rceil$ with $0 < \kappa_2 \leq \bar{\kappa}_2 < 1$;

(c) $2 + 2\kappa_1 < 5\kappa_2$.

Remark 4.2. There are obviously a great deal of feasible options for κ_1, κ_2 and $\bar{\kappa}_2$ satisfying the conditions. Note that condition (c) is quite reasonable since it follows automatically when $\kappa_2 > 0.8$.

Next assumption describes the families of functionals F and G we are studying in the asymptotic distribution of the estimator.

Assumption 4.4

- (a) Suppose that $F(\cdot, \cdot)$ and $G(\cdot, \cdot)$ are in class $\mathcal{T}(HI)$ with homogeneity powers $v(\cdot)$ and $\varrho(\cdot)$ and normal functions $f(\cdot, \cdot)$ and $g(\cdot, \cdot)$ respectively. Let $v(n) = n^\varsigma$ and $\varrho(n) = n^\iota$ with $\varsigma \geq 0$ and $\iota \geq 0$ satisfying $\frac{1}{2}(\kappa_1 - \frac{1}{2}) < \iota < \min\{\frac{5}{4}\kappa_2 - \frac{1}{4} - \frac{1}{2}\kappa_1, \frac{7}{4}\kappa_1 - \frac{1}{4}\}$.
- (b) Suppose also that $F^2(\cdot, \cdot)$, $G^2(\cdot, \cdot)$ and $F(\cdot, \cdot)G(\cdot, \cdot)$ are all in class $\mathcal{T}(HI)$ with homogeneity powers $v^2(\cdot)$, $\varrho^2(\cdot)$ and $v(\cdot)\varrho(\cdot)$, and normal functions $f^2(\cdot, \cdot)$, $g^2(\cdot, \cdot)$ and $f(\cdot, \cdot)g(\cdot, \cdot)$ respectively.
- (c) Suppose that $F(\cdot, \cdot)$ and $G(\cdot, \cdot)$ are in class $\mathcal{T}(HH)$ with homogeneity powers $v_1(\cdot)$, $v_2(\cdot)$ and $\varrho_1(\cdot)$, $\varrho_2(\cdot)$ and normal functions $f(\cdot, \cdot)$ and $g(\cdot, \cdot)$ respectively. Let $v_1(n) = n^{\varsigma_1}$, $v_2(n) = n^{\varsigma_2}$, $\varrho_1(n) = n^{\iota_1}$, and $\varrho_2(n) = n^{\iota_2}$ with $\varsigma_i \geq 0$, $\iota_i \geq 0$, $i = 1, 2$, satisfying that $\iota_1 + \frac{1}{2}\iota_2 < \min\{\frac{5}{4}\kappa_2 - \frac{1}{2}(1 + \kappa_1), \frac{7}{4}\kappa_1 - \frac{1}{2}\}$.
- (d) Suppose also that $F^2(\cdot, \cdot)$, $G^2(\cdot, \cdot)$ and $F(\cdot, \cdot)G(\cdot, \cdot)$ are all in class $\mathcal{T}(HH)$ with homogeneity powers $v_1^2(\cdot)$ and $v_2^2(\cdot)$; $\varrho_1^2(\cdot)$ and $\varrho_2^2(\cdot)$; $v_1(\cdot)\varrho_1(\cdot)$ and $v_2(\cdot)\varrho_2(\cdot)$ as well as normal functions $f^2(\cdot, \cdot)$, $g^2(\cdot, \cdot)$ and $f(\cdot, \cdot)g(\cdot, \cdot)$ respectively.

Remark 4.3. Assumption 4.3 ensures that two upper bounds for ι and $\iota_1 + \frac{1}{2}\iota_2$ are positive. Of course, we can simplify these conditions in (a) and (c) if we impose more constraints on κ_1 and κ_2 . However, these conditions allow more options.

Note that in the proof of the following theorem, whatever conditions for ι and $\iota_1 + \frac{1}{2}\iota_2$ we actually use also involve $\bar{\kappa}_2 - \kappa_2$. Since we may require $\bar{\kappa}_2$ to be much closer to κ_2 such that $\bar{\kappa}_2 - \kappa_2$ is as small as we wish, conditions in (a) and (c) tacitly provide what we need in the proof. Obviously, this does not harm any thing else and applies to the subsequential subsections.

Note also that the ambit for both ι and $\iota_1 + \frac{1}{2}\iota_2$ can be enlarged at price of enhancing the order of differentiability for the coefficient functions in the expansion of m function, as can be seen in the proof of the following theorem.

We are now ready to state the main result in the subsection.

Theorem 4.1. *Suppose that $\{x_{s,n}\}_1^n$ and $\{e_s\}_1^n$ satisfy Assumption B. Let Assumptions 4.1–4.3 hold.*

If Assumption 4.4(a) and (b) hold, then

$$\begin{aligned} & \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max} \|A(\tau, x)\|^2}} (\widehat{m}(\tau, x) - m(\tau, x)) \\ & \rightarrow_D \left(\int_0^1 G_3(t) dL_W(t, 0) \right)^{\frac{1}{2}} N, \end{aligned} \quad (4.11)$$

where $G_3(t) = \int f(t, x)^2 dx$, W is a standard Brownian motion on $[0, 1]$, N is a standard normal random variable independent of W , and L_W is the local-time process of W .

If Assumption 4.4(c) and (d) hold, then

$$\begin{aligned} & \frac{1}{\sqrt{nv_1(n)v_2(\sqrt{n}\sigma_z)}} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max} \|A(\tau, x)\|^2}} (\widehat{m}(\tau, x) - m(\tau, x)) \\ & \rightarrow_D \int_0^1 f(r, W(r)) dU(r), \end{aligned} \quad (4.12)$$

where $(W(r), U(r))$ is the vector of Brownian motions involved in Assumption B.

See Remark B.5 for discussion on the theorem.

4.2 Finite time horizon

Assume time variable t lies in $[0, T]$ with T fixed. In this subsection function m is defined on $[0, T] \times I$. Therefore, conditions on m would be weakened since square integrability on $[0, T]$ is much weaker than that on the half line. We make the following assumptions about $m(t, x)$ in model (4.1).

Assumption 4.5

- (a) Let $D^r m(t, x) \in L^2(I, \rho_r(t, x))$ for any $t \in [0, T]$ and $r = 0, 1, 2$. Moreover, the expansion of $D^2 m(t, Z(t))$ in terms of $Q_{2i}(t, Z(t))$ converges in the sense of mean square uniformly on $[0, T]$.

(b) For each i , $b_i(t, m) = E[m(t, Z_t)Q_i(t, Z_t)]$ and its derivatives of up to third order belong to $L^2[0, T]$.

(c) Furthermore, $\|b_i''(t, m)\|_{L^2[0, T]}$ are bounded uniformly in i .

Remark 4.4. Both Condition (a) and (b) are quite general. Condition (a) ensures the possibility of the expansions of $D^r m(t, Z(t))$. Condition (c) is also reasonable. For example, if function m is polynomial, sine, cosine function, or their combinations, the condition is fulfilled automatically.

Suppose that we have n observations for the process $Y(t)$ on $[0, T]$ and the observations are $Y_{s,n} = Y(t_{s,n})$ at $t_{s,n} = T\frac{s}{n}$ for $s = 1, 2, \dots, n$. At the sampling points, we have the following model

$$Y_{s,n} = m(t_{s,n}, X_{s,n}) + e_s, \quad s = 1, \dots, n, \quad (4.13)$$

where $X_{s,n} = Z(T\frac{s}{n})$ denotes the Lévy process $Z(t)$ at point $t_{s,n}$, $e_s = \varepsilon(T\frac{s}{n})$ ($s = 1, \dots, n$) form an error sequence with mean zero and finite variance.

Note that $X_{s,n} = \frac{s}{n}T\mu + \sqrt{T}\sigma_z \frac{1}{\sqrt{n}} \sum_{i=1}^s w_i$, where $w_i = \frac{\sqrt{n}}{\sqrt{T}\sigma_z} (X_{i,n} - X_{i-1,n} - \frac{1}{n}T\mu)$ form an i.i.d.(0,1) sequence. Let $x_{s,n} = \frac{1}{\sqrt{n}} \sum_{i=1}^s w_i$. It follows from the functional central limit theorem that $x_{s,n}$ converges to a standard Brownian motion in distribution as $n \rightarrow \infty$. It is also clear that $x_{s,n}$ satisfies Assumption A.

Under Assumption 4.5 we can expand $m(t, Z(t))$ at every point $t \in [0, T]$ using basis $\varphi_{jT}(t)Q_i(t, Z(t))$. Let k and p_i be truncation parameters for i and j . Thus, model (4.13) for $s = 1, 2, \dots, n$ becomes

$$Y_{s,n} = \left(\sum_{i=0}^k \sum_{j=0}^{p_i} + \sum_{i=0}^k \sum_{j=p_i+1}^{\infty} + \sum_{i=k+1}^{\infty} \sum_{j=0}^{\infty} \right) b_{ij} \varphi_{jT}(t_{s,n}) Q_i(t_{s,n}, X_{s,n}) + e_s.$$

Equivalently, the matrix form of (4.2) is

$$Y = X\beta + \delta + \gamma + \varepsilon, \quad (4.14)$$

where all notations remain similar to what has been defined in the last subsection so that we avoid reciting them. The SLS estimator of β is given by

$$\hat{\beta} = (X'X)^{-1}X'Y. \quad (4.15)$$

With the help of $\hat{\beta}$ we are able to estimate $m(\cdot, \cdot)$ at (τ, x) where τ is any point in $[0, T]$ and x is any point on the path of $Z(\tau)$. The estimator $\hat{m}(\tau, x)$ of $m(\tau, x)$ is obtained by

replacing β by $\widehat{\beta}$ and abandoning all the residues in the expansion of $m(\tau, x)$. Explicitly, $m(\tau, x) := A'(\tau, x)\beta + \delta(\tau, x) + \gamma(\tau, x)$ and $\widehat{m}(\tau, x) = A'(\tau, x)\widehat{\beta}$, where

$$\begin{aligned} \delta(\tau, x) &= \sum_{i=0}^k \sum_{j=p_i+1}^{\infty} b_{ij} \varphi_{jT}(\tau) Q_i(\tau, x), \quad \gamma(\tau, x) = \sum_{i=k+1}^{\infty} b_i(\tau, m) Q_i(\tau, x), \\ A'(\tau, x) &= (\varphi_{0T}(\tau) Q_0(\tau, x), \dots, \varphi_{p_0T}(\tau) Q_0(\tau, x), \\ &\quad \dots, \varphi_{0T}(\tau) Q_k(\tau, x), \dots, \varphi_{p_kT}(\tau) Q_k(\tau, x)). \end{aligned}$$

We shall investigate the limit of

$$\begin{aligned} \widehat{m}(\tau, x) - m(\tau, x) &= A'(\tau, x)(\widehat{\beta} - \beta) - \delta(\tau, x) - \gamma(\tau, x) \\ &= A'(\tau, x)(X'X)^{-1}X'(\delta + \gamma + \varepsilon) - \delta(\tau, x) - \gamma(\tau, x). \end{aligned} \quad (4.16)$$

Let us define

$$A = \frac{A(\tau, x)A'(\tau, x)}{\|A(\tau, x)\|^2} \quad \text{and} \quad B = (X'X)A(X'X)^{-1}. \quad (4.17)$$

By virtue of Lemma 4.1 and similarity, B has eigenvalues $\lambda_1 = 1$, $\lambda_2 = \dots = \lambda_p = 0$. Let unit column vector α be the left eigenvector of B pertaining to $\lambda_1 = 1$, viz., $\alpha'B = \alpha'$ and $\|\alpha\| = 1$. In accordance with the notation of $A(\tau, x)$, the subscript of α is specified in double-index, that is, $\alpha' = (\alpha_{00}, \dots, \alpha_{0p_0}, \dots, \alpha_{k0}, \dots, \alpha_{kp_k})$.

Let us apply the reshuffling procedure for the set \mathcal{S} from Assumption 4.2 by α and $\frac{1}{\|A(\tau, x)\|}A(\tau, x)$. Denote by $\widetilde{\mathcal{S}}$ and $\bar{\mathcal{S}}$ the resulting sets:

- 1) $\widetilde{\mathcal{S}} = \{\widetilde{a}_0, \dots, \widetilde{a}_i, \dots\}$, and $\bar{\mathcal{S}} = \{\bar{a}_0, \dots, \bar{a}_i, \dots\}$.
- 2) $\widetilde{a}_i = \{\widetilde{a}_{ij}\}$ where $\widetilde{a}_{ij} = \frac{1}{\sqrt{p_{\max}}} \alpha_{ij}$ for $0 \leq i \leq k$ and $0 \leq j \leq p_i$; otherwise, $\widetilde{a}_{ij} = a_{ij}$.
- 3) $\bar{a}_i = \{\bar{a}_{ij}\}$ where $\bar{a}_{ij} = \frac{1}{\sqrt{p_{\max}\|A(\tau, x)\|}} \varphi_{jT}(\tau) Q_i(\tau, x)$ for $0 \leq i \leq k$ and $0 \leq j \leq p_i$; otherwise, $\bar{a}_{ij} = a_{ij}$.

Since Riesz-Fischer theorem is satisfied by both $\widetilde{\mathcal{S}}$ and $\bar{\mathcal{S}}$, there exist two functions, denoted by $F(t, x)$ and $G(t, x)$, such that for all $t \in [0, T]$,

$$F(t, Z(t)) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \widetilde{a}_{ij} \varphi_{jT}(t) Q_i(t, Z(t)), \quad (4.18)$$

$$G(t, Z(t)) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \bar{a}_{ij} \varphi_{jT}(t) Q_i(t, Z(t)). \quad (4.19)$$

Therefore, by virtue of equations in (4.18) and (4.19),

$$\frac{1}{\sqrt{p_{\max}}} \alpha' X' = \mathbf{F}' - \tilde{\delta}' - \tilde{\gamma}', \quad (4.20)$$

$$\frac{1}{\sqrt{p_{\max} \|A(\tau, x)\|}} A(\tau, x)' X' = \mathbf{G}' - \tilde{\delta}' - \tilde{\gamma}', \quad (4.21)$$

where

$$\mathbf{F}' = (F(t_{1,n}, X_{1,n}), \dots, F(t_{n,n}, X_{n,n})), \mathbf{G}' = (G(t_{1,n}, X_{1,n}), \dots, G(t_{n,n}, X_{n,n})),$$

$$\tilde{\delta}' = (\tilde{\delta}_1, \dots, \tilde{\delta}_n), \text{ with } \tilde{\delta}_s = \sum_{i=0}^k \sum_{j=p_i+1}^{\infty} a_{ij} \varphi_{jT}(t_{s,n}) Q_i(t_{s,n}, X_{s,n}),$$

$$\tilde{\gamma}' = (\tilde{\gamma}_1, \dots, \tilde{\gamma}_n), \text{ with } \tilde{\gamma}_s = \sum_{i=k+1}^{\infty} \sum_{j=0}^{\infty} a_{ij} \varphi_{jT}(t_{s,n}) Q_i(t_{s,n}, X_{s,n}).$$

The following proposition demonstrates the finiteness of the second moments of $F(t, Z(t))$ and $G(t, Z(t))$.

Proposition 4.2. *For any $t \in [0, T]$, (a) $E[F^2(t, Z(t))] < \infty$, and (b) $E[G^2(t, Z(t))] < \infty$.*

In order to obtain an asymptotic theory for \hat{m} , we make the following assumptions for the truncation parameters.

Assumption 4.6

- (a) Let $k = [n^{\kappa_1}]$ and $\frac{1}{2} < \kappa_1 < 1$
- (b) Let $p_{\min} = [n^{\bar{\kappa}_2}]$, $p_{\max} = [n^{\kappa_2}]$ with $0 < \kappa_2 \leq \bar{\kappa}_2 < 1$ and $0 \leq \bar{\kappa}_2 - \kappa_2 < 3\kappa_2 - \kappa_1 - 1$.

Clearly, feasible solutions of such truncation parameters do exist. The last assumption is about the functions $F(t, x)$, $G(t, x)$ and $m(\tau, x)$.

Assumption 4.7 Both $F(t, x)$ and $G(t, x)$ are continuous in t and x .

We are now ready to establish the following theorem.

Theorem 4.2. *Suppose that $\{x_{s,n}\}_1^n$ and $\{e_s\}_1^n$ satisfy Assumption B. Under Assumptions 4.5–4.7 we have as $n \rightarrow \infty$,*

$$\begin{aligned} & \frac{1}{\sqrt{n}} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max} \|A(\tau, x)\|^2}} (\hat{m}(\tau, x) - m(\tau, x)) \\ & \rightarrow_D \int_0^1 F(Tr, T\mu r + \sqrt{T}\sigma_z W(r)) dU(r), \end{aligned} \quad (4.22)$$

where $(U(r), W(r))$ is the vector of Brownian motions involved in Assumption B.

See Remark B.6 for the discussion on the theorem.

4.3 Time horizon approaching infinity

We are also interested in the scenario where time variable lies in $[0, T_n]$ and $T_n \rightarrow \infty$ as $n \rightarrow \infty$.

The relationship between T_n and n is crucial for the following development. Both of them are divergent. The divergence of T_n , however, is negligible comparing with that of n , viz., $\frac{T_n}{n} \rightarrow 0$ as $n \rightarrow \infty$. The main reason is that the proposed method requires sufficient information from the path of the process to estimate the coefficients in the expansion.

In addition, such designed framework will help us avoid two drawbacks, that is, on $(0, \infty)$ we could not shrink the time span of observations, whereas on $[0, T]$ with fixed T we ignore considerable information beyond the time zone that may be helpful for our estimation. In technical terms, allowing $T = T_n \rightarrow \infty$ and $\frac{T_n}{n} \rightarrow 0$ amounts to both infill and long span asymptotics. Meanwhile, the two-fold limit theory keeps ones away from the so-called aliasing problem (i.e., different continuous-time processes may be indistinguishable when sampled at discrete time). Phillips (1973) and Hansen and Sargent (1983) were among the first discussing the aliasing phenomenon in the literature. Recent studies include Bandi and Phillips (2003, 2007).

We propose the following assumptions for the function $m(t, x)$ in the model (4.1).

Assumption 4.8

- (a) For every $t > 0$, $m(t, x)$ and $D^r m(t, x)$ are all in $L^2(I, \rho_r(t, x))$, $r = 1, 2, 3$.
- (b) For each i , $b_i(t, m) = E[m(t, Z(t))Q_i(t, Z(t))]$, belongs to $C^3[0, T]$ for any $T > 0$.
- (c) For i large enough, the coefficient functions $b_i(t, D^3 m)$ of $D^3 m(t, Z(t))$ expanded by the system $\{Q_{3i}(t, Z(t))\}$ are such that $\psi(t)^3 b_i^2(t, D^3 m)$ are bounded on $(0, \infty)$ uniformly in i .
- (d) $\|b_i''(t, m)\|_{L^2[0, T]}$ are bounded uniformly in i for any $T > 0$.

Remark 4.5. Since the framework in this subsection is a combination of the first two, the requirements for $m(t, x)$ contain the basic conditions in Assumptions 4.1 and 4.5.

There are many functions that satisfy these four conditions at the same time. For instance, $m(t, x) = t^\eta e^{-ct} P(x)$ with $\eta \geq 1$, $c > 0$ and $P(x)$ being any polynomial of fixed degree; $m(t, x) = \frac{t}{1+t^\eta} \cos x$ with $\eta \geq 3$, and so on.

For the truncation parameters and time span T_n , we make the following assumption.

Assumption 4.9

- (a) Let $k = \lceil n^{\kappa_1} \rceil$, $p_{\min} = \lceil n^{\kappa_2} \rceil$, $p_{\max} = \lceil n^{\kappa_3} \rceil$ and $T_n = \lceil n^{\kappa_3} \rceil$, where $0 < \kappa_i < 1$ ($i = 1, 2, 3$), $\kappa_2 \leq \bar{\kappa}_2 < 1$ and $\kappa_1 > \frac{1}{2}$.
- (b) Let $3\kappa_3 + \kappa_1 + 1 < 3\kappa_2$.

Remark 4.6. Feasible solutions for κ_i ($i = 1, 2, 3$) do exist. For instance, $\kappa_1 = 0.6$, $\kappa_2 = 0.8$ and $\kappa_3 = 0.2$. Meanwhile, condition (b) implies that $\kappa_2 > \frac{1}{2} + \kappa_3$.

Given the observation number n , one can choose $T = T_n$ according to Assumption 4.9. Let us sample on $[0, T_n]$ at equally spaced points: $t_{s,n} = T_n \frac{s}{n}$ ($s = 1, \dots, n$) for model (4.1). Denote by $Y_{s,n}$ the process $Y(t)$ at $t_{s,n}$, $X_{s,n} = Z(t_{s,n})$ for the Lévy process at the discrete points and $e_s = \varepsilon(t_{s,n})$. Observe that $X_{s,n} = \frac{s}{n} T_n \mu + \sqrt{T_n} \sigma_z \frac{1}{\sqrt{n}} \sum_{i=1}^s w_i$ where $w_i = \frac{\sqrt{n}}{\sqrt{T_n} \sigma_z} (X_{i,n} - X_{i-1,n} - \frac{1}{n} T_n \mu)$ form an i.i.d (0,1) sequence. Let $x_{s,n} = \frac{1}{\sqrt{n}} \sum_{i=1}^s w_i$. It therefore follows from the functional central limit theorem that $x_{s,n}$ converges in distribution to a Brownian motion on $[0, 1]$ as $n \rightarrow \infty$. In addition, it is clear that $x_{s,n}$ satisfies Assumption A.

The following procedure is similar to the preceding subsections. The $m(t, Z(t))$ is expanded using an orthonormal basis $\{\varphi_{jT_n}(t) Q_i(t, Z(t))\}$ at each sampling point, and then obtain n equations. The n equations can be written in the following matrix form

$$Y = X\beta + \delta + \gamma + \varepsilon, \tag{4.23}$$

where all notations remain the similar meanings as before, so that we spare our effort to recite them.

The SLS estimator of β is given by

$$\hat{\beta} = (X'X)^{-1} X'Y. \tag{4.24}$$

Obtaining $\hat{\beta}$ enables us to estimate $m(\tau, x)$ for fixed $\tau > 0$ and fixed x on the path of $Z(\tau)$. $\hat{m}(\tau, x)$ is generated from the expansion of $m(\tau, x)$ by superceding β by $\hat{\beta}$ and removing all residues. Whence, we have $\hat{m}(\tau, x) = A'(\tau, x)\hat{\beta}$, where $A'(\tau, x) = (\varphi_{0T_n}(\tau)Q_0(\tau, x), \dots, \varphi_{p_0T_n}(\tau)Q_0(\tau, x), \dots, \varphi_{0T_n}(\tau)Q_k(\tau, x), \dots, \varphi_{p_kT_n}(\tau)Q_k(\tau, x))$. The difference between $\hat{m}(\tau, x)$ and $m(\tau, x)$ is

$$\begin{aligned} \hat{m}(\tau, x) - m(\tau, x) &= A'(\tau, x)(\hat{\beta} - \beta) - \delta(\tau, x) - \gamma(\tau, x) \\ &= A'(\tau, x)(X'X)^{-1} X'(\delta + \gamma + \varepsilon) - \delta(\tau, x) - \gamma(\tau, x), \end{aligned} \tag{4.25}$$

where

$$\delta(\tau, x) = \sum_{i=0}^k \sum_{j=p_i+1}^{\infty} b_{ij} \varphi_{jT_n}(\tau) Q_i(\tau, x); \quad \gamma(\tau, x) = \sum_{i=k+1}^{\infty} b_i(\tau, m) Q_i(\tau, x).$$

Thus, one desired result is the asymptotic distribution of $\widehat{m}(\tau, x) - m(\tau, x)$. To this end, put

$$A = \frac{A(\tau, x)A'(\tau, x)}{\|A(\tau, x)\|^2} \quad \text{and} \quad B = (X'X)A(X'X)^{-1}. \quad (4.26)$$

Once again B has eigenvalues $\lambda_1 = 1, \lambda_2 = \dots = \lambda_p = 0$. Let unit vector α be the left eigenvector of B pertaining to λ_1 . Hence, we have $\alpha'B = \alpha'$ and $\|\alpha\| = 1$. Denote $\alpha' = (\alpha_{00}, \dots, \alpha_{0p_0}, \dots, \alpha_{k0}, \dots, \alpha_{kp_k})$ in accordance with $A(\tau, x)$.

Let us apply the reshuffling procedure for the set \mathcal{S} from Assumption 4.2 by α and $\frac{1}{\|A(\tau, x)\|}A(\tau, x)$. Denote by $\widetilde{\mathcal{S}}$ and $\bar{\mathcal{S}}$ the resulting sets:

- 1) $\widetilde{\mathcal{S}} = \{\widetilde{a}_0, \dots, \widetilde{a}_i, \dots\}$, and $\bar{\mathcal{S}} = \{\bar{a}_0, \dots, \bar{a}_i, \dots\}$.
- 2) $\widetilde{a}_i = \{\widetilde{a}_{ij}\}$ where $\widetilde{a}_{ij} = \sqrt{\frac{T_n}{p_{\max}}} \alpha_{ij}$ for $0 \leq i \leq k$ and $0 \leq j \leq p_i$; otherwise, $\widetilde{a}_{ij} = a_{ij}$.
- 3) $\bar{a}_i = \{\bar{a}_{ij}\}$ where $\bar{a}_{ij} = \sqrt{\frac{T_n}{p_{\max}}} \frac{1}{\|A(\tau, x)\|} \varphi_{jT_n}(\tau) Q_i(\tau, x)$ for $0 \leq i \leq k$ and $0 \leq j \leq p_i$; otherwise, $\bar{a}_{ij} = a_{ij}$.

Due to Riesz-Fischer theorem, for two sequences $\widetilde{\mathcal{S}}$ and $\bar{\mathcal{S}}$, there exist two functions, denoted by $\widetilde{F}(t, x)$ and $\widetilde{G}(t, x)$, such that

$$\widetilde{F}(t, Z(t)) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \widetilde{a}_{ij} \varphi_{jT_n}(t) Q_i(t, Z(t)), \quad (4.27)$$

$$\widetilde{G}(t, Z(t)) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \bar{a}_{ij} \varphi_{jT_n}(t) Q_i(t, Z(t)), \quad (4.28)$$

for any $t \in [0, T_n]$.

In view of the expressions of \widetilde{a}_{ij} and \bar{a}_{ij} , rewrite (4.27) and (4.28) as

$$\sqrt{\frac{T_n}{p_{\max}}} \alpha' X' = \widetilde{\mathbf{F}}' - \widetilde{\delta}' - \widetilde{\gamma}', \quad (4.29)$$

$$\frac{1}{\|A(\tau, x)\|} \sqrt{\frac{T_n}{p_{\max}}} A(\tau, x)' X' = \widetilde{\mathbf{G}}' - \widetilde{\delta}' - \widetilde{\gamma}', \quad (4.30)$$

where

$$\begin{aligned}\tilde{\mathbf{F}}' &= (\tilde{F}(t_{1,n}, X_{1,n}), \dots, \tilde{F}(t_{n,n}, X_{n,n})), \tilde{\mathbf{G}}' = (\tilde{G}(t_{1,n}, X_{1,n}), \dots, \tilde{G}(t_{n,n}, X_{n,n})), \\ \tilde{\delta}' &= (\tilde{\delta}_1, \dots, \tilde{\delta}_n), \text{ with } \tilde{\delta}_s = \sum_{i=0}^k \sum_{j=p_i+1}^{\infty} a_{ij} \varphi_{jT_n}(t_{s,n}) Q_i(t_{s,n}, X_{s,n}), \\ \tilde{\gamma}' &= (\tilde{\gamma}_1, \dots, \tilde{\gamma}_n), \text{ with } \tilde{\gamma}_s = \sum_{i=k+1}^{\infty} \sum_{j=0}^{\infty} a_{ij} \varphi_{jT_n}(t_{s,n}) Q_i(t_{s,n}, X_{s,n}).\end{aligned}$$

Proposition 4.3. *For any $t \in [0, T_n]$, (a) $E[\tilde{G}(t, Z(t))]^2 < \infty$, and (b) $E[\tilde{F}(t, Z(t))]^2 < \infty$.*

Let $\tilde{G}(t, x) = \tilde{G}(t, \mu t + x - \mu t) := G(t, x - \mu t)$ and $\tilde{F}(t, x) = \tilde{F}(t, \mu t + x - \mu t) := F(t, x - \mu t)$. These reforms are because we are working on the centralized underlying process.

Assumption 4.10

- (a) Both $F(t, x)$ and $G(t, x)$ are in Class (HI) with normal functions $f(t, x)$, $g(t, x)$ and homogeneity powers $v(\cdot)$ and $\varrho(\cdot)$ respectively. Let $v(n) = n^\varsigma$ and $\varrho(n) = n^\iota$ satisfying
 - (i) $1 + \kappa_1 + (2\iota + 2.5)\kappa_3 < 3\kappa_2$; (ii) $1 + (2\iota - 0.5)\kappa_3 < 2.5\kappa_1$.
- (b) Suppose further that $F^2(t, x)$, $G^2(t, x)$ and $F(t, x)G(t, x)$ are in Class (HI) with normal functions $f^2(t, x)$, $g^2(t, x)$ and $f(t, x)g(t, x)$ and homogeneity powers $v^2(\cdot)$, $\varrho^2(\cdot)$ and $v(\cdot)\varrho(\cdot)$ respectively.
- (c) Both $F(t, x)$ and $G(t, x)$ are in Class (HH) with normal functions $f(t, x)$, $g(t, x)$ and homogeneity powers $v_1(\cdot)$, $v_2(\cdot)$ and $\varrho_1(\cdot)$, $\varrho_2(\cdot)$ respectively. Let $v_1(n) = n^{\varsigma_1}$, $v_2(n) = n^{\varsigma_2}$ and $\varrho_1(n) = n^{\iota_1}$, $\varrho_2(n) = n^{\iota_2}$ satisfying (i) $1 + \kappa_1 + (2\iota_1 + \iota_2 + 3)\kappa_3 < 3\kappa_2$; (ii) $1 + (2\iota_1 + \iota_2)\kappa_3 < 2.5\kappa_1$.
- (d) Suppose further that $F^2(t, x)$, $G^2(t, x)$ and $F(t, x)G(t, x)$ are in Class (HH) with normal functions $f^2(t, x)$, $g^2(t, x)$ and $f(t, x)g(t, x)$ and homogeneity powers $v_1^2(\cdot)$, $v_2^2(\cdot)$; $\varrho_1^2(\cdot)$, $\varrho_2^2(\cdot)$; $v_1(\cdot)\varrho_1(\cdot)$, $v_2(\cdot)\varrho_2(\cdot)$ respectively.

Remark 4.7. Note that the conditions in (a) and (c) are untidy since we would like to show the original requirement for the parameters.

It is clear that if $0 < \iota < 0.25$, Assumption 4.9 (b) implies the condition (i) of Assumption 4.10 (a); conversely, when $\iota \geq 0.25$ the latter always implies the former. Of

course, there are feasible options for them to satisfy all the requirements. For example, if $\kappa_1 = 0.7$, $\kappa_2 = 0.9$ and $\kappa_3 = 0.1$, then ι can be chosen from $(0, 3.5)$. By the way, if we impose some relationship among κ_i ($i = 1, 2, 3$), such as $\kappa_2 < \frac{7}{6}\kappa_1 + \kappa_3$, (i) implies (ii) in (a).

Let $\zeta = 2\iota_1 + \iota_2$ for the time being. Since $\zeta \geq 0$, the condition (i) in (c) always implies Assumption 4.9 (b). Evidently, if a relationship is imposed among κ_i ($i = 1, 2, 3$), (i) and (ii) in (c) may substitute each other, depending on what relationship is being imposed. Note that there are feasible choices for all parameters. For instance, $\kappa_1 = 0.6$, $\kappa_2 = 0.8$, $\kappa_3 = 0.1$, $\zeta \in (0, 2.5)$.

The following theorem is the main result of this subsection.

Theorem 4.3. *Suppose that $\{x_{s,n}\}_{s=1}^n$ and $\{e_s\}_{s=1}^n$ satisfy Assumption B. Let Assumptions 4.8–4.9 hold.*

If Assumption 4.10 (a) and (b) are true, then

$$\begin{aligned} & \frac{\sqrt[4]{T_n}^3 \sqrt{\sigma_z}}{\sqrt{nv}(T_n)} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max}} \|A(\tau, x)\|^2} (\widehat{m}(\tau, x) - m(\tau, x)) \\ & \rightarrow_D \left(\int_0^1 G_3(u) dL_W(u, 0) \right)^{\frac{1}{2}} N, \end{aligned} \quad (4.31)$$

where $G_3(\cdot) = \int f^2(\cdot, x) dx$, W is a standard Brownian motion on $[0, 1]$ and N is a standard normal random variable independent of W , and L_W is the local-time process of W .

If Assumption 4.10 (c) and (d) are true, then

$$\begin{aligned} & \frac{\sqrt{T_n}}{\sqrt{nv_1}(T_n) v_2(\sqrt{T_n} \sigma_z)} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max}} \|A(\tau, x)\|^2} (\widehat{m}(\tau, x) - m(\tau, x)) \\ & \rightarrow_D \int_0^1 f(r, W(r)) dU(r), \end{aligned} \quad (4.32)$$

where vector $(W(r), U(r))$ of Brownian motions is from Assumption B.

See Remark B.7 for the discussion on the theorem.

5 Conclusion and discussion

We have established orthogonal expansions of Lévy process functionals for both time-homogeneous and time-inhomogeneous ones under consideration of time horizon being

finite and infinite. With a verity of options of process sequence constructed from underlying process and error sequence verifying conventional assumptions, we have studied the convergence of sample mean and sample covariance for four classes of quite general functionals, which are applicable as seen in Section 4. As an application of the expansions and asymptotic theory, we have developed a sophisticated method to tackle the statistical estimation problem in nonlinear and nonstationary continuous-time models and proved the proposed estimators are consistent.

As far as what we noticed, the expansion method for unknown functionals of Lévy process can be used in economics and finance for some relevant research fields. For instance, in economics there are a great deal of models with conditional moment restriction containing unknown functionals in nonstationary process; in finance, more often than not, derivative pricing problem is associated with a functional, much popular nowadays, in a general Lévy process rather than only a Brownian motion. It can be expected that our expansion method is applicable in complete financial market for perfect hedging problem and in incomplete financial market for mean-variance hedging problem.

6 Acknowledgments

The authors acknowledge the financial support from the Australian Research Council Discovery Grants Program under Grant Number: DP1096374.

A Existence of orthogonal polynomial system associated to a Lévy process

A Lévy process $(Z(t), t \geq 0)$ is a stochastic process which has independent increments (for $0 \leq s \leq t$, $Z(t) - Z(s)$ is independent of $\{Z(u), u \leq s\}$), stationary distribution (for $0 \leq s \leq t$, $Z(t) - Z(s)$ is equal in distribution to $Z(t - s)$) and starts almost surely at zero ($\mathbb{P}(Z(0) = 0) = 1$).

Let $\phi(\theta)$ be the characteristic function of $Z(1)$, viz. $\phi(\theta) = E[e^{i\theta Z(1)}]$. It follows from the infinite divisibility of the distribution of $Z(t)$ that $E[e^{i\theta Z(t)}] = (\phi(\theta))^t$. According to Schoutens (2000, p. 50), it can be shown that

$$\frac{\phi'(\theta)}{\phi(\theta)} = i(\mu + \sigma^2 \tau(i\theta)) \tag{A.1}$$

for some function $\tau(\cdot)$ with $\tau(0) = 0$, where $\mu = EZ(1)$ and $\sigma^2 = \text{Var}(Z(1))$.

Let $u(\cdot)$ be the inverse function of $\tau(\cdot)$ and define $\pi(z) = [\phi(-iu(z))]^{-1}$.

Definition A.1 (Lévy-Meixner system). A polynomial set $\{q_i(t, x), i \geq 0, t \geq 0\}$ is called a Lévy-Meixner system if it is defined by a generating function of the form

$$\sum_{i=0}^{\infty} q_i(t, x) \frac{z^i}{i!} = (\pi(z))^t \exp(xu(z)). \quad (\text{A.2})$$

Such $q_i(t, x)$ functions are orthogonal with respect to the distribution $\Psi_t(x)$ of $Z(t)$: $\int q_i(t, x)q_j(t, x)d\Psi_t(x) = \delta_{ij}\tilde{d}_i^2(t)$, where $\tilde{d}_i^2(t)$ is the squared norm of $q_i(t, x)$.

Remark A.1. When process $Z(t)$ is specified as Brownian motion, $q_i(t, x)$ becomes Hermite polynomial with the density of normal distribution $N(0, t)$ being the weight function; when $Z(t)$ is specified as Gamma process, $q_i(t, x)$ will be the Laguerre polynomial system $L_i^{(\alpha t)}(x)$; if $Z(t) = N(t)$ a Poisson process with intensity μ , $q_i(t, x)$ will be the Charlier polynomial system $c_i(\mu t, x)$; if $Z(t)$ is a Pascal process, $q_i(t, x)$ will be the Meixner polynomial system.

In what follows we are about to show some explicit expressions, orthogonality, squared norm for $q_i(t, x)$ and its derivatives. Since their derivations are analogous to those given in Nikiforov and Uvarov (1988), we omit them for brevity.

Let us now consider differential and difference equations of hypergeometric type with parameter $t > 0$:

$$s(t, x)y''(t, x) + v(t, x)y'(t, x) + \lambda(t)y(t, x) = 0; \quad (\text{A.3a})$$

$$s(t, x)\Delta\nabla y(t, x) + v(t, x)\Delta y(t, x) + \lambda(t)y(t, x) = 0, \quad (\text{A.3b})$$

where $s(t, x)$ and $v(t, x)$ are polynomials in x of degree at most 2 and 1 respectively, while $\lambda(t)$ is independent of x . Note that in the equations of (A.3), and in the sequel, all derivatives and differences are conducted with respect to x , not to t .

Remember we denote by D in the text differential or difference operation for notational convenience.

It is known that when $\lambda(t) \equiv \lambda_i(t) = -iv'(t, x) - \frac{i(i-1)}{2}s''(t, x)$, $y(t, x) = y_i(t, x)$, as a solution of (A.3), is a polynomial in x of degree exactly i . In addition, if $\rho(t, x)$ satisfies

$$\begin{aligned} (i) \quad & D(s(t, x)\rho(t, x)) = v(t, x)\rho(t, x), \\ (ii) \quad & s(t, x)\rho(t, x)x^k|_a^b = 0, \quad k = 0, 1, \dots, \end{aligned} \quad (\text{A.4})$$

where a and b are the boundary points of the support of $\rho(t, x)$ (so that the condition (ii) of (A.4) is called boundary condition), we have explicit expressions of $y_i(t, x)$ and $D^k y_i(t, x)$:

$$D^k y_i(t, x) = \frac{A_{ki} B_i}{\rho_k(t, x)} D^{i-k} [\rho_i(t, x)], \quad (k = 0, 1, \dots, i) \quad (\text{A.5})$$

where $A_{0,1} = 1$, $A_{ki} := \prod_{j=0}^{k-1} (\lambda_i(t) - \lambda_j(t))$ for $k = 1, \dots, i$ and $B_i = \frac{1}{A_{ii}} y_i^{(i)}(t, x)$, $\rho_k(t, x) = \rho(t, x + k) \prod_{j=1}^k s(t, x + j)$ in the discrete case and $\rho_k(t, x) = s^k(t, x) \rho(t, x)$ in the continuous case for $k = 1, 2, \dots$, with convention of $\rho_0(t, x) = \rho(t, x)$.

Moreover, such $y_i(t, x)$ and $D^k y_i(t, x)$ are orthogonal on (a, b) with respect to $\rho(t, x)$ and $\rho_k(t, x)$, respectively. We also have with $d_{ki}^2(t)$ standing for the squared norm of $D^k y_i(t, x)$ that

$$d_{ki}^2(t) = d_i^2(t) \prod_{j=0}^{k-1} \eta_{ji}(t), \quad (\text{A.6})$$

for $k \geq 1$ where $d_i^2(t) := d_{0i}^2(t)$, $\eta_{0i}(t) = \lambda_i(t)$ and $\eta_{ji}(t) = \lambda_i(t) - \lambda_j(t)$.

Remark A.2. Because an orthogonal polynomial system is determined uniquely up to a normalizing factor by the interval (a, b) and the weight $\rho(t, x)$, it is not difficult to obtain that $q_i(t, x) = \frac{\tilde{d}_i(t)}{d_i(t)} y_i(t, x)$.

If for Lévy process $Z(t)$, $\rho(t, x)$ satisfies the conditions in (A.4), as stated before, there is an orthogonal polynomial system $y_i(t, x)$ with weight $\rho(t, x)$. Let us define $Q_i(t, x)$ and $Q_{ki}(t, x)$ by

$$Q_i(t, x) = \frac{1}{d_i(t)} y_i(t, x) \quad \text{and} \quad Q_{ki}(t, x) = \frac{1}{d_{ki}(t)} D^k y_i(t, x), \quad (\text{A.7})$$

each of which is an orthonormal polynomial system with either $\rho(t, x)$ or $\rho_k(t, x)$ being its weight. In such a situation, we say that $Z(t)$ admits a *classical orthonormal polynomial system* $Q_i(t, x)$.

Example A.1

(1) If $Z(t) = B(t)$ is a Brownian motion with $\rho(t, x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}$, the corresponding hypergeometric differential equations are

$$t y''(t, x) - x y'(t, x) + i y(t, x) = 0,$$

which have polynomial solution $y_i(t, x) = H_i(x/\sqrt{t})$, where $H_i(\cdot)$ for $i \geq 1$ are Hermite polynomials.

(2) If $Z(t) = G(t)$ is a Gamma process with $\rho(t, x) = \frac{1}{\Gamma(1+\alpha t)} x^{\alpha t} e^{-x}$, the corresponding hypergeometric differential equations are

$$xy''(t, x) + (\alpha t + 1 - x)y'(t, x) + iy(t, x) = 0,$$

which have polynomial solution $y_i(t, x) = L_i^{(\alpha t)}(x)$, where $L_i^{(\cdot)}(\cdot)$ for $i \geq 1$ are Laguerre polynomials.

(3) If $Z(t) = N(t)$ is a Poisson process with $\rho(t, x) = e^{-\mu t} \frac{(\mu t)^x}{x!}$, $x = 0, 1, 2, \dots$, the corresponding hypergeometric difference equations are

$$x\Delta\nabla y(t, x) + (\mu t - x)\Delta y(t, x) + iy(t, x) = 0,$$

which have polynomial solution $y_i(t, x) = c_i^{(\mu t)}(x)$, where $c_i^{(\cdot)}(\cdot)$ for $i \geq 1$ are Charlier polynomials.

Remark A.3. It follows from (A.6) that for any i , $\lambda_i(t) > 0$. That entails that $v'(t, x) < 0$ and $s''(t, x) = 0$ or $v'(t, x) < 0$ and $s''(t, x) < 0$. The former includes three processes in Example A.1, while in the latter, after a transformation $s(t, x)$ can be written as $c^2 - x^2$ with fixed $c > 0$. However, this scenario is beyond the scope of this paper since we are interested in that $Z(t)$ assumes values on infinite interval or set, specifically, \mathbb{R} , \mathbb{R}^+ or \mathbb{N} . Therefore, our development will focus on the case where $v'(t, x) < 0$ and $s''(t, x) = 0$.

Although in our examples $v'(t, x) = -1$, in order to keep the framework as general as possible, we shall always treat $v'(t, x)$ as a negative function of t . Denote $\psi(t) := -[v'_x(t, x)]^{-1} > 0$, which is used frequently in the paper.

Remark A.4. We may also need some asymptotic properties about the orthogonal polynomials. In the sequel, the following inequalities for Hermite polynomials and Laguerre polynomials are useful, which can be found in Nikiforov and Uvarov (1988, p54) for large i ,

$$\frac{1}{d_i} |H_i(x)| \leq C_1 i^{-\frac{1}{4}} \quad \text{and} \quad \frac{1}{d_i} |L_i^{(\alpha)}(x)| \leq C_2 i^{-\frac{1}{4}},$$

where d_i 's are the norm of Hermite and Laguerre polynomials in different inequalities respectively; C_1 and C_2 only depend on fixed x .

In addition, in view of the relation $c_i(\mu, x) = c_x(\mu, i) = x! L_x^{(i-x)}(\mu)$, the above inequality is true for Charlier polynomials as well. Thus, we may assert that within the ambit of our study, all classical orthonormal polynomials $Q_i(t, x)$ satisfy that $|Q_i(t, x)| \leq C i^{-\frac{1}{4}}$ for fixed t and x , where C is independent of i .

B Remarks, justifications and examples

Remark B.1. We now give some examples of expansion of $f(Z(t))$.

(1) $f(Z(t)) = a_0 + a_1 Z(t) + \dots + a_k Z(t)^k$. Obviously, $f(Z(t))$ can be expanded by the first $k+1$ terms of $Q_i(t, Z(t))$ with coefficients $c_i(t) = E[f(Z(t))Q_i(t, Z(t))]$, $i = 0, 1, \dots, k$. We have two particular examples for $Z(t) = B(t)$, a Brownian motion and $Z(t) = N(t)$, a Poisson process with intensity 1:

$$\begin{aligned} B^5(t) &= 15t^{5/2}h_1(t, B(t)) + 10\sqrt{6}t^{5/2}h_3(t, B(t)) + 2\sqrt{30}t^{5/2}h_5(t, B(t)), \\ N^2(t) &= t(1+t)\mathcal{C}_0(t; N(t)) - \sqrt{t}(2t+1)\mathcal{C}_1(t; N(t)) + \sqrt{2t}\mathcal{C}_2(t; N(t)), \end{aligned}$$

where $h_i(t, B(t)) = \frac{1}{\sqrt{i!}}H_i(B(t)/\sqrt{t})$ with $H_i(\cdot)$ being Hermite polynomials, and $\mathcal{C}_i(t; N(t)) = \frac{\sqrt{t}^i}{\sqrt{i!}}c_i(t; n)$ with $c_i(\cdot; \cdot)$ being Charlier polynomials.

(2) $f(Z(t)) = \cos Z(t)$ and $g(Z(t)) = \sin Z(t)$.

$$\cos Z(t) = \sum_{j=0}^{\infty} b_j(t)Q_j(t, Z(t)), \quad \sin Z(t) = \sum_{j=0}^{\infty} c_j(t)Q_j(t, Z(t)),$$

where

$$\begin{aligned} b_j(t) &= \frac{1}{j!}\tilde{d}_j(t)\frac{1}{2}\{\tau(i)^j[\phi(1)]^t + \tau(-i)^j[\phi(-1)]^t\}, \\ c_j(t) &= \frac{1}{j!}\tilde{d}_j(t)\frac{1}{2i}\{\tau(i)^j[\phi(1)]^t - \tau(-i)^j[\phi(-1)]^t\}. \end{aligned}$$

in which τ , ϕ and $\tilde{d}_j(t)$ being defined in Appendix A and i in this example being imaginary unit.

Particularly, with α being a constant and $\beta = 1 - \cos 1$,

$$\begin{aligned} \sin B(t) &= \sum_{k=0}^{\infty} c_{2k+1}h_{2k+1}(t, B(t)) \quad \text{with } c_{2k+1} = (-1)^k \frac{\sqrt{t}^{2k+1}}{\sqrt{(2k+1)!}}e^{-t/2}, \\ \cos B(t) &= \sum_{k=0}^{\infty} c_{2k}h_{2k}(t, B(t)) \quad \text{with } c_{2k} = (-1)^k \frac{t^k}{\sqrt{(2k)!}}e^{-t/2}, \\ \cos N(t) &= \sum_{k=0}^{\infty} (-1)^k \frac{\sqrt{t}^k}{\sqrt{k!}}e^{-t\beta} \sqrt{2\beta}^k \cos(\alpha k + t \sin 1)\mathcal{C}_k(t; N(t)), \\ \sin N(t) &= \sum_{k=0}^{\infty} (-1)^k \frac{\sqrt{t}^k}{\sqrt{k!}}e^{-t\beta} \sqrt{2\beta}^k \sin(\alpha k + t \sin 1)\mathcal{C}_k(t; N(t)). \end{aligned}$$

Remark B.2 (Theorem 2.4). The error of approximation $f_{k,p}(t, Z(t))$ to $f(t, Z(t))$ consists of two types because the expansion is of two-step, that is, the first term in the right hand side of (2.7) is incurred since we abandon the residue in the first step expansion, while the second term is due to giving up the residues in the second step.

Because for each $i : 0 \leq i \leq k$, $\sum_{j=p_{\min}+1}^{\infty} b_j^2(c_i'')$ is an infinitesimal when p_{\min} goes to infinity, for fixed k , $C(k, p)$ is an infinitesimal as well. However, when both k and p_{\min} approach to infinity, $C(k, p)$ could not be infinitesimal any more. One sufficient condition that $C(k, p)$ is bounded is that the norm $\|c_i''(t, f)\|_{L^2[0, T]}$ is uniformly bounded in i , so we always have $\sum_{j=p_{\min}+1}^{\infty} b_j^2(c_i'') \leq \|c_i''(t, f)\|_{L^2[0, T]}^2$.

Remark B.3 (Definition 3.3). (a) If the functions involved in the definition reduce to univariate functions without time variable, i.e., $F(t, x) \equiv F(x)$, $v_1(\xi) = 1$, $f(t, x) \equiv f(x)$ and $R(\xi, \eta; t, x) \equiv R(\eta; x)$ with $q(t) \equiv 1$, $b(\xi) = 1$, it becomes the Class (H) in Park and Phillips (1999, 2001).

(b) In practice, often one of the two dominated terms of R appears. The only appearance of the first term implies that $q(t) = 0$, while the appearance of the second term indicates that $P(x) = 0$.

(c) There are many functions that have asymptotic homogeneity. For example,

(1). $F(t, x) = a_1 t^{m_1} x^{l_1} + \dots + a_k t^{m_k} x^{l_k}$ with $m_1 \geq \dots \geq m_k \geq 0$, $m_1 \geq 1$ and $l_1 \geq \dots \geq l_m \geq 0$, is homogeneous where $f(t, x) = a_1 t^{m_1} x^{l_1}$, $v_1(\xi) = \xi^{m_1}$, $v_2(\eta) = \eta^{l_1}$, and if $m_1 > m_2$, $|R(\xi, \eta; t, x)| \leq A_\xi(t) a(\eta) P(x)$ where $A_\xi(t) = |a_2| \xi^{m_2} t^{m_2} + \dots + |a_k| \xi^{m_k} t^{m_k}$, $a(\eta) = \eta^{l_2}$ and $P(x) = 1 + |x|^{l_2}$. Clearly, $\lim_{\xi \rightarrow \infty} \frac{A_\xi(t)}{v_1(\xi)} = 0$ uniformly in t . If $l_2 < l_1$, $|R(\xi, \eta; t, x)| \leq q(t) b(\xi) B_\eta(x)$ where $q(t) = 1 + t^{m_2}$, $b(\xi) = \xi^{m_2}$, $B_\eta(x) = \eta^{l_2} (1 + |x|^{l_2})$. Palpably, $\lim_{\eta \rightarrow \infty} \frac{\eta^{l_2}}{v_2(\eta)} = 0$ and $1 + |x|^{l_2} \in \mathcal{F}_{LB}^0$.

(2). $F(t, x) = t^\alpha \log(1 + |x|)$ with $\alpha \geq 1$. The normal function $f(t, x) = t^\alpha$ with $v_1(\xi) = \xi^\alpha$ and $v_2(\eta) = \log(\eta)$, while $R(\xi, \eta; t, x) \leq \xi^\alpha t^\alpha \log(1 + |x|)$. Notice that $b(\xi) = \xi^\alpha$, $q(t) = t^\alpha$, $B_\eta(x) = \log(1 + |x|)$ with $\bar{B}(\eta) = 1$ and $\log(1 + |x|) \in \mathcal{F}_{LB}^0$.

(3). $F(t, x) = t^2 x + \sqrt{1 + t^4} \frac{1}{1 + |\ln t|} \sqrt[3]{x}$. Note that $f(t, x) = t^2 x$, $v_1(\xi) = \xi^2$ and $v_2(\eta) = \eta$; while $R(\xi, \eta; t, x) \leq A_\xi(t) a(\eta) P(x)$, where $a(\eta) = \sqrt[3]{\eta}$, $P(x) = \sqrt[3]{x}$, $A_\xi(t) = \bar{A}_\xi(t) Q(\xi t)$ with $\bar{A}_\xi(t) = \sqrt{1 + \xi^4 t^4}$ and $Q(y) = \frac{1}{1 + |\ln y|}$.

(4). $F(t, x) = t^\alpha D(x)$ where $\alpha \geq 1$ and $D(x)$ is a distribution function for any random variable. Then $f(t, x) = t^\alpha I(x \geq 0)$, $v_1(\xi) = \xi^\alpha$, $v_2(\eta) = 1$, $R(\xi, \eta; t, x) < b(\xi) q(t) Q(\eta x)$ where $b(\xi) = \xi^\alpha$, $q(t) = t^\alpha$ and $Q(y) = D(y) I(y < 0) + (1 - D(y)) I(y \geq 0)$, which goes to zero when $y \rightarrow +\infty$.

Remark B.4 (Assumption 4.1). Note that the notations $\rho_r(t, x)$, $\psi(t)$ and $Q_{3i}(t, Z(t))$ are defined in Appendix A.

Condition (a) imposes some basic requirements, under which we can expand not only $m(t, Z(t))$ but also $D^r m(t, x)|_{x=Z(t)}$. Conditions (b) and (c) give the necessary conditions for the coefficient functions in order to obtain some kind of rate of convergence on the expansions.

There are many functionals satisfying all the conditions. (1) Let $m_1(t, x) = t^a e^{-bt} P_k(x)$ with $a \geq 1, b > 0$ and $P_k(x)$ being a polynomial of fixed degree k ($k \geq 1$). $m_1(t, x)$ satisfies Condition (a) due to the boundary condition on $\rho(t, x)$; the reason that $m_1(t, x)$ satisfies Condition (b) is that the coefficients $c_i(t, m_1)$ are all of the form $e^{-bt} q(t)$, where $q(\cdot)$ is a power function in t when $i \leq k$ and zero when $i > k$; Condition (c) is fulfilled because when $i > k$, $c_i(t, m_1) = 0$. (2) $m_2(t, x) = \frac{t^\alpha}{1+t^\beta} \sin(x)$ and $m_3(t, x) = \frac{t^\alpha}{1+t^\beta} \cos(x)$ where $\alpha \geq 1$ and $\beta \geq \alpha + 1.25$. In the Brownian motion case, from the example in the Remark B.1 we have explicit expression of the coefficients $c_i(t, m_2) = (-1)^k \frac{1}{\sqrt{i!}} \frac{t^\alpha \sqrt{t^i}}{1+t^\beta} e^{-t/2}$, for $i = 2k + 1$; 0, for $i = 2k$, where $k = 0, 1, \dots$ and $c_i(t, m_3) = (-1)^k \frac{1}{\sqrt{i!}} \frac{t^\alpha \sqrt{t^i}}{1+t^\beta} e^{-t/2}$ for $i = 2k$; 0, for $i = 2k + 1$, where $k = 0, 1, \dots$. It is not difficult to verify the conditions. (3) In the case where $Z(t) = N(t)$ is a Poisson process with intensity one, $m_4(t, x) = t^\xi 2^{-x}$ where $\xi \geq 2$, $m_5(t, x) = \frac{t^\xi}{1+t^\eta} \sin x$ and $m_6(t, x) = \frac{t^\xi}{1+t^\eta} \cos x$ with $\xi \geq 1$ and $\eta \geq \xi + 1.25$. Since $c_i(t, m_4) = t^\xi e^{-t/2} \frac{1}{2^i} \sqrt{\frac{t^i}{i!}}$, the conditions are easy to be verified for m_4 . Meanwhile, from example 3.1 we can have the explicit expressions of $c_i(t, m_5)$ and $c_i(t, m_6)$ and it is not difficult to verify these conditions too.

Remark B.5 (Theorem 4.1). As can be seen from the proof, the order of the convergence of (4.11) is $\frac{\sqrt[4]{n} \sqrt{p_{\max} \varrho(n)}}{\|A(\tau, x)\|}$. By virtue of the calculation of $\|A(\tau, x)\|$ in the proof, we can estimate

$$n^{\frac{1}{4} + \iota - \frac{1}{2} \kappa_1} \leq \frac{\sqrt[4]{n} \sqrt{p_{\max} \varrho(n)}}{\|A(\tau, x)\|} \leq n^{\frac{1}{4} + \frac{1}{2}(\bar{\kappa}_2 - \kappa_2) + \iota - \frac{1}{2} \kappa_1} \leq n^{\frac{5}{4} \kappa_2 - \kappa_1}.$$

This means when ι reaches its upper bound, the rate of convergence is bounded by $n^{\frac{5}{4} \kappa_2 - \kappa_1}$, while when ι is close to $\frac{1}{2}(\kappa_1 - \frac{1}{2})$, the rate of convergence is very slow.

Meanwhile, the rate of convergence for (4.12) is $\frac{\sqrt{n} \sqrt{p_{\max} \varrho_1(n)} \varrho_2(\sqrt{n} \sigma_x)}{\|A(\tau, x)\|}$, which similarly is between $n^{\frac{1}{2}(1 - \kappa_1) + \iota_1 + \frac{1}{2} \iota_2}$ and $n^{\frac{1}{2}(1 - \kappa_1) + \iota_1 + \frac{1}{2} \iota_2 + \frac{1}{2}(\bar{\kappa}_2 - \kappa_2)} \leq n^{\frac{5}{4} \kappa_2 - \kappa_1}$, and as $\iota_1 + \frac{1}{2} \iota_2$ reaches its upper bound, the rate is bounded by $n^{\frac{5}{4} \kappa_2 - \kappa_1}$, while as $\iota_1 + \frac{1}{2} \iota_2$ closes to $\frac{1}{2}(\kappa_1 - \frac{1}{2})$ the rate of convergence is very slow, the same as in the first situation.

Remark B.6 (Theorem 4.2). As can be seen from the proof, the rate of convergence of

$\widehat{m}(\tau, x) - m(\tau, x)$ is about $\frac{\sqrt{np_{\max}}}{\|A(\tau, x)\|}$. In view of the estimation of $\|A(\tau, x)\|$, the rate is between $n^{\frac{1}{2}(1-\kappa_1)}$ and $n^{\frac{1}{2}(1-\kappa_1)+\frac{1}{2}(\bar{\kappa}_2-\kappa_2)}$. The minimum order is smaller than $\frac{1}{4}$, while the maximum order is slightly bigger than the minimum.

Remark B.7 (Theorem 4.3). The rate of convergence of $\widehat{m}(\tau, x) - m(\tau, x)$ in the first case, as can be seen in its proof, is about $\frac{\sqrt{n}\sqrt{p_{\max}}\varrho(T_n)}{\sqrt[4]{T_n^3}\|A(\tau, x)\|}$, which is between $n^{\frac{1}{2}(1-\kappa_1)+(\iota-\frac{1}{4})\kappa_3}$ and $n^{\frac{1}{2}(1-\kappa_1)+(\iota-\frac{1}{4})\kappa_3+\frac{1}{2}(\bar{\kappa}_2-\kappa_2)}$. The order of the lower bound is less than $\frac{1}{4}$, while the order of the upper bound is less than $\frac{1}{2}$.

In the second case, the convergence rate is $\frac{\sqrt{n}\sqrt{p_{\max}}\varrho_1(T_n)\varrho_2(\sqrt{T_n}\sigma_z)}{\sqrt{T_n}\|A(\tau, x)\|}$ revealed by its proof. Approximately, it is between $n^{\frac{1}{2}(1-\kappa_1)+(\iota_1+\frac{1}{2}\iota_2)\kappa_3}$ and $n^{\frac{1}{2}(1-\kappa_1)+(\iota_1+\frac{1}{2}\iota_2)\kappa_3+\frac{1}{2}(\bar{\kappa}_2-\kappa_2)}$.

Comparing the upper bounds and the lower bounds in two scenarios, roughly speaking, the second situation is faster than the first.

C Lemmas and basic results

PROOF OF LEMMA 2.1. (1) Straightforward verification. (2) For any functions $f, g \in L^2(I, d\Psi_t(x))$, we have, $\langle \mathcal{T}(f), \mathcal{T}(g) \rangle_{\Theta} = E[f(Z(t))g(Z(t))] = \int_I f(x)g(x)d\Psi_t(x) = (f, g)_{L^2(I, d\Psi_t(x))}$.

That means the transformation is inner product preserving. Therefore, $f \neq g \Leftrightarrow \mathcal{T}(f) \neq \mathcal{T}(g)$. Thus \mathcal{T} is one-one. (3) Since \mathcal{T} is linear and $\|\mathcal{T}(f)\| = \|f\|$ for $f \in L^2(I, d\Psi_t(x))$, \mathcal{T} is isomorphism. \square

PROOF OF LEMMA 2.2. Note that Θ is a linear space due to linearity of \mathcal{T} . Because \mathcal{T} is one-to-one and inner product preserving, $\{\xi_n\}$ is a Cauchy sequence in Θ if and only if there is a unique sequence $\{f_n\}$ in $L^2(I, d\Psi_t(x))$ such that $\mathcal{T}(f_n) = \xi_n$, $n = 0, 1, 2, \dots$, and $\{f_n(x)\}$ is a Cauchy sequence in $L^2(I, d\Psi_t(x))$. Therefore, due to the completeness of $L^2(I, d\Psi_t(x))$, Θ is a closed subspace of $L^2(\Omega)$. Hence it is a Hilbert space. \square

PROOF OF LEMMA 2.3. By virtue of the properties of \mathcal{T} that \mathcal{T} is one-to-one, inner product preserving, it is valid. \square

It is known from P351 of Sansone (1959) that $\mathcal{L}_j^{(\alpha)}(t), j = 0, 1, \dots$, form an orthonormal basis in $L^2(\mathbb{R}^+) = \{\varphi(t) : \int_0^\infty \varphi(t)^2 dt < \infty\}$ where for $\alpha \geq 0$, $\mathcal{L}_j^{(\alpha)}(t) := \left(\Gamma(\alpha+1)C_{j+\alpha}^j\right)^{-1/2} t^{\alpha/2} e^{-t/2} L_j^{(\alpha)}(t)$ and $\{L_j^{(\alpha)}(t)\}_0^\infty$ is the generalized Laguerre polynomial system which forms a complete orthogonal sequence with respect to the density $t^\alpha e^{-t}$. Thus, a function in the space $L^2(\mathbb{R}^+)$ can be expanded as: $\varphi(t) = \sum_{j=0}^\infty a_j^{(\alpha)} \mathcal{L}_j^{(\alpha)}(t)$. Designate $a_j = a_j^{(0)}$ for convenience.

Lemma C.1. Suppose that $\varphi(t) \in L^2(\mathbb{R}^+)$ is r -th differentiable such that $t^{\frac{r}{2}}\varphi^{(v)}(t), v = 0, 1, \dots, r$, are in the space $L^2(\mathbb{R}^+)$ as well. Let $\varphi_p(t) = \sum_{j=0}^p a_j \mathcal{L}_j(t)$ be the truncation series of $\varphi(t)$. Then

$$\|\varphi(t) - \varphi_p(t)\|^2 \leq \frac{(p+1-r)!}{(p+1)!} R^2(p), \quad (\text{C.1})$$

$$|\varphi(t) - \varphi_p(t)|^2 \leq \frac{1}{(p-r+1)^{r-1}} \left(\sup_{j \geq p+1} |\mathcal{L}_j(t)| \right)^2 R^2(p), \quad (\text{if } r > 1) \quad (\text{C.2})$$

for sufficient large p , where $R^2(p) = \sum_{j=p+1}^{\infty} [a_{j-r}^{(r)}(\tilde{\varphi})]^2$ is an infinitesimal with $p \rightarrow \infty$ in which $\tilde{\varphi}(t) = t^{r/2} e^{-t/2} [\varphi(t) e^{t/2}]^{(r)}$.

Actually approximation of $\varphi_p(t)$ to $\varphi(t)$ in the above lemma is uniformly. The table on page 699 of Askey and Wainger (1965) shows that, given any $\alpha \geq 0$, there are positive constants C and γ , independent of j and t , such that for all integers $j \geq 0$,

$$|\mathcal{L}_j^{(\alpha)}(t)| \leq \begin{cases} Ct^{\alpha/2} m^{\alpha/2}, & \text{if } 0 < t \leq \frac{1}{m} \\ Ct^{-1/4} m^{-1/4}, & \text{if } \frac{1}{m} < t \leq \frac{m}{2} \\ Cm^{-3/4} (m^{1/3} + |t-m|)^{1/4}, & \text{if } \frac{m}{2} < t \leq \frac{3m}{2} \\ Ce^{-\gamma t}, & \text{if } t > \frac{3m}{2} \end{cases} \quad (\text{C.3})$$

where $m = 4j + 2\alpha + 2$.

We omit the proof of Lemma C.1 since it is a conventional result.

PROOF OF THEOREM 2.1. In view of the facts that Θ is a Hilbert space and $\{Q_i(t, Z(t))\}$ is an orthonormal basis in Θ , it follows. \square

PROOF OF THEOREM 2.3. It follows from Hilbert space theory. \square

PROOF OF THEOREM 2.5. For Λ is a Hilbert space with orthonormal basis $\{Q_i(t, Z(t)) \mathcal{L}_j(t)\}$, it follows immediately. \square

Lemma C.2 (The Occupation Time Formula). Let M_t be a continuous SMG with quadratic variation process $[M]_t$. Then,

$$\int_0^t f(s, M_s) d[M]_s = \int_{-\infty}^{\infty} da \int_0^t f(s, a) dL_M(s, a) \quad (\text{C.4})$$

for every positive Borel measurable function $f(t, x)$.

Observe that with the condition B (c) in Assumption B that $(U_n, W_n) \rightarrow_D (U, W)$ on $D[0, 1]^2$, it follows from the so-called Skorohod-Dudley-Wichura representation theorem that there is a common probability space $(\Omega, \mathcal{F}, \mathcal{P})$ supporting (U_n^0, W_n^0) and (U_n, W_n) such that

$$(U_n^0, W_n^0) =_D (U_n, W_n) \quad \text{and} \quad (U_n^0, W_n^0) \rightarrow_{a.s.} (U, W), \quad (\text{C.5})$$

in $D[0, 1]^2$ with uniform topology.

Lemma C.3. *Let Assumption B hold. We may represent U_n^0 introduced in (C.5) as*

$$U_n^0 \left(\frac{k}{n} \right) = U \left(\frac{\tau_{nk}}{n} \right)$$

with an increasing sequence of stopping times τ_{nk} in $(\Omega, \mathcal{F}, \mathcal{P})$ with $\tau_{n0} = 0$ such that as $n \rightarrow \infty$

$$\sup_{1 \leq k \leq n} \left| \frac{\tau_{nk} - k}{n^\delta} \right| \rightarrow_{a.s.} 0 \quad (\text{C.6})$$

for any $\delta > \max\{\frac{1}{2}, \frac{2}{q}\}$, where q is the moment exponent in Assumption B for $\{e_k\}$.

This lemma is exactly Lemma 2.1 in Park and Phillips (2001). Readers can find the proof there.

To study the convergence of the statistics in Theorem 3.1 we introduce for any $\epsilon > 0$ and $0 \leq r \leq 1$,

$$L_n^{(r)} = \frac{c_n}{n} \sum_{k=1}^{[nr]} f \left(\frac{k}{n}, c_n x_{k,n} \right) \quad \text{and} \quad L_{n,\epsilon}^{(r)} = \frac{c_n}{n} \sum_{k=1}^{[nr]} \int_{-\infty}^{\infty} f \left(\frac{k}{n}, c_n (x_{k,n} + z\epsilon) \right) \phi(z) dz,$$

where $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$. For later use we also define $\phi_\epsilon(z) = \frac{1}{\sqrt{2\pi\epsilon}} \exp\left(-\frac{z^2}{2\epsilon^2}\right)$ for some $\epsilon > 0$.

Lemma C.4. *Suppose that Assumptions C and A (c) hold. Then*

$$\lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} \sup_{0 \leq r \leq 1} E |L_n^{(r)} - L_{n,\epsilon}^{(r)}| = 0. \quad (\text{C.7})$$

Proof. The proof consists of two parts according to $x_{k,n}$ being continuous and discrete respectively in A (c).

The following arguments about the continuous case naturally treat those used for the univariate case in Wang and Phillips (2009a) as a special case.

Denote $Y_{k,n}(z) = f\left(\frac{k}{n}, c_n x_{k,n}\right) - f\left(\frac{k}{n}, c_n(x_{k,n} + z\epsilon)\right)$. We have

$$\begin{aligned} \sup_{0 \leq r \leq 1} E|L_n^{(r)} - L_{n,\epsilon}^{(r)}| &= \sup_{0 \leq r \leq 1} E \left| \frac{c_n}{n} \int_{-\infty}^{\infty} \sum_{k=1}^{[nr]} Y_{k,n}(z) \phi(z) dz \right| \\ &\leq \frac{c_n}{n} \int_{-\infty}^{\infty} \sup_{0 \leq r \leq 1} E \left| \sum_{k=1}^{[nr]} Y_{k,n}(z) \right| \phi(z) dz, \end{aligned}$$

by the fact that $\int \phi(z) dz = 1$. Notice that, by Assumption A (c),

$$\begin{aligned} E|Y_{k,n}(z)| &= \int_{-\infty}^{\infty} \left| f\left(\frac{k}{n}, c_n d_{k,0,n} x\right) - f\left(\frac{k}{n}, c_n d_{k,0,n} x + c_n z\epsilon\right) \right| h_{k,0,n}(x) dx \\ &\leq \frac{K}{c_n d_{k,0,n}} \left[\int_{-\infty}^{\infty} \left| f\left(\frac{k}{n}, x\right) \right| dx + \int_{-\infty}^{\infty} \left| f\left(\frac{k}{n}, x + c_n z\epsilon\right) \right| dx \right] \\ &= \frac{2K}{c_n d_{k,0,n}} G_2\left(\frac{k}{n}\right), \end{aligned} \tag{C.8}$$

where $G_2(\cdot) = \int_{-\infty}^{\infty} |f(\cdot, x)| dx$ and K is the uniform upper bound of the density $h_{l,k,n}$. Accordingly, for each $z \in \mathbb{R}$,

$$\frac{c_n}{n} \sup_{0 \leq r \leq 1} E \left| \sum_{k=1}^{[nr]} Y_{k,n}(z) \right| \leq \frac{c_n}{n} \sum_{k=1}^n \frac{2K}{c_n d_{k,0,n}} G_2\left(\frac{k}{n}\right) = 2K K_2 \frac{1}{n} \sum_{t=1}^n \frac{1}{d_{k,0,n}} < \infty$$

by virtue of (3.3), where $K_2 = \sup_{t \in [0,1]} G_2(t) < \infty$ due to the continuity of $G_2(t)$. It therefore follows from the dominated convergence theorem that, to prove the lemma, it suffices to show that for any fixed z ,

$$\Lambda_n(\epsilon) = \frac{c_n^2}{n^2} \sup_{0 \leq r \leq 1} E \left[\sum_{k=1}^{[nr]} Y_{k,n}(z) \right]^2 \rightarrow 0,$$

as $n \rightarrow \infty$ first and then $\epsilon \rightarrow 0$. Meanwhile, we have

$$\begin{aligned} \Lambda_n(\epsilon) &\leq \frac{c_n^2}{n^2} \sum_{k=1}^n E Y_{k,n}^2(z) + \frac{2c_n^2}{n^2} \sum_{k=1}^{n-1} \sum_{l=k+1}^n |E[Y_{k,n}(z) Y_{l,n}(z)]| \\ &:= \Lambda_{1n}(\epsilon) + \Lambda_{2n}(\epsilon). \end{aligned}$$

We next investigate $\Lambda_{1n}(\epsilon)$ and $\Lambda_{2n}(\epsilon)$ separately.

In view of Assumption A (c), we have as $n \rightarrow \infty$

$$\Lambda_{1n}(\epsilon) = \frac{c_n^2}{n^2} \sum_{k=1}^n E Y_{k,n}^2(z) = \frac{c_n^2}{n^2} \sum_{k=1}^n E \left[f\left(\frac{k}{n}, c_n x_{k,n}\right) - f\left(\frac{k}{n}, c_n(x_{k,n} + z\epsilon)\right) \right]^2$$

$$\begin{aligned}
&= \frac{c_n^2}{n^2} \sum_{k=1}^n \int_{-\infty}^{\infty} \left[f\left(\frac{k}{n}, c_n d_{k,0,n} x\right) - f\left(\frac{k}{n}, c_n d_{k,0,n} x + c_n z \epsilon\right) \right]^2 h_{k,0,n}(x) dx \\
&\leq \frac{c_n^2}{n^2} \sum_{k=1}^n \frac{K}{c_n d_{k,0,n}} \int_{-\infty}^{\infty} \left[f\left(\frac{k}{n}, x\right) - f\left(\frac{k}{n}, x + c_n z \epsilon\right) \right]^2 dx \\
&\leq \frac{4K c_n}{n^2} \sum_{k=1}^n \frac{1}{d_{k,0,n}} \int_{-\infty}^{\infty} \left| f\left(\frac{k}{n}, x\right) \right|^2 dx = \frac{4K c_n}{n^2} \sum_{k=1}^n \frac{1}{d_{k,0,n}} G_3\left(\frac{k}{n}\right) \\
&\leq 4K K_3 \frac{c_n}{n} \frac{1}{n} \sum_{k=1}^n \frac{1}{d_{k,0,n}} \rightarrow 0,
\end{aligned}$$

where $K_3 = \sup_{t \in [0,1]} G_3(t)$ and $G_3(\cdot)$ is continuous on the interval in question.

We then prove that $\Lambda_{2n}(\epsilon) \rightarrow 0$ as $n \rightarrow \infty$. Because

$$\begin{aligned}
\Lambda_{2n}(\epsilon) &= \frac{2c_n^2}{n^2} \sum_{k=1}^{n-1} \sum_{l=k+1}^n |E[Y_{k,n}(z)Y_{l,n}(z)]| \\
&= \frac{2c_n^2}{n^2} \sum_{k=1}^{n-1} \sum_{l=k+1}^n |E[Y_{k,n}(z)E(Y_{l,n}(z)|\mathcal{F}_{k,n})]|,
\end{aligned}$$

For $k < l$, we begin with the following calculation of the conditional expectation:

$$\begin{aligned}
&|E(Y_{l,n}(z)|\mathcal{F}_{k,n})| = \left| E \left[f\left(\frac{l}{n}, c_n x_{l,n}\right) - f\left(\frac{l}{n}, c_n(x_{l,n} + z\epsilon)\right) \middle| \mathcal{F}_{k,n} \right] \right| \\
&= \left| E \left[f\left(\frac{l}{n}, c_n x_{k,n} + c_n(x_{l,n} - x_{k,n})\right) - f\left(\frac{l}{n}, c_n x_{k,n} + c_n(x_{l,n} - x_{k,n}) + c_n z \epsilon\right) \middle| \mathcal{F}_{k,n} \right] \right| \\
&= \left| \int_{-\infty}^{\infty} \left[f\left(\frac{l}{n}, c_n x_{k,n} + c_n d_{l,k,n} y\right) - f\left(\frac{l}{n}, c_n x_{k,n} + c_n d_{l,k,n} y + c_n z \epsilon\right) \right] h_{l,k,n}(y) dy \right| \\
&= \frac{1}{c_n d_{l,k,n}} \left| \int_{-\infty}^{\infty} \left[f\left(\frac{l}{n}, y\right) h_{l,k,n}\left(\frac{y - c_n x_{k,n}}{c_n d_{l,k,n}}\right) - f\left(\frac{l}{n}, y\right) h_{l,k,n}\left(\frac{y - c_n x_{k,n} - c_n \epsilon z}{c_n d_{l,k,n}}\right) \right] dy \right| \\
&= \frac{1}{c_n d_{l,k,n}} \left| \int_{-\infty}^{\infty} f\left(\frac{l}{n}, y\right) \left[h_{l,k,n}\left(\frac{y - c_n x_{k,n}}{c_n d_{l,k,n}}\right) - h_{l,k,n}\left(\frac{y - c_n x_{k,n} - c_n \epsilon z}{c_n d_{l,k,n}}\right) \right] dy \right| \\
&\leq \frac{1}{c_n d_{l,k,n}} \int_{-\infty}^{\infty} \left| f\left(\frac{l}{n}, y\right) \right| |V(y, c_n x_{k,n})| dy,
\end{aligned}$$

where $V(y, c_n x_{k,n}) = h_{l,k,n}\left(\frac{y - c_n x_{k,n}}{c_n d_{l,k,n}}\right) - h_{l,k,n}\left(\frac{y - c_n x_{k,n} - c_n \epsilon z}{c_n d_{l,k,n}}\right)$.

Recall the definition of $\Omega_n(\epsilon)$ in Assumption A(c) and note that a pair (l, k) ($l > k$) belongs to either $\Omega_n(\epsilon^{1/2m_0})$ or its complement. It follows that

$$|E(Y_{l,n}(z)|\mathcal{F}_{k,n})|$$

$$\leq \begin{cases} \frac{2K}{c_n d_{l,k,n}} \int_{-\infty}^{\infty} |f\left(\frac{l}{n}, y\right)| dy = \frac{2K}{c_n d_{l,k,n}} G_2\left(\frac{l}{n}\right), & \text{if } (l, k) \notin \Omega_n, \\ \frac{2K}{c_n d_{l,k,n}} \int_{|y| > \sqrt{c_n}} |f\left(\frac{l}{n}, y\right)| dy + \frac{1}{c_n d_{l,k,n}} \int_{|y| \leq \sqrt{c_n}} |f\left(\frac{l}{n}, y\right)| |V(y, c_n x_{k,n})| dy, & \text{otherwise.} \end{cases}$$

According to Assumption A (c), $\inf_{(l,k) \in \Omega_n(\epsilon^{1/2m_0})} d_{l,k,n} \geq \frac{\sqrt{\epsilon}}{C}$, and at the same time we can choose n large enough such that $\sqrt{c_n} \epsilon > 1$. For $|y| \leq \sqrt{c_n}$ and $|x| \leq \sqrt{c_n} + c_n |z| \epsilon$, when $(l, k) \in \Omega_n(\epsilon^{1/2m_0})$, we have

$$\begin{aligned} |V(y, x)| &= \left| h_{l,k,n} \left(\frac{y-x}{c_n d_{l,k,n}} \right) - h_{l,k,n} \left(\frac{y-x-c_n \epsilon z}{c_n d_{l,k,n}} \right) \right| \\ &\leq \left| h_{l,k,n} \left(\frac{y-x}{c_n d_{l,k,n}} \right) - h_{l,k,n}(0) \right| + \left| h_{l,k,n} \left(\frac{y-x-c_n \epsilon z}{c_n d_{l,k,n}} \right) - h_{l,k,n}(0) \right| \\ &\leq 2 \sup_{|u| < 2C(1+|z|)\sqrt{\epsilon}} |h_{l,k,n}(u) - h_{l,k,n}(0)|. \end{aligned} \quad (\text{C.9})$$

Therefore, when $|y| \leq \sqrt{c_n}$, n is large enough and $(l, k) \in \Omega_n(\epsilon^{1/2m_0})$, we have

$$\begin{aligned} &E|Y_{k,n}(z)| |V(y, c_n x_{k,n})| \\ &= \int_{-\infty}^{\infty} \left| f\left(\frac{k}{n}, c_n d_{k,0,n} x\right) - f\left(\frac{k}{n}, c_n d_{k,0,n} x + c_n z \epsilon\right) \right| |V(y, c_n d_{k,0,n} x)| h_{k,0,n}(x) dx \\ &\leq \frac{K}{c_n d_{k,0,n}} \int_{-\infty}^{\infty} \left| f\left(\frac{k}{n}, x\right) - f\left(\frac{k}{n}, x + c_n z \epsilon\right) \right| |V(y, x)| dx \\ &\leq \frac{K}{c_n d_{k,0,n}} \left[\int_{-\infty}^{\infty} \left| f\left(\frac{k}{n}, x\right) \right| |V(y, x)| dx + \int_{-\infty}^{\infty} \left| f\left(\frac{k}{n}, x + c_n z \epsilon\right) \right| |V(y, x)| dx \right] \\ &= \frac{K}{c_n d_{k,0,n}} \int_{-\infty}^{\infty} \left| f\left(\frac{k}{n}, x\right) \right| [|V(y, x)| + |V(y, x - c_n z \epsilon)|] dx \\ &= \frac{K}{c_n d_{k,0,n}} \left[\int_{|x| > \sqrt{c_n}} + \int_{|x| \leq \sqrt{c_n}} \left| f\left(\frac{k}{n}, x\right) \right| [|V(y, x)| + |V(y, x - c_n z \epsilon)|] dx \right] \\ &\leq \frac{2K^2}{c_n d_{k,0,n}} \int_{|x| > \sqrt{c_n}} \left| f\left(\frac{k}{n}, x\right) \right| dx \\ &\quad + \frac{K}{c_n d_{k,0,n}} \int_{|x| \leq \sqrt{c_n}} \left| f\left(\frac{k}{n}, x\right) \right| [|V(y, x)| + |V(y, x - c_n z \epsilon)|] dx \\ &\leq \frac{2K^2}{c_n d_{k,0,n}} \int_{|x| > \sqrt{c_n}} \left| f\left(\frac{k}{n}, x\right) \right| dx \\ &\quad + \frac{4K}{c_n d_{k,0,n}} \sup_{|u| < 2C(1+|z|)\sqrt{\epsilon}} |h_{l,k,n}(u) - h_{l,k,n}(0)| \int_{|x| \leq \sqrt{c_n}} \left| f\left(\frac{k}{n}, x\right) \right| dx. \end{aligned}$$

We summarise that if $(l, k) \notin \Omega_n$, equation (C.8) yields

$$|E(Y_{k,n}(z)Y_{l,n}(z))| = |E[Y_{k,n}(z)E(Y_{l,n}(z)|\mathcal{F}_{k,n})]|$$

$$\begin{aligned}
&\leq \frac{2K}{c_n d_{l,k,n}} G_2 \left(\frac{l}{n} \right) |E Y_{k,n}(z)| \\
&\leq \frac{4K^2}{c_n^2 d_{l,k,n} d_{k,0,n}} G_2 \left(\frac{l}{n} \right) G_2 \left(\frac{k}{n} \right),
\end{aligned}$$

while if $(l, k) \in \Omega_n$,

$$\begin{aligned}
&|E(Y_{k,n}(z)Y_{l,n}(z))| = |E[Y_{k,n}(z)E(Y_{l,n}(z)|\mathcal{F}_{k,n})]| \\
&\leq E[|Y_{k,n}(z)||E(Y_{l,n}(z)|\mathcal{F}_{k,n})|] \\
&\leq \frac{2K}{c_n d_{l,k,n}} \int_{|y|>\sqrt{c_n}} \left| f \left(\frac{l}{n}, y \right) \right| dy E|Y_{k,n}(z)| \\
&\quad + \frac{1}{c_n d_{l,k,n}} \int_{|y|\leq\sqrt{c_n}} \left| f \left(\frac{l}{n}, y \right) \right| E[|Y_{k,n}(z)||V(y, c_n x_{k,n})|] dy \\
&\leq \frac{4K^2}{c_n^2 d_{l,k,n} d_{k,0,n}} G_2 \left(\frac{k}{n} \right) \int_{|y|>\sqrt{c_n}} \left| f \left(\frac{l}{n}, y \right) \right| dy \\
&\quad + \frac{2K^2}{c_n^2 d_{l,k,n} d_{k,0,n}} \int_{|y|\leq\sqrt{c_n}} \left| f \left(\frac{l}{n}, y \right) \right| dy \int_{|x|>\sqrt{c_n}} \left| f \left(\frac{k}{n}, x \right) \right| dx \\
&\quad + \frac{4K}{c_n^2 d_{l,k,n} d_{k,0,n}} \int_{|y|\leq\sqrt{c_n}} \left| f \left(\frac{l}{n}, y \right) \right| dy \int_{|x|\leq\sqrt{c_n}} \left| f \left(\frac{k}{n}, x \right) \right| dx \\
&\quad \times \sup_{|u|<2C(1+|z|)\sqrt{\epsilon}} |h_{l,k,n}(u) - h_{l,k,n}(0)|.
\end{aligned}$$

Finally, we have

$$\begin{aligned}
|\Lambda_{2n}(\epsilon)| &\leq \frac{2c_n^2}{n^2} \left(\sum_{l>k, (l,k)\notin\Omega_n} + \sum_{(l,k)\in\Omega_n} \right) E|Y_{s,n}(z)Y_{t,n}(z)| \\
&\leq \frac{2c_n^2}{n^2} \sum_{k=(1-\epsilon)n}^n \sum_{l=k+1}^n E|Y_{k,n}(z)Y_{l,n}(z)| + \frac{2c_n^2}{n^2} \sum_{k=1}^{n-1} \sum_{l=k+1}^{k+\epsilon n} E|Y_{k,n}(z)Y_{l,n}(z)| \\
&\quad + \frac{2c_n^2}{n^2} \sum_{k=1}^{\epsilon n} \sum_{l=k+1}^n E|Y_{k,n}(z)Y_{l,n}(z)| + \frac{2c_n^2}{n^2} \sum_{k=1}^{n-1} \sum_{l=k+\epsilon n}^n E|Y_{k,n}(z)Y_{l,n}(z)| \\
&\leq 8K^2 K_2^2 \frac{1}{n} \sum_{k=(1-\epsilon)n}^n \frac{1}{d_{k,0,n}} \max_{1\leq k\leq n-1} \frac{1}{n} \sum_{l=k+1}^n \frac{1}{d_{l,k,n}} \\
&\quad + 8K^2 K_2^2 \frac{1}{n} \sum_{k=1}^{n-1} \frac{1}{d_{k,0,n}} \max_{1\leq k\leq n-1} \frac{1}{n} \sum_{l=k+1}^{k+\epsilon n} \frac{1}{d_{l,k,n}} \\
&\quad + 8K^2 K_2^2 \frac{1}{n} \sum_{k=1}^{\epsilon n} \frac{1}{d_{k,0,n}} \max_{1\leq k\leq n-1} \frac{1}{n} \sum_{l=k+1}^n \frac{1}{d_{l,k,n}}
\end{aligned}$$

$$\begin{aligned}
& + 8K^2K_2 \int_{|y| > \sqrt{c_n}} c_f(y) dy \frac{1}{n^2} \sum_{k=1}^{n-1} \frac{1}{d_{k,0,n}} \max_{1 \leq k \leq n-1} \sum_{l=k+1}^n \frac{1}{d_{l,k,n}} \\
& + 4K^2K_2 \int_{|x| > \sqrt{c_n}} c_f(x) dx \frac{1}{n^2} \sum_{k=1}^{n-1} \frac{1}{d_{k,0,n}} \max_{1 \leq k \leq n-1} \sum_{l=k+1}^n \frac{1}{d_{l,k,n}} \\
& + 8K^2K_2^2 \frac{1}{n^2} \sum_{k=1}^{n-1} \frac{1}{d_{k,0,n}} \max_{1 \leq k \leq n-1} \sum_{l=k+1}^n \frac{1}{d_{l,k,n}} \sup_{|u| < 2C(1+|z|)\sqrt{\epsilon}} |h_{l,k,n}(u) - h_{l,k,n}(0)|,
\end{aligned}$$

in which we have used Assumption C (c) that $|f(\frac{l}{n}, y)| \leq c_f(y)$ and the fact that

$$\int_{|y| \leq \sqrt{c_n}} \left| f\left(\frac{l}{n}, y\right) \right| dy \leq G_2\left(\frac{l}{n}\right) \leq K_2.$$

In view of Assumptions A (c) and C, by virtue of the dominated convergence theorem, $\Lambda_{2n}(\epsilon) \rightarrow 0$ as $n \rightarrow \infty$ and then $\epsilon \rightarrow 0$. This finishes the proof of the continuous case.

The proof of the discrete case is quite similar to that of the continuous case. Some critical steps are shown as follows.

Let $\mathcal{A}_{k,n}$ be the set of points that $x_{k,n}$ assumes. Suppose the points are equally distributed on \mathbb{R} with distance Δ . In what follows, define $\mathcal{B}_{k,n} := c_n d_{k,0,n} \mathcal{A}_{k,n} := \{c_n d_{k,0,n} a : a \in \mathcal{A}_{k,n}\}$. Then,

$$\begin{aligned}
E|Y_{k,n}(z)| & = E \left| f\left(\frac{k}{n}, c_n x_{k,n}\right) - f\left(\frac{k}{n}, c_n(x_{k,n} + z\epsilon)\right) \right| \\
& = \sum_{x \in \mathcal{A}_{k,n}} \left| f\left(\frac{k}{n}, c_n d_{k,0,n} x\right) - f\left(\frac{k}{n}, c_n(d_{k,0,n} x + z\epsilon)\right) \right| P_{k,0,n}(x) \\
& = \sum_{x \in \mathcal{B}_{k,n}} \left| f\left(\frac{k}{n}, x\right) - f\left(\frac{k}{n}, x + c_n z\epsilon\right) \right| P_{k,0,n}\left(\frac{x}{c_n d_{k,0,n}}\right) \\
& \leq \sum_{x \in \mathcal{B}_{k,n}} \left| f\left(\frac{k}{n}, x\right) \right| + \sum_{x \in \mathcal{B}_{k,n}} \left| f\left(\frac{k}{n}, x + c_n z\epsilon\right) \right| \\
& = \frac{1}{c_n d_{k,0,n} \Delta} \sum_{x \in \mathcal{B}_{k,n}} \left| f\left(\frac{k}{n}, x\right) \right| c_n d_{k,0,n} \Delta \\
& \quad + \frac{1}{c_n d_{k,0,n} \Delta} \sum_{x \in \mathcal{B}_{k,n}} \left| f\left(\frac{k}{n}, x + c_n z\epsilon\right) \right| c_n d_{k,0,n} \Delta \\
& \leq \frac{1}{c_n d_{k,0,n} \Delta} \left(\int \left| f\left(\frac{k}{n}, x\right) \right| dx + \int \left| f\left(\frac{k}{n}, x + c_n z\epsilon\right) \right| dx \right) \\
& = \frac{2}{c_n d_{k,0,n} \Delta} \int \left| f\left(\frac{k}{n}, x\right) \right| dx = \frac{2}{c_n d_{k,0,n} \Delta} G_2\left(\frac{k}{n}\right)
\end{aligned}$$

$$\leq \frac{2}{c_n d_{k,0,n} \Delta} K_2,$$

where we may modify the function f , e.g. $f^o(\cdot, x) = \max_{y \geq x} |f(\cdot, y)|$ for $x > 0$ to get the inequality in the derivation and note that the result above is similar to (C.8). Following the same arguments as before, to complete the proof, it suffices to show both $\Lambda_{1n}(\epsilon)$ and $\Lambda_{2n}(\epsilon)$ converge to zero. Nevertheless, $\Lambda_{1n}(\epsilon) \rightarrow 0$ is easy to obtain, while the key step in the proof of $\Lambda_{2n}(\epsilon) \rightarrow 0$ is the evaluation of the following conditional expectation.

$$\begin{aligned} |E(Y_{l,n}(z)|\mathcal{F}_{k,n})| &= \left| E \left[f \left(\frac{l}{n}, c_n x_{l,n} \right) - f \left(\frac{l}{n}, c_n (x_{l,n} + z\epsilon) \right) \middle| \mathcal{F}_{k,n} \right] \right| \\ &= \left| E \left[f \left(\frac{l}{n}, c_n x_{k,n} + c_n (x_{l,n} - x_{k,n}) \right) - f \left(\frac{l}{n}, c_n x_{k,n} + c_n (x_{l,n} - x_{k,n}) + c_n z\epsilon \right) \middle| \mathcal{F}_{k,n} \right] \right| \\ &= \left| \int \left[f \left(\frac{l}{n}, c_n x_{k,n} + c_n d_{l,k,n} y \right) - f \left(\frac{l}{n}, c_n x_{k,n} + c_n d_{l,k,n} y + c_n z\epsilon \right) \right] dF_{l,k,n}(y) \right| \\ &= \left| \int f \left(\frac{l}{n}, y \right) dF_{l,k,n} \left(\frac{y - c_n x_{k,n}}{c_n d_{l,k,n}} \right) - \int f \left(\frac{l}{n}, y \right) dF_{l,k,n} \left(\frac{y - c_n x_{k,n} - c_n \epsilon z}{c_n d_{l,k,n}} \right) \right| \\ &= \left| \int f \left(\frac{l}{n}, y \right) d \left[F_{l,k,n} \left(\frac{y - c_n x_{k,n}}{c_n d_{l,k,n}} \right) - F_{l,k,n} \left(\frac{y - c_n x_{k,n} - c_n \epsilon z}{c_n d_{l,k,n}} \right) \right] \right| \\ &= \left| \int f \left(\frac{l}{n}, y \right) dQ(y, c_n x_{k,n}) \right|, \end{aligned}$$

where $Q(y, c_n x_{k,n}) = F_{l,k,n} \left(\frac{y - c_n x_{k,n}}{c_n d_{l,k,n}} \right) - F_{l,k,n} \left(\frac{y - c_n x_{k,n} - c_n \epsilon z}{c_n d_{l,k,n}} \right)$.

Thus,

$$\begin{aligned} &|E(Y_{l,n}(z)|\mathcal{F}_{k,n})| \\ &\leq \begin{cases} \frac{2}{c_n d_{l,k,n} \Delta} \int |f \left(\frac{l}{n}, y \right)| dy = \frac{2}{c_n d_{l,k,n}} G_2 \left(\frac{l}{n} \right), & \text{if } (l, k) \notin \Omega_n, \\ \frac{2}{c_n d_{l,k,n} \Delta} \int_{|y| > \sqrt{c_n}} |f \left(\frac{l}{n}, y \right)| dy + \left| \int_{|y| \leq \sqrt{c_n}} f \left(\frac{l}{n}, y \right) dQ(y, c_n x_{k,n}) \right|, & \text{if } (l, k) \in \Omega_n. \end{cases} \end{aligned}$$

Then the important ingredient is to deal with the following expectation.

$$\begin{aligned} &E|Y_{k,n}| \left| \int_{|y| \leq \sqrt{c_n}} f \left(\frac{l}{n}, y \right) dQ(y, c_n x_{k,n}) \right| \\ &= E \left| f \left(\frac{k}{n}, c_n x_{k,n} \right) - f \left(\frac{k}{n}, c_n (x_{k,n} + z\epsilon) \right) \right| \left| \int_{|y| \leq \sqrt{c_n}} f \left(\frac{l}{n}, y \right) dQ(y, c_n x_{k,n}) \right| \\ &= \int \left| f \left(\frac{k}{n}, c_n d_{k,0,n} x \right) - f \left(\frac{k}{n}, c_n (d_{k,0,n} x + z\epsilon) \right) \right| \\ &\quad \times \left| \int_{|y| \leq \sqrt{c_n}} f \left(\frac{l}{n}, y \right) dQ(y, c_n d_{k,0,n} x) \right| dF_{k,0,n}(x) \end{aligned}$$

$$\begin{aligned}
&= \int \left| f\left(\frac{k}{n}, x\right) - f\left(\frac{k}{n}, x + c_n z \epsilon\right) \right| \left| \int_{|y| \leq \sqrt{c_n}} f\left(\frac{l}{n}, y\right) dQ(y, x) \right| dF_{k,0,n}\left(\frac{x}{c_n d_{k,0,n}}\right) \\
&\leq \int \left| f\left(\frac{k}{n}, x\right) \right| \left| \int_{|y| \leq \sqrt{c_n}} f\left(\frac{l}{n}, y\right) dQ(y, x) \right| dF_{k,0,n}\left(\frac{x}{c_n d_{k,0,n}}\right) \\
&\quad + \int \left| f\left(\frac{k}{n}, x + c_n z \epsilon\right) \right| \left| \int_{|y| \leq \sqrt{c_n}} f\left(\frac{l}{n}, y\right) dQ(y, x) \right| dF_{k,0,n}\left(\frac{x}{c_n d_{k,0,n}}\right) \\
&= \int \left| f\left(\frac{k}{n}, x\right) \right| \left| \int_{|y| \leq \sqrt{c_n}} f\left(\frac{l}{n}, y\right) dQ(y, x) \right| dF_{k,0,n}\left(\frac{x}{c_n d_{k,0,n}}\right) \\
&\quad + \int \left| f\left(\frac{k}{n}, x\right) \right| \left| \int_{|y| \leq \sqrt{c_n}} f\left(\frac{l}{n}, y\right) dQ(y, x - c_n z \epsilon) \right| dF_{k,0,n}\left(\frac{x - c_n z \epsilon}{c_n d_{k,0,n}}\right) \\
&= \int_{|x| \leq \sqrt{c_n}} \left| f\left(\frac{k}{n}, x\right) \right| \left| \int_{|y| \leq \sqrt{c_n}} f\left(\frac{l}{n}, y\right) dQ(y, x) \right| dF_{k,0,n}\left(\frac{x}{c_n d_{k,0,n}}\right) \\
&\quad + \int_{|x| > \sqrt{c_n}} \left| f\left(\frac{k}{n}, x\right) \right| \left| \int_{|y| \leq \sqrt{c_n}} f\left(\frac{l}{n}, y\right) dQ(y, x) \right| dF_{k,0,n}\left(\frac{x}{c_n d_{k,0,n}}\right) \\
&\quad + \int_{|x| \leq \sqrt{c_n}} \left| f\left(\frac{k}{n}, x\right) \right| \left| \int_{|y| \leq \sqrt{c_n}} f\left(\frac{l}{n}, y\right) dQ(y, x - c_n z \epsilon) \right| dF_{k,0,n}\left(\frac{x - c_n z \epsilon}{c_n d_{k,0,n}}\right) \\
&\quad + \int_{|x| > \sqrt{c_n}} \left| f\left(\frac{k}{n}, x\right) \right| \left| \int_{|y| \leq \sqrt{c_n}} f\left(\frac{l}{n}, y\right) dQ(y, x - c_n z \epsilon) \right| dF_{k,0,n}\left(\frac{x - c_n z \epsilon}{c_n d_{k,0,n}}\right) \\
&:= \sum_1^4 T_i(l, k; n).
\end{aligned}$$

For $T_1(l, k; n)$ is similar to $T_3(l, k; n)$, and $T_2(l, k; n)$ is similar to $T_4(l, k; n)$, we only explain $T_1(l, k; n)$ and $T_2(l, k; n)$.

$$\begin{aligned}
T_1(l, k; n) &= \int_{|x| \leq \sqrt{c_n}} \left| f\left(\frac{k}{n}, x\right) \right| \left| \int_{|y| \leq \sqrt{c_n}} f\left(\frac{l}{n}, y\right) dQ(y, x) \right| dF_{k,0,n}\left(\frac{x}{c_n d_{k,0,n}}\right) \\
&= \sum_{\substack{|x| \leq \sqrt{c_n} \\ x \in \mathcal{B}_{k,n}}} \left| f\left(\frac{k}{n}, x\right) \right| \left| \sum_{\substack{|y| \leq \sqrt{c_n} \\ y \in x + c_n z \epsilon \mathcal{B}_{l,k,n}}} f\left(\frac{l}{n}, y\right) P(y, x) \right| P_{k,0,n}\left(\frac{x}{c_n d_{k,0,n}}\right),
\end{aligned}$$

where $\mathcal{B}_{l,k,n} = \{c_n d_{l,k,n} a : a \in \mathcal{A}_{l,k,n}\}$, in which $\mathcal{A}_{l,k,n}$ is the set of points that $(x_{l,n} - x_{k,n})/d_{l,k,n}$ assumes; meanwhile, $P(y, x) = P_{l,k,n}\left(\frac{y - c_n x_{k,n}}{c_n d_{l,k,n}}\right) - P_{l,k,n}\left(\frac{y - c_n x_{k,n} - c_n \epsilon z}{c_n d_{l,k,n}}\right)$.

Notice that when $|x| \leq \sqrt{c_n}$, $|y| \leq \sqrt{c_n}$ and $(l, k) \in \Omega_n(\epsilon)$, we have

$$\begin{aligned} |P(y, x)| &= \left| P_{l,k,n} \left(\frac{y - c_n x_{k,n}}{c_n d_{l,k,n}} \right) - P_{l,k,n} \left(\frac{y - c_n x_{k,n} - c_n \epsilon z}{c_n d_{l,k,n}} \right) \right| \\ &= \left| F_{l,k,n} \left(\frac{y - c_n x_{k,n}}{c_n d_{l,k,n}} \right) - F_{l,k,n}^- \left(\frac{y - c_n x_{k,n}}{c_n d_{l,k,n}} \right) \right. \\ &\quad \left. - F_{l,k,n} \left(\frac{y - c_n x_{k,n} - c_n \epsilon z}{c_n d_{l,k,n}} \right) + F_{l,k,n}^- \left(\frac{y - c_n x_{k,n} - c_n \epsilon z}{c_n d_{l,k,n}} \right) \right| \\ &\leq 4 \sup_{|u| < 2C(1+|z|)\sqrt{\epsilon}} |F_{l,k,n}(u) - F_{l,k,n}(0)|. \end{aligned}$$

Here $F_{l,k,n}^-(\cdot)$ denotes the left limit of the function at the point.

Therefore, we have

$$T_1(l, k; n) \leq \frac{4}{c_n d_{k,0,n} \Delta} \frac{1}{c_n d_{l,k,n} \Delta} G_2 \left(\frac{k}{n} \right) G_2 \left(\frac{l}{n} \right) \sup_{|u| < 2C(1+|z|)\sqrt{\epsilon}} |F_{l,k,n}(u) - F_{l,k,n}(0)|.$$

Regarding $T_2(l, k; n)$, we directly have

$$\begin{aligned} T_2(l, k; n) &\leq \frac{2}{c_n d_{k,0,n} \Delta} \frac{1}{c_n d_{l,k,n} \Delta} \int_{|x| \geq \sqrt{c_n}} \left| f \left(\frac{k}{n}, x \right) \right| dx \int_{|y| \leq \sqrt{c_n}} \left| f \left(\frac{k}{n}, y \right) \right| dy \\ &\leq \frac{2}{c_n d_{k,0,n} \Delta} \frac{1}{c_n d_{l,k,n} \Delta} G_2 \left(\frac{l}{n} \right) \int_{|x| \geq \sqrt{c_n}} \left| f \left(\frac{k}{n}, x \right) \right| dx. \end{aligned}$$

As can be seen, every term in $\Lambda_{2n}(\epsilon)$ has the similar evaluation in both continuous and discrete cases, so that we obtain the vanish of $\Lambda_{2n}(\epsilon)$. As yet, the whole proof is finished. \square

Lemma C.5. *Let Assumption C hold. Then we have for any fixed $\epsilon > 0$,*

$$L_{n,\epsilon}^{(r)} - \int_{-\infty}^{\infty} \frac{1}{n} \sum_{k=1}^{[nr]} f \left(\frac{k}{n}, y \right) \phi_{\epsilon}(x_{k,n}) dy \rightarrow_{a.s.} 0$$

uniformly in $r \in [0, 1]$ as $n \rightarrow \infty$.

Proof. Observe that

$$\begin{aligned} L_{n,\epsilon}^{(r)} &= \frac{c_n}{n} \sum_{k=1}^{[nr]} \int_{-\infty}^{\infty} f \left(\frac{k}{n}, c_n(x_{t,n} + z\epsilon) \right) \phi(z) dz \\ &= \frac{c_n}{n} \sum_{k=1}^{[nr]} \int_{-\infty}^{\infty} f \left(\frac{k}{n}, y \right) \phi \left(\frac{y - c_n x_{t,n}}{c_n \epsilon} \right) \frac{1}{c_n \epsilon} dy \end{aligned}$$

$$= \int_{-\infty}^{\infty} \frac{1}{n} \sum_{k=1}^{\lfloor nr \rfloor} f\left(\frac{k}{n}, y\right) \phi_{\epsilon}\left(\frac{y}{c_n} - x_{t,n}\right) dy.$$

It follows that for any $M > 0$,

$$\begin{aligned} & \left| L_{n,\epsilon}^{(r)} - \int_{-\infty}^{\infty} \frac{1}{n} \sum_{k=1}^{\lfloor nr \rfloor} f\left(\frac{k}{n}, y\right) \phi_{\epsilon}(x_{k,n}) dy \right| \\ & \leq \int_{-\infty}^{\infty} \frac{1}{n} \sum_{k=1}^n \left| f\left(\frac{k}{n}, y\right) \right| \left| \phi_{\epsilon}\left(\frac{y}{c_n} - x_{t,n}\right) - \phi_{\epsilon}(x_{k,n}) \right| dy \\ & = \int_{|y|>M} + \int_{|y|\leq M} \frac{1}{n} \sum_{k=1}^n \left| f\left(\frac{k}{n}, y\right) \right| \left| \phi_{\epsilon}\left(\frac{y}{c_n} - x_{t,n}\right) - \phi_{\epsilon}(x_{k,n}) \right| dy \\ & := \Gamma_{1n} + \Gamma_{2n}. \end{aligned}$$

Notice that,

$$\Gamma_{1n} \leq \frac{2}{\sqrt{2\pi}\epsilon} \int_{|y|>M} \frac{1}{n} \sum_{k=1}^n \left| f\left(\frac{k}{n}, y\right) \right| dy \leq \frac{2}{\sqrt{2\pi}\epsilon} \int_{|y|>M} c_f(y) dy,$$

using Assumption C(c). Due to the integrability of $c_f(y)$ on \mathbb{R} , one can choose large enough M such that $\Gamma_{1n} < \varepsilon$ for any given $\varepsilon > 0$.

Moreover, since $\phi'_{\epsilon}(x) = -\frac{x}{\sqrt{2\pi}\epsilon^3} e^{-x^2/2\epsilon^2}$ and $|\phi'_{\epsilon}(x)|$ is bounded by $\frac{1}{\sqrt{2\pi}\epsilon^2}$ on \mathbb{R} , we have

$$\left| \phi_{\epsilon}\left(\frac{y}{c_n} - x_{t,n}\right) - \phi_{\epsilon}(x_{t,n}) \right| = \left| \phi'_{\epsilon}(\xi) \left(-\frac{y}{c_n}\right) \right| \leq \frac{|y|}{\sqrt{2\pi}\epsilon\epsilon^2 c_n},$$

where ξ is in between $x_{t,n} - \frac{y}{c_n}$ and $x_{t,n}$. Therefore,

$$\begin{aligned} \Gamma_{2n} & \leq \int_{|y|\leq M} \frac{1}{n} \sum_{k=1}^n \left| f\left(\frac{k}{n}, y\right) \right| \frac{|y|}{\sqrt{2\pi}\epsilon\epsilon^2 c_n} dy \\ & \leq \frac{M}{\sqrt{2\pi}\epsilon\epsilon^2 c_n} \frac{1}{n} \sum_{k=1}^n \int_{|y|\leq M} \left| f\left(\frac{k}{n}, y\right) \right| dy \leq \frac{M}{\sqrt{2\pi}\epsilon\epsilon^2 c_n} \frac{1}{n} \sum_{k=1}^n G_2\left(\frac{k}{n}\right). \end{aligned}$$

As $\frac{1}{n} \sum_{k=1}^n G_2\left(\frac{k}{n}\right) \leq K_2$ and $c_n \rightarrow \infty$ as $n \rightarrow \infty$, $\Gamma_{2n} \rightarrow 0$. The assertion follows. \square

PROOF OF LEMMA 3.1. Straightforward verification. \square

PROOF OF LEMMA 4.1. It is evident that V is both symmetric and nonnegative definite, so that for $i = 1, \dots, p$, λ_i are all real and $\lambda_i \geq 0$. Moreover, $\sum_{i=1}^p \lambda_i = \text{tr}(V) =$

$\text{tr}(v'v) = \text{tr}(vv') = \|v\|^2 = 1$. Nonetheless, $\lambda_1 = 1$ since $Vv' = v'vv' = v'$. Whence, the assertion follows. \square

PROOF OF PROPOSITION 4.1. Straightforward calculation. \square

PROOF OF PROPOSITION 4.2. Straightforward calculation. \square

PROOF OF PROPOSITION 4.3. Straightforward calculation. \square

D Proofs of the main results

PROOF OF THEOREM 2.2. We begin with the calculation of the coefficients $c_i(t, f) = E[f(Z(t))Q_i(t, Z(t))]$ where $Q_i(t, x) = \frac{1}{d_i(t)}y_i(t, x)$. If $Z(t)$ is continuous with density function $\rho(t, x)$, the polynomials $y_i(t, x)$ orthogonal with respect to $\rho(t, x)$ satisfy the differential equation

$$s(t, x)y_i''(t, x) + v(t, x)y_i'(t, x) + \lambda_i(t)y_i(t, x) = 0,$$

where $s(t, x) > 0$, $v(t, x)$ and $\rho(t, x)$ satisfy conditions (A.4) and $\lambda_i(t) = -iv'(t, x)$.

The self-adjoint form of the equation is

$$(s(t, x)\rho(t, x)y_i'(t, x))' + \lambda_i(t)\rho(t, x)y_i(t, x) = 0.$$

Multiplying by $f(x)$, integrating by part on (a, b) , we have

$$\begin{aligned} & f(x)s(t, x)\rho(t, x)y_i'(t, x)|_a^b - \int_a^b s(t, x)\rho(t, x)y_i'(t, x)f'(x)dx \\ &= -\lambda_i(t) \int_a^b \rho(t, x)y_i(t, x)f(x)dx. \end{aligned} \tag{D.1}$$

Let us prove that $f(x)s(t, x)\rho(t, x)y_i'(t, x)|_a^b = 0$.

Suppose $\lim_{x \rightarrow b} f(x)s(t, x)\rho(t, x)y_i'(t, x) = b_i \neq 0$. Then, as $x \rightarrow b$,

$$f(x) \approx \frac{b_i}{s(t, x)\rho(t, x)y_i'(t, x)} \Rightarrow f^2(x)\rho(t, x) \sim \frac{b_i^2}{s^2(t, x)\rho(t, x)[y_i'(t, x)]^2}.$$

(1) Note that $b = +\infty$. Because of boundary condition, $f^2(x)\rho(t, x)$ will go to positive infinity as $x \rightarrow \infty$, which leads to the infiniteness of the integral $\int_a^b f^2(x)\rho(t, x)dx$.

The above discussion applies to the situation where $a = -\infty$ as well.

(2) When a is finite, according to Nikiforov and Uvarov (1988, p21),

$$s(t, x) \sim x - a, \quad \text{and} \quad \rho(t, x) \sim (x - a)^\alpha, \quad \text{where } \alpha > -1.$$

Hence, when $x \rightarrow +a$,

$$f(x) \sim \frac{1}{(x-a)^{1+\alpha}}, \quad \text{and} \quad f^2(x)\rho(t, x) \sim \frac{1}{(x-a)^{2+\alpha}}$$

which implies the infiniteness of $\int_a^b f^2(x)\rho(t, x)dx$.

Thus, the relation (D.1) reduces to

$$\int_a^b \rho_1(t, x)y_i'(t, x)f'(x)dx = \lambda_i(t) \int_a^b \rho(t, x)y_i(t, x)f(x)dx,$$

which is exactly the following relationship: $d_{1i}(t)c_i(t, f') = \lambda_i(t)d_i(t)c_i(t, f)$, or equivalently

$$c_i(t, f) = \frac{d_{1i}(t)}{\lambda_i(t)d_i(t)}c_i(t, f').$$

We can iterate the relation until r -th derivative,

$$\begin{aligned} c_i(t, f) &= \frac{d_{ri}(t)}{d_i(t)\lambda_i(t)\eta_{1i}(t)\cdots\eta_{r-1,i}(t)}c_i(t, f^{(r)}) \\ &= \sqrt{\psi(t)}^r \sqrt{\frac{(i-r)!}{i!}}c_i(t, f^{(r)}), \end{aligned} \tag{D.2}$$

where we have used the relationship (A.6) with $\eta_{ji}(t) = \lambda_i(t) - \lambda_j(t) = -v'(t, x)(i-j) = \frac{1}{\psi(t)}(i-j)$.

If $Z(t)$ is a discrete variable for each $t > 0$, $\rho(t, x)$ is the probability distribution of $Z(t)$. The polynomials $y_i(t, x)$ orthogonal with respect to $\rho(t, x)$ satisfy the following difference equation

$$s(t, x)\Delta\nabla y_i(t, x) + v(t, x)\Delta y_i(t, x) + \lambda_i(t)y_i(t, x) = 0$$

where $s(t, x) > 0$, $s''(t, x) = 0$ and $v'(t, x) < 0$. The self-adjoint form of the difference equation is

$$\Delta(s(t, x)\rho(t, x)\nabla y_i(t, x)) + \lambda_i(t)\rho(t, x)y_i(t, x) = 0.$$

Multiplying by $f(x)$ and summing up over the support of $\rho(t, x)$,

$$\sum_m f(x_m)\Delta(s(t, x_m)\rho(t, x_m)\nabla y_i(t, x_m)) = -\lambda_i(t) \sum_m f(x_m)\rho(t, x_m)y_i(t, x_m).$$

Summation by parts gives

$$f(x)s(t, x)\rho(t, x)\nabla y_i(t, x)|_a^b - \sum_m s(t, x_{m+1})\rho(t, x_{m+1})\nabla y_i(t, x_{m+1})\Delta f(x_m)$$

$$= -\lambda_i(t) \sum_m f(x_m) \rho(t, x_m) y_i(t, x_m).$$

It is easy to prove that $f(x)s(t, x)\rho(t, x)\nabla y_i(t, x)|_a^b = 0$ similar to the continuous case. Note that $\nabla y_i(t, x_{m+1}) = \Delta y_i(t, x_m)$ and $s(t, x_{m+1})\rho(t, x_{m+1}) = \rho_1(t, x_m)$. Therefore,

$$\sum_m \rho_1(t, x_m) \Delta y_i(t, x_m) \Delta f(x_m) = \lambda_i(t) \sum_j f(x_m) \rho(t, x_m) y_i(t, x_m)$$

which reads $c_i(t, f) = \frac{d_{1i}(t)}{\lambda_i(t)d_i(t)} c_i(t, \Delta f)$. We iterate this relationship and obtain again (D.2) with derivative being substituted by difference. Adopting our operator D , we have

$$c_i(t, f) = \sqrt{\psi(t)}^r \sqrt{\frac{(i-r)!}{i!}} c_i(t, D^r f). \quad (\text{D.3})$$

Now we are ready to obtain the result. Using (D.3),

$$\begin{aligned} \|f(Z(t)) - f_k(Z(t))\|_{\Theta}^2 &= \sum_{i=k+1}^{\infty} c_i^2(t, f) = \sum_{i=k+1}^{\infty} [v(t)]^r \frac{(i-r)!}{i!} c_i^2(t, D^r f) \\ &\leq [\psi(t)]^r \frac{(k-r)!}{k!} \sum_{i=k+1}^{\infty} c_i^2(t, D^r f) = \frac{1}{k^r} R_k^2(t, D^r f), \end{aligned}$$

where $R_k^2(t, D^r f) = (1+o(1))[\psi(t)]^r \sum_{i=k+1}^{\infty} c_i^2(t, D^r f) \rightarrow 0$ as $k \rightarrow \infty$ for every $t > 0$. \square

PROOF OF THEOREM 2.4. It follows from the orthogonality that

$$\|f(t, Z(t)) - f_{k,p}(t, Z(t))\|_{\Xi}^2 = \sum_{i=0}^k \sum_{j=p_i+1}^{\infty} c_{ij}^2 + \sum_{i=k+1}^{\infty} \sum_{j=0}^{\infty} c_{ij}^2.$$

Since $c_i(t, f) \in C^2[0, T]$, the expansion of $c_i''(t, f)$ in terms of $\varphi_{jT}(t)$ is convergent uniformly on $[0, T]$ (see Davis, 1963, p142). We thus have $c_{ij} = \left(\frac{T}{j\pi}\right)^2 b_j(c_i'')$, where $b_j(c_i'')$ stands for the j -th coefficient in the expansion of $c_i''(t, f)$ in terms of $\varphi_{jT}(t)$. Then

$$\begin{aligned} \sum_{i=0}^k \sum_{j=p_i+1}^{\infty} c_{ij}^2 &= \sum_{i=0}^k \sum_{j=p_i+1}^{\infty} \left(\frac{T}{j\pi}\right)^4 b_j^2(c_i'') \leq \left(\frac{T}{\pi}\right)^4 \sum_{i=0}^k \frac{1}{p_i^4} \sum_{j=p_i+1}^{\infty} b_j^2(c_i'') \\ &\leq \frac{T^4}{\pi^4} \frac{k}{p_{\min}^4} \sum_{j=p_{\min}+1}^{\infty} b_j^2(c_i'') \leq C(k, p) T^4 \frac{k}{p_{\min}^4}, \end{aligned}$$

where $C(k, p) = \pi^{-4} \max_{0 \leq i \leq k} \sum_{j=p_{\min}+1}^{\infty} b_j^2(c_i'')$.

On the other hand, using (D.3) we have

$$\sum_{i=k+1}^{\infty} \sum_{j=0}^{\infty} c_{ij}^2 = \sum_{i=k+1}^{\infty} \|c_i(t, f)\|_{L^2[0, T]}^2 = \sum_{i=k+1}^{\infty} \frac{(i-r)!}{i!} \left\| c_i \left(t, \sqrt{v(t)}^r D^r f \right) \right\|^2$$

$$\leq \frac{(k-r)!}{k!} \sum_{i=k+1}^{\infty} \left\| c_i \left(t, \sqrt{v(t)}^r D^r f \right) \right\|_{L^2[0,T]}^2 = \frac{1}{k^r} R_k^2. \quad \square$$

PROOF OF THEOREM 2.6. It follows from the orthogonality that

$$\|f(t, Z(t)) - f_{k,p}(t, Z(t))\|_{\Lambda}^2 = \sum_{i=0}^k \sum_{j=p_i+1}^{\infty} b_{ij}^2 + \sum_{i=k+1}^{\infty} \sum_{j=0}^{\infty} b_{ij}^2.$$

Due to (D.3), $b_i(t, f) = \sqrt{v(t)}^{r_1} \sqrt{\frac{(i-r_1)!}{i!}} b_i(t, D^{r_1} f)$. Accordingly,

$$\begin{aligned} \sum_{i=k+1}^{\infty} \sum_{j=0}^{\infty} b_{ij}^2 &= \sum_{i=k+1}^{\infty} \|b_i(t, f)\|_{L^2(\mathbb{R}^+)}^2 = \sum_{i=k+1}^{\infty} \frac{(i-r_1)!}{i!} \|b_i(t, \sqrt{v(t)}^{r_1} D^{r_1} f)\|^2 \\ &\leq \frac{(k+1-r_1)!}{(k+1)!} \sum_{i=k+1}^{\infty} \|b_i(t, \sqrt{v(t)}^{r_1} D^{r_1} f)\|^2 \\ &= (1+o(1)) \frac{1}{k^{r_1}} \sum_{i=k+1}^{\infty} \|b_i(t, \sqrt{v(t)}^{r_1} D^{r_1} f)\|^2 = \frac{1}{k^{r_1}} R_1^2(k), \end{aligned}$$

where $R_1^2(k) = (1+o(1)) \sum_{i=k+1}^{\infty} \|b_i(t, \sqrt{v(t)}^{r_1} D^{r_1} f)\|_{L^2(\mathbb{R}^+)}^2$.

On the other hand, according to (C.1) in Lemma C.1,

$$\sum_{j=p_i+1}^{\infty} b_{ij}^2 \leq \frac{(p_i+1-r_2)!}{(p_i+1)!} \sum_{j=p_i+1}^{\infty} [a_{j-r_2}^{(r_2)}(\tilde{b}_i(t))]^2$$

where $\tilde{b}_i(t) = t^{r_2/2} e^{-t/2} [b_i(t, f) e^{t/2}]^{(r_2)}$ and $a_{j-r_2}^{(r_2)}(\tilde{b}_i(t))$ are the coefficients of the expansion of $\tilde{b}_i(t)$ in terms of $\mathcal{L}_j^{(r_2)}(t)$. Thus,

$$\begin{aligned} \sum_{i=0}^k \sum_{j=p_i+1}^{\infty} b_{ij}^2 &\leq \sum_{i=0}^k \frac{(p_i+1-r_2)!}{(p_i+1)!} \sum_{j=p_i+1}^{\infty} [a_{j-r_2}^{(r_2)}(\tilde{b}_i(t))]^2 \\ &\leq k \frac{(p_{\min}+1-r_2)!}{(p_{\min}+1)!} \sum_{j=p_{\min}+1}^{\infty} [a_{j-r_2}^{(r_2)}(\tilde{b}_i(t))]^2 \leq C(k, p) \frac{k}{p_{\min}^{r_2}}. \end{aligned}$$

where $C(k, p) = (1+o(1)) \max_{0 \leq i \leq k} \sum_{j=p_{\min}+1}^{\infty} [a_{j-r_2}^{(r_2)}(\tilde{b}_i(t))]^2$. This finishes the proof. \square

PROOF OF THEOREM 3.1. In view of Lemmas C.4 and C.5, we start to investigate the convergence of

$$\int_{-\infty}^{\infty} \frac{1}{n} \sum_{k=1}^{\lfloor nr \rfloor} f\left(\frac{k}{n}, y\right) \phi_{\epsilon}(x_{k,n}) dy.$$

It follows from Assumptions A (a) and C, the continuous mapping theorem and the occupation time formula in Lemma C.2 that

$$\begin{aligned}
& \int_{-\infty}^{\infty} \frac{1}{n} \sum_{k=1}^{[nr]} f\left(\frac{k}{n}, y\right) \phi_{\epsilon}(x_{k,n}) dy \\
&= \frac{1}{n} \sum_{k=1}^{[nr]} \phi_{\epsilon}(x_{k,n}) \int_{-\infty}^{\infty} f\left(\frac{k}{n}, y\right) dy = \frac{1}{n} \sum_{k=1}^{[nr]} \phi_{\epsilon}(x_{k,n}) G_1\left(\frac{k}{n}\right) \\
&= \int_0^r G_1\left(\frac{[nt]}{n}\right) \phi_{\epsilon}(x_{[nt],n}) dt - \frac{1}{n} G_1(0) \phi_{\epsilon}(0) + \frac{1}{n} G_1\left(\frac{[nr]}{n}\right) \phi_{\epsilon}(x_{[nr],n}) \\
&\rightarrow_D \int_0^r G_1(t) \phi_{\epsilon}(W(t)) dt \quad \text{as } n \rightarrow \infty \\
&= \int_{-\infty}^{\infty} dy \int_0^r G_1(t) \phi_{\epsilon}(y) dL_W(t, y) \\
&= \int_{-\infty}^{\infty} dy \int_0^r G_1(t) \phi(y) dL_W(t, \epsilon y) \quad \text{and then as } \epsilon \rightarrow 0 \\
&\rightarrow_{a.s.} \int_{-\infty}^{\infty} dy \int_0^r G_1(t) \phi(y) dL_W(t, 0) \\
&= \int_0^r G_1(t) dL_W(t, 0).
\end{aligned}$$

This finishes the proof of (3.6). To prove (3.7), we need only to show the following:

$$\begin{aligned}
& \sup_{0 \leq r \leq 1} \left| \int_{-\infty}^{\infty} \frac{1}{n} \sum_{k=1}^{[nr]} f\left(\frac{k}{n}, y\right) \phi_{\epsilon}(x_{k,n}) dy - \int_0^r G_1(t) \phi_{\epsilon}(W(t)) dt \right| \quad (D.4) \\
&= \sup_{0 \leq r \leq 1} \left| \frac{1}{n} \sum_{k=1}^{[nr]} \phi_{\epsilon}(x_{k,n}) G_1\left(\frac{k}{n}\right) - \int_0^r G_1(t) \phi_{\epsilon}(W(t)) dt \right| \\
&= \sup_{0 \leq r \leq 1} \left| \int_0^r G_1\left(\frac{[nt]}{n}\right) \phi_{\epsilon}(x_{[nt],n}) dt - \frac{1}{n} G_1(0) \phi_{\epsilon}(0) + \frac{1}{n} G_1\left(\frac{[nr]}{n}\right) \phi_{\epsilon}(x_{[nr],n}) \right. \\
&\quad \left. - \int_0^r G_1(t) \phi_{\epsilon}(W(t)) dt \right| \\
&\leq \int_0^1 \left| G_1\left(\frac{[nt]}{n}\right) \phi_{\epsilon}(x_{[nt],n}) - G_1(t) \phi_{\epsilon}(W(t)) \right| dt + \frac{A}{n},
\end{aligned}$$

where A comes from the bounds of G_1 on $[0, 1]$ and ϕ_{ϵ} on \mathbb{R} . It follows that

$$\begin{aligned}
& \int_0^1 \left| G_1\left(\frac{[nt]}{n}\right) \phi_{\epsilon}(x_{[nt],n}) - G_1(t) \phi_{\epsilon}(W(t)) \right| dt \\
&\leq \int_0^1 \left| G_1\left(\frac{[nt]}{n}\right) - G_1(t) \right| |\phi_{\epsilon}(x_{[nt],n})| dt + \int_0^1 |G_1(t)| |\phi_{\epsilon}(x_{[nt],n}) - \phi_{\epsilon}(W(t))| dt
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{\sqrt{2\pi\epsilon}} \int_0^1 \left| G_1\left(\frac{[nt]}{n}\right) - G_1(t) \right| dt + \max_{0 \leq t \leq 1} |G_1(t)| \int_0^1 |\phi_\epsilon(x_{[nt],n}) - \phi_\epsilon(W(t))| dt \\
&\leq \frac{1}{\sqrt{2\pi\epsilon}} \int_0^1 \left| G_1\left(\frac{[nt]}{n}\right) - G_1(t) \right| dt + \frac{1}{\sqrt{2\pi\epsilon\epsilon^2}} \max_{0 \leq t \leq 1} |G_1(t)| \sup_{0 \leq t \leq 1} |x_{[nt],n} - W(t)|.
\end{aligned}$$

Hence, using the dominated convergence theorem and Assumption A(b), as $n \rightarrow \infty$, (D.4) converges in probability to zero. Then the assertion follows as $\epsilon \rightarrow 0$.

Now we turn to prove (3.8). Define, for $\frac{\tau_{n,i-1}}{n} < t \leq \frac{\tau_{n,i}}{n}$,

$$\begin{aligned}
M_n(t) &= \sqrt{c_n} \sum_{k=1}^{i-1} f\left(\frac{k}{n}, c_n W_n^0\left(\frac{k}{n}\right)\right) \left(U\left(\frac{\tau_{nk}}{n}\right) - U\left(\frac{\tau_{n,k-1}}{n}\right) \right) \\
&\quad + \sqrt{c_n} f\left(\frac{i}{n}, c_n W_n^0\left(\frac{i}{n}\right)\right) \left(U(t) - U\left(\frac{\tau_{n,i-1}}{n}\right) \right), \tag{D.5}
\end{aligned}$$

where τ_{nk} ($k = 1, \dots, n$) are the stopping times in Lemma C.3. It follows that, for any $n \geq 1$, $M_n(t)$ is a continuous martingale with respect to the filtration $\mathcal{F}_n(t) := \sigma(W_n^0(\frac{1}{n}), \dots, W_n^0(\frac{i}{n}), U(s) \mid s \leq t, \frac{\tau_{n,i-1}}{n} < t \leq \frac{\tau_{n,i}}{n})$. We can then derive that

$$\sqrt{\frac{c_n}{n}} \sum_{k=1}^{[nr]} f\left(\frac{k}{n}, c_n x_{k,n}\right) e_k \stackrel{D}{=} M_n\left(\frac{\tau_{n,i}}{n}\right), \quad \text{if } \frac{\tau_{n,i-1}}{n} < r \leq \frac{\tau_{n,i}}{n}, \tag{D.6}$$

and deduct from (C.6) that

$$\sup_{1 \leq k \leq n} \left| \left(\frac{\tau_{nk}}{n} - \frac{\tau_{n,k-1}}{n} \right) - \frac{1}{n} \right| = o(1), \quad a.s.. \tag{D.7}$$

The quadratic variation process $[M_n]$ of $M_n(t)$ is that

$$\begin{aligned}
[M_n]_t &= c_n \sum_{k=1}^{i-1} f^2\left(\frac{k}{n}, c_n W_n^0\left(\frac{k}{n}\right)\right) \left(\frac{\tau_{nk}}{n} - \frac{\tau_{n,k-1}}{n} \right) + c_n f^2\left(\frac{i}{n}, c_n W_n^0\left(\frac{i}{n}\right)\right) \left(t - \frac{\tau_{n,i-1}}{n} \right) \\
&= \frac{c_n}{n} \sum_{k=1}^{i-1} f^2\left(\frac{k}{n}, c_n W_n^0\left(\frac{k}{n}\right)\right) (1 + o_{a.s.}(1)) + c_n f^2\left(\frac{i}{n}, c_n W_n^0\left(\frac{i}{n}\right)\right) \left(t - \frac{\tau_{n,i-1}}{n} \right).
\end{aligned}$$

Because

$$\begin{aligned}
&E \left| c_n f^2\left(\frac{i}{n}, c_n W_n^0\left(\frac{i}{n}\right)\right) \left(t - \frac{\tau_{n,i-1}}{n} \right) \right| \leq \frac{c_n}{n} E f^2\left(\frac{i}{n}, c_n W_n^0\left(\frac{i}{n}\right)\right) \\
&= \frac{c_n}{n} \int_{-\infty}^{\infty} f^2\left(\frac{i}{n}, c_n d_{i,0,n} x\right) h_{i,0,n}(x) dx \leq \frac{K}{n d_{i,0,n}} \int_{-\infty}^{\infty} f^2\left(\frac{i}{n}, x\right) dx \\
&= \frac{K}{n d_{i,0,n}} G_3\left(\frac{i}{n}\right) \leq \frac{K K_3}{n d_{i,0,n}} \rightarrow 0,
\end{aligned}$$

where $K_3 = \max_{0 \leq t \leq 1} G_3(t)$ and by Assumption A (C), $nd_{i,0,n} \rightarrow \infty$, we have that $c_n f^2\left(\frac{i}{n}, c_n x_{i,n}\right) \left(r - \frac{\tau_{n,i-1}}{n}\right) \rightarrow_P 0$.

It therefore follows from (3.6) that

$$[M_n]_t \rightarrow_D \int_0^t G_3(a) dL_W(a, 0), \quad (\text{D.8})$$

as $n \rightarrow \infty$.

Moreover, the covariance process $[M_n, W]$ of (M_n, W) is

$$\begin{aligned} [M_n, W]_t &= \sqrt{c_n} \sum_{k=1}^{i-1} f\left(\frac{k}{n}, c_n W_n^0\left(\frac{k}{n}\right)\right) \left(\frac{\tau_{nk}}{n} - \frac{\tau_{n,k-1}}{n}\right) \sigma_{uw} \\ &\quad + \sqrt{c_n} f\left(\frac{i}{n}, c_n W_n^0\left(\frac{i}{n}\right)\right) \left(t - \frac{\tau_{n,i-1}}{n}\right) \sigma_{uw} \\ &= \sigma_{uw} (1 + o_{a.s.}(1)) \frac{\sqrt{c_n}}{n} \sum_{k=1}^{i-1} f\left(\frac{k}{n}, c_n W_n^0\left(\frac{k}{n}\right)\right) \\ &\quad + \sigma_{uw} \sqrt{c_n} f\left(\frac{i}{n}, c_n W_n^0\left(\frac{i}{n}\right)\right) \left(t - \frac{\tau_{n,i-1}}{n}\right), \end{aligned}$$

for any $t \in [0, 1]$, where $\sigma_{uw} = \text{Cov}(U, W)$. Meanwhile, using argument in Example 25.7 on Billingsley (1995, p332),

$$\left| \frac{\sqrt{c_n}}{n} \sum_{k=1}^{i-1} f\left(\frac{k}{n}, c_n W_n^0\left(\frac{k}{n}\right)\right) \right| \leq \frac{1}{\sqrt{c_n}} \frac{c_n}{n} \sum_{k=1}^n \left| f\left(\frac{k}{n}, c_n W_n^0\left(\frac{k}{n}\right)\right) \right| \rightarrow_P 0,$$

because $c_n \rightarrow \infty$ and using (3.6), we have

$$\frac{c_n}{n} \sum_{k=1}^n \left| f\left(\frac{k}{n}, c_n W_n^0\left(\frac{k}{n}\right)\right) \right| \rightarrow_D \int_0^1 G_2(t) dL_W(t, 0).$$

Additionally, $\left| \sqrt{c_n} f\left(\frac{i}{n}, c_n W_n^0\left(\frac{i}{n}\right)\right) \left(t - \frac{\tau_{n,i-1}}{n}\right) \right| \leq \frac{\sqrt{c_n}}{n} \left| f\left(\frac{i}{n}, c_n W_n^0\left(\frac{i}{n}\right)\right) \right| \rightarrow_P 0$ by the similar argument as above. Thus,

$$[M_n, W]_{T_n(t)} \rightarrow_P 0, \quad (\text{D.9})$$

where $T_n(t) = \inf\{s \in [0, 1], [M_n]_s > t\}$ be the sequence of time changes. Then, in virtue of DDS (Dambis, Dubins-Schwarz) theorem (see, for example, Revuz and Yor, 1999, p181), it follows that the process defined by

$$B^n(t) = M_n(T_n(t))$$

becomes a so-called DDS Brownian motion. Also, $M_n(t) = B^n([M_n]_t)$, and it follows from Theorem 2.3 of Revuz and Yor (1999, p524) that (W, B^n) converges in distribution jointly to two independent Brownian motions (W, B) . Therefore, we have as $n \rightarrow \infty$

$$\begin{aligned} & \sqrt{\frac{c_n}{n}} \sum_{k=1}^{[nr]} f\left(\frac{k}{n}, c_n x_{k,n}\right) e_k \stackrel{D}{=} M_n\left(\frac{\tau_{n,i}}{n}\right) \\ & = M_n(r) + o_{a.s.}(1) = B^n([M_n]_r) + o_{a.s.}(1) \\ & \rightarrow_D B\left(\int_0^r G_3(a) dL_W(a, 0)\right) = \left(\int_0^r G_3(a) dL_W(a, 0)\right)^{\frac{1}{2}} B(1). \end{aligned}$$

This finishes the whole proof. □

PROOF OF THEOREM 3.2. It follows from the definition of the class $\mathcal{T}(HI)$ that

$$\begin{aligned} \frac{c_n}{n\nu(n)} \sum_{s=1}^{[nr]} F(s, c_n x_{s,n}) &= \frac{c_n}{n} \sum_{s=1}^{[nr]} f\left(\frac{s}{n}, c_n x_{s,n}\right) + \frac{c_n}{n\nu(n)} \sum_{s=1}^{[nr]} R_n\left(\frac{s}{n}, c_n x_{s,n}\right) \\ &:= \Pi_1 + \Pi_2. \end{aligned}$$

As suggested by Theorem 3.1, under Condition A(a) and A(c) in Assumption A, we have $\Pi_1 \rightarrow_D \int_0^r G_1(t) dL_W(t, 0)$, while under Condition A(b) and A(c), we have $\Pi_1 \rightarrow_P \int_0^r G_1(t) dL_W(t, 0)$ uniformly in r . In order to complete (3.9) and (3.10), it thus suffices to prove that $\Pi_2 \rightarrow_P 0$ uniformly in r under Condition A(c) by virtue of the properties of convergence in distribution and in probability (see Theorem 25.4 on Billingsley (1995, p332) for weak convergence and any text book for convergence in probability. We do not mention this any more in the sequel).

If $F(t, x)$ is in the class $\mathcal{T}(HI_1)$ and $q_n(t)/\nu(n) \rightarrow 0$ uniformly in $t \in [0, 1]$ as $n \rightarrow \infty$, for a given $\epsilon > 0$, when n is large, $0 < q_n(t)/\nu(n) < \epsilon$ for all t . We then have from Assumption A(c) that

$$\begin{aligned} \sup_{0 \leq r \leq 1} E|\Pi_2| &\leq \frac{c_n}{n\nu(n)} \sup_{0 \leq r \leq 1} \sum_{s=1}^{[nr]} E \left| R_n\left(\frac{s}{n}, c_n x_{s,n}\right) \right| \\ &\leq \frac{c_n}{n\nu(n)} \sup_{0 \leq r \leq 1} \sum_{s=1}^{[nr]} q_n\left(\frac{s}{n}\right) E[P(c_n x_{s,n})] \\ &\leq \epsilon \frac{c_n}{n} \sum_{s=1}^n \int_{-\infty}^{\infty} P(c_n d_{s,0,n} x) h_{s,0,n}(x) dx \end{aligned}$$

$$\begin{aligned}
&= \epsilon \frac{1}{n} \sum_{s=1}^n \frac{1}{d_{s,0,n}} \int_{-\infty}^{\infty} P(x) h_{s,0,n} \left(\frac{1}{c_n d_{s,0,n}} x \right) dx \\
&\leq \epsilon K \int_{-\infty}^{\infty} P(x) dx \frac{1}{n} \sum_{s=1}^n \frac{1}{d_{s,0,n}},
\end{aligned}$$

where K is the uniform upper bound of the densities $h_{l,k,n}(x)$. Thus, the desired result follows from (3.3) as $n \rightarrow \infty$ and $\epsilon \rightarrow 0$.

If $F(t, x)$ is in the class $\mathcal{T}(HI_2)$, $|R_n(\frac{s}{n}, c_n x_{s,n})| \leq q_n(t) Q(nt) P(c_n x_{s,n})$ with $P(x)$ integrable, $\lim_{n \rightarrow \infty} q_n(t)/v(n) = l(t)$ which is bounded on $[0,1]$ and $Q(y)$ that is bounded on any compact interval and $\lim_{y \rightarrow +\infty} Q(y) = 0$. We have when n is large, $q_n(t)/v(n) = l(t)(1 + o(1))$ and for a given $\epsilon > 0$, there exists $s_0 > 0$ such that $0 < Q(s) < \epsilon$ whenever $s > s_0$. Whence,

$$\begin{aligned}
\sup_{0 \leq r \leq 1} E|\Pi_2| &\leq \frac{c_n}{nv(n)} \sup_{0 \leq r \leq 1} \sum_{s=1}^{[nr]} E \left| R_n \left(\frac{s}{n}, c_n x_{s,n} \right) \right| \\
&\leq \frac{c_n}{nv(n)} \sup_{0 \leq r \leq 1} \sum_{s=1}^{[nr]} q_n \left(\frac{s}{n} \right) Q(s) E[P(c_n x_{s,n})] \\
&\leq \frac{c_n}{nv(n)} \sum_{s=1}^n q_n \left(\frac{s}{n} \right) Q(s) E[P(c_n x_{s,n})] \\
&\leq \frac{c_n}{n} \max_{0 \leq t \leq 1} l(t) \sum_{s=1}^n Q(s) \int_{-\infty}^{\infty} P(c_n d_{s,0,n} x) h_{s,0,n}(x) dx \\
&\leq K \max_{0 \leq t \leq 1} l(t) \int_{-\infty}^{\infty} P(x) dx \frac{1}{n} \sum_{s=1}^n \frac{1}{d_{s,0,n}} Q(s) \\
&= K \max_{0 \leq t \leq 1} l(t) \int_{-\infty}^{\infty} P(x) dx \left(\frac{1}{n} \sum_{s=1}^{s_0} \frac{Q(s)}{d_{s,0,n}} + \frac{1}{n} \sum_{s=s_0}^n \frac{Q(s)}{d_{s,0,n}} \right) \\
&\leq K \max_{0 \leq t \leq 1} l(t) \int_{-\infty}^{\infty} P(x) dx \left(K(s_0) \frac{1}{n} \sum_{s=1}^{s_0} \frac{1}{d_{s,0,n}} + \epsilon \frac{1}{n} \sum_{s=1}^n \frac{1}{d_{s,0,n}} \right) \\
&\rightarrow 0,
\end{aligned}$$

as $n \rightarrow \infty$ and then $\epsilon \rightarrow 0$ due to (3.2) and (3.3) where $K(s_0)$ is the maximum over $Q(1), \dots, Q(s_0)$. This finishes the proof of (3.9) and (3.10). Now we turn to prove (3.11).

By virtue of the definition of the class $\mathcal{T}(HI)$,

$$\sqrt{\frac{c_n}{n}} \frac{1}{v(n)} \sum_{s=1}^{[nr]} F(s, c_n x_{s,n}) e_s$$

$$\begin{aligned}
&= \sqrt{\frac{c_n}{n}} \sum_{s=1}^{[nr]} f\left(\frac{s}{n}, c_n x_{s,n}\right) e_s + \sqrt{\frac{c_n}{n}} \frac{1}{v(n)} \sum_{s=1}^{[nr]} R_n\left(\frac{s}{n}, c_n x_{s,n}\right) e_s \\
&:= \Pi_3 + \Pi_4.
\end{aligned}$$

It follows from Theorem 3.1 that

$$\Pi_3 \rightarrow_D \left(\int_0^r G_3(t) dL_W(t, 0) \right)^{\frac{1}{2}} N,$$

as $n \rightarrow \infty$. Hence, it is sufficient to show $\Pi_4 \rightarrow_P 0$ to complete the proof.

The structure of martingale difference of $(e_s, \mathcal{F}_{n.s})$ and the adaptivity between e_s and $x_{s,n}$ give

$$E[\Pi_4]^2 = \sigma_e^2 \frac{c_n}{n} \frac{1}{v(n)^2} \sum_{s=1}^{[nr]} ER_n^2\left(\frac{s}{n}, c_n x_{s,n}\right).$$

In a very similar fashion as in the proof of $\Pi_2 \rightarrow_P 0$ we can show that $E[\Pi_4]^2 \rightarrow 0$. This finishes the proof. \square

PROOF OF THEOREM 3.3. Notice firstly that under Assumption B, with the argument in Lemma C.3, we may rewrite any statistic about U_n and W_n equivalently in distribution into an expression of U_n^0 and W_n^0 , so that we can obtain the weak convergence of the statistic by studying the latter with almost sure convergence of $(U_n^0, W_n^0) \rightarrow_{a.s.} (U, W)$. Such schedule of consideration is referred to as the *embedding schedule*¹ in the sequel. It therefore is reasonable in the sequel to assume without loss of generality that $(U_n, W_n) \rightarrow_{a.s.} (U, W)$ in order to avoid notational complication. We begin to prove the result in (3.15). We first write that

$$\begin{aligned}
\frac{1}{n} \sum_{s=1}^n f\left(\frac{s}{n}, x_{s,n}\right) &= \frac{1}{n} \sum_{s=1}^n f\left(\frac{s-1}{n} + \frac{1}{n}, W_n\left(\frac{s-1}{n} + \frac{1}{n}\right)\right) \\
&= \sum_{s=1}^n \int_{\frac{s-1}{n}}^{\frac{s}{n}} f(r + o(1), W_n(r + o(1))) dr = \int_0^1 f(r + o(1), W_n(r + o(1))) dr.
\end{aligned}$$

Thus, to complete the result in (3.15), it suffices to show

$$\int_0^1 f(r + o(1), W_n(r + o(1))) dr \rightarrow_{a.s.} \int_0^1 f(r, W(r)) dr.$$

¹We emphasize that the embedding schedule applies in the subsequent proofs. We shall mention it without showing the details whenever it is used.

Because of the condition (b) in regularity definition, there exists a constant $c > 0$ such that $f(t, x)$ is continuous in x whenever $|x| > c$. Let $J = [-c - 2, c + 2]$. For any given $\epsilon > 0$, it follows from the regularity of f that there exist continuous functions $\underline{f}_\epsilon(r, x)$, $\bar{f}_\epsilon(r, x)$ in x and $\delta > 0$ such that whenever $|x - y| < \delta$ on J , for each $r \in [0, 1]$,

$$\underline{f}_\epsilon(r, x) \leq f(r, y) \leq \bar{f}_\epsilon(r, x).$$

Note that when $x = y \in J$, we always have $\underline{f}_\epsilon(r, x) \leq f(r, x) \leq \bar{f}_\epsilon(r, x)$.

Since $\sup_{0 \leq r \leq 1} |W_n(r) - W(r)| = o_{a.s.}(1)$, let n large enough such that $\sup_{0 \leq r \leq 1} |W_n(r) - W(r)| < \frac{1}{2}\delta$ almost surely. Without loss of generality, assume that $\delta < 1$.

Observe that for large n , $|W_n(r + o(1)) - W(r)| \leq |W_n(r + o(1)) - W(r + o(1))| + |W(r + o(1)) - W(r)| < \delta$ almost surely uniformly in r where we exploit the fact that Brownian motion sample path is almost surely continuous, hence almost surely uniformly continuous on $[0, 1]$.

Denote $A(r) = \{|W(r)| < c + 1\}$. It follows that on $A(r)$, when n is large, $W_n(r + o(1)) \in J$, $W(r) \in J$; while on $\bar{A}(r)$, $|W_n(r + o(1))| > c$, $|W(r)| > c$.

Notice that from Condition (a) of regularity,

$$\begin{aligned} & \left| \int_0^1 f(r + o(1), W_n(r + o(1))) dr - \int_0^1 f(r, W(r)) dr \right| \\ & \leq \int_0^1 |f(r + o(1), W_n(r + o(1))) - f(r, W_n(r + o(1)))| dr \\ & \quad + \left| \int_0^1 [f(r, W_n(r + o(1))) - f(r, W(r))] dr \right| \\ & \leq o(1) \int_0^1 L(W_n(r + o(1))) dr + \left| \int_0^1 [f(r, W_n(r + o(1))) - f(r, W(r))] dr \right|. \end{aligned}$$

However,

$$\begin{aligned} & \left| \int_0^1 [f(r, W_n(r + o(1))) - f(r, W(r))] dr \right| \\ & \leq \int_0^1 |f(r, W_n(r + o(1))) - f(r, W(r))| I(A(r)) dr \\ & \quad + \left| \int_0^1 [f(r, W_n(r + o(1))) - f(r, W(r))] I(\bar{A}(r)) dr \right| \\ & \leq \int_0^1 |\bar{f}_\epsilon(r, W(r)) - \underline{f}_\epsilon(r, W(r))| I(A(r)) dr \\ & \quad + \left| \int_0^1 [f(r, W_n(r + o(1))) - f(r, W(r))] I(\bar{A}(r)) dr \right| \end{aligned}$$

$$:= \Delta_1 + \Delta_2.$$

where $I(\cdot)$ is the indicator function.

Moreover, it follows from the occupation time formula for bivariate Brownian functional that

$$\begin{aligned} \Delta_1 &= \int_0^1 [\bar{f}_\epsilon(r, W(r)) - \underline{f}_\epsilon(r, W(r))] I(|W(r)| < c + 1) dr \\ &= \int_{-c-1}^{c+1} da \int_0^1 [\bar{f}_\epsilon(r, a) - \underline{f}_\epsilon(r, a)] dL_W(r, a) \\ &\leq \int_J \sup_{0 \leq r \leq 1} [\bar{f}_\epsilon(r, a) - \underline{f}_\epsilon(r, a)] da \int_0^1 dL_W(r, a) \\ &= \int_J L_W(1, a) \sup_{0 \leq r \leq 1} [\bar{f}_\epsilon(r, a) - \underline{f}_\epsilon(r, a)] da \\ &\leq \sup_{a \in J} L_W(1, a) \int_J \sup_{0 \leq r \leq 1} [\bar{f}_\epsilon(r, a) - \underline{f}_\epsilon(r, a)] da \xrightarrow{a.s.} 0, \end{aligned}$$

as $\epsilon \rightarrow 0$, due to regularity of f and $\sup_{a \in J} L_W(1, a) \leq 1$ almost surely.

Furthermore, because $f(r, \cdot)$ is continuous on $|x| > c$, the continuous mapping theorem implies that $\Delta_2 \rightarrow 0$ a.s.

Regarding of $\int_0^1 L(W_n(r + o(1))) dr$, since $L(\cdot)$ satisfies Condition (b) and (c) in regularity, similar derivation as above yields it approaches to $\int_0^1 L(W(r)) dr$ almost surely. Hence, the proof of (3.15) is completed.

We are ready to prove (3.16). Once again the embedding schedule described in the first part permits us to derive it under a stronger condition that $(U_n, W_n) \xrightarrow{a.s.} (W, U)$.

Let us write

$$\begin{aligned} &\frac{1}{\sqrt{n}} \sum_{s=1}^n f\left(\frac{s}{n}, x_{s,n}\right) e_s = \sum_{s=1}^n f\left(\frac{s}{n}, x_{s,n}\right) \frac{1}{\sqrt{n}} e_s \\ &= \sum_{s=1}^n f\left(\frac{s-1}{n} + o(1), W_n\left(\frac{s-1}{n} + o(1)\right)\right) \left(U_n\left(\frac{s}{n}\right) - U_n\left(\frac{s-1}{n}\right)\right) \\ &= \sum_{s=1}^n \int_{\frac{s-1}{n}}^{\frac{s}{n}} f(r + o(1), W_n(r + o(1))) dU_n(r) \\ &= \int_0^1 f(r + o(1), W_n(r + o(1))) dU_n(r) := \sum_{k=1}^4 \Pi_k, \end{aligned}$$

where

$$\Pi_1 = \int_0^1 [f(r + o(1), W_n(r + o(1))) - f(r, W_n(r + o(1)))] dU_n(r)$$

$$\begin{aligned}
\Pi_2 &= \int_0^1 [f(r, W_n(r + o(1))) - f_\epsilon(r, W_n(r + o(1)))] dU_n(r) \\
\Pi_3 &= \int_0^1 f_\epsilon(r, W_n(r + o(1))) dU_n(r) - \int_0^1 f_\epsilon(r, W(r)) dU(r) \\
\Pi_4 &= \int_0^1 f_\epsilon(r, W(r)) dU(r),
\end{aligned}$$

in which denote $f_\epsilon(r, x) = \bar{f}_\epsilon(r, x)$ or $\underline{f}_\epsilon(r, x)$ for notational convenience. Observe that $(f_\epsilon(r, W_n(r + o(1))), U_n(r)) \rightarrow (f_\epsilon(r, W(r)), U(r))$ almost surely due to continuity in x of f_ϵ . It follows from Theorem 2.2 in Kurtz and Protter (1991) that $\Pi_3 \rightarrow_P 0$ as $n \rightarrow \infty$.

Therefore, in order to finish the proof, we need to show (1) $\Pi_1 \rightarrow_P 0$ when $n \rightarrow \infty$; (2) for all large n , $\Pi_2 \rightarrow_P 0$ and $\Pi_4 \rightarrow_P \int_0^1 f(r, W(r)) dU(r)$ when $\epsilon \rightarrow 0$. Let us investigate them term by term.

It follows from Assumption B and regularity that

$$\begin{aligned}
E[\Pi_1]^2 &= E \left\{ \int_0^1 [f(r + o(1), W_n(r + o(1))) - f(r, W_n(r + o(1)))] dU_n(r) \right\}^2 \\
&= \sigma_\epsilon^2 E \int_0^1 [f(r + o(1), W_n(r + o(1))) - f(r, W_n(r + o(1)))]^2 dr \\
&\leq o(1) \sigma_\epsilon^2 E \int_0^1 L^2(W_n(r + o(1))) dr \rightarrow 0,
\end{aligned}$$

as $n \rightarrow \infty$ because we have $\int_0^1 L^2(W_n(r + o(1))) dr \rightarrow_{a.s.} \int_0^1 L^2(W(r)) dr$ similar to the counterpart in first part, and by virtue of the regularity, $L^2(W_n(r))$ can be dominated by $L_\epsilon^2(W(r))$ when n is large for some $\epsilon > 0$ and $L_\epsilon(\cdot)$ is continuous, $E \int_0^1 L^2(W_n(r)) dr \rightarrow E \int_0^1 L^2(W(r)) dr < \infty$. This finishes the proof of (1).

The convergence of Π_2 and Π_4 can be proven at the same time if we show

$$\int_0^1 [f(r, W_n(r + o(1))) - f_\epsilon(r, W_n(r + o(1)))] dU_n(r) \rightarrow_P 0,$$

as $\epsilon \rightarrow 0$ for all large n including $n = \infty$ that means conventionally $(U_\infty(r), V_\infty(r)) = (U(r), V(r))$.

Let real c be defined as before. All notations $\epsilon, \delta, J, A(r), \bar{f}_\epsilon(t, x)$ and $\underline{f}_\epsilon(t, x)$ keep the same meanings as in the first part. In view of regularity condition (b), we may find $\bar{f}_\epsilon(r, x)$ and $\underline{f}_\epsilon(r, x)$ such that they are continuous in x on \mathbb{R} for each $r \in [0, 1]$, since beyond $[-c, c]$, we can take $\bar{f}_\epsilon(r, x) = \underline{f}_\epsilon(r, x) = f(t, x)$ and due to this reason, $\bar{f}_\epsilon(r, x) - \underline{f}_\epsilon(r, x)$ is bounded on \mathbb{R} . Consequently, $\sup_{r \in [0, 1]} (\bar{f}_\epsilon(r, x) - \underline{f}_\epsilon(r, x))$ is bounded

on \mathbb{R} because it is continuous and beyond $[-c, c]$ it is zero. Let C be the upper bound of $\sup_{r \in [0,1]} [\bar{f}_\epsilon(r, x) - \underline{f}_\epsilon(r, x)]$.

By the adaptivity of $(U_n(r), W_n(r + o(1)))$, for large n ,

$$\begin{aligned}
& E \left\{ \int_0^1 [f(r, W_n(r + o(1))) - f_\epsilon(r, W_n(r + o(1)))] dU_n(r) \right\}^2 \\
&= \sigma_\epsilon^2 E \int_0^1 [f(r, W_n(r + o(1))) - f_\epsilon(r, W_n(r + o(1)))]^2 dr \\
&= \sigma_\epsilon^2 E \int_0^1 [f(r, W_n(r + o(1))) - f_\epsilon(r, W_n(r + o(1)))]^2 I(A(r)) dr \\
&\leq \sigma_\epsilon^2 E \int_0^1 [\bar{f}_\epsilon(r, W_n(r + o(1))) - \underline{f}_\epsilon(r, W_n(r + o(1)))]^2 I(A(r)) dr \\
&\rightarrow_{a.s.} \sigma_\epsilon^2 E \int_0^1 [\bar{f}_\epsilon(r, W(r)) - \underline{f}_\epsilon(r, W(r))]^2 I(A(r)) dr,
\end{aligned}$$

by virtue of continuity and boundedness of $\bar{f}_\epsilon(t, x) - \underline{f}_\epsilon(t, x)$ in x and the fact that indicator function is bounded as $n \rightarrow \infty$. Observe that by the occupation formula

$$\begin{aligned}
& \int_0^1 [\bar{f}_\epsilon(r, W(r)) - \underline{f}_\epsilon(r, W(r))]^2 I(|W(r)| \leq c + 1) dr \\
&= \int_{-\infty}^{\infty} da \int_0^1 [\bar{f}_\epsilon(r, a) - \underline{f}_\epsilon(r, a)]^2 I(|a| \leq c + 1) dL_W(r, a) \\
&= \int_{-c-1}^{c+1} da \int_0^1 [\bar{f}_\epsilon(r, a) - \underline{f}_\epsilon(r, a)]^2 dL_W(r, a) \\
&\leq C \int_J \sup_{0 \leq r \leq 1} [\bar{f}_\epsilon(r, a) - \underline{f}_\epsilon(r, a)] da \int_0^1 dL_W(r, a) \\
&\leq C \sup_a L_W(1, a) \int_J \sup_{0 \leq r \leq 1} [\bar{f}_\epsilon(r, a) - \underline{f}_\epsilon(r, a)] da \rightarrow_{a.s.} 0,
\end{aligned}$$

as $\epsilon \rightarrow 0$.

It follows from the dominated convergence theorem that $\Pi_2 \rightarrow_P 0$ and Π_4 converges to the desired variable in probability as $\epsilon \rightarrow 0$. This finishes the proof. \square

PROOF OF THEOREM 3.4. Observe that, like preceding proofs, the embedding schedule allow us to work under a stronger condition $(W_n, U_n) \rightarrow (W, U)$ almost surely but still achieve the weak convergence for the assertion.

It follows from the asymptotic homogeneity of F function that

$$\frac{1}{nv_1(n)v_2(c_n)} \sum_{s=1}^n F(s, c_n x_{s,n})$$

$$= \frac{1}{n} \sum_{s=1}^n f\left(\frac{s}{n}, x_{s,n}\right) + \frac{1}{nv_1(n)v_2(c_n)} \sum_{s=1}^n R(n, c_n; s, c_n x_{s,n}).$$

Note that $f(t, x)$ is regular and thus by the proof (not the result) of Theorem 3.3,

$$\frac{1}{n} \sum_{s=1}^n f\left(\frac{s}{n}, x_{s,n}\right) \rightarrow_{a.s.} \int_0^1 f(r, W(r)) dr,$$

as $n \rightarrow \infty$.

In order to complete the proof of (3.18), it thus suffices to show

$$\frac{1}{nv_1(n)v_2(c_n)} \sum_{s=1}^n R(n, c_n; s, c_n x_{s,n}) \rightarrow_{a.s.} 0.$$

Let $\lim_{n \rightarrow \infty} \frac{a(c_n)}{v_2(c_n)} = a$ and $\lim_{n \rightarrow \infty} \frac{b(n)}{v_1(n)} = b$. Let $K = [s_{\min} - 1, s_{\max} + 1]$ with $s_{\min} = \inf_{r \in [0,1]} W(r)$ and $s_{\max} = \sup_{r \in [0,1]} W(r)$. Note that almost surely K is a finite compact interval.

It follows from the definition that as n is large,

$$\begin{aligned} & \frac{1}{nv_1(n)v_2(c_n)} \sum_{s=1}^n |R(n, c_n; s, c_n x_{s,n})| \\ & \leq \frac{a(c_n)}{nv_1(n)v_2(c_n)} \sum_{s=1}^n A_n\left(\frac{s}{n}\right) P(x_{s,n}) + \frac{b(n)}{nv_1(n)v_2(c_n)} \sum_{s=1}^n q\left(\frac{s}{n}\right) B_{c_n}(x_{s,n}) \\ & = \frac{a(1+o(1))}{nv_1(n)} \sum_{s=1}^n A_n\left(\frac{s}{n}\right) P(x_{s,n}) + \frac{b(1+o(1))}{nv_2(c_n)} \sum_{s=1}^n q\left(\frac{s}{n}\right) B_{c_n}(x_{s,n}) \\ & := \Pi_1 + \Pi_2. \end{aligned}$$

If $\limsup_{n \rightarrow \infty} \frac{A_n\left(\frac{s}{n}\right)}{v_1(n)} = 0$ uniformly in s , then for any given $\epsilon > 0$, when n is large enough, $0 < \frac{A_n\left(\frac{s}{n}\right)}{v_1(n)} < \epsilon$. Thus,

$$0 \leq \Pi_1 < \epsilon a(1+o(1)) \frac{1}{n} \sum_{s=1}^n P(x_{s,n}) \leq \epsilon a(1+o(1)) \|P\|_K \rightarrow_{a.s.} 0,$$

as $n \rightarrow \infty$ and $\epsilon \rightarrow 0$ since $x_{s,n} = W_n(r) \in K$ due to convergence of $W_n(r)$ to $W(r)$ almost surely and $\|P\|_K$, the bound of $P(x)$ on K (in the sequel similar notations have the similar meaning), is almost surely finite. Thus, $\Pi_1 \rightarrow 0$, a.s..

If $v_1(n) \rightarrow \infty$ as $n \rightarrow \infty$ and $A_n(t) = A(t)$ which is Riemann integrable on $[0, 1]$, then

$$0 < \Pi_1 = \frac{a(1+o(1))}{nv_1(n)} \sum_{s=1}^n A\left(\frac{s}{n}\right) P(x_{s,n})$$

$$\leq \frac{a(1+o(1))}{v_1(n)} \|P\|_K \frac{1}{n} \sum_{s=1}^n A\left(\frac{s}{n}\right) \rightarrow_{a.s.} 0,$$

since as $n \rightarrow \infty$, $\frac{1}{v_1(n)} \rightarrow 0$, $\frac{1}{n} \sum_{s=1}^n A\left(\frac{s}{n}\right) \rightarrow \int_0^1 A(t)dt < \infty$ and $\|P\|_K < \infty$ a.s.. We have $\Pi_1 \rightarrow_{a.s.} 0$ as well.

If $A_n(t) = \bar{A}_n(t)Q(nt)$ with $\limsup_{n \rightarrow \infty} \frac{\bar{A}_n(t)}{v_1(n)} = l(t)$ bounded on $[0,1]$ and $Q(y)$ is bounded on \mathbb{R} as well as $\lim_{y \rightarrow +\infty} Q(y) = 0$, then for any given $\epsilon > 0$, there exists a positive integer s_0 such that when $y > s_0$, $0 < Q(y) < \epsilon$. Therefore,

$$\begin{aligned} 0 < \Pi_1 &= \frac{a(1+o(1))}{nv_1(n)} \sum_{s=1}^n \bar{A}_n\left(\frac{s}{n}\right) Q(s)P(x_{s,n}) \\ &\leq \frac{a(1+o(1))}{n} \sum_{s=1}^n l\left(\frac{s}{n}\right) Q(s)P(x_{s,n}) \\ &\leq a(1+o(1)) \max_{0 \leq t \leq 1} l(t) \|P\|_K \frac{1}{n} \sum_{s=1}^n Q(s) \\ &\leq a(1+o(1)) \max_{0 \leq t \leq 1} l(t) \|P\|_K \left[\frac{1}{n} \sum_{s=1}^{s_0} Q(s) + \frac{1}{n} \sum_{s=s_0}^n \epsilon \right] \rightarrow_{a.s.} 0, \end{aligned}$$

as $n \rightarrow \infty$ and $\epsilon \rightarrow 0$. Thus, $\Pi_1 \rightarrow_{a.s.} 0$ too.

We are now in a position to show $\Pi_2 \rightarrow_{a.s.} 0$.

If $B_{c_n}(x_{s,n}) = \bar{B}(c_n)V(x_{s,n})$ with $\limsup_{n \rightarrow \infty} \frac{\bar{B}(c_n)}{v_2(c_n)} = 0$, $0 \leq q(t) \leq M_q < \infty$ on $[0,1]$ and $V(x)$ is locally bounded, then for any given $\epsilon > 0$, when n is large, $0 < \frac{\bar{B}(c_n)}{v_2(c_n)} < \epsilon$. Thus,

$$\begin{aligned} 0 < \Pi_2 &= \frac{b(1+o(1))}{nv_2(c_n)} \sum_{s=1}^n q\left(\frac{s}{n}\right) \bar{B}(c_n)V(x_{s,n}) \\ &\leq \epsilon b(1+o(1)) \|V\|_K M_q \rightarrow_{a.s.} 0, \end{aligned}$$

as $n \rightarrow \infty$ and $\epsilon \rightarrow 0$. Thus, $\Pi_2 \rightarrow_{a.s.} 0$.

If $B_{c_n}(x_{s,n}) = \bar{B}(c_n)V(c_n x_{s,n})$ where $\limsup_{n \rightarrow \infty} \frac{\bar{B}(c_n)}{v_2(c_n)} = l < \infty$ and $V(y)$ is bounded and vanishes at infinity, viz., $\lim_{y \rightarrow \infty} V(y) = 0$, then when n is large, $\frac{\bar{B}(c_n)}{v_2(c_n)} = l(1+o(1))$ and when $|y| > y_0$ for some positive y_0 and a given $\epsilon > 0$, $|V(y)| < \epsilon$. Therefore,

$$\begin{aligned} 0 < \Pi_2 &= \frac{b(1+o(1))\bar{B}(c_n)}{nv_2(c_n)} \sum_{s=1}^n q\left(\frac{s}{n}\right) V(c_n x_{s,n}) \\ &= bl(1+o(1)) \frac{1}{n} \sum_{s=1}^n q\left(\frac{s}{n}\right) V(c_n x_{s,n}) \end{aligned}$$

$$\begin{aligned}
&= bLM_q(1 + o(1)) \frac{1}{n} \sum_{s=1}^n V(c_n x_{s,n}) \\
&\quad \times [I(|c_n x_{s,n}| \leq y_0) + I(|c_n x_{s,n}| > y_0)] \\
&\leq bLM_q(1 + o(1)) \left(\|V\| \frac{1}{n} \sum_{s=1}^n I(c_n |x_{s,n}| \leq y_0) + \epsilon \right) \\
&= bLM_q(1 + o(1)) \left(\|V\| \int_0^1 I(c_n |W_n(r + o(1))| \leq y_0) dr + \epsilon \right).
\end{aligned}$$

Observe that for $\epsilon > 0$,

$$\begin{aligned}
&\{c_n |W_n(r + o(1))| \leq y_0\} \\
&= \left\{ |W_n(r + o(1))| \leq \frac{y_0}{c_n}, |W(r)| \leq \frac{y_0}{c_n} + \epsilon \right\} \\
&\quad \cup \left\{ |W_n(r + o(1))| \leq \frac{y_0}{c_n}, |W(r)| > \frac{y_0}{c_n} + \epsilon \right\} \\
&\subset \left\{ |W(r)| \leq \frac{y_0}{c_n} + \epsilon \right\} \cup \{|W_n(r + o(1)) - W(r)| > \epsilon\}.
\end{aligned}$$

Thus,

$$\begin{aligned}
I\{c_n |W_n(r + o(1))| \leq y_0\} &\leq I \left\{ |W(r)| \leq \frac{1}{c_n} y_0 + \epsilon \right\} \\
&\quad + I \{|W_n(r + o(1)) - W(r)| > \epsilon\}.
\end{aligned}$$

However, as $n \rightarrow \infty$, for every $r \in [0, 1)$,

$$\begin{aligned}
&\left\{ |W(r)| \leq \frac{1}{c_n} y_0 + \epsilon \right\} \downarrow \{|W(r)| \leq \epsilon\}, \quad \text{and} \\
&\{|W_n(r + o(1)) - W(r)| > \epsilon\} \downarrow \emptyset,
\end{aligned}$$

which imply that

$$\begin{aligned}
&I \left\{ |W(r)| \leq \frac{1}{c_n} y_0 + \epsilon \right\} \rightarrow_{a.s.} I \{|W(r)| \leq \epsilon\} \\
&I \left\{ \sup_{0 \leq r \leq 1} |W_n(r + o(1)) - W(r)| > \epsilon \right\} \rightarrow_{a.s.} 0.
\end{aligned}$$

It follows from the dominated convergence theorem that

$$0 \leq \int_0^1 I(c_n |W_n(r + o(1))| \leq y_0) dr$$

$$\begin{aligned} &\leq \int_0^1 I \left\{ |W(r)| \leq \frac{1}{c_n} y_0 + \epsilon \right\} dr + \int_0^1 I \{ |W_n(r + o(1)) - W(r)| > \epsilon \} dr \\ &\xrightarrow{a.s.} \int_0^1 I \{ |W(r)| \leq \epsilon \} dr. \end{aligned}$$

Then, as $\epsilon \rightarrow 0$, $I \{ |W(r)| \leq \epsilon \} \xrightarrow{a.s.} I \{ |W(r)| = 0 \} = 0$ almost surely except $r = 0$. Once again, the dominated convergence theorem implies that

$$\int_0^1 I \{ |W(r)| \leq \epsilon \} dr \xrightarrow{a.s.} \int_0^1 I \{ |W(r)| = 0 \} dr = 0, \quad a.s.$$

Hence, $\Pi_2 \xrightarrow{a.s.} 0$ as $n \rightarrow \infty$ first and then $\epsilon \rightarrow 0$. This finishes the proof of (3.18). We are now ready to prove (3.19).

It follows from the asymptotic homogeneity of $F(\cdot, \cdot)$ that

$$\begin{aligned} &\frac{1}{\sqrt{n}v_1(n)v_2(c_n)} \sum_{s=1}^n F(s, c_n x_{s,n}) e_s \\ &= \frac{1}{\sqrt{n}} \sum_{s=1}^n f\left(\frac{s}{n}, x_{s,n}\right) e_s + \frac{1}{\sqrt{n}v_1(n)v_2(c_n)} \sum_{s=1}^n R(n, c_n; s, c_n x_{s,n}) e_s \\ &:= \Pi_3 + \Pi_4. \end{aligned}$$

According to Theorem 3.3,

$$\Pi_3 = \frac{1}{\sqrt{n}} \sum_{s=1}^n f\left(\frac{s}{n}, x_{s,n}\right) e_s \xrightarrow{D} \int_0^1 f(r, W(r)) dU(r).$$

It thus suffices to show that with the help of the embedding schedule, $\Pi_4 \xrightarrow{P} 0$ as $n \rightarrow \infty$ in order to finish the proof. Using martingale structure of $(e_s, \mathcal{F}_{n,s})$ we have

$$\begin{aligned} E[\Pi_4]^2 &= \frac{1}{nv_1(n)^2 v_2(c_n)^2} E \left[\sum_{s=1}^n R(n, c_n; s, c_n x_{s,n}) e_s \right]^2 \\ &= \frac{\sigma_e^2}{nv_1(n)^2 v_2(c_n)^2} \sum_{s=1}^n E R^2(n, c_n; s, c_n x_{s,n}) \\ &\leq \frac{2\sigma_e^2 a^2(c_n)}{nv_1(n)^2 v_2(c_n)^2} \sum_{s=1}^n A_n^2\left(\frac{s}{n}\right) E[P(x_{s,n})]^2 \\ &\quad + \frac{2\sigma_e^2 b^2(n)}{nv_1(n)^2 v_2(c_n)^2} \sum_{s=1}^n q^2\left(\frac{s}{n}\right) E[B_{c_n}(x_{s,n})]^2 \\ &:= \Pi_{41} + \Pi_{42}. \end{aligned}$$

In a similar way to the convergence of Π_1 and Π_2 , we can show that Π_{41} and Π_{42} converge to zero. This finishes the proof. \square

PROOF OF THEOREM 4.1. We first prove (4.11). By virtue of (4.6) and $\alpha'B = \alpha'$, we may write

$$\begin{aligned}
& \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max}} \|A(\tau, x)\|^2} (\widehat{m}(\tau, x) - m(\tau, x)) \\
&= \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max}} \|A(\tau, x)\|^2} \\
&\quad \times [A'(\tau, x)(X'X)^{-1}X'(\delta + \gamma + \varepsilon) - \delta(\tau, x) - \gamma(\tau, x)] \\
&= \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}\sqrt{p_{\max}}} \alpha' X'(\delta + \gamma + \varepsilon) \\
&\quad - \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max}} \|A(\tau, x)\|^2} [\delta(\tau, x) + \gamma(\tau, x)] \\
&:= \sum_{i=1}^3 \Pi_i - \Pi_4,
\end{aligned}$$

where

$$\begin{aligned}
\Pi_1 &= \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}\sqrt{p_{\max}}} \alpha' X' \delta, & \Pi_2 &= \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}\sqrt{p_{\max}}} \alpha' X' \gamma, \\
\Pi_3 &= \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}\sqrt{p_{\max}}} \alpha' X' \varepsilon, & \Pi_4 &= \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max}} \|A(\tau, x)\|^2} [\delta(\tau, x) + \gamma(\tau, x)].
\end{aligned}$$

We are about to show that $\Pi_i \rightarrow_P 0$, $i = 1, 2, 4$, and Π_3 converges to the desired variable in distribution as $n \rightarrow \infty$.

Firstly, it follows from (4.9) that

$$\begin{aligned}
\Pi_3 &= \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}} (\widetilde{\mathbf{F}}' - \widetilde{\delta}' - \widetilde{\gamma}') \varepsilon = \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}} (\widetilde{\mathbf{F}}' \varepsilon - \widetilde{\delta}' \varepsilon - \widetilde{\gamma}' \varepsilon) \\
&= \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}} \left(\sum_{s=1}^n \widetilde{F}(s, X_s) e_s + \sum_{s=1}^n \widetilde{\delta}_s e_s + \sum_{s=1}^n \widetilde{\gamma}_s e_s \right) \\
&= \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}} \left(\sum_{s=1}^n \widetilde{F}(s, s\mu + \sqrt{n}\sigma_z x_{s,n}) e_s + \sum_{s=1}^n \widetilde{\delta}_s e_s + \sum_{s=1}^n \widetilde{\gamma}_s e_s \right) \\
&= \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}} \sum_{s=1}^n F(s, \sqrt{n}\sigma_z x_{s,n}) e_s + \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}} \sum_{s=1}^n \widetilde{\delta}_s e_s + \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}} \sum_{s=1}^n \widetilde{\gamma}_s e_s \\
&:= \Pi_{31} + \Pi_{32} + \Pi_{33}.
\end{aligned}$$

In view of the third part of Theorem 3.2 with $c_n = \sqrt{n}\sigma_z$ and Assumption 4.4 (a), (b), we have

$$\Pi_{31} \rightarrow_D \left(\int_0^1 G_3(t) dL_W(t, 0) \right)^{\frac{1}{2}} N, \quad (\text{D.10})$$

where $G_3(t) = \int f(t, x)^2 dx$, W is a standard Brownian motion on $[0, 1]$, N is a standard normal random variable independent of W , and L_W is the local-time process of W .

Meanwhile, the martingale difference structure $(e_s, \mathcal{F}_{n,s})$ and the adaptivity of $x_{s+1,n}$ with $\mathcal{F}_{n,s}$ yield

$$E(\Pi_{32})^2 = \frac{\sigma_z \sigma_e^2}{\sqrt{nv(n)^2}} \sum_{s=1}^n E \tilde{\delta}_s^2 \quad \text{and} \quad E(\Pi_{33})^2 = \frac{\sigma_z \sigma_e^2}{\sqrt{nv(n)^2}} \sum_{s=1}^n E \tilde{\gamma}_s^2. \quad (\text{D.11})$$

As for the first part of (D.11), using the expression of $\tilde{\delta}_s$ we have

$$\begin{aligned} E(\Pi_{32})^2 &= \frac{\sigma_z \sigma_e^2}{\sqrt{nv(n)^2}} \sum_{s=1}^n E \tilde{\delta}_s^2 \\ &= \frac{\sigma_z \sigma_e^2}{\sqrt{nv(n)^2}} \sum_{s=1}^n E \left(\sum_{i=0}^k \sum_{j=p_i+1}^{\infty} a_{ij} \mathcal{L}_j(s) Q_i(s, X_s) \right)^2 \\ &= \frac{\sigma_z \sigma_e^2}{\sqrt{nv(n)^2}} \sum_{s=1}^n \sum_{i=0}^k \left(\sum_{j=p_i+1}^{\infty} a_{ij} \mathcal{L}_j(s) \right)^2 \\ &\leq \frac{\sigma_z \sigma_e^2}{\sqrt{nv(n)^2}} \sum_{s=1}^n \sum_{i=0}^k \left(\max_{j \geq p_i} |\mathcal{L}_j(s)| \right)^2 \left(\sum_{j=p_i+1}^{\infty} |a_{ij}| \right)^2 \\ &\leq o(1) \frac{\sigma_z \sigma_e^2}{\sqrt{nv(n)^2}} \sum_{s=1}^n \frac{1}{\sqrt{s}} \sum_{i=0}^k \frac{1}{\sqrt{p_i}} \frac{1}{p_i^2} \leq o(1) \frac{\sigma_z \sigma_e^2}{v(n)^2} \frac{k}{p_{\min}^{5/2}} \\ &= o(1) \frac{\sigma_z \sigma_e^2}{v(n)^2} n^{\kappa_1 - \frac{5}{2}\kappa_2} \rightarrow 0, \end{aligned}$$

as $n \rightarrow \infty$ due to Assumptions 4.2 and 4.3, and the upper bound of $|\mathcal{L}_j(t)|$ in (C.3). Thus, $\Pi_{32} \rightarrow_P 0$.

Similarly, using the expression of $\tilde{\gamma}_s$ we have

$$\begin{aligned} E(\Pi_{33})^2 &= \frac{\sigma_z \sigma_e^2}{\sqrt{nv(n)^2}} \sum_{s=1}^n E \tilde{\gamma}_s^2 \\ &= \frac{\sigma_z \sigma_e^2}{\sqrt{nv(n)^2}} \sum_{s=1}^n E \left(\sum_{i=k+1}^{\infty} \sum_{j=0}^{\infty} a_{ij} \mathcal{L}_j(s) Q_i(s, X_s) \right)^2 \end{aligned}$$

$$\begin{aligned}
&= \frac{\sigma_z \sigma_e^2}{\sqrt{nv(n)^2}} \sum_{s=1}^n \sum_{i=k+1}^{\infty} \left(\sum_{j=0}^{\infty} a_{ij} \mathcal{L}_j(s) \right)^2 \\
&\leq C^2 \frac{\sigma_z \sigma_e^2}{\sqrt{nv(n)^2}} \sum_{s=1}^n \sum_{i=k+1}^{\infty} \left(\sum_{j=0}^{\infty} |a_{ij}| \right)^2 \\
&\leq o(1) \frac{\sigma_z \sigma_e^2}{\sqrt{nv(n)^2}} \sum_{s=1}^n \frac{1}{k} = o(1) \frac{\sigma_z \sigma_e^2}{v(n)^2} \frac{\sqrt{n}}{k} \\
&= o(1) \frac{\sigma_z \sigma_e^2}{v(n)^2} n^{\frac{1}{2} - \kappa_1} \rightarrow 0,
\end{aligned}$$

due to Assumption 4.2 and 4.3, and simply using $|\mathcal{L}_j(s)| \leq C$ for any j . Hence, $\Pi_{33} \rightarrow_P 0$.

Whence, we obtain

$$\Pi_3 \rightarrow_D \left(\int_0^1 G_3(t) dL_W(t, 0) \right)^{\frac{1}{2}} N. \quad (\text{D.12})$$

Secondly, for Π_1 and Π_2 , from (4.9) it follows that

$$\begin{aligned}
\Pi_1 &= \frac{\sqrt{\sigma_z}}{\sqrt{nv(n)} \sqrt{p_{\max}}} \alpha' X' \delta = \frac{\sqrt{\sigma_z}}{\sqrt{nv(n)}} (\tilde{\mathbf{F}}' \delta - \tilde{\delta}' \delta - \tilde{\gamma}' \delta), \\
\Pi_2 &= \frac{\sqrt{\sigma_z}}{\sqrt{nv(n)} \sqrt{p_{\max}}} \alpha' X' \gamma = \frac{\sqrt{\sigma_z}}{\sqrt{nv(n)}} (\tilde{\mathbf{F}}' \gamma - \tilde{\delta}' \gamma - \tilde{\gamma}' \gamma).
\end{aligned}$$

Thus, by Cauchy-Schwarz inequality, in order to obtain $\Pi_1 \rightarrow_P 0$ and $\Pi_2 \rightarrow_P 0$, we only need to show $\|\delta\| \rightarrow_P 0$ and $\|\gamma\| \rightarrow_P 0$ since the convergence of (D.11) indicates that

$$\frac{1}{\sqrt{n}} \|\tilde{\delta}\|^2 \rightarrow_P 0, \quad \text{and} \quad \frac{1}{\sqrt{n}} \|\tilde{\gamma}\|^2 \rightarrow_P 0, \quad (\text{D.13})$$

and because of Theorem 3.2 and Assumption 4.4 (b), we have

$$\begin{aligned}
\frac{\sigma_z}{\sqrt{nv(n)^2}} \|\tilde{\mathbf{F}}'\|^2 &= \frac{\sigma_z}{\sqrt{nv(n)^2}} \sum_{s=1}^n \tilde{F}^2(s, X_s) \\
&= \frac{\sigma_z}{\sqrt{nv(n)^2}} \sum_{s=1}^n F^2(s, \sqrt{n} \sigma_x x_{s,n}) \rightarrow_D \int_0^1 \int_{-\infty}^{\infty} f^2(r, x) dx dL_W(r, 0).
\end{aligned} \quad (\text{D.14})$$

In fact, by orthogonality of Q_i and C.2 in Lemma C.1 with $r = 3$,

$$\begin{aligned}
E \|\delta\|^2 &= E \sum_{s=1}^n \delta_s^2 = E \sum_{s=1}^n \left(\sum_{i=0}^k \sum_{j=p_i+1}^{\infty} c_{ij} \mathcal{L}_j(s) Q_i(s, X_s) \right)^2 \\
&= \sum_{s=1}^n \sum_{i=0}^k \left(\sum_{j=p_i+1}^{\infty} c_{ij} \mathcal{L}_j(s) \right)^2 = \sum_{s=1}^n \sum_{i=0}^k \left(c_i(t, m) - \sum_{j=0}^{p_i+1} c_{ij} \mathcal{L}_j(s) \right)^2
\end{aligned}$$

$$\begin{aligned}
&\leq \sum_{s=1}^n \sum_{i=0}^k o(1) \frac{1}{\sqrt{sp_i}} \frac{1}{p_i^2} \leq o(1) \frac{k}{p_{\min}^{\frac{5}{2}}} \sum_{s=1}^n \frac{1}{\sqrt{s}} \\
&= o(1) n^{\frac{1}{2} + \kappa_1 - \frac{5}{2} \kappa_2} \rightarrow 0
\end{aligned}$$

as $n \rightarrow \infty$, where once again we have used the upper bound for $|\mathcal{L}_j(s)|$ for j sufficient large and Assumption 4.3.

In addition, in view of (D.3), $c_i(t, m) = \sqrt{\psi(t)^r} \sqrt{\frac{(i-r)!}{i!}} c_i(t, D^r m)$, it follows from Assumption 4.1 (a) and (c) with $r = 3$ that

$$\begin{aligned}
E\|\gamma\|^2 &= E \sum_{s=1}^n \gamma_s^2 = E \sum_{s=1}^n \left(\sum_{i=k+1}^{\infty} c_i(s, m) Q_i(s, X_s) \right)^2 \\
&= \sum_{s=1}^n \sum_{i=k+1}^{\infty} (c_i(s, m))^2 = \sum_{s=1}^n \sum_{i=k+1}^{\infty} \psi(s)^3 \frac{(i-3)!}{i!} c_i(s, D^3 m)^2 \\
&= \sum_{i=k+1}^{\infty} \frac{(i-3)!}{i!} \sum_{s=1}^n \psi(s)^3 c_i(s, D^3 m)^2 \leq An \frac{1}{k^2} (1 + o(1)) \rightarrow 0,
\end{aligned}$$

where A is the uniform bound of $\psi(s)^3 c_i(s, D^3 m)^2$ on account of Assumption 4.1 and we have used Assumption 4.3.

Now, we are ready to prove that $\Pi_4 \rightarrow_P 0$ as $n \rightarrow \infty$. We may rewrite

$$\begin{aligned}
\Pi_4 &= \frac{\sqrt{\sigma_z}}{\sqrt[4]{nv(n)}} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max}} \|A(\tau, x)\|^2} [\delta(\tau, x) + \gamma(\tau, x)] \\
&= \frac{\sqrt{\sigma_z}}{\sqrt{nv(n)} \varrho(n)} \frac{\alpha' X' X A(\tau, x)}{p_{\max} \|A(\tau, x)\|} \times \frac{\sqrt[4]{n} \sqrt{p_{\max}} \varrho(n)}{\|A(\tau, x)\|} [\delta(\tau, x) + \gamma(\tau, x)] \\
&:= \Pi_{41} \times (\Pi_{42} + \Pi_{43}).
\end{aligned}$$

We shall demonstrate that Π_{41} converges to a random variable in distribution and both Π_{42} and Π_{43} approach to zero as $n \rightarrow \infty$, which guarantee that $\Pi_4 \rightarrow_P 0$ (see Example 25.7 on Billingsley (1995, p332). We do not mention this in what follows). To begin with, it follows from (4.9) and (4.10) that

$$\begin{aligned}
\Pi_{41} &= \frac{\sqrt{\sigma_z}}{\sqrt{nv(n)} \varrho(n)} (\tilde{\mathbf{F}}' - \tilde{\delta}' - \tilde{\gamma}') (\tilde{\mathbf{G}} - \tilde{\delta} - \tilde{\gamma}) \\
&= \frac{\sqrt{\sigma_z}}{\sqrt{nv(n)} \varrho(n)} (\tilde{\mathbf{F}}' \tilde{\mathbf{G}} - \tilde{\delta}' \tilde{\mathbf{G}} - \tilde{\gamma}' \tilde{\mathbf{G}} - \tilde{\mathbf{F}}' \tilde{\delta} - \tilde{\mathbf{F}}' \tilde{\gamma} + \|\tilde{\delta}\|^2 + \|\tilde{\gamma}\|^2 + 2\tilde{\gamma}' \tilde{\delta}).
\end{aligned}$$

Therefore, by virtue of (D.13) and (D.14), to find out the limit of Π_{41} , our remaining task is to prove the convergence of $\frac{\sigma_z}{\sqrt{n} \varrho(n)^2} \|\tilde{\mathbf{G}}\|^2$ and $\frac{\sigma_z}{\sqrt{nv(n)} \varrho(n)} \tilde{\mathbf{F}}' \tilde{\mathbf{G}}$ due to Cauchy-Schwarz inequality.

In fact, similarly to (D.14), due to Assumption 4.4(b), It follows from Theorem 3.2 that

$$\begin{aligned}\frac{\sigma_z}{\sqrt{nv(n)^2}} \|\tilde{\mathbf{G}}\|^2 &\rightarrow_D \int_0^1 \int_{-\infty}^{\infty} g^2(t, x) dx dL_W(t, 0), \\ \frac{\sigma_z}{\sqrt{nv(n)\varrho(n)}} \tilde{\mathbf{F}}' \tilde{\mathbf{G}} &\rightarrow_D \int_0^1 \int_{-\infty}^{\infty} f(t, x) g(t, x) dx dL_W(t, 0).\end{aligned}$$

Whence, $\Pi_{41} \rightarrow_D \int_0^1 \int_{-\infty}^{\infty} f(t, x) g(t, x) dx dL_W(t, 0)$ as $n \rightarrow \infty$.

For the vanish of Π_{42} and Π_{43} , we first estimate $\|A(\tau, x)\|$:

$$\|A(\tau, x)\|^2 = \sum_{i=0}^k \sum_{j=0}^{p_i} \mathcal{L}_j^2(\tau) Q_i^2(\tau, x) = e^{-\tau} \sum_{i=0}^k Q_i^2(\tau, x) \sum_{j=0}^{p_i} L_j^2(\tau),$$

which leads to $O(1)kp_{\min} \leq \|A(\tau, x)\|^2 \leq O(1)kp_{\max}$ where we have invoked the fact that $\sum_{i=0}^k H_i^2(x) = O(1)k$ uniformly in x for any orthogonal polynomial $H_i(x)$ on any compact interval (see Alexits, 1961, p295).

Accordingly, due to Assumption 4.1(b), using in C.2 in Lemma C.1 with $r = 3$ and the bound for \mathcal{L}_j in (C.3) gives

$$\begin{aligned}|\Pi_{42}| &= \frac{\sqrt[4]{n}\sqrt{p_{\max}\varrho(n)}}{\|A(\tau, x)\|} |\delta(\tau, x)| = \frac{\sqrt[4]{n}\sqrt{p_{\max}\varrho(n)}}{\|A(\tau, x)\|} \left| \sum_{i=0}^k \sum_{j=p_i+1}^{\infty} c_{ij} \mathcal{L}_j(\tau) Q_i(\tau, x) \right| \\ &\leq \frac{\sqrt[4]{n}\sqrt{p_{\max}\varrho(n)}}{\|A(\tau, x)\|} \sum_{i=0}^k |Q_i(\tau, x)| \left| \sum_{j=p_i+1}^{\infty} c_{ij} \mathcal{L}_j(\tau) \right| \\ &\leq \frac{\sqrt[4]{n}\sqrt{p_{\max}\varrho(n)}}{\|A(\tau, x)\|} \left(\sum_{i=0}^k Q_i^2(\tau, x) \right)^{\frac{1}{2}} \left[\sum_{i=0}^k \left(\sum_{j=p_i+1}^{\infty} c_{ij} \mathcal{L}_j(\tau) \right)^2 \right]^{\frac{1}{2}} \\ &\leq O(1) \frac{\sqrt[4]{n}\sqrt{p_{\max}n^\iota}}{\sqrt{kp_{\min}}} \sqrt{k} \left[\sum_{i=0}^k \frac{1}{\sqrt{\tau} p_i} \frac{o(1)}{p_i^2} \right]^{\frac{1}{2}} \leq o(1) \frac{n^{\frac{1}{4}+\iota} \sqrt{p_{\max}k}^{\frac{1}{2}}}{\sqrt{p_{\min}p_{\min}^{\frac{5}{4}}}} \\ &= o(1) n^{\frac{1}{4}+\iota+\frac{1}{2}\kappa_1+\frac{1}{2}(\bar{\kappa}_2-\kappa_2)-\frac{5}{4}\kappa_2} \rightarrow 0,\end{aligned}$$

where we have used Assumption 4.4 (a) for the parameters involved.

Meanwhile, on account of Assumption 6.1 using (D.3) with $r = 3$ and the estimation for Q_i in Remark A.4, we have

$$|\Pi_{43}| = \frac{\sqrt[4]{n}\sqrt{p_{\max}\varrho(n)}}{\|A(\tau, x)\|} |\gamma(\tau, x)| = \frac{\sqrt[4]{n}\sqrt{p_{\max}\varrho(n)}}{\|A(\tau, x)\|} \left| \sum_{i=k+1}^{\infty} c_i(\tau, m) Q_i(\tau, x) \right|$$

$$\begin{aligned}
&= \frac{\sqrt[4]{n}\sqrt{p_{\max}}\varrho(n)}{\|A(\tau, x)\|} \left| \sum_{i=k+1}^{\infty} \frac{\sqrt{\psi(\tau)}^3}{\sqrt{i(i-1)(i-2)}} c_i(\tau, D^3 m) Q_i(\tau, x) \right| \\
&\leq o(1) \frac{\sqrt[4]{n}\sqrt{p_{\max}}\varrho(n)}{\|A(\tau, x)\|} \left(\sum_{i=k+1}^{\infty} \frac{1}{i(i-1)(i-2)} Q_i(\tau, x)^2 \right)^{\frac{1}{2}} \\
&\leq o(1) \frac{\sqrt[4]{n}n^t \sqrt{p_{\max}}}{\sqrt{k}p_{\min}} \left(\sum_{i=k+1}^{\infty} \frac{1}{i(i-1)(i-2)\sqrt{i}} \right)^{\frac{1}{2}} \\
&\leq o(1) \frac{\sqrt[4]{n}n^t \sqrt{p_{\max}}}{\sqrt{k}p_{\min}} \frac{1}{k^{5/4}} = o(1)n^{\frac{1}{4}+\iota+\frac{1}{2}(\bar{\kappa}_2-\kappa_2)-\frac{7}{4}\kappa_1} \rightarrow 0,
\end{aligned}$$

because of Assumption 4.4 (a). The proof of (4.11) is completed.

We are now in a position to prove (4.12). In view of (4.6) and $\alpha' B = \alpha'$,

$$\begin{aligned}
&\frac{1}{\sqrt{n}v_1(n)v_2(\sqrt{n}\sigma_z)} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max}}\|A(\tau, x)\|^2} (\hat{m}(\tau, x) - m(\tau, x)) \\
&= \frac{1}{\sqrt{n}v_1(n)v_2(\sqrt{n}\sigma_z)} \frac{\alpha' X'}{\sqrt{p_{\max}}} (\delta + \gamma + \varepsilon) \\
&\quad - \frac{1}{\sqrt{n}v_1(n)v_2(\sqrt{n}\sigma_z)} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max}}\|A(\tau, x)\|^2} [\delta(\tau, x) + \gamma(\tau, x)] \\
&:= \sum_{i=1}^3 \Gamma_i - \Gamma_4,
\end{aligned}$$

where we by Γ_i ($i = 1, 2, 3, 4$) signify that

$$\begin{aligned}
\Gamma_1 &= \frac{1}{\sqrt{n}v_1(n)v_2(\sqrt{n}\sigma_z)} \frac{\alpha' X'}{\sqrt{p_{\max}}} \delta = \frac{1}{\sqrt{n}v_1(n)v_2(\sqrt{n}\sigma_z)} (\tilde{\mathbf{F}}' - \tilde{\delta}' - \tilde{\gamma}') \delta, \\
\Gamma_2 &= \frac{1}{\sqrt{n}v_1(n)v_2(\sqrt{n}\sigma_z)} \frac{\alpha' X'}{\sqrt{p_{\max}}} \gamma = \frac{1}{\sqrt{n}v_1(n)v_2(\sqrt{n}\sigma_z)} (\tilde{\mathbf{F}}' - \tilde{\delta}' - \tilde{\gamma}') \gamma, \\
\Gamma_3 &= \frac{1}{\sqrt{n}v_1(n)v_2(\sqrt{n}\sigma_z)} \frac{\alpha' X'}{\sqrt{p_{\max}}} \varepsilon = \frac{1}{\sqrt{n}v_1(n)v_2(\sqrt{n}\sigma_z)} (\tilde{\mathbf{F}}' - \tilde{\delta}' - \tilde{\gamma}') \varepsilon, \\
\Gamma_4 &= \frac{1}{\sqrt{n}v_1(n)v_2(\sqrt{n}\sigma_z)} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max}}\|A(\tau, x)\|^2} [\delta(\tau, x) + \gamma(\tau, x)].
\end{aligned}$$

Observe that because of Assumption 4.4(c) and (d), we have

$$\begin{aligned}
&\frac{1}{\sqrt{n}v_1(n)v_2(\sqrt{n}\sigma_z)} \tilde{\mathbf{F}}' \varepsilon = \frac{1}{\sqrt{n}v_1(n)v_2(\sqrt{n}\sigma_z)} \sum_{s=1}^n \tilde{F}(s, X_s) e_s \\
&= \frac{1}{\sqrt{n}v_1(n)v_2(\sqrt{n}\sigma_z)} \sum_{s=1}^n F(s, \sqrt{n}\sigma_z x_{s,n}) e_s \rightarrow_D \int_0^1 f(r, W(r)) dU(r),
\end{aligned} \tag{D.15}$$

and

$$\begin{aligned} \frac{1}{nv_1(n)^2 v_2(\sqrt{n}\sigma_z)^2} \|\tilde{\mathbf{F}}\|^2 &= \frac{1}{nv_1(n)^2 v_2(\sqrt{n}\sigma_z)^2} \sum_{s=1}^n \tilde{F}(s, X_s)^2 \\ &= \frac{1}{nv_1(n)^2 v_2(\sqrt{n}\sigma_z)^2} \sum_{s=1}^n F(s, \sqrt{n}\sigma_z x_{s,n})^2 \rightarrow_D \int_0^1 f^2(r, W(r)) dr, \end{aligned} \quad (\text{D.16})$$

as $n \rightarrow \infty$ by Theorem 3.4.

Note that in first part we have shown that

$$\frac{1}{\sqrt{n}} \|\tilde{\delta}'\|^2 \rightarrow_P 0, \quad \frac{1}{\sqrt{n}} \|\tilde{\gamma}'\|^2 \rightarrow_P 0, \quad \|\delta\|^2 \rightarrow_P 0, \quad (\text{D.17a})$$

$$\|\gamma\|^2 \rightarrow_P 0, \quad \frac{1}{\sqrt[4]{n}} \tilde{\delta}' \varepsilon \rightarrow_P 0, \quad \frac{1}{\sqrt[4]{n}} \tilde{\gamma}' \varepsilon \rightarrow_P 0. \quad (\text{D.17b})$$

All results in (D.17) remain true since all conditions for δ , γ , $\tilde{\delta}$, $\tilde{\gamma}$ and ε have not changed. Therefore, (D.15), (D.16) and (D.17) imply that $\Gamma_1 \rightarrow_P 0$ and $\Gamma_2 \rightarrow_P 0$, as well as $\Gamma_3 \rightarrow_D \int_0^1 f(r, W(r)) dU(r)$ as $n \rightarrow \infty$. Thus, our remaining task is to prove $\Gamma_4 \rightarrow_P 0$ as $n \rightarrow \infty$.

To this end, let us rewrite

$$\begin{aligned} \Gamma_4 &= \frac{1}{\sqrt{n} v_1(n) v_2(\sqrt{n}\sigma_z)} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max}} \|A(\tau, x)\|^2} [\delta(\tau, x) + \gamma(\tau, x)] \\ &:= \Gamma_{41} \times (\Gamma_{42} + \Gamma_{43}), \end{aligned}$$

where we denote

$$\begin{aligned} \Gamma_{41} &= \frac{1}{nv_1(n) v_2(\sqrt{n}\sigma_z) \varrho_1(n) \varrho_2(\sqrt{n}\sigma_z)} \frac{\alpha' X' X A(\tau, x)}{p_{\max} \|A(\tau, x)\|}, \\ \Gamma_{42} &= \frac{\sqrt{n} \sqrt{p_{\max}} \varrho_1(n) \varrho_2(\sqrt{n}\sigma_z)}{\|A(\tau, x)\|} \delta(\tau, x), \quad \text{and} \\ \Gamma_{43} &= \frac{\sqrt{n} \sqrt{p_{\max}} \varrho_1(n) \varrho_2(\sqrt{n}\sigma_z)}{\|A(\tau, x)\|} \gamma(\tau, x). \end{aligned}$$

It follows from (4.9) and (4.10) that

$$\begin{aligned} \Gamma_{41} &= \frac{1}{nv_1(n) v_2(\sqrt{n}\sigma_z) \varrho_1(n) \varrho_2(\sqrt{n}\sigma_z)} (\tilde{\mathbf{F}}' - \tilde{\delta}' - \tilde{\gamma}') (\tilde{\mathbf{G}}' - \tilde{\delta}' - \tilde{\gamma}') \\ &= \frac{1}{nv_1(n) v_2(\sqrt{n}\sigma_z) \varrho_1(n) \varrho_2(\sqrt{n}\sigma_z)} \\ &\quad \times (\tilde{\mathbf{F}}' \tilde{\mathbf{G}} - \tilde{\delta}' \tilde{\mathbf{G}} - \tilde{\gamma}' \tilde{\mathbf{G}} - \tilde{\mathbf{F}}' \tilde{\delta} - \tilde{\mathbf{F}}' \tilde{\gamma} + \|\tilde{\delta}\|^2 + \|\tilde{\gamma}\|^2 + 2\tilde{\gamma}' \tilde{\delta}). \end{aligned}$$

Once again, due to Theorem 3.4 and Assumption 4.4 (d), similar to (D.16), we have

$$\begin{aligned} \frac{1}{n\varrho_1(n)^2\varrho_2(\sqrt{n}\sigma_z)^2}\|\tilde{\mathbf{G}}\|^2 &\rightarrow_D \int_0^1 g^2(r, W(r))dr, \\ \frac{1}{nv_1(n)v_2(\sqrt{n}\sigma_z)\varrho_1(n)\varrho_2(\sqrt{n}\sigma_z)}\tilde{\mathbf{F}}'\tilde{\mathbf{G}} &\rightarrow_D \int_0^1 f(r, W(r))g(r, W(r))dr. \end{aligned} \quad (\text{D.18})$$

Thus, Cauchy-Schwarz inequality as well as (D.15), (D.16), (D.17) and (D.18) suggest that Γ_{41} converges to the last limit in distribution.

We are ready to prove both $\Gamma_{42} \rightarrow 0$ and $\Gamma_{43} \rightarrow 0$, as $n \rightarrow \infty$. Because $\delta(\tau, x)$ and $\gamma(\tau, x)$ remain the same as in the first part, similar to Π_{42} and Π_{43} , we have

$$\begin{aligned} |\Gamma_{42}| &= \frac{\sqrt{n}\sqrt{p_{\max}}\varrho_1(n)\varrho_2(\sqrt{n}\sigma_z)}{\|A(\tau, x)\|} |\delta(\tau, x)| \\ &\leq o(1) \frac{\sqrt{n}\sqrt{p_{\max}}n^{\iota_1}n^{\frac{1}{2}\iota_2}}{\sqrt{k}p_{\min}} \sqrt{k} \left[\sum_{i=0}^k \frac{1}{\sqrt{\tau}p_i} \frac{o(1)}{p_i^2} \right]^{\frac{1}{2}} \\ &\leq o(1) \sqrt{n}n^{\frac{1}{2}(\bar{\kappa}_2 - \kappa_2)} n^{\iota_1 + \frac{1}{2}\iota_2} \sqrt{k} p_{\min}^{-\frac{5}{4}} \\ &= o(1) n^{\frac{1}{2} + \frac{1}{2}(\bar{\kappa}_2 - \kappa_2) + \iota_1 + \frac{1}{2}\iota_2 + \frac{1}{2}\kappa_1 - \frac{5}{4}\kappa_2} \rightarrow 0, \\ |\Gamma_{43}| &= \frac{\sqrt{n}\sqrt{p_{\max}}\varrho_1(n)\varrho_2(\sqrt{n}\sigma_z)}{\|A(\tau, x)\|} |\gamma(\tau, x)| \\ &\leq o(1) \frac{\sqrt{n}\sqrt{p_{\max}}n^{\iota_1}n^{\frac{1}{2}\iota_2}}{\sqrt{k}p_{\min}} \frac{1}{k^{5/4}} \\ &= o(1) n^{\frac{1}{2} + \frac{1}{2}(\bar{\kappa}_2 - \kappa_2) + \iota_1 + \frac{1}{2}\iota_2 - \frac{7}{4}\kappa_1} \rightarrow 0, \end{aligned}$$

due to Assumption 4.4 (c). This finishes the proof. \square

PROOF OF THEOREM 4.2. We here only give an outline, as it is similar to the proof of Theorem 4.1. It follows from (4.16) that

$$\begin{aligned} &\frac{1}{\sqrt{n}} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max}} \|A(\tau, x)\|^2} (\hat{m}(\tau, x) - m(\tau, x)) \\ &= \frac{1}{\sqrt{n}\sqrt{p_{\max}}} \alpha' X' (\delta + \gamma + \varepsilon) - \frac{1}{\sqrt{n}} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max}} \|A(\tau, x)\|^2} [\delta(\tau, x) + \gamma(\tau, x)] \\ &:= \Pi_1 - \Pi_2. \end{aligned}$$

We shall show that Π_1 converges to the desired variable in distribution and $\Pi_2 \rightarrow_P 0$ as $n \rightarrow \infty$.

In what follows, we shall adopt the embedding schedule that allows us to work on a stronger condition: $(W_n, U_n) \rightarrow (W, U)$ almost surely but still gets a weak convergence.

In view of (4.20), Π_1 can be rephrased as $\Pi_1 = \frac{1}{\sqrt{n}}(\mathbf{F}' - \tilde{\delta}' - \tilde{\gamma}')(\delta + \gamma + \varepsilon)$. In order to complete the convergence of Π_1 , we are about to demonstrate that

$$\frac{1}{\sqrt{n}}\mathbf{F}'\varepsilon \rightarrow_P \int_0^1 F(Tr, T\mu r + \sqrt{T}\sigma_z W(r))dU(r), \quad (\text{D.19})$$

and all the other terms in Π_1 converge to zero in probability.

In fact, (D.19) is valid because from Assumption B it follows that

$$\begin{aligned} \frac{1}{\sqrt{n}}\mathbf{F}'\varepsilon &= \frac{1}{\sqrt{n}} \sum_{s=1}^n F(t_{s,n}, X_{s,n})e_s \\ &= \frac{1}{\sqrt{n}} \sum_{s=1}^n F\left(\frac{s}{n}T, \frac{s}{n}T\mu + \sqrt{T}\sigma_z x_{s,n}\right) e_s \\ &= \sum_{s=1}^n F\left(\frac{s}{n}T, \frac{s}{n}T\mu + \sqrt{T}\sigma_z W_n\left(\frac{s}{n}\right)\right) \left(U_n\left(\frac{s}{n}\right) - U_n\left(\frac{s-1}{n}\right)\right) \\ &= \int_0^1 F(rT + o(1), rT\mu + o(1) + \sqrt{T}\sigma_z W_n(r + o(1)))dU_n(r), \end{aligned}$$

and since $(W_n(r + o(1)), U_n(r)) \rightarrow_{a.s.} (W(r), U(r))$, it follows from the continuity of $F(\cdot, \cdot)$ that

$$\begin{aligned} &(F(rT + o(1), rT\mu + o(1) + \sqrt{T}\sigma_x W_n(r + o(1))), U_n(r)) \\ &\rightarrow_{a.s.} (F(rT, rT\mu + \sqrt{T}\sigma_x W(r)), U(r)). \end{aligned}$$

Invoking Theorem 2.2 in Kurtz and Protter (1991) yields the result.

In addition, using the martingale difference structure of $(e_s, \mathcal{F}_{n,s})$ and adaptivity of $x_{s+1,n}$ with $\mathcal{F}_{n,s}$ yields

$$\frac{1}{n}E(\tilde{\delta}'\varepsilon)^2 = \sigma_e^2 \frac{1}{n}E\|\tilde{\delta}'\|^2 \quad \text{and} \quad \frac{1}{n}E(\tilde{\gamma}'\varepsilon)^2 = \sigma_e^2 \frac{1}{n}E\|\tilde{\gamma}'\|^2.$$

Therefore, by virtue of Cauchy-Schwarz inequality, in order to prove the convergence of other terms in Π_1 it is sufficient to show that as $n \rightarrow \infty$,

$$\|\delta\|^2 \rightarrow_P 0, \quad \|\gamma\|^2 \rightarrow_P 0, \quad \frac{1}{n}E\|\tilde{\delta}'\|^2 \rightarrow 0, \quad \frac{1}{n}E\|\tilde{\gamma}'\|^2 \rightarrow 0, \quad (\text{D.20})$$

since

$$\begin{aligned} \frac{1}{n}\|\mathbf{F}'\|^2 &= \frac{1}{n} \sum_{s=1}^n F^2\left(\frac{s}{n}T, \frac{s}{n}T\mu + \sqrt{T}\sigma_z x_{s,n}\right) \\ &\rightarrow_{a.s.} \int_0^1 F^2(Tr, T\mu r + \sqrt{T}\sigma_z W(r))dr, \end{aligned} \quad (\text{D.21})$$

using the continuity of $F(\cdot, \cdot)$ and $W_n(r) \rightarrow_{a.s.} W(r)$.

Let us prove the results in (D.20) one by one. Firstly, because of Assumptions 4.5 (b), $b_i(t) := b_i(t, m)$ is differentiable up to third order, hence all expansions of $b_i(t)$, $b'_i(t)$ and $b''_i(t)$ in terms of $\varphi_{jT}(t)$ are convergent uniformly on $[0, T]$. Whence $b_{ij} = \left(\frac{T}{j\pi}\right)^2 c_j(b''_i)$ where $c_j(b''_i)$ stands for the j -th coefficient in the expansion of $b''_i(t)$. We have

$$\begin{aligned} E\|\delta\|^2 &= \sum_{s=1}^n \sum_{i=1}^k \left(\sum_{j=p_i+1}^{\infty} b_{ij} \varphi_{jT}(t_{s,n}) \right)^2 \leq \frac{2T^3}{\pi^4} \sum_{s=1}^n \sum_{i=1}^k \left(\sum_{j=p_i+1}^{\infty} \frac{1}{j^2} |c_j(b''_i)| \right)^2 \\ &\leq \frac{2T^3}{\pi^4} \sum_{s=1}^n \sum_{i=1}^k \sum_{j=p_i+1}^{\infty} \frac{1}{j^4} \sum_{j=p_i+1}^{\infty} |c_j(b''_i)|^2 \\ &\leq o(1) \frac{nk}{p_{\min}^3} = o(1) n^{1+\kappa_1-3\kappa_2} \rightarrow 0 \end{aligned}$$

as $n \rightarrow \infty$, which implies $\|\delta\|^2 \rightarrow_P 0$.

Secondly, by virtue of (D.3) with $r = 2$,

$$\begin{aligned} E\|\gamma\|^2 &= \sum_{s=1}^n E[\gamma_s^2] = \sum_{s=1}^n E \left(\sum_{i=k+1}^{\infty} b(t_{s,n}, m) Q_i(t_{s,n}, X_{s,n}) \right)^2 \\ &= \sum_{s=1}^n \sum_{i=k+1}^{\infty} b_i^2(t_{s,n}, m) = \sum_{s=1}^n \sum_{i=k+1}^{\infty} \psi(t_{s,n})^2 \frac{(i-2)!}{i!} b_i^2(t_{s,n}, D^2m) \\ &\leq \frac{1}{k^2} \sum_{s=1}^n \psi(t_{s,n})^2 \sum_{i=k+1}^{\infty} b_i^2(t_{s,n}, D^2m) \leq o(1) \frac{1}{k^2} \sum_{s=1}^n \psi(t_{s,n})^2 \\ &\leq o(1) \frac{n}{k^2} \max_{0 \leq t \leq T} \psi(t)^2 = o(1) n^{1-2\kappa_1} \rightarrow 0, \end{aligned}$$

due to Assumption 4.6, and we have invoked Assumption 4.5 that the convergence of expansion of D^2m is uniformly on $[0, T]$, $\sum_{i=k+1}^{\infty} b_i^2(t_{s,n}, D^2m) = o(1)$ independent of s ; in addition, in the scope of this study, $\psi(t) \in C[0, T]$. Therefore, it is bounded on the interval. Whence, $\|\gamma\|^2 \rightarrow_P 0$.

Thirdly, similar to the counterparts in the proof of Theorem 4.1, it follows that $\frac{1}{n} E\|\tilde{\delta}'\|^2 \rightarrow 0$ and $\frac{1}{n} E\|\tilde{\gamma}\|^2 \rightarrow 0$.

Now we are in a position to prove $\Pi_2 \rightarrow_P 0$ as $n \rightarrow \infty$. Since Π_2 can be rephrased as

$$\Pi_2 = \frac{1}{n} \frac{\alpha' X' X A(\tau, x)}{p_{\max} \|A(\tau, x)\|} \cdot \frac{\sqrt{n} \sqrt{p_{\max}}}{\|A(\tau, x)\|} [\delta(\tau, x) + \gamma(\tau, x)] := \Pi_{21} \cdot (\Pi_{22} + \Pi_{23}),$$

we shall show that Π_{21} converges to some random variable in probability and $\Pi_{22}, \Pi_{23} \rightarrow 0$.

To begin with, by virtue of (4.20) and (4.21) we have

$$\Pi_{21} = \frac{1}{n} \frac{\alpha' X' X A(\tau, x)}{p_{\max} \|A(\tau, x)\|} = \frac{1}{n} (\mathbf{F}' - \tilde{\delta}' - \tilde{\gamma}') (\mathbf{G} - \tilde{\delta} - \tilde{\gamma})$$

Noting that (D.20) and (D.21) imply that $\frac{1}{n} \mathbf{F}' \tilde{\delta} \rightarrow_P 0$, $\frac{1}{n} \mathbf{F}' \tilde{\gamma} \rightarrow_P 0$, $\frac{1}{n} \tilde{\delta}' \tilde{\gamma} \rightarrow_P 0$, $\frac{1}{n} \|\tilde{\delta}'\|^2 \rightarrow_P 0$ and $\frac{1}{n} \|\tilde{\gamma}'\|^2 \rightarrow_P 0$, it suffices to show that the convergence of $\frac{1}{n} \mathbf{F}' \mathbf{G}$ and $\frac{1}{n} \|\mathbf{G}\|^2$ as $n \rightarrow \infty$.

In effect, similar to (D.21),

$$\begin{aligned} \frac{1}{n} \mathbf{F}' \mathbf{G} &= \frac{1}{n} \sum_{s=1}^n F(t_{s,n}, X_{s,n}) G(t_{s,n}, X_{s,n}) \\ &\rightarrow_{a.s.} \int_0^1 F(Tr, T\mu r + \sqrt{T} \sigma_z W(r)) G(Tr, T\mu r + \sqrt{T} \sigma_z W(r)) dr, \\ \frac{1}{n} \|\mathbf{G}\|^2 &= \frac{1}{n} \sum_{s=1}^n G^2(t_{s,n}, X_{s,n}) \rightarrow_{a.s.} \int_0^1 G^2(Tr, T\mu r + \sqrt{T} \sigma_z W(r)) dr, \end{aligned}$$

by continuous mapping theorem as $n \rightarrow \infty$.

Therefore, as $n \rightarrow \infty$,

$$\Pi_{21} \rightarrow_P \int_0^1 F(Tr, T\mu r + \sqrt{T} \sigma_z W(r)) G(Tr, T\mu r + \sqrt{T} \sigma_z W(r)) dr.$$

Consider the convergence of Π_{22} and Π_{23} . It is known that $O(1)kp_{\min} \leq \|A(\tau, x)\|^2 \leq O(1)kp_{\max}$. It follows from Assumption 4.5 that

$$\begin{aligned} |\Pi_{22}| &= \frac{\sqrt{np_{\max}}}{\|A(\tau, x)\|} |\delta(\tau, x)| \leq \frac{\sqrt{np_{\max}}}{\|A(\tau, x)\|} \sum_{i=0}^k |Q_i(\tau, x)| \left| \sum_{j=p_i+1}^{\infty} b_{ij} \varphi_{jT}(\tau) \right| \\ &\leq O(1) \frac{\sqrt{np_{\max}}}{\sqrt{kp_{\min}}} \left(\sum_{i=0}^k |Q_i(\tau, x)|^2 \right)^{\frac{1}{2}} \left(\sum_{i=0}^k \frac{2T^3}{\pi^4} \left| \sum_{j=p_i+1}^{\infty} \frac{1}{j^2} |c_j(b_i'')| \right|^2 \right)^{\frac{1}{2}} \\ &\leq O(1) \frac{\sqrt{np_{\max}}}{\sqrt{kp_{\min}}} \sqrt{k} \left(\sum_{i=0}^k \sum_{j=p_i+1}^{\infty} \frac{1}{j^4} \sum_{j=p_i+1}^{\infty} |c_j(b_i'')|^2 \right)^{\frac{1}{2}} \\ &\leq o(1) \frac{\sqrt{np_{\max}} \sqrt{k}}{\sqrt{p_{\min} p_{\min}^{3/2}}} = o(1) n^{\frac{1}{2} + \frac{1}{2} \kappa_1 + \frac{1}{2} (\bar{\kappa}_2 - \kappa_2) - \frac{3}{2} \kappa_2} \rightarrow 0 \end{aligned}$$

as $n \rightarrow \infty$ using Assumption 4.6.

Additionally, using (D.3) with $r = 2$ and the asymptotic property of Q_i in Remark A.4 gives

$$\begin{aligned}
|\Pi_{23}| &= \frac{\sqrt{np_{\max}}}{\|A(\tau, x)\|} |\gamma(\tau, x)| = \frac{\sqrt{np_{\max}}}{\|A(\tau, x)\|} \left| \sum_{i=k+1}^{\infty} b_i(\tau, m) Q_i(\tau, x) \right| \\
&= \frac{\sqrt{np_{\max}}}{\|A(\tau, x)\|} \left| \sum_{i=k+1}^{\infty} \frac{\psi(\tau)}{\sqrt{i(i-1)}} b_i(\tau, D^2 m) Q_i(\tau, x) \right| \\
&\leq O(1) \frac{\sqrt{np_{\max}}}{\sqrt{kp_{\min}}} \left(\sum_{i=k+1}^{\infty} |b_i(\tau, D^2 m)|^2 \right)^{\frac{1}{2}} \left(\sum_{i=k+1}^{\infty} \frac{1}{i(i-1)\sqrt{i}} \right)^{\frac{1}{2}} \\
&= o(1) \frac{\sqrt{np_{\max}}}{\sqrt{kp_{\min}}} \frac{1}{k^{3/4}} = o(1) n^{\frac{1}{2} + \frac{1}{2}(\bar{\kappa}_2 - \kappa_2) - \frac{5}{4}\kappa_1} \rightarrow 0
\end{aligned}$$

as $n \rightarrow \infty$ by virtue of Assumption 4.6. The proof is finished. \square

PROOF OF THEOREM 4.3. Here is a sketch of the proof. Let us prove the first part to begin. It follows from (4.25) that

$$\begin{aligned}
&\frac{\sqrt[4]{T_n}^3 \sqrt{\sigma_z}}{\sqrt{nv(T_n)}} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max}} \|A(\tau, x)\|^2} (\hat{m}(\tau, x) - m(\tau, x)) \\
&= \frac{\sqrt[4]{T_n}^3 \sqrt{\sigma_z}}{\sqrt{nv(T_n)} \sqrt{p_{\max}}} \alpha' X' (\delta + \gamma + \varepsilon) \\
&\quad - \frac{\sqrt[4]{T_n}^3 \sqrt{\sigma_z}}{\sqrt{nv(T_n)}} \frac{\alpha' X' X A(\tau, x)}{\sqrt{p_{\max}} \|A(\tau, x)\|^2} [\delta(\tau, x) + \gamma(\tau, x)] \\
&:= \Pi_1 - \Pi_2.
\end{aligned}$$

We are about to show that Π_1 converges to the desired random variable in distribution and $\Pi_2 \rightarrow_P 0$ as $n \rightarrow \infty$.

Using (4.29) we can write

$$\Pi_1 = \frac{\sqrt[4]{T_n} \sqrt{\sigma_z}}{\sqrt{nv(T_n)}} (\tilde{\mathbf{F}}' - \tilde{\delta}' - \tilde{\gamma}') (\delta + \gamma + \varepsilon).$$

It follows from Theorem 3.2 and Assumption 4.10 (a) that

$$\begin{aligned}
\frac{\sqrt[4]{T_n} \sqrt{\sigma_z}}{\sqrt{nv(T_n)}} \tilde{\mathbf{F}}' \varepsilon &= \frac{\sqrt[4]{T_n} \sqrt{\sigma_z}}{\sqrt{nv(T_n)}} \sum_{s=1}^n \tilde{F}(t_{s,n}, X_{s,n}) e_s \\
&= \frac{\sqrt[4]{T_n} \sqrt{\sigma_z}}{\sqrt{nv(T_n)}} \sum_{s=1}^n F\left(\frac{s}{n} T_n, \sqrt{T_n} \sigma_z x_{s,n}\right) e_s \rightarrow_D \left(\int_0^1 G_3(u) dL_W(u, 0) \right)^{\frac{1}{2}} N,
\end{aligned}$$

where $G_3(\cdot) = \int f^2(\cdot, x)dx$, W is a standard Brownian motion on $[0, 1]$, N is a standard normal random variable independent of W and L_W is the local-time process of W . In addition, it follows once again from Theorem 3.2 and Assumption 4.10 (b) that

$$\frac{\sqrt{T_n}\sigma_z}{nv(T_n)^2} \|\tilde{\mathbf{F}}'\|^2 \rightarrow_D \int_0^1 \int_{-\infty}^{\infty} f^2(t, x) dx dL_W(t, 0). \quad (\text{D.22})$$

Also, since $x_{s,n}$ and e_s ($s = 1, \dots, n$) satisfy Assumption B, using martingale difference structure, we have

$$\frac{\sqrt{T_n}\sigma_z}{nv(T_n)^2} E(\tilde{\delta}'\varepsilon)^2 = \frac{\sqrt{T_n}\sigma_z\sigma_e^2}{nv(T_n)^2} E\|\tilde{\delta}'\|^2, \quad \frac{\sqrt{T_n}\sigma_z}{nv(T_n)^2} E(\tilde{\gamma}'\varepsilon)^2 = \frac{\sqrt{T_n}\sigma_z\sigma_e^2}{nv(T_n)^2} E\|\tilde{\gamma}'\|^2.$$

Therefore, in order to complete the convergence of Π_1 , by virtue of Cauchy-Schwarz inequality, it suffices to demonstrate that as $n \rightarrow \infty$,

$$\|\delta\|^2 \rightarrow_P 0, \quad \|\gamma\|^2 \rightarrow_P 0, \quad (\text{D.23a})$$

$$\frac{\sqrt{T_n}}{nv(T_n)^2} E\|\tilde{\delta}'\|^2 \rightarrow 0, \quad \frac{\sqrt{T_n}}{nv(T_n)^2} E\|\tilde{\gamma}'\|^2 \rightarrow 0. \quad (\text{D.23b})$$

To begin with the proof of (D.23a), similar to the counterpart in the proof of Theorem 4.2,

$$\begin{aligned} E\|\delta\|^2 &= \sum_{s=1}^n E \left(\sum_{i=0}^k \sum_{j=p_i+1}^{\infty} b_{ij} \varphi_{jT_n}(t_{s,n}) Q_i(t_{s,n}, X_{s,n}) \right)^2 \\ &= \sum_{s=1}^n \sum_{i=0}^k \left(\sum_{j=p_i+1}^{\infty} b_{ij} \varphi_{jT_n}(t_{s,n}) \right)^2 \leq o(1) T_n^3 \sum_{s=1}^n \sum_{i=0}^k \frac{1}{p_i^3} \\ &\leq o(1) T_n^3 \frac{nk}{p_{\min}^3} = o(1) n^{1+3\kappa_3+\kappa_1-3\kappa_2} \rightarrow 0, \end{aligned}$$

by virtue of Assumption 4.9, which in turn implies $\|\delta\|^2 \rightarrow_P 0$.

And, it follows from (D.3) with $r = 3$ that as $n \rightarrow \infty$

$$\begin{aligned} E\|\gamma\|^2 &= \sum_{s=1}^n E \left(\sum_{i=k+1}^{\infty} b_i(t_{s,n}, m) Q_i(t_{s,n}, X_{s,n}) \right)^2 \\ &= \sum_{s=1}^n \sum_{i=k+1}^{\infty} b_i^2(t_{s,n}, m) = \sum_{s=1}^n \sum_{i=k+1}^{\infty} \frac{\psi^3(t_{s,n})}{i(i-1)(i-2)} b_i^2(t_{s,n}, D^3 m) \\ &= \sum_{i=k+1}^{\infty} \frac{1}{i(i-1)(i-2)} \sum_{s=1}^n \psi^3(t_{s,n}) b_i^2(t_{s,n}, D^3 m) \end{aligned}$$

$$\leq An \sum_{i=k+1}^{\infty} \frac{1}{i(i-1)(i-2)} \leq A(1+o(1))\frac{n}{k^2} = An^{1-2\kappa_1} \rightarrow 0$$

by Assumption 4.8, 4.9, where A is the uniform bound of $\psi(t)^3 b_i^2(t, D^3 m)$.

Meanwhile, similar to its counterpart in the preceding theorem we can show that (D.23b) is true. Thus, we are now in a position to prove $\Pi_2 \rightarrow_P 0$ as $n \rightarrow \infty$.

Notice that Π_2 can be rephrased as $\Pi_2 := \Pi_{21} \times (\Pi_{22} + \Pi_{23})$, where

$$\begin{aligned} \Pi_{21} &= \frac{\sqrt{T_n} \sigma_z}{nv(T_n) \varrho(T_n) p_{\max} \|A(\tau, x)\|} \frac{T_n}{\alpha' X' X A(\tau, x)}, \\ \Pi_{22} &= \frac{\sqrt{n} \sqrt{p_{\max}} \varrho(T_n)}{\sqrt{\sigma_z} \sqrt[4]{T_n}^3 \|A(\tau, x)\|} \delta(\tau, x), \quad \Pi_{23} = \frac{\sqrt{n} \sqrt{p_{\max}} \varrho(T_n)}{\sqrt{\sigma_z} \sqrt[4]{T_n}^3 \|A(\tau, x)\|} \gamma(\tau, x). \end{aligned}$$

To complete the convergence of Π_2 , we are going to show that Π_{21} converges to some random variable in distribution, while both Π_{22} and Π_{23} are convergent to zero.

It follows from (4.29) and (4.30) that

$$\begin{aligned} \Pi_{21} &= \frac{\sqrt{T_n} \sigma_z}{nv(T_n) \varrho(T_n) p_{\max} \|A(\tau, x)\|} \alpha' X' X A(\tau, x) \\ &= \frac{\sqrt{T_n} \sigma_z}{nv(T_n) \varrho(T_n)} (\tilde{\mathbf{F}}' - \tilde{\delta}' - \tilde{\gamma}') (\tilde{\mathbf{G}} - \tilde{\delta} - \tilde{\gamma}). \end{aligned}$$

Nevertheless, (D.22) and (D.23b) as well as Cauchy-Schwarz inequality suggest that in order to obtain the limit of Π_{21} , one only needs to find the limits of

$$\frac{\sqrt{T_n} \sigma_z}{nv(T_n) \varrho(T_n)} \tilde{\mathbf{F}}' \tilde{\mathbf{G}} \quad \text{and} \quad \frac{\sqrt{T_n} \sigma_z}{n \varrho(T_n)^2} \|\tilde{\mathbf{G}}\|^2.$$

Indeed, similar to (D.22), by Theorem 3.2 and Assumption 4.10 (b),

$$\begin{aligned} \frac{\sqrt{T_n} \sigma_z}{nv(T_n) \varrho(T_n)} \tilde{\mathbf{F}}' \tilde{\mathbf{G}} &\rightarrow_D \int_0^1 \int_{-\infty}^{\infty} f(t, x) g(t, x) dx dL_W(t, 0), \\ \frac{\sqrt{T_n} \sigma_z}{n \varrho(T_n)^2} \|\tilde{\mathbf{G}}'\|^2 &\rightarrow_D \int_0^1 \int_{-\infty}^{\infty} g^2(t, x) dx dL_W(t, 0). \end{aligned}$$

Hence, we have $\Pi_{21} \rightarrow_D \int_0^1 \int_{-\infty}^{\infty} f(t, x) g(t, x) dx dL_W(t, 0)$.

As for Π_{22} , it can be deduced that $O(1)kp_{\min} \leq T_n \|A(\tau, x)\|^2 \leq O(1)kp_{\max}$ and $b_{ij} = \left(\frac{T_n}{j\pi}\right)^2 c_j(b_i'')$ where $c_j(b_i'')$ is the j -th coefficient in the expansion of b_i'' . Hence,

$$|\Pi_{22}| = \frac{\sqrt{n} \sqrt{p_{\max}} \varrho(T_n)}{\sqrt{\sigma_z} \sqrt[4]{T_n}^3 \|A(\tau, x)\|} |\delta(\tau, x)|$$

$$\begin{aligned}
&= \frac{\sqrt{n}\sqrt{p_{\max}}\varrho(T_n)}{\sqrt{\sigma_z}\sqrt[4]{T_n^3}\|A(\tau, x)\|} \left| \sum_{i=0}^k \sum_{j=p_i+1}^{\infty} b_{ij}\varphi_{jT_n}(\tau)Q_i(\tau, x) \right| \\
&\leq \frac{\sqrt{n}\sqrt{p_{\max}}\varrho(T_n)}{\sqrt{\sigma_z}\sqrt[4]{T_n^3}\|A(\tau, x)\|} \left(\sum_{i=0}^k |Q_i(\tau, x)|^2 \right)^{\frac{1}{2}} \left(\sum_{i=0}^k \left| \sum_{j=p_i+1}^{\infty} b_{ij}\varphi_{jT_n}(\tau) \right|^2 \right)^{\frac{1}{2}} \\
&\leq o(1) \frac{\sqrt{n}\sqrt{p_{\max}T_n}\varrho(T_n)}{\sqrt{k}p_{\min}\sqrt[4]{T_n^3}} \sqrt{k} \left(\sum_{i=0}^k \frac{T_n^3}{p_i^3} \right)^{\frac{1}{2}} \\
&\leq o(1) \frac{\sqrt{n}p_{\max}T_n^\iota}{\sqrt{p_{\min}}\sqrt[4]{T_n^3}} \frac{\sqrt{k}\sqrt{T_n^3}}{\sqrt{p_{\min}}^3} \\
&= o(1)n^{\frac{1}{2}+\frac{1}{2}\kappa_1+(\iota+\frac{5}{4})\kappa_3+\frac{1}{2}(\bar{\kappa}_2-\kappa_2)-\frac{3}{2}\kappa_2} \rightarrow 0,
\end{aligned}$$

as $n \rightarrow \infty$ where we have exploited the condition in Assumption 4.10 (a) for ι and truncation parameters.

Analogously, by (D.3) with $r = 2$,

$$\begin{aligned}
|\Pi_{23}| &= \frac{\sqrt{n}\sqrt{p_{\max}}\varrho(T_n)}{\sqrt{\sigma_z}\sqrt[4]{T_n^3}\|A(\tau, x)\|} |\gamma(\tau, x)| \\
&= \frac{\sqrt{n}\sqrt{p_{\max}}\varrho(T_n)}{\sqrt{\sigma_z}\sqrt[4]{T_n^3}\|A(\tau, x)\|} \left| \sum_{i=k+1}^{\infty} b_i(\tau)Q_i(\tau, x) \right| \\
&= \frac{\sqrt{n}\sqrt{p_{\max}}\varrho(T_n)}{\sqrt{\sigma_z}\sqrt[4]{T_n^3}\|A(\tau, x)\|} \left| \sum_{i=k+1}^{\infty} \frac{\psi(\tau)}{\sqrt{i(i-1)}} b_i(\tau, D^2m)Q_i(\tau, x) \right| \\
&\leq o(1) \frac{\sqrt{n}p_{\max}T_n^\iota}{\sqrt[4]{T_n}\sqrt{k}p_{\min}} \left(\sum_{i=k+1}^{\infty} \frac{1}{i(i-1)\sqrt{i}} \right)^{\frac{1}{2}} \leq o(1) \frac{\sqrt{n}p_{\max}T_n^\iota}{\sqrt[4]{T_n}\sqrt{k}p_{\min}} \frac{1}{k^{3/4}} \\
&= o(1)n^{\frac{1}{2}+\frac{1}{2}(\bar{\kappa}_2-\kappa_2)+(\iota-\frac{1}{4})\kappa_3-\frac{5}{4}\kappa_1} \rightarrow 0,
\end{aligned}$$

as $n \rightarrow \infty$ in view of Assumption 4.10 (a).

Up to now, the first part of the theorem is finished. In what follows we shall prove the second part. It follows from (4.25) and $\alpha'B = \alpha'$ that

$$\begin{aligned}
&\frac{\sqrt{T_n}}{\sqrt{n}v_1(T_n)v_2(\sqrt{T_n}\sigma_z)} \frac{\alpha'X'XA(\tau, x)}{\sqrt{p_{\max}}\|A(\tau, x)\|^2} (\widehat{m}(\tau, x) - m(\tau, x)) \\
&= \frac{1}{\sqrt{n}v_1(T_n)v_2(\sqrt{T_n}\sigma_z)} \sqrt{\frac{T_n}{p_{\max}}} \alpha'X'(\delta + \gamma + \varepsilon) \\
&\quad - \frac{\sqrt{T_n}}{\sqrt{n}v_1(T_n)v_2(\sqrt{T_n}\sigma_z)} \frac{\alpha'X'XA(\tau, x)}{\sqrt{p_{\max}}\|A(\tau, x)\|^2} [\delta(\tau, x) + \gamma(\tau, x)]
\end{aligned}$$

$$:= \Pi_3 - \Pi_4.$$

We are about to show that Π_3 is convergent in distribution to the desired stochastic integral and $\Pi_4 \rightarrow_P 0$.

Observe that in view of (4.29) we have

$$\Pi_3 = \frac{1}{\sqrt{n}v_1(T_n)v_2(\sqrt{T_n}\sigma_z)} (\tilde{\mathbf{F}}' - \tilde{\delta}' - \tilde{\gamma}')(\delta + \gamma + \varepsilon).$$

It follows from Theorem 3.4 that

$$\begin{aligned} \frac{1}{\sqrt{n}v_1(T_n)v_2(\sqrt{T_n}\sigma_z)} \tilde{\mathbf{F}}' \varepsilon &= \frac{1}{\sqrt{n}v_1(T_n)v_2(\sqrt{T_n}\sigma_z)} \sum_{s=1}^n \tilde{F}(t_{s,n}, X_{s,n}) e_s \\ &= \frac{1}{\sqrt{n}v_1(T_n)v_2(\sqrt{T_n}\sigma_z)} \sum_{s=1}^n F\left(\frac{s}{n}T_n, \sqrt{T_n}\sigma_z x_{s,n}\right) e_s \\ &\rightarrow_D \int_0^1 f(r, W(r)) dU(r), \end{aligned} \tag{D.24}$$

on account of Assumption 4.10 (c) where $(W(r), U(r))$ is in Assumption B the limit of $(W_n(r), U_n(r))$.

Also, for the same reason we have

$$\frac{1}{nv_1^2(T_n)v_2^2(\sqrt{T_n}\sigma_z)} \|\tilde{\mathbf{F}}'\|^2 \rightarrow_D \int_0^1 f^2(r, W(r)) dr, \tag{D.25}$$

on account of Assumption 4.10 (d).

Notice that the parameters k, p_i , the expressions of $\tilde{\delta}'$ and $\tilde{\gamma}'$ as well as δ and γ remain unchanged as in the first part. Whence, (D.23) is still valid now with a modification that $v(T_n)$ is superceded by $v_1(T_n)$. Therefore, Cauchy-Schwarz inequality and (D.25) imply all the terms of Π_3 except for (D.24) are approaching to zero in probability, hence $\Pi_3 \rightarrow_D \int_0^1 f(r, W(r)) dU(r)$.

Now we are ready to prove that $\Pi_4 \rightarrow_P 0$. Rewrite $\Pi_4 := \Pi_{41} \times (\Pi_{42} + \Pi_{43})$, where

$$\begin{aligned} \Pi_{41} &= \frac{T_n}{nv_1(T_n)v_2(\sqrt{T_n}\sigma_z)\varrho_1(T_n)\varrho_2(\sqrt{T_n}\sigma_z)} \frac{\alpha' X' X A(\tau, x)}{p_{\max} \|A(\tau, x)\|}, \\ \Pi_{42} &= \frac{\sqrt{n}\sqrt{p_{\max}}\varrho_1(T_n)\varrho_2(\sqrt{T_n}\sigma_z)}{\sqrt{T_n}\|A(\tau, x)\|} \delta(\tau, x), \\ \Pi_{43} &= \frac{\sqrt{n}\sqrt{p_{\max}}\varrho_1(T_n)\varrho_2(\sqrt{T_n}\sigma_z)}{\sqrt{T_n}\|A(\tau, x)\|} \gamma(\tau, x). \end{aligned}$$

It follows from (4.29) that

$$\begin{aligned} \Pi_{41} &= \frac{1}{nv_1(T_n)v_2(\sqrt{T_n}\sigma_z)\varrho_1(T_n)\varrho_2(\sqrt{T_n}\sigma_z)} \\ &\quad \times (\tilde{\mathbf{F}}'\tilde{\mathbf{G}} - \tilde{\delta}'\tilde{\mathbf{G}} - \tilde{\gamma}'\tilde{\mathbf{G}} - \tilde{\mathbf{F}}'\tilde{\delta} - \tilde{\mathbf{F}}'\tilde{\gamma} + \|\tilde{\delta}\|^2 + \|\tilde{\gamma}\|^2 + 2\tilde{\gamma}'\tilde{\delta}). \end{aligned}$$

Once again, due to Theorem 3.4, Assumption 4.10 (d), we similarly have

$$\frac{1}{n\varrho_1(T_n)^2\varrho_2(\sqrt{T_n}\sigma_z)^2}\|\tilde{\mathbf{G}}\|^2 \rightarrow_D \int_0^1 g^2(r, W(r))dr. \quad (\text{D.26})$$

Thus, Cauchy-Schwarz inequality as well as (D.23b), (D.25), (D.26) suggests that all the terms in Π_{41} except for the one containing $\tilde{\mathbf{F}}'\tilde{\mathbf{G}}$ converge in probability to zero. Hence, to find out the limit of Γ_{41} it suffices to find that of the term involving $\tilde{\mathbf{F}}'\tilde{\mathbf{G}}$. In fact,

$$\frac{1}{nv_1(T_n)v_2(\sqrt{T_n}\sigma_z)\varrho_1(T_n)\varrho_2(\sqrt{T_n}\sigma_z)}\tilde{\mathbf{F}}'\tilde{\mathbf{G}} \rightarrow_D \int_0^1 f(r, W(r))g(r, W(r))dr,$$

as $n \rightarrow \infty$ by Theorem 3.4 and Assumption 4.10 (d), so that Π_{41} converges to the same limit as above in distribution.

Now let us turn to prove both $\Pi_{42} \rightarrow 0$ and $\Pi_{43} \rightarrow 0$, as $n \rightarrow \infty$. Recall that $O(1)kp_{\min} \leq T_n\|A(\tau, x)\|^2 \leq O(1)kp_{\max}$. Because $\delta(\tau, x)$ and $\gamma(\tau, x)$ are the same as in the first part,

$$\begin{aligned} |\Pi_{42}| &= \frac{\sqrt{n}\sqrt{p_{\max}}\varrho_1(T_n)\varrho_2(\sqrt{T_n}\sigma_z)}{\sqrt{T_n}\|A(\tau, x)\|}|\delta(\tau, x)| \\ &\leq o(1)\frac{\sqrt{n}\sqrt{p_{\max}}T_n^{\nu_1}T_n^{\frac{1}{2}\nu_2}}{\sqrt{kp_{\min}}}\sqrt{k}\left(\sum_{i=0}^k T_n^3 \frac{1}{p_i^3}\right)^{\frac{1}{2}} \\ &\leq o(1)n^{\frac{1}{2}+\frac{1}{2}\kappa_1+\frac{1}{2}(\bar{\kappa}_2-\kappa_2)+(\nu_1+\frac{1}{2}\nu_2+\frac{3}{2})\kappa_3-\frac{3}{2}\kappa_2} \rightarrow 0, \\ |\Pi_{43}| &= \frac{\sqrt{n}\sqrt{p_{\max}}\varrho_1(T_n)\varrho_2(\sqrt{T_n}\sigma_z)}{\sqrt{T_n}\|A(\tau, x)\|}|\gamma(\tau, x)| \\ &\leq o(1)\frac{\sqrt{n}\sqrt{p_{\max}}T_n^{\nu_1}T_n^{\frac{1}{2}\nu_2}}{\sqrt{kp_{\min}}}\frac{1}{k^{3/4}} \\ &= o(1)n^{\frac{1}{2}+\frac{1}{2}(\bar{\kappa}_2-\kappa_2)+(\nu_1+\frac{1}{2}\nu_2)\kappa_3-\frac{5}{4}\kappa_1} \rightarrow 0, \end{aligned}$$

as $n \rightarrow \infty$ by Assumption 4.10 (c) for the parameters. This finishes the proof. \square

References

- Ai, C. and Chen, X. (2003). Efficient estimation of models with conditional moment restrictions containing unknown functions. *Econometrica*, 71:1795–1843.
- Ai, C. and Chen, X. (2007). Estimation of possibly misspecified semiparametric conditional moment restriction models with different conditioning variables . *Journal of Econometrics*, 141:5–43.
- Alexits, G. (1961). *Convergence Problems of Orthogonal Series*. International Series of Monographs in Pure and Applied Mathematics. Pergamon Press, New York.
- Askey, R. and Wainger, S. (1965). Mean convergence of expansions in Laguerre and Hermite series. *American Journal of Mathematics*, 87:695–708.
- Bandi, F. and Phillips, P. C. B. (2003). Fully nonparametric estimation of scalar diffusion models. *Econometrica*, 71:241–283.
- Bandi, F. and Phillips, P. C. B. (2007). A simple approach to the parametric estimation of potentially nonstationary diffusions . *Journal of Econometrics*, 137:354–395.
- Billingsley, P. (1995). *Probability and Measure*. JOHN WILEY & SONS, Inc., New York.
- Cai, Z., Li, Q., and Park, J. (2009). Functional-coefficient cointegration models for nonstationary time series data. *Journal of Econometrics*, 148:101–113.
- Chen, X. and Ludvigson, S. C. (2009). Land of addicts? an empirical investigation of habit-based asset pricing models. *Journal of Applied Econometrics*, 24:1057–1093.
- Davis, H. F. (1963). *Fourier Series and Orthogonal Functions*. Allyn and Bacon, Boston.
- Dudley, R. M. (2003). *Real Analysis and Probability*. Cambridge Studies in Advanced Mathematics 74. Cambridge University Press, Cambridge, U.K.
- Gao, J. (2007). *Nonlinear Time Series: semiparametric and nonparametric methods*. Monographs on Statistics and Applied Probability. Chapman & Hall, New York.
- Gao, J. and Phillips, P. C. B. (2011). Semiparametric estimation in multivariate non stationary time series models. Working Paper at <http://www.buseco.monash.edu.au/ebs/pubs/wpapers/2011/wp17-11.pdf>.

- Hansen, L. P. and Sargent, T. J. (1983). The dimensionality of the aliasing problem in models with rational spectral densities. *Econometrica*, 50:377–387.
- Karlsen, H. A., Mykelbust, T., and Tjøstheim, D. (2007). Nonparametric estimation in a nonlinear cointegration type model. *The Annals of Statistics*, 35:252–299.
- Karlsen, H. A. and Tjøstheim, D. (2001). Nonparametric estimation in null recurrent time series. *The Annals of Statistics*, 29:372–416.
- Kurtz, T. G. and Protter, P. (1991). Weak limit theorems for stochastic integrals and stochastic differential equations. *The Annals of Probability*, 19(3):1035–1070.
- Li, Q. and Racine, J. (2007). *Nonparametric Econometrics: Theory and Practice*. Princeton University Press, Princeton.
- Nikiforov, A. F. and Uvarov, V. B. (1988). *Special Functions of Mathematical Physics*. Birkhauser, Boston.
- Park, J. Y. and Phillips, P. C. B. (1999). Asymptotics for nonlinear transformations of integrated time series. *Econometric Theory*, 15:269–298.
- Park, J. Y. and Phillips, P. C. B. (2001). Nonlinear regression with integrated time series. *Econometrica*, 69(1):117–161.
- 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. (1987). Time series regression with a unit root. *Econometrica*, 55(2):277–301.
- Phillips, P. C. B. (2009). Local time theory and spurious regressions. *Econometric Theory*, 25:1466–1497.
- Phillips, P. C. B. and Park, J. Y. (1998). Nonstationary density estimation and kernel autoregression. Cowles Foundation Discussion Paper No.1181, Yale University.
- Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion*. A Series of Comprehensive Studies in Mathematics 293. Springer-Verlag.

- Sansone, G. (1959). *Orthogonal Functions*. Pure and Applied Mathematics. Interscience Publishers, Inc., New York.
- Schoutens, W. (2000). *Stochastic Processes and Orthogonal Polynomials*. Springer-Verlag.
- Schoutens, W. (2003). *Levy Processes in Finance: Pricing Financial Derivatives*. Wiley Series in Probability and Statistics. WILEY.
- Wang, Q. and Phillips, P. C. B. (2009a). Asymptotic theory for local time density estimation and nonparametric cointegreting regression. *Econometric Theory*, 25:710–738.
- Wang, Q. and Phillips, P. C. B. (2009b). Structure nonparametric cointegrating regression. *Econometrica*, 77:1901–1948.
- Xiao, Z. (2009). Functional-coefficient cointegration models. *Journal of Econometrics*, 152:81–92.