Hindawi
Journal of Applied Mathematics
Volume 2020, Article ID 6086983, 17 pages
https://doi.org/10.1155/2020/6086983

Research Article

Partial Derivative Estimation for Underlying Functional-Valued Process in a Unified Framework

Yunbei Ma,[1] **Fanyin Zhou**,[1] **and Xuan Luo**[2]

[1]*School of Statistics and Research Center of Statistics, Southwestern University of Finance and Economics, Chengdu, China*
[2]*Sichuan Southwest Vocational College of Civil Aviation, Chengdu, China*

Correspondence should be addressed to Yunbei Ma; myb@swufe.edu.cn

Received 7 July 2019; Accepted 6 September 2019; Published 13 January 2020

Academic Editor: Syed Abdul Mohiuddine

We consider functional data analysis when the observations at each location are functional rather than scalar. When the dynamic of underlying functional-valued process at each location is of interest, it is desirable to recover partial derivatives of a sample function, especially from sparse and noise-contaminated measures. We propose a novel approach based on estimating derivatives of eigenfunctions of marginal kernels to obtain a representation for functional-valued process and its partial derivatives in a unified framework in which the number of locations and number of observations at each location for each individual can be any rate relative to the sample size. We derive almost sure rates of convergence for the procedures and further establish consistency results for recovered partial derivatives.

1. Introduction

With the rapid advance in computational and analytical technology, many time-dynamic processes are monitored and recoded continuously during a time interval or intermittently at several discrete time points. Functional data analysis (FDA) is a powerful tool to deal with the analysis and theory of data that are in the form of functions, images and shapes, or more general objects. Traditional functional data typically consist of a random sample of independent real-valued functions, which can be viewed as the realization of a one-dimensional stochastic process. In this field of research, a general introduction of the available methods can be found in Ramsay and Silverman [1] and Wang et al. [2].

Many recent developments in FDA concern multivariate functional data and spatially indexed functional data. Chen and Müller [3] introduced a methodology for repeatedly observed and thus dependent functional data, covering the case where the recordings of the curves are scheduled on a regular and dense grid with often sparse and random recording times.

We consider special situations where the observations at each location are functional rather than scalar. For $\mathcal{S} \subset \mathbb{R}^{p_1}$ and $\mathcal{T} \subset \mathbb{R}^{p_2}$, we consider the stochastic process $X: \mathcal{T} \longrightarrow L^2(\mathcal{S})$ and denote its value at time $t \in \mathcal{T}$ by $X(\cdot, t)$, which is a square integrable random function with argument $s \in \mathcal{S}$. Chen et al. [4] proposed marginal FPCA and product FPCA models for $X(s, t)$ and developed estimating methods and theoretical results under designs that are dense and regular in s. In practice, we may deal with functional data which are dense and random at the s direction. Under these cases, a presmoothing of individual curve at each location is necessary. However, in practice, it is possible that we are faced with sparse and random designs in s. Moreover, it is also possible that curves at some locations are densely observed, while curves at other locations are sparsely observed. In this paper, we aim to recover X(s,t) by estimating the multivariate mean function, the marginal covariance function, and then the FPCA in a unified framework. This unified framework allows the number of locations and the number of observations at each location for each individual to be any rate relative to the sample size. Thus, the proposed procedure avoids a challenging issue of classifying which scenario we are faced with and hence deciding which methodology to use when dealing with real data.

On the other hand, it is often of interest to recover derivatives of a sample of random functions, especially when the dynamics of underlying processes is of interest. Since currently available statistical methods for estimating derivatives require densely observed data, it is quiet challenging to recover derivatives from sparse functional data with noise-contaminated measurements. Liu and Müller [5] proposed an approach based on estimating derivatives of eigenfunctions to obtain a representation for derivatives of a sample of sparsely observed one-dimensional functions. Our further work in this paper is aimed at recovering partial derivatives of underlying functional-valued process at each location, that is the dth partial derivatives of $X(s,t)$ with respect to s, which is denoted as $X^{(d,0)}(s,t) = (\partial^d/\partial s^d)X(s,t)$. The whole procedure is also in a unified framework in which multiple functional data can be either densely or sparsely observed.

The article is organized as follows. In Section 2, we introduce the model and all estimation procedures for recovering both functional-valued process and its partial derivatives. We establish the uniformly almost sure convergence rates of the procedures in Section 3, where we also discuss the rates corresponding to some special scenarios. Some relative issues to our proposed procedures are discussed in Section 4. In Section 5, simulation studies are conducted to evaluate the performance of our procedures. All technique lemmas and all proofs are included in Appendix.

2. Models and Estimation

2.1. Representations. Consider process $X(s,t)$ with mean $\mu(s,t) = E[X(s,t)]$ for all $s \in \mathcal{S} \subset \mathbb{R}^{p_1}$ and $t \in \mathcal{T} \subset \mathbb{R}^{p_2}$ and covariance function

$$\begin{aligned} C\{(s,t),(u,v)\} &= E[X(s,t)X(u,v)] - \mu(s,t)\mu(u,v) \\ &= E\left[X^c(s,t)X^c(u,v)\right]. \end{aligned} \tag{1}$$

Chen et al. [4] proposed a representation as

$$\begin{aligned} X(s,t) &= \mu(s,t) + \sum_{j=1}^{\infty} \xi_j(t)\psi_j(s) = \mu(s,t) \\ &\quad + \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \zeta_{ijk}\phi_{jk}(t)\psi_j(s), \end{aligned} \tag{2}$$

where $\{\psi\colon j \ge 1\}$ are the eigenfunctions of the operator in $L^2(\mathcal{S})$ with kernel

$$G_{\mathcal{S}}(s,u) = \int_{\mathcal{T}} C\{(s,t),(u,t)\}\mathrm{d}t, \tag{3}$$

and $\{\xi_j(t)\colon j \ge 1\}$ are the random coefficients of the expansion of the centred processes $X^c(\cdot,t)$ in $\psi_j(s)$ and

$$\xi_j(t) = \sum_{k=1}^{\infty} \zeta_{ijk}\phi_{jk}(t), \tag{4}$$

is the Karhunen–Loeve expansion of the random functions $\xi_j(t)$ in $L^2(\mathcal{T})$. Here, for each $j \ge 1$, $\{\phi_{jk}, k \ge 1\}$ are the eigenfunctions of the operator with kernel

$$\Gamma_{\mathcal{T}}(t,v) = E\left[\xi_j(t)\xi_j(v)\right], \tag{5}$$

and $\{\zeta_{ijk}, k \ge 1\}$ are the FPC scores of $\xi_j(t)$.

Based on the representation of $X(s,t)$ shown as (2), we can write $X^{(d,0)}(s,t)$ as

$$X^{(d,0)}(s,t) = \mu^{(d,0)}(s,t) + \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \zeta_{ijk}\phi_{jk}(t)\psi_j^{(d)}(s), \tag{6}$$

where $\mu^{(d,0)}(s,t)$ is the dth partial derivative of $\mu(s,t)$ with respect to s and $\psi_j^{(d)}(\cdot)$ is the dth derivative of $\psi_j(\cdot)$ on $\mathcal{S}$.

Denote $\lambda_1 \ge \lambda_2 \ge \cdots \ge 0$ are eigenvalues of $G(u,s)$, and then the eigenfunctions ψ_j are the solutions of the eigenequations $\int_{\mathcal{S}} G_{\mathcal{S}}(u,s)\psi_j(u)\mathrm{d}s = \lambda_j\psi_j(s)$, under side conditions of norm 1 and orthogonality on all previous eigenfunctions. Upon taking the dth derivative with respect to s from both side of these eigenequations,

$$\frac{\partial^d}{\partial s^d}\int_{\mathcal{S}} G_{\mathcal{S}}(s,u)\psi_j(u)\mathrm{d}s = \lambda_j\psi_j^{(d)}(s). \tag{7}$$

If $(\partial^d/\partial s^d)G_{\mathcal{S}}(s,u)\psi_j(u)$ exists for all $u, s \in \mathcal{S}$, and

$$\left|\frac{\partial^d}{\partial s^d} G_{\mathcal{S}}(s,u)\psi_j(u)\right|, \tag{8}$$

is bounded and integrable for all $s, u \in \mathcal{S}$, interchanging integrations and differentiation leads to

$$\psi_j^{(d)}(s) = \frac{1}{\lambda_j}\int_{\mathcal{S}} \frac{\partial^d}{\partial s^d} G_{\mathcal{S}}(s,u)\psi_j(u)\mathrm{d}s. \tag{9}$$

One can then estimate derivatives by approximating $X^{(d,0)}(s,t)$ with a truncated representation

$$\widetilde{X}^{(d,0)}(s,t) = \mu^{(d,0)}(s,t) + \sum_{j=1}^{J}\sum_{k=1}^{K_j} \zeta_{ijk}\phi_{jk}(t)\psi_j^{(d)}(s), \tag{10}$$

with finite $K_j \ge 1$, $j = 1, 2, \cdots, J$, and $J \ge 1$.

2.2. Estimation. Time-indexed functional data consist of a sample of n independent subjects or units. For the ith subject, suppose we observe

$$\begin{aligned} Y_{iml} &= X\left(S_{iml}, T_{im}\right) + \varepsilon_{iml}, \\ &\quad i = 1, \ldots, n, m = 1, \ldots, M_i, l = 1, \ldots, L_{im}, \end{aligned} \tag{11}$$

which means that, on each time point T_{im}, ∞, $X_i(\cdot, T_{im})$ is recorded at a grid of functional points S_{iml}, $l = 1, \ldots, L_{im}$. Here, ε_{iml} are additional measurement errors, assumed to be iid with mean zero and finite variance σ^2. We also assume that ε_{iml} are independent of all X_i, T_{im}, and S_{iml}.

Our approach is based on the local-polynomial smoother [6].

Step 1. Estimation of the mean function and the partial derivatives of mean functions.

For fixed $(s,t) \in \mathcal{S} \times \mathcal{T}$ and some bandwidths h_1 and h_2,

$$\begin{aligned}\mu(S_{iml}, T_{im}) &\approx \mu(s,t) + h_1\mu^{(1,0)}(s,t)\left(\frac{S_{iml}-s}{h_1}\right) + \cdots \\ &+ \frac{h_1^{d+1}}{(d+1)!}\mu^{(d+1,0)}(s,t)\left(\frac{S_{iml}-s}{h_1}\right)^{d+1} \\ &+ h_2\mu^{(0,1)}(s,t)\left(\frac{T_{im}-t}{h_2}\right) \equiv \sum_{k=0}^{d+1}\theta_{k0}\left(\frac{S_{iml}-s}{h_1}\right)^k \\ &+ \theta_{01}\left(\frac{T_{im}-t}{h_2}\right).\end{aligned} \quad (12)$$

Let $\theta = (\theta_{00}, \theta_{10}, \ldots, \theta_{(d+1)0}, \theta_{01})^T$, where $\theta_{l0} = h_1^l\mu^{(l,0)}(s,t)/l!, l = 0,1,\ldots,d+1$ and $\theta_{01} = \mu^{(0,1)}(s,t)$. Then, we obtain a smoothing estimator $\widehat{\theta} = (\widehat{\theta}_{00}, \widehat{\theta}_{10}, \ldots, \widehat{\theta}_{(d+1)0}, \widehat{\theta}_{01})^T$ as

$$\begin{aligned}\widehat{\theta} &= \arg\min \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}}\sum_{l=1}^{L_{im}}K_{h_1}(S_{iml}-s)K_{h_2}(T_{im}-t) \\ &\cdot\left[Y_{iml} - \sum_{k=0}^{d+1}\theta_{k0}\left(\frac{S_{iml}-s}{h_1}\right)^k - \theta_{01}\left(\frac{T_{im}-t}{h_2}\right)\right]^2,\end{aligned} \quad (13)$$

where $K(\cdot)$ is a symmetric probability density function on $[0,1]$ and $K_h(\cdot) = (1/h)K(\cdot/h)$. Here, the kernel function $K(\cdot)$ can be different at different occasions. Then, local estimators of $\mu(s,t)$ and $\mu^{(d,0)}(s,t)$ are given by

$$\begin{aligned}\widehat{\mu}(s,t) &= \widehat{\theta}_{00}, \\ \widehat{\mu}^{(d,0)}(s,t) &= d!h_1^{-d}\widehat{\theta}_{d0},\end{aligned} \quad (14)$$

respectively.

Step 2. Estimation of $G_S(s,u)$, $G_{\mathbf{S}}^{(d,0)}(s,u)$, and σ^2.

Note that

$$\begin{aligned}G_{\mathcal{S}}(s,u) &= \int_{\mathcal{T}} C\{(s,t),(u,t)\}dt \\ &= \int_{\mathcal{T}}\{D(s,u,t) - \mu(s,t)\mu(u,t)\}dt,\end{aligned} \quad (15)$$

where $D(s,u,t) = E[X(s,t)X(u,t)]$. On the contrary, if $(\partial^d/\partial s^d)C\{(s,t),(u,t)\}$ exists for all $(s,u,t) \in \mathcal{S}^2 \times \mathcal{T}$ and $|(\partial^d/\partial s^d)C\{(s,t),(u,t)\}|$ is bounded and integrable for all $(s,u,t) \in \mathcal{S}^2 \times \mathcal{T}$, then

$$\begin{aligned}G_{\mathcal{S}}^{(d,0)}(s,u) &= \int_{\mathcal{T}}\frac{\partial^d}{\partial s^d}C\{(s,t),(u,t)\}dt \\ &= \int_{\mathcal{T}}\left\{D^{(d,0,0)}(s,u,t) - \mu^{(d,0)}(s,t)\mu(u,t)\right\}dt,\end{aligned} \quad (16)$$

where $D^{(d,0,0)}(s,u,t) = (\partial^d/\partial s^d)D(s,u,t)$. Thus, in order to estimate $G_{\mathcal{S}}(s,u)$ and $G_{\mathcal{S}}^{(d,0)}(s,u)$, we estimate $D(s,u,t)$ and $D^{(d,0,0)}(s,u,t)$ first.

To this end, we estimate $D(s,u,t)$ and $D^{(d,0,0)}(s,u,t)$ based on the following procedures. For fixed $(s,u,t) \in \mathcal{S}^2 \times \mathcal{T}$ and some bandwidths h_3, h_4, and h_5 (for $d=0$, we choose $h_3 = h_4$),

$$\begin{aligned}D(S_{iml}, S_{iml'}, T_{im}) &\approx D(s,u,t) + h_3D^{(1,0,0)}(s,u,t)\left(\frac{S_{iml}-s}{h_3}\right) \\ &+ \cdots + \frac{h_3^{d+1}}{(d+1)!}D^{(d+1,0,0)}(s,u,t) \\ &\cdot\left(\frac{S_{iml}-s}{h_3}\right)^{d+1} + h_4D^{(0,1,0)}(s,u,t) \\ &\cdot\left(\frac{S_{iml'}-u}{h_4}\right) + h_5D^{(0,0,1)}(s,u,t) \\ &\cdot\left(\frac{T_{im}-t}{h_5}\right) \equiv \sum_{k=0}^{d+1}\theta_{k00}^*\left(\frac{S_{iml}-s}{h_3}\right)^k \\ &+ \theta_{010}^*\left(\frac{S_{iml'}-u}{h_4}\right) + \theta_{001}^*\left(\frac{T_{im}-t}{h_5}\right).\end{aligned} \quad (17)$$

Let $\theta^* = (\theta_{000}^*, \theta_{100}^*, \ldots, \theta_{(d+1)00}^*, \theta_{010}^*, \theta_{001}^*)^T$, where $\theta_{l00}^* = h_3^lD^{(l,0,0)}(s,u,t)/l!, l = 1,\ldots,d+1$, $\theta_{010}^* = D^{(0,1,0)}(s,u,t)$, and $\theta_{001}^* = D^{(0,0,1)}(s,u,t)$. Then, we obtain an estimator $\widehat{\theta}^* = (\widehat{\theta}_{000}^*, \widehat{\theta}_{100}^*, \ldots, \widehat{\theta}_{(d+1)00}^*, \widehat{\theta}_{010}^*, \widehat{\theta}_{001}^*)^T$ as

$$\begin{aligned}\widehat{\theta}^* &= \arg\min \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}^*}\sum_{l\neq l'}^{L_{im}}K_{h_3}(S_{iml}-s)K_{h_4} \\ &\cdot(S_{iml'}-u)K_{h_5}(T_{im}-t)\left[Y_{iml}Y_{iml'} - \sum_{k=0}^{d+1}\theta_{k00}^*\left(\frac{S_{iml}-s}{h_3}\right)^k\right. \\ &\left.- \theta_{010}^*\left(\frac{S_{iml'}-u}{h_4}\right) - \theta_{001}^*\left(\frac{T_{im}-t}{h_5}\right)\right]^2,\end{aligned} \quad (18)$$

where $L_{im}^* = (L_{im}-1)L_{im}$. Then, smoothing estimators of $D(s,u,t)$ and $D^{(d,0,0)}(s,u,t)$ are given by

$$\begin{aligned}\widehat{D}(s,u,t) &= \widehat{\theta}_{000}^*, \\ \widehat{D}^{(d,0,0)}(s,u,t) &= d!h_3^{-d}\widehat{\theta}_{d00}^*,\end{aligned} \quad (19)$$

respectively. We then can obtain that

$$\widehat{G}_{\mathcal{S}}(s,u) = \int_{\mathcal{T}} \left\{ \widehat{D}(s,u,t) - \widehat{\mu}(s,t)\widehat{\mu}(u,t) \right\} dt,$$
$$\widehat{G}_{\mathcal{S}}^{(d,0)}(s,u) = \int_{\mathcal{T}} \left\{ \widehat{D}^{(d,0,0)}(s,u,t) - \widehat{\mu}^{(d,0)}(s,t)\widehat{\mu}(u,t) \right\} dt. \quad (20)$$

To estimate σ^2, we first estimate $V(s,t) = D(s,s,t) + \sigma^2$ by $\widehat{V}(s,t) = \widehat{a}_0$, where

$$(\widehat{a}_0, \widehat{a}_1, \widehat{a}_2) = \arg\min \frac{1}{n} \sum_{i=1}^{n} \frac{1}{M_i} \sum_{m=1}^{M_i} \frac{1}{L_{im}} \sum_{l=1}^{L_{im}} K_{h_{v1}} \cdot (S_{iml} - s) K_{h_{v2}}(T_{im} - t) \cdot \left[Y_{iml}^2 - a_0 - a_1 (S_{iml} - u) - a_2 (T_{im} - t) \right]^2, \quad (21)$$

with some bandwidths h_{v1} and h_{v2}. $D(s,s,t)$ is estimated in the same way as estimating $D(s,u,t)$, but with $d = 0$ and $h_3 = h_4$, we then estimate σ^2 by

$$\widehat{\sigma}^2 = \frac{1}{|\mathcal{S}||\mathcal{T}|} \int_{\mathcal{T}} \int_{\mathcal{S}} \left\{ \widehat{V}(s,t) - \widehat{D}(s,s,t) \right\} ds\, dt, \quad (22)$$

where $|\mathcal{S}|$ and $|\mathcal{T}|$ are Lebesgue measures of $\mathcal{S}$ and $\mathcal{T}$, respectively.

Remark 1. In practice, the empirical estimator of $G_{\mathcal{S}}(s,u)$ [4] can be used and remains at convergence rate $(\log n/n)^{1/2}$ for dense and regular designs in s; that is, all $X_i(\cdot,t)$s are observed at $\{s_l\}_{l=1}^{l}$ and $\max(s_l - s_{l-1}) = O(n^{-1})$. On the contrary, by presmoothing for individual curves, the empirical estimator of $G_{\mathcal{S}}(s,u)$ is also applicable for dense and random designs for s, and as designs get denser, the overall convergence rate $(\log n/n)^{1/2}$ remains under appropriate regularity conditions. Under these circumstances, a further smoothing estimator of $G^{(d,0)}(s,u)$ can also be obtained based on empirical estimators of $G_{\mathcal{S}}(s,u)$. Similar results hold for the estimation of σ^2.

However, in practice, it is possible that some sample curves are densely observed, while others are sparsely observed at the s direction. Moreover, in dealing with real data, it may even be difficult to classify which scenario we are faced with and hence to decide which methodology to use.

Step 3. Estimation of eigenfunctions $\psi_j(s)$ and $\psi_j^{(d)}(s)$ and eigenvalues λ_j of the operator in $L^2(S)$ with kernel $G_S(s,u)$, as well as estimation of FPC functions $\xi_{ij}(t) = \int_S X_i^c(s,t)\psi_j(s)ds$.

The estimated eigenfunctions $\widehat{\psi}_j(s)$ and estimated eigenvalues $\widehat{\lambda}_j$ can be obtained by standard methods of computing the eigenvalues and eigenfunctions of an integral operator with a symmetric kernel. Then, we have

$$\widehat{\psi}_j^{(d)}(s) = \widehat{\lambda}_j^{-1} \int_{\mathcal{S}} \widehat{G}_{\mathcal{S}}^{(d,0)}(s,u)\widehat{\psi}_j(u)du. \quad (23)$$

For designs that are dense in s, one can obtain $\widehat{\xi}_{ij}(t)$ by interpolating numerical approximations of the integrals:

$$\widehat{\xi}_{ij}(T_{im}) = \int [X_i(s,T_{im}) - \widehat{\mu}(s,T_{im})]\widehat{\psi}_j(s)ds. \quad (24)$$

On the contrary, for designs that are sparse in s, one can estimate $\xi_{ij}(t)$ by the PACE approach [7].

Step 4. Estimation of eigenfunctions $\phi_{jk}(t)$ of the operator with kernel $\Gamma_T(t,v)$ and FPCs $\zeta_{ijk} = \int_T \xi_{ij}(t)\phi_{jk}(t)dt$.

This is a standard FPCA of one-dimensional processes $\{\widehat{\xi}_{ij}(t), j \geq 1\}$. For each fixed j, one obtains estimates for the FPCs ζ_{ijk} and eigenfunctions $\phi_{jk}(t)$ for designs that are dense in t [1] and for designs that are sparse in t [7]. One can also adapt the approach of Li and Hsing [8], which is suitable for both sparse and dense functional data, to one-dimensional processes $\{\widehat{\xi}_{ij}(t), j \geq 1\}$.

In this step, for each j, we are able to approximate $\widehat{\xi}_{ij}(t)$ by $\sum_{k=1}^{K_j} \widehat{\zeta}_{ijk}\widehat{\phi}_{jk}(t)$.

After selecting appropriate numbers of included components J and K_j, $j = 1, \ldots, J$, we obtain the overall representation:

$$\widehat{X}(s,t) = \widehat{\mu}(s,t) + \sum_{j=1}^{J} \sum_{k=1}^{K_j} \widehat{\zeta}_{ijk}\widehat{\phi}_{jk}(t)\widehat{\psi}_j(s), \quad (25)$$

$$\widehat{X}^{(d,0)}(s,t) = \widehat{\mu}^{(d,0)}(s,t) + \sum_{j=1}^{J} \sum_{k=1}^{K_j} \widehat{\zeta}_{ijk}\widehat{\phi}_{jk}(t)\widehat{\psi}_j^{(d)}(s). \quad (26)$$

The included number of components J and K_j, $j = 1, \ldots, J$, can be selected via a variety of methods, including fraction of variance explained (FVE) criterion [8], leave-one-curve-out cross-validation [9], pseudo-AIC [10], or pseudo-BIC [7, 11]. We will illustrate these procedures in Section 4.

3. Asymptotic Theory

We first define the notations and conditions to be used. Assume that M_i and L_{im} may depend on n as well, namely, $M_i = M_{in}$ and $L_{im} = L_{imn}$. However, for simplicity, we continue to use the notation M_i and L_{im}. Define

$$\gamma_{nkl} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{M_i^k} \sum_{m=1}^{M_i} \frac{1}{L_{im}^l}, \quad k = 1, 2,\ l = 0, 1, 2. \quad (27)$$

For any bandwidths h_1, h_2, and h_3, we also define

$$\delta_{n1}(h_1,h_2) = \sqrt{(nh_1h_2)^{-1}(\gamma_{n21}+2\gamma_{n20}h_1+2\gamma_{n11}h_2+4h_1h_2)\log n},$$
$$\delta_{n2}(h_1,h_2,h_3) = \sqrt{(nh_1h_2h_3)^{-1}(\gamma_{n22}+4\gamma_{n20}h_1h_2+2\gamma_{n12}h_3+8h_1h_2h_3)\log n}. \quad (28)$$

From now on, without loss of generality, we assume that the domain $\mathcal{S}\times\mathcal{T}$ of the process is $[0,a]\times[0,b]$. Some assumptions needed for the asymptotic theory are as follows. We use $0<C<\infty$ as a generic constant that can take different values at different places.

Now, we state the assumptions:

Assumption 1. All second-order partial derivatives of $\mu(s,t)$ are uniformly continuous and bounded on $[0,a]\times[0,b]$. Furthermore, $\mu^{(d+2,0)}(s,t)$ exists and is uniformly continuous and bounded on $[0,a]\times[0,b]$.

Assumption 2. All second-order partial derivatives of $D(s,u,t)$ are uniformly continuous and bounded on $[0,a]^2\times[0,b]$. Furthermore, $D^{(d+2,0,0)}(s,u,t)$ exists and is uniformly continuous and bounded on $[0,a]^2\times[0,b]$.

Assumption 3. Let $f_S(\cdot)$ and $f_T(\cdot)$ be the density distribution functions of S and T, respectively. Assume both of $f_S(\cdot)$ and $f_T(\cdot)$ are lower bounded away from 0 and $\sup_{[0,a]}f_S(\cdot)\le M_S$ and $\sup_{[0,b]}f_T(\cdot)\le M_T$ for some positive constants M_S and M_T.

Assumption 4. Let $f(\cdot,\cdot)$ be the joint density distribution function of (S,T) and $f_2(\cdot,\cdot,\cdot)$ be the joint density distribution function of (S_1,S_2,T). Both $f(\cdot,\cdot)$ and $f_2(\cdot,\cdot,\cdot)$ are upper bounded and lower bounded away from 0. Furthermore, assume that both $f(\cdot,\cdot)$ and $f_2(\cdot,\cdot,\cdot)$ have continuous and bounded second-order derivatives uniformly on their domains.

Assumption 5. $E(|\varepsilon_{iml}|^{\lambda_1})<\infty$ and $E(\sup|X(s,t)|^{\lambda_1})<\infty$ for some $\lambda_1\in(2,\infty)$. $h_1,h_2\longrightarrow 0$ as $n\longrightarrow\infty$, and $h_1^{-1}h_2^{-1}(\gamma_{n21}+2\gamma_{n20}h_1+2\gamma_{n11}h_2+4h_1h_2)^{-1}(\log n/n)^{1-2/\lambda_1}=o(1)$.

Assumption 6. $E(|\varepsilon_{iml}|^{2\lambda_2})<\infty$ and $E(\sup|X(s,t)|^{2\lambda_2})<\infty$ for some $\lambda_2\in(2,\infty)$. $h_3,h_4,h_5\longrightarrow 0$ as $n\longrightarrow\infty$ and $h_3^{-1}h_4^{-1}h_5^{-1}(\gamma_{n22}+4\gamma_{n20}h_3h_4+2\gamma_{n12}h_5+8h_3h_4h_5)^{-1}(\log n/n)^{1-2/\lambda_2}=o(1)$.

Assumption 7. $E(|\varepsilon_{iml}|^{2\lambda_3})<\infty$ and $E(\sup|X(s,t)|^{2\lambda_3})<\infty$ for some $\lambda_3\in(2,\infty)$. $h_{v1},h_{v2}\longrightarrow 0$ as $n\longrightarrow\infty$ and $h_{v1}^{-1}h_{v2}^{-1}(\gamma_{n21}+2\gamma_{n20}h_{v1}+2\gamma_{n11}h_{v2}+4h_{v1}h_{v2})^{-1}(\log n/n)^{1-2/\lambda_3}=o(1)$.

Assumption 8. Assume the autocovariance operator Δ_S generated by $G_{\mathcal{S}}(s,u)$ is positive definite, such that

$$\sup_s\int G_{\mathcal{S}}^{(d,0)}(s,u)\psi_j(u)\mathrm{d}u<\infty,$$
$$\iint\left|G_{\mathcal{S}}^{(d,0)}(s,u)\right|\mathrm{d}s\,\mathrm{d}u<\infty. \quad (29)$$

and with eigenfunctions satisfying $\sup_{s\in[0,a]}|\psi_j(s)|=O(1)$.

Assumption 9. Let $K(\cdot)$ be a symmetric probability density function on $[0,1]$, $\nu_k=\int u^kK(u)\mathrm{d}u$. Let

$$\Gamma=\begin{pmatrix}1 & \nu_1 & \cdots & \nu_{(d+1)}\\ \nu_1 & \nu_2 & \cdots & \nu_{(d+2)}\\ \vdots & & & \\ \nu_{(d+1)} & \nu_{(d+2)} & \cdots & \nu_{(2d+2)}\end{pmatrix}. \quad (30)$$

Assume that Γ is a nonsingular matrix.

Assumptions (1) and (2) are regular smoothness conditions on the mean function μ and the covariance function D. Since we do not impose any parametric structure on the distribution of X, assumptions (3) and (4) are required for the derivation of uniform convergence. The moment conditions in (5)–(7) are similar to that in (C.5)–(C.7) of Li and Hsing [8] and hold rather generally. Assumptions (8) is similar to condition (B4) in Liu and Müller [5] and is needed for Theorem 4. When $d=1$, the standard normal distribution function is an example for a kernel satisfying (13).

3.1. Uniform Convergence Rates of $\hat{\mu}(s,t)$ and $\hat{\mu}^{(d,0)}(s,t)$.

We establish the uniform convergence rates of $\hat{\mu}(s,t)$ and $\hat{\mu}^{(d,0)}(s,t)$. First, we give some definitions and notations. For $p=0,1,\dots,2d+2$, and $q=0,1,2$, let

$$S_{npq}(s,t)=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}}\sum_{l=1}^{L_{im}}K_{h_1}(S_{iml}-s)K_{h_2}(T_{im}-t)\cdot\left(\frac{S_{iml}-s}{h_1}\right)^p\left(\frac{T_{im}-t}{h_2}\right)^q,$$
$$\widetilde{S}_{npq}(s,t)=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}}\sum_{l=1}^{L_{im}}Y_{iml}K_{h_1}(S_{iml}-s)K_{h_2}\cdot(T_{im}-t)\left(\frac{S_{iml}-s}{h_1}\right)^p\left(\frac{T_{im}-t}{h_2}\right)^q. \quad (31)$$

By some simple algebra, we can obtain that $\hat{\theta}$ satisfies

$$\widetilde{S}_n(s,t)=S_n(s,t)\hat{\theta}, \quad (32)$$

where $\widetilde{S}_n(s,t)=(\widetilde{S}_{n00}(s,t),\widetilde{S}_{n10}(s,t),\dots,\widetilde{S}_{n(d+1)0}(s,t),\widetilde{S}_{n01}(s,t))^T$ and

$$S_n(s,t)=\begin{pmatrix} S_{n00} & S_{n10} & \cdots & S_{n(d+1)0} & S_{n01} \\ S_{n10} & S_{n20} & \cdots & S_{n(d+2)0} & S_{n11} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ S_{n(d+1)0} & S_{n(d+2)0} & \cdots & S_{n(2d+2)0} & S_{n(d+1)1} \\ S_{n01} & S_{n11} & \cdots & S_{n(d+1)1} & S_{n02} \end{pmatrix}(s,t). \tag{33}$$

It then follows that $\widehat{\theta}=S_n^{-1}(s,t)\widetilde{S}_n(s,t)$ and $\widehat{\theta}-\theta=S_n^{-1}(s,t)[\widetilde{S}_n(s,t)-S_n(s,t)\theta]$.

Theorem 1. *Under assumptions (1), (3)–(5), and (9), let* $\gamma_{1d}(h_1,h_2)=h_1^{(d+2)}+h_1h_2+h_2^2$, *then*

$$\begin{aligned}\sup_{s\in\mathcal{S},t\in\mathcal{T}}|\widehat{\mu}(s,t)-\mu(s,t)| &= O_{a.s}(\delta_{n1}(h_1,h_2) + \gamma_{1d}(h_1,h_2)),\\ \sup_{s\in\mathcal{S},t\in\mathcal{T}}\left|\widehat{\mu}^{(d,0)}(s,t)-\mu^{(d,0)}(s,t)\right| &= O_{a.s}\left(h_1^{-d}\delta_{n1}(h_1,h_2) + h_1^{-d}\gamma_{1d}(h_1,h_2)\right).\end{aligned} \tag{34}$$

Remark 2. We discuss special cases of Theorem 1 under dense or sparse designs.

(1) For the designs that are sparse in both s and t: if both $\max_{1\le i\le n}M_i$ and $\max_{1\le i\le n,1\le m\le M_i}L_{im}$ are bounded, then $\gamma_{nkl}=O(1)$ for all $k=1,2$ and $l=0,1,2$. Thus, Theorem 1 implies that

$$\begin{aligned}\sup_{s\in\mathcal{S},t\in\mathcal{T}}|\widehat{\mu}(s,t)-\mu(s,t)| &= O_{a.s}\left(\sqrt{(nh_1h_2)^{-1}\log n} + \gamma_{1d}(h_1,h_2)\right),\\ \sup_{s\in\mathcal{S},t\in\mathcal{T}}\left|\widehat{\mu}^{(d,0)}(s,t)-\mu^{(d,0)}(s,t)\right| &= O_{a.s}\left(h_1^{-d}\sqrt{(nh_1h_2)^{-1}\log n} + h_1^{-d}\gamma_{1d}(h_1,h_2)\right).\end{aligned} \tag{35}$$

By choosing h_1 and h_2 satisfying that $\gamma_{1d}(h_1,h_2)=O(h_1h_2)=O((\log n/n)^{1/3})$, $\sup_{s\in\mathcal{S},t\in\mathcal{T}}|\widehat{\mu}(s,t)-\mu(s,t)|$ can achieve its optimal convergence rate as $O((\log n/n)^{1/3})$. Moreover, $\sup_{s\in\mathcal{S},t\in\mathcal{T}}|\widehat{\mu}^{(d,0)}(s,t)-\mu^{(d,0)}(s,t)|$ can also achieve its optimal convergence rate as $O_{a.s.}((\log n/n)^{2/(3d+6)})$ by further choosing $h_2=O(h_1^{d+1})$, that is $h_2=O(h_1^{d+1})=O((\log n/n)^{(d+1)/(3d+6)})$.

(2) For the designs that are dense in both s and t: if $\min_{1\le i\le n}M_i=M_n$ and $\min_{1\le i\le n,1\le m\le M_i}L_{im}=L_n$, satisfying $M_n^{-1}\lesssim h_2$ and $L_n^{-1}\lesssim h_1$, then for all $k=1,2$ and $l=0,1,2$, $\gamma_{nkl}\asymp M_n^{1-k}L_n^{-l}$ and $\delta_{n1}(h_1,h_2)\asymp\sqrt{\log n/n}$. Thus, Theorem 1 yields that

$$\begin{aligned}\sup_{s\in\mathcal{S},t\in\mathcal{T}}|\widehat{\mu}(s,t)-\mu(s,t)| &= O_{a.s}\left(\sqrt{\log n/n} + \gamma_{1d}(h_1,h_2)\right),\\ \sup_{s\in\mathcal{S},t\in\mathcal{T}}\left|\widehat{\mu}^{(d,0)}(s,t)-\mu^{(d,0)}(s,t)\right| &= O_{a.s}\left(h_1^{-d}\sqrt{\log n/n} + h_1^{-d}\gamma_{1d}(h_1,h_2)\right).\end{aligned} \tag{36}$$

Moreover, if the Assumption (5) is replaced by a strong version, in which we assume that $\sup_{s,t}|X(s,t)|$ and ε_{iml} are bounded, then if h_1 and h_2 satisfying $h_1h_2=O(\log n/n)^{1/2}$ and $h_2=O(h_1^{d+1})$, Theorem 1 implies that

$$\begin{aligned}\sup_{s\in\mathcal{S},t\in\mathcal{T}}|\widehat{\mu}(s,t)-\mu(s,t)| &= O_{a.s}\left(\sqrt{\log n/n}\right),\\ \sup_{s\in\mathcal{S},t\in\mathcal{T}}\left|\widehat{\mu}^{(d,0)}(s,t)-\mu^{(d,0)}(s,t)\right| &= O_{a.s}\left((\log n/n)^{1/(d+2)}\right).\end{aligned} \tag{37}$$

(3) For the designs that are sparse in s and dense in t: if $\max_{1\le i\le n,1\le m\le M_i}L_{im}$ is bounded and $\min_{1\le i\le n}M_i=M_n$, where $M_n^{-1}\lesssim h_2$, then $\delta_{n1}(h_1,h_2)\asymp\sqrt{\log n/nh_1}$, that is

$$\begin{aligned}\sup_{s\in\mathcal{S},t\in\mathcal{T}}|\widehat{\mu}(s,t)-\mu(s,t)| &= O_{a.s}\left(\sqrt{\log n/nh_1} + \gamma_{1d}(h_1,h_2)\right),\\ \sup_{s\in\mathcal{S},t\in\mathcal{T}}\left|\widehat{\mu}^{(d,0)}(s,t)-\mu^{(d,0)}(s,t)\right| &= O_{a.s}\left(h_1^{-d}\sqrt{\log n/nh_1} + h_1^{-d}\gamma_{1d}(h_1,h_2)\right).\end{aligned} \tag{38}$$

Furthermore, if $h_1=O((\log n/n)^{1/(2d+5)})$ and $h_2=O(h_1^{d+1})$, then Theorem 1 shows that

$$\begin{aligned}\sup_{s\in\mathcal{S},t\in\mathcal{T}}|\widehat{\mu}(s,t)-\mu(s,t)| &= O_{a.s}\left((\log n/n)^{(d+2)/(2d+5)}\right),\\ \sup_{s\in\mathcal{S},t\in\mathcal{T}}\left|\widehat{\mu}^{(d,0)}(s,t)-\mu^{(d,0)}(s,t)\right| &= O_{a.s}\left((\log n/n)^{2/(2d+5)}\right).\end{aligned} \tag{39}$$

(4) For the designs that are dense in s and sparse in t: if $\max_{1\le i\le n}M_i$ is bounded and $\min_{1\le i\le n,1\le m\le M_i}L_{im}=L_n$, where $L_n^{-1}\lesssim h_1$, then $\delta_{n1}(h_1,h_2)\asymp\sqrt{\log n/nh_2}$, that is

$$\sup_{s\in\mathcal{S},t\in\mathcal{T}} |\widehat{\mu}(s,t)-\mu(s,t)| = O_{a.s}\Big(\sqrt{\log n/nh_2} + \gamma_{1d}(h_1,h_2)\Big),$$
$$\sup_{s\in\mathcal{S},t\in\mathcal{T}} \left|\widehat{\mu}^{(d,0)}(s,t)-\mu^{(d,0)}(s,t)\right| = O_{a.s}\Big(h_1^{-d}\sqrt{\log n/nh_2} + h_1^{-d}\gamma_{1d}(h_1,h_2)\Big). \quad (40)$$

Furthermore, if $h_1 = O((\log n/n)^{1/(3d+5)})$ and $h_2 = O(h_1^{d+1})$, then Theorem 1 shows that

$$\sup_{s\in\mathcal{S},t\in\mathcal{T}} |\widehat{\mu}(s,t)-\mu(s,t)| = O_{a.s}\Big((\log n/n)^{(d+2)/(3d+5)}\Big),$$
$$\sup_{s\in\mathcal{S},t\in\mathcal{T}} \left|\widehat{\mu}^{(d,0)}(s,t)-\mu^{(d,0)}(s,t)\right| = O_{a.s}\Big((\log n/n)^{2/(3d+5)}\Big). \quad (41)$$

3.2. Uniform Convergence Rates of $\widehat{G}(s,u)$, $\widehat{G}^{(d,0)}(s,u)$, and $\widehat{\sigma^2}$. We next establish the convergence rates of $\widehat{G}(s,u)$, $\widehat{G}^{(d,0)}(s,u)$, and $\widehat{\sigma^2}$. Since $\widehat{G}(s,u)=\int_{\mathcal{T}}\{\widehat{D}(s,u,t)-\widehat{\mu}(s,t)\widehat{\mu}(u,t)\}dt$, integrating $\widehat{D}(s,u,t)$ and $\widehat{\mu}(s,t)\widehat{\mu}(u,t)$ over t results in extrasmoothing, which leads to a faster convergence rate $\delta_{n1}(h_1,1)$ and $\delta_{n2}(h_3,h_4,1)$ than $\delta_{n1}(h_1,1)$ and $\delta_{n2}(h_3,h_4,1)$, respectively. The similar conclusion holds for $\widehat{G}^{(d,0)}(s,u)$.

For $p=0,1,\ldots,2d+2$, $q=0,1,2$, and $r=0,1,2$, let

$$R_{npqr}(s,u,t) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}^*}\sum_{l\neq l'}^{L_{im}} K_{h_3}(S_{iml}-s) \cdot K_{h_4}(S_{iml'}-u)K_{h_5}(T_{im}-t)\left(\frac{S_{iml}-s}{h_3}\right)^p \cdot \left(\frac{S_{iml'}-u}{h_4}\right)^q\left(\frac{T_{im}-t}{h_5}\right)^r,$$
$$\widetilde{R}_{npqr}(s,u,t) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}^*}\sum_{l\neq l'}^{L_{im}} Y_{iml}Y_{iml'}K_{h_3}(S_{iml}-s) \cdot K_{h_4}(S_{iml'}-u)K_{h_5}(T_{im}-t)\left(\frac{S_{iml}-s}{h_3}\right)^p \cdot \left(\frac{S_{iml'}-u}{h_4}\right)^q\left(\frac{T_{im}-t}{h_5}\right)^r, \quad (42)$$

then we can obtain that

$$\widehat{\theta}^* = R_n^{-1}(s,u,t)\widetilde{R}_n(s,u,t), \quad (43)$$

with $\widetilde{R}_n(s,u,t) = (R_{n000}(s,u,t), R_{n100}(s,u,t),\ldots, R_{n(d+1)00}(s,u,t), R_{n010}(s,u,t), R_{n001}(s,u,t))^T$ and

$$R_n(s,u,t) = \begin{pmatrix} R_{n000} & R_{n100} & \cdots & R_{n(d+1)00} & R_{n010} & R_{n001} \\ R_{n100} & R_{n200} & \cdots & R_{n(d+2)00} & R_{n110} & R_{n101} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ R_{n(d+1)00} & R_{n(d+2)00} & \cdots & R_{n(2d+2)00} & R_{n(d+1)10} & R_{n(d+1)01} \\ R_{n010} & R_{n110} & \cdots & R_{n(d+1)10} & R_{n020} & R_{n011} \\ R_{n001} & R_{n100} & \cdots & R_{n(d+1)01} & R_{n011} & R_{n002} \end{pmatrix}(s,u,t). \quad (44)$$

Theorem 2. *Under assumptions (1)–(6) and (9), let $\gamma_{2d}(h_3,h_4,h_5) = h_3^{d+2}+h_4^2+h_5^2+h_3h_4+h_3h_5$, then*

$$\sup_{s,u\in\mathcal{S}} |\widehat{G}(s,u)-G(s,u)| = O_{a.s.}\big(\delta_{n2}(h_3,h_4,1) + \gamma_{2d}(h_3,h_4,h_5) + \delta_{n1}(h_1,1) + \gamma_{1d}(h_1,h_2)\big), \quad (45)$$

$$\sup_{s,u\in\mathcal{S}} \left|\widehat{G}^{(d,0)}(s,u)-G^{(d,0)}(s,u)\right| = O_{a.s}\Big(h_3^{-d}\big[\delta_{n2}(h_3,h_4,1) + \gamma_{2d}(h_3,h_4,h_5) + \delta_{n1}(h_1,1) + \gamma_{1d}(h_1,h_2)\big]\Big). \quad (46)$$

Remark 3. We discuss special cases of Theorem 2. Actually, whether the design is dense or sparse in t, the convergence rate in Theorem 2 is not affected. Hence, we only discuss different designs with respect to s.

(1) For the designs that are sparse in s: if $\max_{1\le i\le n, 1\le m\le M_i} L_{im}$ is bounded, then

$$\delta_{n1}(h_1,1) \asymp \sqrt{\log n/nh_1},\quad \delta_{n2}(h_3,h_4,1) \asymp \sqrt{\log n/nh_3h_4}. \tag{47}$$

On the contrary, since

$$\delta_{n2}(h_3,h_4,1) \asymp \delta_{n1}(h_3,h_4),\quad \gamma_{2d}(h_3,h_4,h_5) \asymp \gamma_{1d}(h_3,h_4) + \gamma_{1d}(h_3,h_5), \tag{48}$$

we choosing $h_3 = O((\log n/n)^{1/(3d+6)})$ and $h_4 \asymp h_5 = O(h_3^{d+1})$. Either h_1 or h_2 is chosen as in Remark 2(1) or Remark 2(3), Theorem 2 implies that

$$\sup_{s,u\in\mathcal{S}} |\widehat{G}(s,u) - G(s,u)| = O_{a.s.}\left((\log n/n)^{1/3}\right),\quad \sup_{s,u\in\mathcal{S}} \left|\widehat{G}^{(d,0)}(s,u) - G^{(d,0)}(s,u)\right| = O_{a.s}\left((\log n/n)^{2/(3d+6)}\right). \tag{49}$$

(2) For the designs that are dense in s: if $\min_{1\le i\le n, 1\le m\le M_i} L_{im} = L_n$, then

$$\delta_{n1}(h_1,1) \asymp \sqrt{(L_n^{-1}+h_1)\log n/nh_1},\quad \delta_{n2}(h_3,h_4,1) \asymp \sqrt{(L_n^{-2}+h_3h_4)\log n/nh_3h_4}. \tag{50}$$

Assume that $L_n^{-1} = o(h_1)$ and $L_n^{-2} = o(h_3h_4)$, where $h_1 \asymp h_3 = O((\log n/n)^{1/(2d+4)})$ and $h_4 \asymp h_5 \asymp h_2 = O(h_1^{d+1})$. If the Assumption (5) is replaced by a strong version, in which we assume that $\sup_{s,t}|X(s,t)|$ and ε_{iml} are bounded, then Theorem 2 implies that

$$\sup_{s,u\in\mathcal{S}} |\widehat{G}(s,u) - G(s,u)| = O_{a.s.}\left(\sqrt{\log n/n}\right),\quad \sup_{s,u\in\mathcal{S}} \left|\widehat{G}^{(d,0)}(s,u) - G^{(d,0)}(s,u)\right| = O_{a.s}\left((\log n/n)^{1/(d+2)}\right). \tag{51}$$

Theorem 3. *Under assumptions (1)–(7) and (9),*

$$\widehat{\sigma}^2 - \sigma^2 = O_{a.s.}\left(\delta_{n2}(h_3,1,1) + h_3^2 + h_5^2 + h_{v1}^2 + h_{v2}^2\right). \tag{52}$$

Remark 4. Same as Remark 3, we discuss the convergence rate of $\widehat{\sigma}^2$ under special cases.

(1) For the designs that are sparse in s: if $\max_{1\le i\le n, 1\le m\le M_i} L_{im}$ is bounded, then

$$\delta_{n2}(h_3,1,1) \asymp \sqrt{\log n/nh_3}, \tag{53}$$

which results in

$$\widehat{\sigma}^2 - \sigma^2 = O_{a.s.}\left(\sqrt{\log n/nh_3} + h_3^2 + h_5^2 + h_{v1}^2 + h_{v2}^2\right). \tag{54}$$

(2) For the designs that are dense in s: if $\min_{1\le i\le n, 1\le m\le M_i} L_{im} = L_n$, satisfying that $L_n^{-1} \lesssim h_3^{1/2}$, then

$$\delta_{n2}(h_3,1,1) \asymp \sqrt{\log n/n}. \tag{55}$$

If $h_3, h_5, h_{v1}, h_{v2} \lesssim (\log n/n)^{1/4}$, then

$$\widehat{\sigma}^2 - \sigma^2 = O_{a.s.}\left(\sqrt{\log n/n}\right). \tag{56}$$

3.3. Uniform Convergence Rates in FPCA. We next establish the convergence rates in FPCA. Let J be a fixed positive constant.

Theorem 4. *Under assumptions (1)–(9), for $j \le J$,*

(1) $\|\widehat{\psi}_j - \psi_j\| = O_{a.s.}(\delta_{n2}(h_4,1,1) + \delta_{n2}^2(h_3,h_4,1) + \gamma_{2d}(h_3,h_4,h_5) + \delta_{n1}(h_1,1) + \gamma_{1d}(h_1,h_2))$

(2) $\widehat{\lambda}_j - \lambda_j = O_{a.s.}((\log n/n)^{1/2} + \delta_{n2}^2(h_4,1,1) + \gamma_{2d}(h_3,h_4,h_5) + \delta_{n1}^2(h_1,1) + \gamma_{1d}(h_1,h_2))$

(3) $\sup_{s\in[0,a]} |\widehat{\psi}_j(s) - \psi_j(s)| = O_{a.s.}(\log n/n)^{1/2} + \delta_{n2}(h_4,1,1) + \delta_{n2}^2(h_3,h_4,1) + \gamma_{2d}(h_3,h_4,h_5) + \delta_{n1}(h_1,1) + \gamma_{1d}(h_1,h_2)$

(4) $\sup_{s\in[0,a]} \|\widehat{\psi}_j^{(d)}(s) - \psi_j^{(d)}(s)\| = O_{a.s.}((\log n/n)^{1/2} + \delta_{n2}^2(h_4,1,1) + \delta_{n1}^2(h_1,1) + h_3^{-d}(\delta_{n2}(h_4,1,1) + \gamma_{2d}(h_3,h_4,h_5) + \delta_{n1}(h_1,1) + \gamma_{1d}(h_1,h_2)))$

(5) For each $1 \le i \le n$,

$$\max_{1\le m\le M_i} \left|\widehat{\xi}_{ij}(T_{im}) - \xi_{ij}(T_{im})\right| \xrightarrow{a.s.} 0. \tag{57}$$

The consistency of $\widehat{\xi}_{ij}(t)$ guarantees the appropriacy of estimation procedures in Step 4. The proof of the theorems will be given in the Appendix.

4. Relative Issues

In this section, we discussed a few issues that are related to the implementation of our proposed methods.

4.1. Selection of Bandwidths. The performance of the estimators depends on the choice of bandwidths for $\mu(\cdot,\cdot)$ and $D(\cdot,\cdot,\cdot)$, and the best bandwidths vary with Ms and Ls. The bandwidth selection problem turns out to be very challenging and hence an important problem for future research. For lack of a better approach, we suggest picking the bandwidths by minimizing the integrated mean square error (IMSE). That is, for each function above, one calculated the IMSE over a range of h and selected the one that minimizes the IMSE.

4.2. Selection of Ks and J in the Overall Representations (25) and (26). In practice, the choice of the numbers of components J and K_js to be included in (25) can be based on the leave-one-curve-out cross-validation method [9] or by the fraction of variance explained (FVE) by the first J components [4]. One can also adopt AIC [10] or BIC [11] type of criteria, see Yao et al. [7] for one-dimensional function data.

For bivariate functional data, a pseudo-Gaussian log-likelihood is given by

$$\widehat{L}(J)=\sum_{i=1}^{n}\left\{-\frac{N_i}{2}\log(2\pi)-\frac{N_i}{2}\log\widehat{\sigma}^2-\frac{1}{2\widehat{\sigma}^2}\sum_{m=1}^{M_i}\sum_{l=1}^{L_{im}}\left[Y_{iml}-\widehat{\mu}(S_{iml},T_{im})-\sum_{j=1}^{J}\widehat{\xi}_j(T_{im})\widehat{\psi}_j(S_{iml})\right]^2\right\}, \tag{58}$$

where $N_i=\sum_{m=1}^{M_i}L_{im}$. One can choose J through minimizing $\text{aic}(J)=-\widehat{L}(J)+J$ (resp., $\text{bic}(J)=-\widehat{L}(J)+J\log n$) with respect to J.

For each $1\le j\le J$, define

$$\widehat{L}_j(K_j)=\sum_{i=1}^{n}\left\{-\frac{M_i}{2}\log(2\pi)-\frac{M_i}{2}\log\widehat{\sigma}^2-\frac{1}{2\widehat{\sigma}^2}\sum_{m=1}^{M_i}\left[\widehat{\xi}_{ij}(T_{im})-\sum_{k=1}^{K_j}\widehat{\zeta}_{ijk}\widehat{\phi}_{jk}(T_{im})\right]^2\right\}. \tag{59}$$

The number of components K_j is selected by minimizing $\text{aic}(K_j)=-\widehat{L}_j(K_j)+K_j$ (resp., $\text{bic}(K_j)=-\widehat{L}_j(K_j)+K_j\log n$) with respect to K_j.

Appendix

This is a five-part appendix organized as follows. Appendix A states some technical lemmas are needed for our main results. The proofs of these lemmas are not included here as they are lengthy and tedious. We provide them in an online supplementary material available online. Appendices B–E provide the proofs of Theorems 1–4, respectively.

A. Technical Lemmas

Some technical lemmas needed for our main results are shown as follows.

Lemma A.1. *Let $Z_{iml}=X_i(S_{iml},T_{im})$ or ε_{iml} for $i=1,\dots,n$, $m=1,\dots,M_i$, and $l=1,\dots,L_{im}$. Suppose for some $\lambda\in(2,\infty)$ that*

$$E\left(\sup_{s\in\mathcal{S},t\in\mathcal{T}}\left|X_i(S_{iml},T_{im})\right|^{\lambda}\right)<\infty,\qquad E|\varepsilon|^{\lambda}<\infty. \tag{A.1}$$

Define

$$\begin{aligned}D_n((s_1,s_2),(t_1,t_2))&=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}}\sum_{l=1}^{L_{im}}Z_{iml}I\\&\cdot(S_{iml}\in[s_1\wedge s_2,s_1\vee s_2])I\\&\cdot((T_{im}\in[t_1\wedge t_2,t_1\vee t_2]),\end{aligned} \tag{A.2}$$

and $D((s_1,s_2),(t_1,t_2))=E[D_n((s,s+u),(t,t+v))]$. Let c_n and c_n' be any positive sequences tending to 0 and $\beta_n=c_nc_n'(\gamma_{n21}+\gamma_{n20}c_n+\gamma_{n11}c_n'+c_nc_n')$, then if $\beta_n^{-1}(\log n/n)^{1-2/\lambda}=o(1)$, we have

$$\begin{aligned}&\sup_{\substack{s\in[0,a]\\t\in[0,b]}}\sup_{\substack{|u|\le c_n\\|v|\le c_n'}}\left|D_n((s,s+u),(t,t+v))-D((s,s+u),(t,t+v))\right|\\&\quad=O_{a.s.}\left(n^{-1/2}(\beta_n\log n)^{1/2}\right).\end{aligned} \tag{A.3}$$

The proof of Lemma A.1 is provided in the supplementary material for saving space.

Lemma A.2. *Let Z_{iml} be as in Lemma A.1 and assume that (A.1) holds. For bandwidths h_1 and h_2 and nonnegative integers p and q, let*

$$\begin{aligned}D_{npq}(s,t)&=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}}\sum_{l=1}^{L_{im}}Z_{iml}K_{h_1}\\&\cdot(S_{iml}-s)K_{h_2}(T_{im}-t)\left(\frac{S_{iml}-s}{h_1}\right)^p\left(\frac{T_{im}-t}{h_2}\right)^q.\end{aligned} \tag{A.4}$$

Let $\beta_n = h_1h_2(\gamma_{n21} + 2\gamma_{n20}h_1 + 2\gamma_{n11}h_2 + 4h_1h_2)$, *assume that* $h_1 \longrightarrow 0$, $h_2 \longrightarrow 0$, *and* $\beta_n^{-1}(\log n/n)^{1-2/\lambda} = o(1)$, *then we have*

$$\sqrt{nh_1^2h_2^2/(\beta_n\log n)}\sup_{\substack{s\in[0,a]\\ t\in[0,b]}}\left|D_{npq}(s,t) - E\left[D_{npq}(s,t)\right]\right| = O_{a.s.}(1). \tag{A.5}$$

The proof of Lemma A.2 is provided in the supplementary material for saving space.

Lemma A.3. *Let* $Z_{imll'} = X_i(S_{iml}, T_{im})X_i(S_{iml'}, T_{im})$, $X_i(S_{iml}, T_{im})\varepsilon_{iml'}$, *or* $\varepsilon_{iml}\varepsilon_{iml'}$ *for* $i = 1, \dots, n$, $m = 1, \dots, M_i$, *and* $l, l' = 1, \dots, L_{im}$. *Suppose for some* $\lambda \in (2, \infty)$ *that*

$$E\left(\sup_{s\in\mathcal{S},t\in\mathcal{T}}\left|X_i(S_{iml}, T_{im})\right|^{2\lambda}\right) < \infty, \qquad E|\varepsilon|^{2\lambda} < \infty. \tag{A.6}$$

Define

$$\begin{aligned} Q_n((s_1,s_2),(u_1,u_2),(t_1,t_2)) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}^*}\sum_{l\neq l'}^{L_{im}} Z_{imll'}I \\ \cdot (S_{iml}\in[s_1\wedge s_2, s_1\vee s_2])I \\ \cdot (S_{iml'}\in[u_1\wedge u_2, u_1\vee u_2])I \\ \cdot (T_{im}\in[t_1\wedge t_2, t_1\vee t_2]), \end{aligned} \tag{A.7}$$

and $Q((s_1, s_2), (u_1, u_2), (t_1, t_2)) = E[Q_n((s_1, s_2), (u_1, u_2), (t_1, t_2))]$. *Let* c_n, c_n', *and* c_n'' *be any positive sequences tending to 0 and* $\beta_n = c_nc_n'c_n''(\gamma_{n22} + \gamma_{n20}c_nc_n' + \gamma_{n12}c_n'' + c_nc_n'c_n'')$, *then if* $\beta_n^{-1}(\log n/n)^{1-2/\lambda} = o(1)$, *we have*

$$\begin{aligned} &\sup_{\substack{s,u\in[0,a]\\ t\in[0,b]}}\sup_{\substack{|v_1\le c_n|\\ |v_2\le c_n'|}}\left|Q_n((s,s+v_1),(u,u+v_2),(t,t+v_3))\right.\\ &\quad\left.-Q((s,s+v_1),(u,u+v_2),(t,t+v_3))\right|\\ &\quad= O_{a.s.}\left(n^{-1/2}(\beta_n\log n)^{1/2}\right). \end{aligned} \tag{A.8}$$

The proof of Lemma A.3 is provided in the supplementary material for saving space.

Lemma A.4. *Let* $Z_{imll'}$ *be as in Lemma A.3 and assume that (A.6) holds. For bandwidths* h_3, h_4, *and* h_5 *and nonnegative integers* p, q, *and* r, *let*

$$\begin{aligned} Q_{npqr}(s,u,t) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}^*}\sum_{l\neq l'}^{L_{im}} Z_{imll'}K_{h_3,p}(S_{iml}-s) \\ \cdot K_{h_4,q}(S_{iml'}-u)K_{h_5,r}(T_{im}-t), \end{aligned} \tag{A.9}$$

where $K_{h,p}(\cdot) = (\cdot/h)^pK_h(\cdot)$. *Let* $\beta_n = h_3h_4h_5(\gamma_{n22} + 4\gamma_{n20}h_3h_4 + 2\gamma_{n12}h_5 + 8h_3h_4h_5)$; *assume that* $h_3 \longrightarrow 0$, $h_4 \longrightarrow 0$, $h_5 \longrightarrow 0$, *and* $\beta_n^{-1}(\log n/n)^{1-2/\lambda} = o(1)$, *then we have*

$$\begin{aligned} &\sqrt{nh_3^2h_4^2h_5^2/(\beta_n\log n)}\sup_{\substack{s,u\in[0,a]\\ t\in[0,b]}}\left|Q_{npqr}(s,u,t) - E\left[Q_{npqr}(s,u,t)\right]\right|\\ &\quad= O_{a.s.}(1). \end{aligned} \tag{A.10}$$

The proof of Lemma A.4 is provided in the supplementary material for saving space.

Lemma A.5. *Let* Z_{iml} *be as in Lemma A.1 and assume that (A.1) holds. For bandwidths* h_1 *and nonnegative integers* p, *let*

$$\begin{aligned} \overline{D}_n &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}}\sum_{l=1}^{L_{im}} Z_{iml},\\ \widetilde{D}_{np}(s) &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}}\sum_{l=1}^{L_{im}} Z_{iml}K_{h_1}(S_{iml}-s)\left(\frac{S_{iml}-s}{h_1}\right)^p. \end{aligned} \tag{A.11}$$

Then, we have

$$\sqrt{n/(\log n)}\left|\overline{D}_n - E\overline{D}_n\right| = O_{a.s.}(1). \tag{A.12}$$

Let $\beta_n = h_1(\gamma_{n21} + 2\gamma_{n20} + 2\gamma_{n11}h_2 + 4h_2)$, *assume that* $h_1 \longrightarrow 0$ *and* $\beta_n^{-1}(\log n/n)^{1-2/\lambda} = o(1)$, *then we have*

$$\sqrt{nh_1^2/(\beta_n\log n)}\sup_{s\in[0,a]}\left|\widetilde{D}_{np}(s) - E\left[\widetilde{D}_{np}(s)\right]\right| = O_{a.s.}(1). \tag{A.13}$$

The proof of Lemma A.5 is provided in the supplementary material for saving space.

Lemma A.6. *Let* $Z_{imll'}$ *be as in Lemma A.3 and assume that (A.6) holds. For bandwidths* h_3 *and* h_4 *and nonnegative integers* p *and* q, *let*

$$\begin{aligned} \widetilde{Q}_{npq}(s,u) &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}^*}\sum_{l\neq l'}^{L_{im}} Z_{imll'}K_{h_3,p}\\ &\quad\cdot(S_{iml}-s)K_{h_4,q}(S_{iml'}-u),\\ \check{Q}_{np}(s) &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}^*}\sum_{l\neq l'}^{L_{im}} Z_{imll'}K_{h_3,p}(S_{iml}-s),\\ \overline{Q}_n &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}^*}\sum_{l\neq l'}^{L_{im}} Z_{imll'}K_{h_3}(S_{iml}-S_{iml'}+uh_3), \end{aligned} \tag{A.14}$$

where $K_{h,p}(\cdot) = (\cdot/h)^pK_h(\cdot)$.

Let $\beta_n = h_3h_4(\gamma_{n22} + 4\gamma_{n20}h_3h_4 + 2\gamma_{n12} + 8h_3h_4)$, *assume that* $h_3 \longrightarrow 0$, $h_4 \longrightarrow 0$, *and* $\beta_n^{-1}(\log n/n)^{1-2/\lambda} = o(1)$, *then we have*

$$\sqrt{nh_3^2h_4^2/(\beta_n\log n)}\sup_{s,u\in[0,a]}\left|\widetilde{Q}_{npq}(s,u)-E\left[\widetilde{Q}_{npq}(s,u)\right]\right|=O_{a.s.}(1). \tag{A.15}$$

Let $\widetilde{\beta}_n = h_3(\gamma_{n22}+4\gamma_{n20}h_3+2\gamma_{n12}+8h_3)$, *assume that* $h_3 \longrightarrow 0$ *and* $\widetilde{\beta}_n^{-1}(\log n/n)^{1-2/\lambda}=o(1)$, *then we have*

$$\sqrt{nh_3^2/(\widetilde{\beta}_n\log n)}\sup_{s,u\in[0,a]}\left|\check{Q}_{np}(s)-E\left[\widetilde{Q}_{np}(s)\right]\right|=O_{a.s.}(1), \tag{A.16}$$

$$\sqrt{nh_3^2/(\widetilde{\beta}_n\log n)}\left|\overline{Q}_n-E\overline{Q}_n\right|=O_{a.s.}(1). \tag{A.17}$$

The proof of Lemma A.6 is provided in the supplementary material for saving space.

B. Proof of Theorem 1

Recall that $\widehat{\theta}=S_n^{-1}(s,t)\widetilde{S}_n(s,t)$ and $\widehat{\theta}-\theta=S_n^{-1}(s,t)[\widetilde{S}_n(s,t)-S_n(s,t)\theta]$. Note that we can write

$$\widetilde{S}_n(s,t)-S_n(s,t)\theta=\left(S_{n00}^*(s,t),S_{n10}^*(s,t),\ldots,S_{n(d+1)0}^*(s,t),S_{n01}^*(s,t)\right)^T, \tag{B.1}$$

where for $p=0,1,\ldots,(d+1)$, $q=0,1$,

$$\begin{aligned}S_{npq}^*(s,t)=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}}\sum_{l=1}^{L_{im}}K_{h_1,p}(S_{iml}-s)K_{h_2,q}(T_{im}-t)\Bigg\{\varepsilon_{iml}+\left[X_i(S_{iml},T_{im})-\mu(S_{iml},T_{im})\right]\\+\left[\mu(S_{iml},T_{im})-\mu(s,t)-\sum_{k=1}^{d+1}\frac{1}{k!}\mu^{(k,0)}(s,t)(S_{iml}-s)^k-\mu^{(0,1)}(s,t)(T_{im}-t)\right]\Bigg\}.\end{aligned} \tag{B.2}$$

By Taylor's expansion and Lemma A.2, uniformly in $(s,t)\in[0,a]\times[0,b]$,

$$\begin{aligned}S_{npq}^*(s,t)&=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}}\sum_{l=1}^{L_{im}}K_{h_1,p}(S_{iml}-s)K_{h_2,q}(T_{im}-t)\\&\quad\{\varepsilon_{iml}+[X_i(S_{iml},T_{im})-\mu(S_{iml},T_{im})]\}\\&\quad+O\left(h_1^{(d+2)}+h_1h_2+h_2^2\right)\\&=O_{a.s.}\left(n^{-1/2}h_1^{-1}h_2^{-1}(\beta_n\log n)^{1/2}+h_1^{(d+2)}+h_1h_2+h_2^2\right),\end{aligned} \tag{B.3}$$

where $\beta_n=h_1h_2(\gamma_{n21}+2\gamma_{n20}h_1+2\gamma_{n11}h_2+4h_1h_2)$, from which we conclude that, uniformly in $(s,t)\in[0,a]\times[0,b]$,

$$\widetilde{S}_n(s,t)-S_n(s,t)\theta=O_{a.s.}\left(\delta_{n1}(h_1,h_2)+h_1^{(d+2)}+h_1h_2+h_2^2\right). \tag{B.4}$$

We next consider $S_n^{-1}(s,t)$. For any interior point $(s,t)\in[h_1,a-h_1]\times[h_2,b-h_2]$,

$$\begin{aligned}E\left[S_{npq}(s,t)\right]&=\int\int K(u)K(v)u^pv^qf(s+h_1u,t+h_2v)\mathrm{d}u\,\mathrm{d}v\\&=f(s,t)\nu_p\nu_q+O(h_1)\nu_{p+1}\nu_q+O(h_2)\nu_p\nu_{q+1}\\&\quad+O\left(h_1^2\right)\nu_{p+2}\nu_q+O\left(h_2^2\right)\nu_p\nu_{q+2}\\&\quad+O(h_1h_2)\nu_{p+1}\nu_{q+1},\end{aligned} \tag{B.5}$$

where $f(\cdot,\cdot)$ is the joint density of (S,T) and $\nu_p=\int u^pK(u)\mathrm{d}u$. Since $K(\cdot)$ is symmetric, we can further obtain that

$$E\left[S_{npq}(s,t)\right]=\begin{cases}f(s,t)\nu_p\nu_q+O\left(h_1^2+h_2^2\right), & \text{if both } p \text{ and } q \text{ are even numbers,}\\ O(h_1h_2)\nu_{p+1}\nu_{q+1}, & \text{if both } p \text{ and } q \text{ are odd numbers,}\\ O(h_2)\nu_p\nu_{q+1}, & \text{if } p \text{ is even and } q \text{ is odd numbers,}\\ O(h_1)\nu_{p+1}\nu_q, & \text{if } p \text{ is odd and } q \text{ is even numbers,}\end{cases} \tag{B.6}$$

and hence, uniformly for $(s,t) \in [h_1, a-h_1] \times [h_2, b-h_2]$,

$$E[S_n(s,t)] = f(s,t)\operatorname{diag}(\Gamma, \nu_2) + O(h_1+h_2) \equiv f(s,t)\Omega_1^{-1} + O(h_1+h_2). \tag{B.7}$$

Thus, uniformly for $(s,t) \in [h_1, a-h_1] \times [h_2, b-h_2]$, $\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta} = O_{a.s.}(\delta_{n1}(h_1,h_2) + h_1^{(d+2)} + h_1h_2 + h_2^2)$. The same rate can be achieved for boundary points. Note that

$$\mu(s,t) = \theta_{00} \text{ and } \mu^{(d,0)}(s,t) = h_1^{-d} d!\theta_{10}, \tag{B.8}$$

and thus Theorem 1 holds.

C. Proof of Theorem 2

Recall that

$$\widehat{G}_{\mathcal{S}}(s,u) = \int_{\mathcal{T}} \left\{\widehat{D}(s,u,t) - \widehat{\mu}(s,t)\widehat{\mu}(u,t)\right\} dt,$$
$$\widehat{G}_{\mathcal{S}}^{(d,0)}(s,u) = \int_{\mathcal{T}} \left\{\widehat{D}^{(d,0,0)}(s,u,t) - \widehat{\mu}^{(d,0)}(s,t)\widehat{\mu}(u,t)\right\} dt. \tag{C.1}$$

To bound $\sup_{(s,u)\in[0,a]^2} |\widehat{G}_{\mathcal{S}}(s,u) - G_{\mathcal{S}}(s,u)|$ and $\sup_{(s,u)\in[0,a]^2} |\widehat{G}_{\mathcal{S}}^{(d,0)}(s,u) - G_{\mathcal{S}}^{(d,0)}(s,u)|$, we consider $\widehat{D}(s,u,t) - D(s,u,t)$ and $\widehat{D}^{(d,0,0)}(s,u,t) - D^{(d,0,0)}(s,u,t)$ first.

Note that

$$\widehat{\theta}^* = R_n^{-1}(s,u,t)\widetilde{R}_n(s,u,t),$$
$$\widehat{\theta}^* - \theta^* = R_n^{-1}(s,u,t)\left[\widetilde{R}_n(s,u,t) - R_n(s,u,t)\theta^*\right]. \tag{C.2}$$

Now write

$$R_n^*(s,u,t) = \widetilde{R}_n(s,u,t) - R_n(s,u,t)\theta^* = \left(R_{n000}^*(s,u,t), \dots, R_{n(d+1)00}^*(s,u,t), R_{n010}^*(s,u,t), R_{n001}^*(s,u,t)\right)^T, \tag{C.3}$$

where for $p = 0,1,\dots,(d+1)$, $q = 0,1$ and $r = 0,1$.

$$\begin{aligned} R_{npqr}^*(s,u,t) = {} & \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}^*}\sum_{l\neq l'}^{L_{im}} K_{h_3,p}(S_{iml}-s)K_{h_4,q}(S_{iml'}-u)K_{h_5,r}(T_{im}-t)\{Y_{iml}Y_{iml'} - D(S_{iml},S_{iml'},T_{im}) + D(S_{iml},S_{iml'},T_{im}) \\ & - D(s,u,t) - \sum_{k=1}^{d+1}\frac{1}{k!}D^{(k,0,0)}(s,u,t)(S_{iml}-s)^k - D^{(0,1,0)}(s,u,t)(S_{iml'}-u) - D^{(0,0,1)}(s,u,t)(T_{im}-t)\Big\}. \end{aligned} \tag{C.4}$$

By Taylor's expansion and Lemma A.4, uniformly in $(s,u,t) \in [0,a]^2 \times [0,b]$,

$$\begin{aligned} R_{npqr}^*(s,u,t) = {} & \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}^*}\sum_{l\neq l'}^{L_{im}} K_{h_3,p}(S_{iml}-s)K_{h_4,q}(S_{iml'}-u)K_{h_5,r}(T_{im}-t)\left[Y_{iml}Y_{iml'} - D(S_{iml},S_{iml'},T_{im})\right] \\ & + O\left(h_3^{d+2} + h_4^2 + h_5^2 + h_3h_4 + h_3h_5\right) \\ = {} & O_{a.s.}\left(n^{-1/2}h_3^{-1}h_4^{-1}h_5^{-1}(\beta_n \log n)^{1/2} + h_3^{d+2} + h_4^2 + h_5^2 + h_3h_4 + h_3h_5\right), \end{aligned} \tag{C.5}$$

where $\beta_n = h_3h_4h_5(\gamma_{n22} + 4\gamma_{n20}h_3h_4 + 2\gamma_{n12}h_5 + 8h_3h_4h_5)$. Thus, uniformly in $(s,u,t) \in [0,a]^2 \times [0,b]$,

$$\widetilde{R}_n(s,u,t) - R_n(s,u,t)\theta^* = O_{a.s.}\left(\delta_{n2}(h_3,h_4,h_5) + \gamma_{2d}(h_3,h_4,h_5)\right). \tag{C.6}$$

We next consider $R_n^{-1}(s,u,t)$. For any interior point $(s,u,t) \in [h_3, a-h_3] \times [h_4, a-h_4] \times [h_5, b-h_5]$,

$$\begin{aligned} E\left[R_{npqr}(s,u,t)\right] &= \iiint K(v)K(w)K(z)v^p w^q z^r f_2 \cdot (s+h_3v, u+h_4w, t+h_5z)\, dv\, dw\, dz \\ &= f_2(s,u,t)\nu_p\nu_q\nu_r + O(h_3+h_4+h_5), \end{aligned} \tag{C.7}$$

where $f_2(\cdot,\cdot,\cdot)$ is the joint density function of S_1, S_2, and T, and is bounded away form 0. Hence, uniformly for $(s,u,t) \in [h_3, a-h_3] \times [h_4, a-h_4] \times [h_5, b-h_5]$,

$$E[R_n(s,u,t)] = f_2(s,u,t)\mathrm{diag}(\Gamma,\nu_2,\nu_2) + O(h_3+h_4+h_5) \equiv f_2(s,u,t)\Omega_2^{-1} + O(h_3+h_4+h_5), \tag{C.8}$$

and thus,

$$\widehat{\theta}^* - \theta^* = O_{a.s.}\left(\delta_{n2}(h_3,h_4,h_5) + \gamma_{2d}(h_3,h_4,h_5)\right). \tag{C.9}$$

The same rate can achieve for boundary points, from which we conclude that, uniformly for $(s,u,t) \in [0,a]^2 \times [0,b]$,

$$\widehat{D}(s,u,t) - D(s,u,t) = O_{a.s.}\left(\delta_{n2}(h_3,h_4,h_5) + \gamma_{2d}(h_3,h_4,h_5)\right), \tag{C.10}$$

$$\widehat{D}^{(d,0,0)}(s,u,t) - D^{(d,0,0)}(s,u,t) = O_{a.s.}\left(h_3^{-d}\delta_{n2}(h_3,h_4,h_5) + h_3^{-d}\gamma_{2d}(h_3,h_4,h_5)\right). \tag{C.11}$$

Now we consider $\widehat{G}_{\mathcal{S}}(s,u)$ and $\widehat{G}_{\mathcal{S}}^{(d,0)}(s,u)$. Recall that

$$\widehat{G}_{\mathcal{S}}(s,u) = \int_{\mathcal{T}}\left\{\widehat{D}(s,u,t) - \widehat{\mu}(s,t)\widehat{\mu}(u,t)\right\}dt,$$
$$\widehat{G}_{\mathcal{S}}^{(d,0)}(s,u) = \int_{\mathcal{T}}\left\{\widehat{D}^{(d,0,0)}(s,u,t) - \widehat{\mu}^{(d,0)}(s,t)\widehat{\mu}(u,t)\right\}dt. \tag{C.12}$$

Since

$$\widehat{G}_{\mathcal{S}}(s,u) - G_{\mathcal{S}}(s,u) = \int_{\mathcal{T}}\left\{\widehat{D}(s,u,t) - D(s,u,t)\right\}dt - \int_{\mathcal{T}}\left\{\widehat{\mu}(s,t)\widehat{\mu}(u,t) - \mu(s,t)\mu(u,t)\right\}dt, \tag{C.13}$$

we consider the first and the second terms on the right-hand side of (C.13) separately. Now recall that $\widehat{D}(s,u,t) - D(s,u,t)$ is the first component of $R_n^{-1}(s,u,t)R_n^*(s,u,t)$, and then by (C.8) that uniformly for $(s,u,t) \in [0,a]^2 \times [0,b]$,

$$\widehat{D}(s,u,t) - D(s,u,t) = \left\{f_2^{-1}(s,u,t)\left(\omega_{11},\ldots,\omega_{1,d+3}\right) + O_{a.s}\left(\delta_{n2}(h_3,h_4,h_5) + h_3 + h_4 + h_5\right)\right\}R_n^*(s,u,t), \tag{C.14}$$

where ω_{jk} is the (j,k)th component of Ω_2. Note that

$$\int f_2^{-1}(s,u,t)R_{npqr}^*(s,u,t)dt \le \left[\inf_{s,u\in[0,a],t\in[0,b]} f_2(s,u,t)\right]^{-1}\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}^*}\sum_{l\neq l'}^{L_{im}}\int K_{h_5,r}(T_{im}-t)dt\,K_{h_3,p}(S_{iml}-s)K_{h_4,q}(S_{iml'}-u) \cdot \left[Y_{iml}Y_{iml'} - D(S_{iml},S_{iml'},T_{im})\right]\right] + O_{a.s}\left(\gamma_{2d}(h_3,h_4,h_5)\right)\Big\}. \tag{C.15}$$

Since by Lemma A.6, uniformly for $(s,u) \in [0,a]^2$,

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}^*}\sum_{l\neq l'}^{L_{im}}\left[\int K_{h_5,r}(T_{im}-t)dt\,K_{h_3,p}(S_{iml}-s)K_{h_4,q}(S_{iml'}-u)\left[Y_{iml}Y_{iml'} - D(S_{iml},S_{iml'},T_{im})\right]\right]\\&= \frac{-\nu_r}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L_{im}^*}\sum_{l\neq l'}^{L_{im}}K_{h_3,p}(S_{iml}-s)K_{h_4,q}(S_{iml'}-u)\left[Y_{iml}Y_{iml'} - D(S_{iml},S_{iml'},T_{im})\right]\\&= O_{a.s.}\left(\delta_{n2}(h_3,h_4,1)\right),\end{aligned} \tag{C.16}$$

it follows from (C.14) that, uniformly for $(s,u) \in [0,a]^2$,

$$\int_{\mathcal{T}}\left\{\widehat{D}(s,u,t) - D(s,u,t)\right\}dt = O_{a.s.}\left(\delta_{n2}(h_3,h_4,1) + \gamma_{2d}(h_3,h_4,h_5)\right). \tag{C.17}$$

We next look into the second term on the right-hand side of (C.13). By Lemma A.5 and the similar derivation leading to (C.17), uniformly for $(s,u) \in [0,a]^2$,

$$\begin{aligned}\left|\iint_{\mathcal{T}} \{\widehat{\mu}(s,t)\widehat{\mu}(u,t) - \mu(s,t)\mu(u,t)\}dt\right| &\le \left|\iint_{\mathcal{T}} \widehat{\mu}(s,t)\{\widehat{\mu}(u,t) - \mu(u,t)\}dt\right| + \left|\iint_{\mathcal{T}} \{\widehat{\mu}(s,t) - \mu(s,t)\}\mu(u,t)dt\right| \\ &\le O(1) \sup_{(s,t)\in[0,a]\times[0,b]} |\widehat{\mu}(s,t)| \left|\int [\widehat{\mu}(u,t) - \mu(u,t)]dt\right| \\ &+ O(1) \sup_{(u,t)\in[0,a]\times[0,b]} |\mu(u,t)| \left|\int [\widehat{\mu}(s,t) - \mu(s,t)]dt\right| \\ &= O_{a.s}(\delta_{n1}(h_1,1) + \gamma_{1d}(h_1,h_2)).\end{aligned} \tag{C.18}$$

Thus, combining (C.13), (C.17), and (C.18) leading to (E.6), which is

$$\begin{aligned}\sup_{(s,u)\in[0,a]^2} \left|\widehat{G}_{\mathcal{S}}(s,u) - G_{\mathcal{S}}(s,u)\right| = O_{a.s.}(\delta_{n2}(h_3,h_4,1) \\ + \gamma_{2d}(h_3,h_4,h_5) + \delta_{n1}(h_1,1) \\ + \gamma_{1d}(h_1,h_2)).\end{aligned} \tag{C.19}$$

By mirroring the derivations above, we can prove (E.8), which is

$$\begin{aligned}\sup_{(s,u)\in[0,a]^2} \left|\widehat{G}_{\mathcal{S}}^{(d,0)}(s,u) - G_{\mathcal{S}}^{(d,0)}(s,u)\right| = O_{a.s.}(h_3^{-d}[\delta_{n2}(h_3,h_4,1) \\ + \gamma_{2d}(h_3,h_4,h_5) + \delta_{n1}(h_1,1) \\ + \gamma_{1d}(h_1,h_2)]).\end{aligned} \tag{C.20}$$

Theorem 2 holds.

D. Proof of Theorem 3

Let Δ be the integral operator with kernel $\widehat{G}_{\mathcal{S}}(s,u) - G_{\mathcal{S}}(s,u)$. The following Lemma A.7 is needed for the proof of Theorem 3 and Theorem 4.

Lemma A.7. *For any bounded measurable function ψ on $[0,a]$,*

$$\begin{aligned}\sup_{u\in[0,a]} |(\Delta\psi)(u)| O_{a.s.}(\delta_{n2}(h_4,1,1) \\ + \gamma_{2d}(h_3,h_4,h_5) + \delta_{n1}(h_1,1) \\ + \gamma_{1d}(h_1,h_2)).\end{aligned} \tag{D.1}$$

The proof of Lemma A.7 is provided in the supplementary material.

Proof of Theorem 3. Since for all $(s,t) \in [0,a] \times [0,b]$, $V(s,t) = D(s,s,t) + \sigma^2$,

$$\begin{aligned}\widehat{\sigma}^2 - \sigma^2 &= \frac{1}{ab}\int_0^a \int_0^b \left\{\widehat{V}(s,t) - V(s,t)\right\}dt\, ds \\ &- \frac{1}{ab}\int_0^a \int_0^b \left\{\widehat{D}(s,s,t) - D(s,s,t)\right\}dt\, ds.\end{aligned} \tag{D.2}$$

Note that $\int_0^a \int_0^b \{\widehat{D}(s,s,t) - D(s,s,t)\}dt\, ds$ is a special case of B_{n1} in the proof of Lemma A.7 with $h_3 = h_4$, $d = 0$, and $\psi(\cdot) \equiv 1$, then according to the proof of Lemma A.7, we can obtain that $\int_0^a \int_0^b \{\widehat{D}(s,s,t) - D(s,s,t)\}dt\, ds$ has the same rate as $\int_0^a \int_0^b R_n^*(s,s,t)dt\, ds$ with $d = 0$ and $h_3 = h_4$.

According to the expression (C.5) of $R_{npqr}^*(s,u,t)$, we can obtain that

$$\begin{aligned}\int_0^a \int_0^b R_{npqr}^*(s,s,t)dt\, ds = \Bigg\{\frac{1}{n}\sum_{i=1}^{n} \frac{1}{M_i}\sum_{m=1}^{M_i} \frac{1}{L_{im}^*}\sum_{l\neq l'}^{L_{im}} \Bigg[\int_0^a K_{h_3,p}(S_{iml} - s)K_{h_3,q}(S_{iml'} - s)ds \int_0^b K_{h_5,r}(T_{im} - t)dt \\ \cdot [Y_{iml}Y_{iml'} - D(S_{iml}, S_{iml'}, T_{im})]\Bigg] + O_{a.s}(\gamma_{20}(h_3,h_3,h_5))\Bigg\}.\end{aligned} \tag{D.3}$$

Obviously, $\int_0^a K_{h_3,p}(S_{iml}-s)K_{h_3,q}(S_{iml'}-s)ds = O(h_3^{-1})$ and $\int_0^b K_{h_5,r}(T_{im}-t)dt = -\nu_r$, which together with Lemma A.6 leads to that

$$\int_0^a\int_0^b R^*_{npqr}(s,s,t)dt\,ds = \int\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{m=1}^{M_i}\frac{1}{L^*_{im}}\sum_{l\neq l'}^{L_{im}}\left[Y_{iml}Y_{iml'} - D(S_{iml},S_{iml'},T_{im})\right]K_{h_3}(S_{iml}-S_{iml'}+uh_3)\right\}du$$
$$+O_{a.s}(\gamma_{20}(h_3,h_3,h_5))$$
$$= O_{a.s}(\delta_{n3}(h_3,1,1)+\gamma_{20}(h_3,h_3,h_5)), \quad \text{(D.4)}$$

which is also the rate of $\int_0^a\int_0^b\{\widehat{D}(s,s,t)-D(s,s,t)\}dt\,ds$.

On the contrary, based on the definition of $\widehat{V}(s,t)$, by applying the similar proof of Lemma A.1 and A.2, Lemma 5, and Theorem 2 to $\int_0^a\int_0^b\widehat{V}(s,t)dt\,ds$, it is easy to show that

$$\int_0^a\int_0^b\{\widehat{V}(s,t)-V(s,t)\}dt\,ds = O_{a.s.}\left(\sqrt{\log n/n}+\gamma_{10}(h_{\nu_1},h_{\nu_2})\right). \quad \text{(D.5)}$$

Since $\gamma_{10}(h_{\nu_1},h_{\nu_2}) = h_{\nu1}^2+h_{\nu2}^2$ and $\gamma_{20}(h_3,h_3,h_5) = h_3^2+h_5^2$, combing (D.2)–(D.5) leads to Theorem 3. □

E. Proof of Theorem 4

(1) By the L^2 expansion [12] and Bessel's inequality, we have for some constant $C>0$:

$$\left\|\widehat{\psi}_j-\psi_j\right\| \le C\left(\left\|\Delta\psi_j\right\|+\|\Delta\|^2\right), \quad \text{(E.1)}$$

where $\|\Delta\| = \left\{\int\int[\widehat{G}_{\mathcal{S}}(s,u)-G_{\mathcal{S}}(s,u)]^2ds\,du\right\}^{1/2}$ is the Hilbert–Schmidt norm of Δ. Then, it follows form Theorem 2 and Lemma A.7 that

$$\left\|\Delta\psi_j\right\| = O_{a.s.}(\delta_{n2}(h_4,1,1)+\gamma_{2d}(h_3,h_4,h_5)$$
$$+\delta_{n1}(h_1,1)+\gamma_{1d}(h_1,h_2)), \quad \text{(E.2)}$$

$$\|\Delta\|^2 = O_{a.s.}\left(\delta_{n2}^2(h_3,h_4,1)+\gamma_{2d}^2(h_3,h_4,h_5)\right.$$
$$\left.+\delta_{n1}^2(h_1,1)+\gamma_{1d}^2(h_1,h_2)\right), \quad \text{(E.3)}$$

and hence

$$\left\|\widehat{\psi}_j-\psi_j\right\| = O_{a.s.}\left(\delta_{n2}(h_4,1,1)+\delta_{n2}^2(h_3,h_4,1)\right.$$
$$\left.+\gamma_{2d}(h_3,h_4,h_5)+\delta_{n1}(h_1,1)+\gamma_{1d}(h_1,h_2)\right). \quad \text{(E.4)}$$

Theorem 4(1) holds.

(2) By (4.9) in Hall et al. [13],

$$\widehat{\lambda}_j-\lambda_j = \int\int\left[\widehat{G}_{\mathcal{S}}(s,u)-G_{\mathcal{S}}(s,u)\right]\psi_j(s)\psi_j(u)ds\,du+O\left(\left\|\Delta\psi_j\right\|^2\right)$$
$$= \int_0^a\int_0^a\int_0^b[\widehat{D}(s,u,t)-D(s,u,t)]dt\,\psi(s)ds\,\psi_j(u)du$$
$$-\int_0^a\int_0^a\int_0^b[\widehat{\mu}(s,t)\widehat{\mu}(u,t)-\mu(s,t)\mu(u,t)]dt\,\psi(s)ds\,\psi_j(u)du+O\left(\left\|\Delta\psi_j\right\|^2\right)$$
$$\equiv \widetilde{B}_{n1}-\widetilde{B}_{n2}+O\left(\left\|\Delta\psi_j\right\|^2\right). \quad \text{(E.5)}$$

Similarly to the argument leading to the rate of B_{n1} in Lemma A.7, we can obtain that

$$\widetilde{B}_{n1} = O_{a.s}\left([\log n/n]^{1/2}+\gamma_{2d}(h_3,h_4,h_5)\right). \quad \text{(E.6)}$$

Next, we write

$$\widetilde{B}_{n2} = \int_0^a\int_0^a\int_0^b\{\widehat{\mu}(s,t)\widehat{\mu}(u,t)-\mu(s,t)\mu(u,t)\}dt\,\psi(s)ds\,\psi_j(u)du$$
$$\le \int_0^a\int_0^a\int_0^b\widehat{\mu}(s,t)\{\widehat{\mu}(u,t)-\mu(u,t)\}dt\,\psi_j(s)ds\,\psi_j(u)du$$
$$+\int_0^a\int_0^b\{\widehat{\mu}(s,t)-\mu(s,t)\}\mu(u,t)dt\,\psi_j(s)ds\,\psi_j(u)du, \quad \text{(E.7)}$$

and similarly, we can show that

$$\widetilde{B}_{n2} = O_{a.s}\left([\log n/n]^{1/2} + \gamma_{1d}(h_1, h_2)\right). \tag{E.8}$$

By combining (E.2) and (E.5)–(E.8), Theorem 4(2) is proved.

(3) For any $u \in [0, a]$,

$$\begin{aligned}\widehat{\lambda}_j\widehat{\psi}_j(u) - \lambda_j\psi_j(u) &= \int \widehat{G}_{\mathcal{S}}(s,u)\widehat{\psi}_j(s)\mathrm{d}s - \int G_{\mathcal{S}}(s,u)\psi_j(s)\mathrm{d}s \\ &= \int \left[\widehat{G}_{\mathcal{S}}(s,u) - G_{\mathcal{S}}(s,u)\right]\psi_j(s)\mathrm{d}s \\ &\quad + \int \widehat{G}_{\mathcal{S}}(s,u)\left[\widehat{\psi}_j(s) - \psi_j(s)\right]\mathrm{d}s.\end{aligned} \tag{E.9}$$

By the Cauchy–Schwarz inequality, uniformly for all $u \in [0, a]$,

$$\begin{aligned}&\left|\int \widehat{G}_{\mathcal{S}}(s,u)\left[\widehat{\psi}_j(s) - \psi_j(s)\right]\mathrm{d}s\right| \\ &\leq \left[\int \widehat{G}^2_{\mathcal{S}}(s,u)\mathrm{d}s\right]^{1/2}\left\|\widehat{\psi}_j - \psi_j\right\| \\ &\leq \sup_{s,u\in[0,a]} \widehat{G}^2_{\mathcal{S}}(s,u)\left\|\widehat{\psi}_j - \psi_j\right\| = O_{a.s.}\left(\left\|\widehat{\psi}_j - \psi_j\right\|\right).\end{aligned} \tag{E.10}$$

Thus, by Lemma A.7, we have uniformly for all $u \in [0, a]$:

$$\begin{aligned}\widehat{\lambda}_j\widehat{\psi}_j(u) - \lambda_j\psi_j(u) = O_{a.s.}\big(&\delta_{n2}(h_4, 1, 1) + \delta^2_{n2}(h_3, h_4, 1) \\ &+ \gamma_{2d}(h_3, h_4, h_5) + \delta_{n1}(h_1, 1) \\ &+ \gamma_{1d}(h_1, h_2)\big).\end{aligned} \tag{E.11}$$

By the triangle inequality and Theorem 4(2),

$$\begin{aligned}\sup_{u\in[0,a]} \lambda_j\left|\widehat{\psi}_j(u) - \psi_j(u)\right| &\leq \sup_{u\in[0,a]}\left|\widehat{\lambda}_j\widehat{\psi}_j(u) - \lambda_j\psi_j(u)\right| \\ &\quad + \left|\widehat{\lambda}_j - \lambda_j\right| \sup_{u\in[0,a]}\left|\widehat{\psi}_j(u)\right| \\ &= O_{a.s.}\big((\log n/n)^{1/2} + \delta_{n2}(h_4, 1, 1) \\ &\quad + \delta^2_{n2}(h_3, h_4, 1) + \gamma_{2d}(h_3, h_4, h_5) \\ &\quad + \delta_{n1}(h_1, 1) + \gamma_{1d}(h_1, h_2)\big).\end{aligned} \tag{E.12}$$

Theorem 4(3) holds.

(4) Note that for all $s \in [0, a]$,

$$\begin{aligned}\widehat{\psi}^{(d)}_j(s) - \psi^{(d)}_j(s) &= \widehat{\lambda}^{-1}_j\int_{\mathcal{S}}\widehat{G}^{(d,0)}_{\mathcal{S}}(s,u)\widehat{\psi}_j(u)\mathrm{d}u - \lambda^{-1}_j\int_{\mathcal{S}}G^{(d,0)}_{\mathcal{S}}(s,u)\psi_j(u)\mathrm{d}u \\ &= \widehat{\lambda}^{-1}_j\left[\int_{\mathcal{S}}\widehat{G}^{(d,0)}_{\mathcal{S}}(s,u)\widehat{\psi}_j(u)\mathrm{d}u - \int_{\mathcal{S}}G^{(d,0)}_{\mathcal{S}}(s,u)\psi_j(u)\mathrm{d}u\right] + \left[\lambda^{-1}_j - \widehat{\lambda}^{-1}_j\right]\int_{\mathcal{S}}G^{(d,0)}_{\mathcal{S}}(s,u)\psi_j(u)\mathrm{d}u \\ &= \widehat{\lambda}^{-1}_j\left\{\int\left[\widehat{G}^{(d,0)}_{\mathcal{S}}(s,u) - G^{(d,0)}_{\mathcal{S}}(s,u)\right]\psi_j(u)\mathrm{d}u + \int\widehat{G}^{(d,0)}_{\mathcal{S}}(s,u)\left[\widehat{\psi}_j(u) - \psi_j(u)\right]\mathrm{d}u\right\} + \left[\lambda^{-1}_j - \widehat{\lambda}^{-1}_j\right]\int_{\mathcal{S}}G^{(d,0)}_{\mathcal{S}}(s,u)\psi_j(u)\mathrm{d}u.\end{aligned} \tag{E.13}$$

By the Cauchy–Schwarz inequality, uniformly for all $s \in [0, a]$,

$$\begin{aligned}&\left|\int\widehat{G}^{(d,0)}_{\mathcal{S}}(s,u)\left[\widehat{\psi}_j(u) - \psi_j(u)\right]\mathrm{d}u\right| \\ &\leq \sup_{s,u\in[0,a]}\left[\widehat{G}^{(d,0)}_{\mathcal{S}}(s,u)\right]^2\left\|\widehat{\psi}_j - \psi_j\right\| = O_{a.s.}\left(\left\|\widehat{\psi}_j - \psi_j\right\|\right).\end{aligned} \tag{E.14}$$

On the contrary, by the similar argument in the proof of Lemma A.7, we can show that, uniformly for all $s \in [0, a]$,

$$\left|\int\left[\widehat{G}^{(d,0)}_{\mathcal{S}}(s,u) - G^{(d,0)}_{\mathcal{S}}(s,u)\right]\psi_j(u)\mathrm{d}u\right| = O_{a.s.}\left(h_3^{-d}\left(\delta_{n2}(h_4, 1, 1) + \gamma_{2d}(h_3, h_4, h_5) + \delta_{n1}(h_1, 1) + \gamma_{1d}(h_1, h_2)\right)\right). \tag{E.15}$$

It then follows from Theorem 4(2) that

$$\sup_{s\in[0,a]}\left|\widehat{\psi}_j^{(d)}(s)-\psi_j^{(d)}(s)\right| = O_{a.s.}\Big(\left(\log n/n\right)^{1/2}+\delta_{n2}^2\left(h_4,1,1\right)+\delta_{n1}^2\left(h_1,1\right)+h_3^{-d}\left(\delta_{n2}\left(h_4,1,1\right)+\gamma_{2d}\left(h_3,h_4,h_5\right)\right.$$
$$\left.+\delta_{n1}\left(h_1,1\right)+\gamma_{1d}\left(h_1,h_2\right)\right)\Big), \quad (E.16)$$

which proves Theorem 4(4).

(5) The uniform consistency of $\widehat{\xi}_{ij}(t)$ is straightforward, and the detailed discussions are omitted.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Ma and Zhou's research were partially supported by the Fundamental Research Funds for the Central Universities (nos. JBK140507 and JBK1806002) of China. Zhou was partially supported by the National Natural Science Foundation of China (NSFC) (no. 11571282).

Supplementary Materials

Some technical lemmas needed for our main results are stated and proved in the Supplementary Material. (*Supplementary Materials*)

References

[1] J. Ramsay and B. Silverman, *Functional Data Analysis*, Springer, New York, NY, USA, 2005.

[2] J.-L. Wang, J.-M. Chiou, and H.-G. Müller, "Functional data analysis," *Annual Review of Statistics and Its Application*, vol. 3, no. 1, pp. 257–295, 2016.

[3] K. Chen and H.-G. Müller, "Modeling repeated functional observations," *Journal of the American Statistical Association*, vol. 107, no. 500, pp. 1599–1609, 2012.

[4] K. Chen, P. Delicado, and H.-G. Müller, "Modelling function-valued stochastic processes, with applications to fertility dynamics," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 79, no. 1, pp. 177–196, 2017.

[5] B. Liu and H.-G. Müller, "Estimating derivatives for samples of sparsely observed functions, with application to online auction dynamics," *Journal of the American Statistical Association*, vol. 104, no. 486, pp. 704–717, 2009.

[6] J. Fan and I. Gijbels, *Analysis of Longitudinal Data with Semiparametric Estimation of Covariance Function*, Chapman & Hall, New York, NY, USA, 1995.

[7] F. Yao, H.-G. Müller, and J.-L. Wang, "Functional data analysis for sparse longitudinal data," *Journal of the American Statistical Association*, vol. 100, no. 470, pp. 577–590, 2005.

[8] Y. Li and T. Hsing, "Uniform convergence rates for non-parametric regression and principal component analysis in functional/longitudinal data," *The Annals of Statistics*, vol. 38, no. 6, pp. 3321–3351, 2010.

[9] J. A. Rice and B. W. Silverman, "Estimating the mean and covariance structure nonparametrically when the data are curves," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 53, no. 1, pp. 233–243, 1991.

[10] R. Shibata, "An optimal selection of regression variables," *Biometrika*, vol. 68, no. 1, pp. 45–54, 1981.

[11] G. Schearz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461–464, 1978.

[12] P. Hall and M. Hosseini-Nasab, "On properties of functional principal components analysis," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 1, pp. 109–126, 2006.

[13] P. Hall, H.-G. Müller, and J.-L. Wang, "Properties of principal component methods for functional and longitudinal data analysis," *The Annals of Statistics*, vol. 34, no. 3, pp. 1493–1517, 2006.


Advances in
Operations Research
Advances in
Decision Sciences
Journal of
Applied Mathematics
The Scientific
World Journal
Journal of
Probability and Statistics
International
Journal of
Mathematics and
Mathematical
Sciences
Journal of
Optimization
Hindawi
Submit your manuscripts at
www.hindawi.com
International Journal of
Engineering
Mathematics
International Journal of
Analysis
Journal of
Complex Analysis
Advances in
Numerical Analysis
Mathematical Problems
in Engineering
International Journal of
Differential Equations
Discrete Dynamics in
Nature and Society
International Journal of
Stochastic Analysis
Journal of
Mathematics
Journal of
Function Spaces
Abstract and
Applied Analysis
Advances in
Mathematical Physics