Technical Report RM 689
Department of Statistics and Probability
Michigan State University
East Lansing, MI 48824

Technical Report on

Spline Confidence Envelopes For Covariance Function In Dense Functional/Longitudinal Data

By

Guanqun Cao
Michigan State University

Li Wang
The University of Georgia

Yehua Li
The University of Georgia

Lijian Yang
Michigan State University and Soochow University

Spline Confidence Envelopes For Covariance Function In Dense Functional/Longitudinal Data*

Guanqun Cao[1], Li Wang[2], Yehua Li[2] and Lijian Yang[1,3]

Michigan State University[1], University of Georgia[2] and Soochow University[3]

Abstract

We consider nonparametric estimation of the covariance function for dense functional data using tensor product B-splines. The proposed estimator is computationally more efficient than the kernel-based methods. We develop both local and global asymptotic distributions for the proposed estimator, and show that our estimator is as efficient as an oracle estimator where the true mean function is known. Simultaneous confidence envelopes are developed based on asymptotic theory to quantify the variability in the covariance estimator and to make global inferences on the true covariance. Monte Carlo simulation experiments provide strong evidence that corroborates the asymptotic theory.

Keyword: B spline, confidence envelope, covariance function, functional data, Karhunen-Loève L^2 representation, longitudinal data.

1 Introduction

Covariance estimation is crucial in both functional and longitudinal data analysis. For longitudinal data, people have found that a good estimation of the covariance function

*Cao and Yang's research was supported in part by NSF awards DMS 0706518, 1007594, a Dissertation Continuation Fellowship from Michigan State University and a grant from the Risk Management Institute, National University of Singapore. Wang's research was supported in part by NSF award DMS 0905730. Li's research was supported in part by NSF award DMS 0806131.

improves the efficiency of model estimation ([20, 5]). In functional data analysis (see [19] for an overview), covariance estimation plays a critical role in functional principal component analysis ([7, 11, 14, 22, 23, 27, 25]), functional generalized linear models ([1, 13, 16, 24]), and other functional nonlinear models ([12, 15]). Other related work on functional data analysis includes [6, 17].

Following [19], the data that we consider are a collection of trajectories $\{\eta_i(x)\}_{i=1}^n$ which are i.i.d. realizations of a smooth random function $\eta(x)$, defined on a continuous interval $\mathcal{X}$. Assume that $\{\eta(x), x \in \mathcal{X}\}$ is a L^2 process, i.e. $E\int_{\mathcal{X}}^2 \eta(x)dx < +\infty$, and define the mean and covariance functions as $m(x) = E\{\eta(x)\}$ and $G(x, x') = \mathrm{cov}\{\eta(x), \eta(x')\}$. The covariance function is a symmetric nonnegative-definite function with a spectral decomposition, $G(x, x') = \sum_{k=1}^{\infty} \lambda_k \psi_k(x)\psi_k(x')$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$, are the eigenvalues, and $\{\psi_k(x)\}_{k=1}^{\infty}$ are the corresponding eigenfunctions and form an orthonormal basis of $L^2(\mathcal{X})$. By the standard Karhunen-Loève representation, $\eta_i(x) = m(x) + \sum_{k=1}^{\infty} \xi_{ik}\phi_k(x)$, where the random coefficients ξ_{ik} are uncorrelated with mean 0 and variance 1, and the functions $\phi_k = \sqrt{\lambda_k}\psi_k$. In what follows, we assume that $\lambda_k = 0$, for $k > \kappa$, where κ is a positive integer, thus $G(x, x') = \sum_{k=1}^{\kappa} \phi_k(x)\phi_k(x')$.

We consider a typical functional data setting where $\eta_i(\cdot)$ is recorded on a regular grid in $\mathcal{X}$, and the measurements are contaminated with measurement errors. Without loss of generality, we take $\mathcal{X} = [0, 1]$. Then the observed data are $Y_{ij} = \eta_i(X_{ij}) + \sigma(X_{ij})\varepsilon_{ij}$, for $1 \leq i \leq n$, $1 \leq j \leq N$, where $X_{ij} = j/N$, ε_{ij} are i.i.d. random errors with $E(\varepsilon_{11}) = 0$ and $E(\varepsilon_{11}^2) = 1$, and $\sigma^2(x)$ is the variance function of the measurement errors. By the Karhunen-Loève representation, the observed data can be written as

$$Y_{ij} = m(j/N) + \sum\nolimits_{k=1}^{\kappa} \xi_{ik}\phi_k(j/N) + \sigma(j/N)\varepsilon_{ij}. \tag{1}$$

We model $m(\cdot)$ and $G(\cdot, \cdot)$ as nonparametric functions, and hence $\{\lambda_k\}_{k=1}^{\kappa}$, $\{\phi_k(\cdot)\}_{k=1}^{\kappa}$ and $\{\xi_{ik}\}_{k=1}^{\kappa}$ are unknown and need to be estimated.

There are some important recent works on nonparametric covariance estimation and principal component analysis in functional data setting described above, for example [7, 14, 23]. These papers are based on kernel smoothing methods and only contain results on convergence rates. So far, there is no theoretical work on simultaneous or uniform confidence envelopes for $G(x, x')$. Nonparametric simultaneous confidence regions are powerful tools for making global inference on functions, see [2, 8, 26] for related theory and applications. The fact that simultaneous confidence regions have not been established for functional data is certainly not due to lack of interesting applications, but to the greater technical difficulty to formulate such regions and to establish their theoretical properties.

In this paper, we consider dense functional data, which requires $N \gg n^{\theta}$ as $n \to \infty$ for some $\theta > 0$, and propose to estimate $G(x, x')$ by tensor product B-splines. In contrast with the existing work on nonparametric covariance estimation, which are mostly based on kernel smoothing ([7, 14, 23]), our proposed spline estimator is much more efficient in terms of computation. The reason is that the kernel smoothers are evaluated pointwise, while for the spline estimator, we only need to solve for a small number of spline coefficients to have an explicit expression for the whole function. For smoothing a two dimensional covariance surface with a moderate sample size, the kernel smoother might take up to one hour, while our spline estimator only takes a few seconds. Computation efficiency is a huge advantage for the spline methods in analyzing large data sets and in performing simulation studies. See [10] for more discussions on the computational merits of spline methods.

We show that the estimation error in the mean function is asymptotically negligible in estimating the covariance function, and our covariance estimator is as efficient as an oracle estimator where the true mean function is known. We derive both local and global asymptotic distribution for the proposed spline covariance estimator. Especially, based on the asymptotic distribution of the maximum deviation of the estimator, we propose a new simultaneous confidence envelope for the covariance function, which can be used to visualize the variability of the covariance estimator and to make global inferences on the shape of the true covariance.

We organize our paper as follows. In Section 2 we describe the spline covariance estimator and state our main theoretical results. In Section 3 we provide further insights into the error structure of our spline estimator. Section 4 describes the actual steps to implement the confidence envelopes. We present simulation studies in Section 5. Proofs of the technical lemmas are in the Appendix.

2 Main results

To describe the tensor product spline estimator of the covariance functions, we first introduce some notation. Denote a sequence of equally-spaced points $\{t_J\}_{J=1}^{N_{\mathrm{s}}}$, called interior knots which divide the interval $[0, 1]$ into $(N_{\mathrm{s}} + 1)$ subintervals $I_J = [t_J, t_{J+1})$, $J = 0, ..., N_{\mathrm{s}} - 1$, $I_{N_{\mathrm{s}}} = [t_{N_{\mathrm{s}}}, 1]$. Let $h_{\mathrm{s}} = 1/(N_{\mathrm{s}} + 1)$ be the distance between neighboring knots. Let $\mathcal{H}^{(p-2)} = \mathcal{H}^{(p-2)}[0, 1]$ be the polynomial spline space of order p, which consists all $p-2$ times continuously differentiable functions on $[0, 1]$ that are polynomials of degree $p - 1$ on each interval. The J^{th} B-spline of order p is denoted by $B_{J,p}$ as in [3]. Thus we define the tensor product spline space as

$$\mathcal{H}^{(p-2),2}[0,1]^2 \equiv \mathcal{H}^{(p-2),2} = \mathcal{H}^{(p-2)} \otimes \mathcal{H}^{(p-2)} = \left\{ \sum_{J,J'=1-p}^{N_s} b_{JJ'p} B_{J,p}(x) B_{J',p}(x'), b_{JJ'p} \in R, x, x' \in [0,1] \right\}.$$

If the mean function $m(x)$ was known, one could compute the errors

$$U_{ij} \equiv Y_{ij} - m(j/N) = \sum_{k=1}^{\kappa} \xi_{ik}\phi_k(j/N) + \sigma(j/N)\varepsilon_{ij}, \ 1 \le i \le n, \ 1 \le j \le N.$$

Denote $\bar{U}_{\cdot jj'} = n^{-1}\sum_{i=1}^{n} U_{ij}U_{ij'}$, $1 \le j \ne j' \le N$, one can then define the "oracle estimator" of the covariance function

$$\tilde{G}_{p_2}(\cdot,\cdot) = \underset{g(\cdot,\cdot)\in\mathcal{H}^{(p_2-2),2}}{\operatorname{argmin}} \sum_{1\le j\ne j'\le N} \left\{\bar{U}_{\cdot jj'} - g(j/N, j'/N)\right\}^2, \tag{2}$$

using tensor product splines of order $p_2 \ge 2$. Since the mean function $m(x)$ is unavailable when one analyzes data, one can use instead the spline smoother of $m(x)$, i.e.,

$$\hat{m}_{p_1}(\cdot) = \underset{g(\cdot)\in\mathcal{H}^{(p_1-2)}}{\operatorname{argmin}} \sum_{i=1}^{n}\sum_{j=1}^{N} \{Y_{ij} - g(j/N)\}^2, \ p_1 \ge 1.$$

To mimic the above oracle smoother, we define

$$\hat{G}_{p_1,p_2}(\cdot,\cdot) = \underset{g(\cdot,\cdot)\in\mathcal{H}^{(p_2-2),2}}{\operatorname{argmin}} \sum_{1\le j\ne j'\le N} \left\{\hat{\bar{U}}_{\cdot jj',p_1} - g(j/N, j'/N)\right\}^2, \tag{3}$$

where $\hat{\bar{U}}_{\cdot jj',p_1} = n^{-1}\sum_{i=1}^{n} \hat{U}_{ijp_1}\hat{U}_{ij'p_1}$ with $\hat{U}_{ijp_1} = Y_{ij} - \hat{m}_{p_1}(j/N)$.

Let N_{s_1} be the number of interior knots for mean estimation, and N_{s_2} be the number of interior knots for $\hat{G}_{p_1,p_2}(x,x')$ in each coordinate. In other words, we have $N_{s_2}^2$ interior knots for the tensor product spline space $\mathcal{H}^{(p_2-2),2}$. For any $\nu \in (0,1]$, we denote $C^{q,\nu}[0,1]$ as the space of ν-Hölder continuous functions on $[0,1]$,
$C^{q,\nu}[0,1] = \left\{\phi : \sup_{x\ne x', x,x'\in[0,1]} \frac{\left|\phi^{(q)}(x) - \phi^{(q)}(x')\right|}{|x-x'|^{\nu}} < +\infty\right\}$. We now state the technical assumptions as follows:

(A1) *The regression function* $m \in C^{p_1-1,1}[0,1]$.

(A2) *The standard deviation function* $\sigma(x) \in C^{0,\nu}[0,1]$. *For any* $k = 1, 2, \ldots \kappa$, $\phi_k(x) \in C^{p_2-1,\nu}[0,1]$. *Also* $\sup_{(x,x')\in[0,1]^2} G(x,x') < C$, for some positive constant C and $\min_{x\in[0,1]} G(x,x) > 0$.

(A3) *The number of knots* N_{s_1} *and* N_{s_2} *satisfy* $n^{1/(4p_1)} \ll N_{s_1} \ll N$, $n^{1/(2p_2)} \ll N_{s_2} \ll \min\left(N^{1/3}, n^{1/3}\right)$ and $N_{s_2} \ll N_{s_1}^{p_1}$.

(A4) *The number* κ *of nonzero eigenvalues is finite. The variables* $(\xi_{ik})_{i=1,k=1}^{\infty,\kappa}$ and $(\varepsilon_{ij})_{i=1,j=1}^{\infty,\infty}$ *are independent. In addition,* $E\varepsilon_{11} = 0$, $E\varepsilon_{11}^2 = 1$, $E\xi_{1k} = 0$, $E\xi_{1k}^2 = 1$ *and* $\max_{1\leq k\leq\kappa} E\left|\xi_{1k}\right|^{\delta_1} < +\infty$, $E\left|\varepsilon_{11}\right|^{\delta_2} < +\infty$, *for some* $\delta_1, \delta_2 > 4$.

Assumptions (A1)-(A4) are standard in the spline smoothing literature; see [9], for instance. In particular, (A1) and (A2) guarantee the orders of the bias terms of the spline smoothers for $m(x)$ and $\boldsymbol{\phi}_k(x)$. Assumption (A3) is a weak assumption to ensure the order of the bias and noise terms provided in Section 3. Assumption (A4) is necessary for strong approximation. More discussion about the assumptions is in Section 4. The next proposition provides that the tensor product spline estimator $\hat{G}_{p_1,p_2}$ is uniformly close to the oracle smoother at the rate of $o_p\left(n^{-1/2}\right)$.

Proposition **2.1.** *Under Assumptions (A1)-(A4), one has*

$$\sup_{(x,x')\in[0,1]^2} n^{1/2}\left|\hat{G}_{p_1,p_2}(x,x') - \tilde{G}_{p_2}(x,x')\right| = o_p(1).$$

This proposition allows one to construct confidence envelopes for G based on the tensor product spline estimator $\hat{G}_{p_1,p_2}$ rather than unavailable "oracle estimator" $\tilde{G}_{p_2}$. The next theorem provides a pointwise asymptotic approximation to the mean squared error of $\hat{G}_{p_1,p_2}(x,x')$.

Theorem **2.1.** *Under Assumptions (A1)-(A4),*

$$nE[\hat{G}_{p_1,p_2}(x,x') - G(x,x')]^2 = V(x,x') + o(1),$$

where $V(x,x') = G(x,x')^2 + G(x,x)\,G(x',x') + \sum_{k=1}^{\kappa}\phi_k^2(x)\,\phi_k^2(x')\left(E\xi_{1k}^4 - 3\right)$.

To obtain the quantile of the distribution of $n^{1/2}\left|\hat{G}_{p_1,p_2}(x,x') - G(x,x')\right| V^{-1/2}(x,x')$, one defines

$$\zeta_Z(x,x') = \left\{\sum_{k\neq k'}^{\kappa} Z_{kk'}\phi_k(x)\,\phi_{k'}(x') + \sum_{k=1}^{\kappa}\phi_k(x)\,\phi_k(x')\,Z_k\left(E\xi_{1k}^4 - 1\right)^{1/2}\right\} V^{-1/2}(x,x'), \tag{4}$$

where $Z_{kk'} = Z_{k'k}$ and Z_k are i.i.d. $N(0,1)$ random variables. Hence, for any $(x,x') \in [0,1]^2$, $\zeta_Z(x,x')$ is a standardized Gaussian field such that $E\zeta_Z(x,x') = 0$, $E\zeta_Z^2(x,x') = 1$.

Define $Q_{1-\alpha}$ as the $100\,(1-\alpha)^{th}$ percentile of the absolute maxima distribution of $\zeta_Z\,(x,x')$, $\forall\,(x,x') \in [0,1]^2$, i.e.

$$P\left\{\sup_{(x,x')\in[0,1]^2} \left|\zeta_Z\,(x,x')\right| \le Q_{1-\alpha}\right\} = 1-\alpha, \forall \alpha \in (0,1)\,.$$

The following addresses the simultaneous envelopes for $G(x,x')$.

Theorem **2.2.** *Under Assumptions (A1)-(A4), for any $\alpha \in (0,1)$,*

$$\lim_{n\to\infty} P\left\{\sup_{(x,x')\in[0,1]^2} n^{1/2}\left|\hat{G}_{p_1,p_2}(x,x') - G(x,x')\right| V^{-1/2}\,(x,x') \le Q_{1-\alpha}\right\} = 1-\alpha,$$

$$\lim_{n\to\infty} P\left\{n^{1/2}\left|\hat{G}_{p_1,p_2}(x,x') - G(x,x')\right| V^{-1/2}\,(x,x') \le Z_{1-\alpha/2}\right\} = 1-\alpha, \quad \forall (x,x') \in [0,1]^2,$$

where $Z_{1-\alpha/2}$ is the $100\,(1-\alpha/2)^{th}$ percentile of the standard normal distribution.

Remark 1. *Although this covariance function estimator cannot be guaranteed to be positive definite, it tends to the true positive definite covariance function in probability.*

The next result follows directly from Theorem 2.2.

Corollary **2.1.** *Under Assumptions (A1)-(A4), as $n \to \infty$, an asymptotic $100\,(1-\alpha)\,\%$ confidence envelope for $G(x,x')$, $\forall (x,x') \in [0,1]^2$ is*

$$\hat{G}_{p_1,p_2}(x,x') \pm n^{-1/2} Q_{1-\alpha} V^{1/2}\,(x,x')\,,\ \forall \alpha \in (0,1)\,, \tag{5}$$

while an asymptotic $100\,(1-\alpha)\,\%$ pointwise confidence envelope for $G(x,x')$, $\forall (x,x') \in [0,1]^2$ is

$$\hat{G}_{p_1,p_2}(x,x') \pm n^{-1/2} Z_{1-\alpha/2} V^{1/2}\,(x,x')\,,\ \forall \alpha \in (0,1)\,.$$

Remark 2. *Although the above confidence envelopes for $G(x,x')$ is most useful, one can also construct asymptotic $100\,(1-\alpha)\,\%$ confidence ceiling and floor for $G(x,x')$ as*

$$\hat{G}_{p_1,p_2}(x,x') + n^{-1/2} Q_{U,1-\alpha} V^{1/2}\,(x,x')\,,\ \forall \alpha \in (0,1)\,,$$

$$\hat{G}_{p_1,p_2}(x,x') - n^{-1/2} Q_{U,1-\alpha} V^{1/2}\,(x,x')\,,\ \forall \alpha \in (0,1)\,, \tag{6}$$

respectively, in which $Q_{U,1-\alpha}$ satisfies that $P\left\{\sup_{(x,x')\in[0,1]^2} \zeta_Z\,(x,x') \le Q_{U,1-\alpha}\right\} = 1-\alpha$, $\forall \alpha \in (0,1)$.

3 Error structure for the spline covariance estimator

To gain a deeper understanding on the behavior of the spline covariance estimator, we provide an asymptotic decomposition for the estimation error $\hat{G}_{p_1,p_2}(x,x') - G(x,x')$. We first introduce some additional notation. For any Lebesgue measurable function ϕ on a domain $\mathcal{D}$, denote $\|\phi\|_\infty = \sup_{\mathbf{x}\in\mathcal{D}} |\phi(\mathbf{x})|$. In this paper, $\mathcal{D} = [0,1]$ or $[0,1]^2$. For any bivariate Lebesgue measurable function ϕ and φ, define their theoretical and empirical inner products as

$$\langle \phi, \varphi \rangle = \int_0^1 \int_0^1 \phi(x,x')\varphi(x,x')\,dxdx', \quad \langle \phi, \varphi \rangle_{2,N} = N^{-2} \sum_{1\le j\ne j'\le N} \phi(j/N, j'/N)\varphi(j/N, j'/N),$$

with the corresponding theoretical and empirical L^2 norms defined as
$\|\phi\|_2^2 = \int_0^1\int_0^1 \phi^2(x,x')\,dxdx'$ and $\|\phi\|_{2,N}^2 = N^{-2}\sum_{1\le j\ne j'\le N} \phi^2(j/N, j'/N)$ respectively.

For simplicity, denote $B_{JJ',p_2}(x,x') = B_{J,p_2}(x)\,B_{J',p_2}(x')$ and

$$\boldsymbol{B}_{p_2}(x,x') = \begin{array}{l} \left(B_{1-p_2,1-p_2,p_2}(x,x'), \ldots, B_{N_{s_2},1-p_2,p_2}(x,x'),\right. \\ \left.\ldots, B_{1-p_2,N_{s_2},p_2}(x,x'), \ldots, B_{N_{s_2},N_{s_2},p_2}(x,x')\right)^{\mathrm{T}} \end{array},$$

where $\sup_{(x,x')\in[0,1]^2} \|\boldsymbol{B}_{p_2}(x,x')\|_\infty \le 1$. Further denote

$$\mathbf{V}_{p_2,2} = \left(\langle B_{JJ',p_2}, B_{J''J''',p_2}\rangle\right)_{J,J',J'',J'''=1-p_2}^{N_{s_2}}, \quad \hat{\mathbf{V}}_{p_2,2} = \left(\langle B_{JJ',p_2}, B_{J''J''',p_2}\rangle_{2,N}\right)_{J,J',J'',J'''=1-p_2}^{N_{s_2}} \tag{7}$$

as the theoretical and empirical inner product matrices of $\{B_{JJ',p_2}(x,x')\}_{J,J'=1-p_2}^{N_{s_2}}$.

Next we discuss the tensor product spline space $\mathcal{H}^{(p_1-2),2}$ and the representation of the tensor product spline estimator $\hat{G}_{p_1,p_2}$ in (3) more carefully. Denote the design matrix $\mathbf{X}$ as

$$\mathbf{X} = \left\{\boldsymbol{B}_{p_2}\left(\frac{2}{N},\frac{1}{N}\right), \ldots, \boldsymbol{B}_{p_2}\left(1,\frac{1}{N}\right), \ldots, \boldsymbol{B}_{p_2}\left(\frac{1}{N},1\right), \ldots, \boldsymbol{B}_{p_2}\left(1-\frac{1}{N},1\right)\right\}^{\mathrm{T}}.$$

We rewrite $\hat{G}_{p_1,p_2}(x,x')$ in (3) as

$$\hat{G}_{p_1,p_2}(x,x') \equiv \hat{\boldsymbol{\beta}}_{p_1,p_2}^{\mathrm{T}} \boldsymbol{B}_{p_2}(x,x'), \tag{8}$$

where $\hat{\boldsymbol{\beta}}_{p_1,p_2}$ is the collector of the estimated spline coefficients by solving the following least squares

$$\hat{\boldsymbol{\beta}}_{p_1,p_2} = \underset{\mathbf{b}_{p_2}\in R^{(N_s+p_2)^2}}{\operatorname{argmin}} \sum_{1\le j\ne j'\le N} \left\{\hat{\hat{U}}_{\cdot jj',p_1} - \mathbf{b}_{p_2}^{\mathrm{T}} B_{JJ'p_2}(j/N)\right\}^2.$$

By elementary algebra, one obtains

$$\hat{G}_{p_1,p_2}(x,x') = \boldsymbol{B}_{p_2}^{\mathrm{T}}(x,x')\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\hat{\hat{\mathbf{U}}}_{p_1},$$

$$\tilde{G}_{p_2}(x,x') = \boldsymbol{B}_{p_2}^{\mathrm{T}}(x,x')\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\bar{\mathbf{U}}, \tag{9}$$

where $\hat{\hat{\mathbf{U}}}_{p_1} = \left(\hat{\hat{U}}_{\cdot 21,p_1},\ldots,\hat{\hat{U}}_{\cdot N1,p_1},\ldots,\hat{\hat{U}}_{\cdot 1N,p_1},\ldots,\hat{\hat{U}}_{\cdot(N-1)N,p_1}\right)^{\mathrm{T}}$ and $\bar{\mathbf{U}} = (\bar{U}_{\cdot 21},\ldots,\bar{U}_{\cdot N1},\ldots,\bar{U}_{\cdot 1N},\ \ldots,\bar{U}_{\cdot(N-1)N})^{\mathrm{T}}$.

By the definitions of empirical inner product and $\hat{\mathbf{V}}_{p_2,2}$ in (7), one has

$$\mathbf{X}^{\mathrm{T}}\mathbf{X} = N^2\hat{\mathbf{V}}_{p_2,2}, \qquad \mathbf{X}^{\mathrm{T}}\bar{\mathbf{U}} = \sum_{1\le j\ne j'\le N}\boldsymbol{B}_{p_2}\left(j/N,j'/N\right)\bar{U}_{\cdot jj'}.$$

By Proposition 2.1, decomposing the error in $\hat{G}_{p_1,p_2}(x,x')$ is asymptotically equivalent to decomposing $\tilde{G}_{p_2}(x,x') - G(x,x')$. Therefore, below we decompose $\bar{U}_{\cdot jj'}$ into four parts, where

$$\begin{aligned}
\bar{U}_{1jj'} &= n^{-1}\sum_{i=1}^{n}\sum_{k\ne k'}^{\kappa}\xi_{ik}\xi_{ik'}\phi_k\left(j/N\right)\phi_{k'}\left(j'/N\right),\\
\bar{U}_{2jj'} &= n^{-1}\sum_{k=1}^{\kappa}\sum_{i=1}^{n}\xi_{ik}^2\phi_k\left(j/N\right)\phi_k\left(j'/N\right),\\
\bar{U}_{3jj'} &= n^{-1}\sum_{i=1}^{n}\sigma\left(j/N\right)\sigma\left(j'/N\right)\varepsilon_{ij}\varepsilon_{ij'},\\
\bar{U}_{4jj'} &= n^{-1}\sum_{i=1}^{n}\left\{\sum_{k=1}^{\kappa}\xi_{ik}\phi_k\left(j/N\right)\sigma\left(j'/N\right)\varepsilon_{ij'} + \sum_{k=1}^{\kappa}\xi_{ik}\phi_k\left(j'/N\right)\sigma\left(j/N\right)\varepsilon_{ij}\right\}.
\end{aligned}$$

Denote $\bar{\mathbf{U}}_i = \left\{\bar{U}_{ijj'}\right\}_{1\le j\ne j'\le N}$, $\tilde{\mathcal{U}}_{ip_2}\left(x,x'\right) = \boldsymbol{B}_{p_2}^{\mathrm{T}}(x,x')\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\bar{\mathbf{U}}_i$ for $i=1,2,3,4$. Then $\tilde{G}_{p_2}(x,x')$ yields the following decompotion

$$\tilde{G}_{p_2}(x,x') = \tilde{\mathcal{U}}_{1p_2}\left(x,x'\right) + \tilde{\mathcal{U}}_{2p_2}\left(x,x'\right) + \tilde{\mathcal{U}}_{3p_2}\left(x,x'\right) + \tilde{\mathcal{U}}_{4p_2}\left(x,x'\right).$$

Define

$$\mathcal{U}_1(x,x') = n^{-1}\sum_{i=1}^{n}\sum_{k\ne k'}^{\kappa}\xi_{ik}\xi_{ik'}\phi_k\left(x\right)\phi_{k'}\left(x'\right), \tag{10}$$

$$\mathcal{U}_2(x,x') = G\left(x,x'\right) + \sum_{k=1}^{\kappa}\left\{\phi_k\left(x\right)\phi_k\left(x'\right)\left(n^{-1}\sum_{i=1}^{n}\xi_{ik}^2 - 1\right)\right\}. \tag{11}$$

In the following proposition, we illustrate the facts that $\tilde{\mathcal{U}}_{1p_2}\left(x,x'\right)$ and $\tilde{\mathcal{U}}_{2p_2}\left(x,x'\right)$ are the dominating terms, which converge uniformly to $\mathcal{U}_1(x,x')$ and $\mathcal{U}_2(x,x')$ respectively, while $\tilde{\mathcal{U}}_{3p_2}\left(x,x'\right)$ and $\tilde{\mathcal{U}}_{4p_2}\left(x,x'\right)$ are negligible noise terms.

Proposition **3.1.** *Under Assumptions (A2)-(A4), one has*

$$\left\|\tilde{\mathcal{U}}_{1p_2} + \tilde{\mathcal{U}}_{2p_2} - \mathcal{U}_1 - \mathcal{U}_2\right\|_{\infty} + \left\|\tilde{\mathcal{U}}_{3p_2}\right\|_{\infty} + \left\|\tilde{\mathcal{U}}_{4p_2}\right\|_{\infty} = o_p\left(n^{-1/2}\right). \tag{12}$$

4 Implementation

In this section, we describe the procedure to implement the confidence envelopes. Given the data set $(j/N, Y_{ij})_{j=1,i=1}^{N,n}$, the number of interior knots N_{s_1} for $\hat{m}_{p_1}(x)$ is taken to be $[n^{1/(4p_1)}\log n]$, where $[a]$ denotes the integer part of a. Meanwhile, the spline estimator $\hat{G}_{p_1,p_2}(x,x')$ is obtained by (8) with the number of interior knots $N_{s_2} = [n^{1/(2p_2)}\log\log n]$. These choices of knots satisfy condition (A3) in our theory.

To construct the confidence envelopes, one needs to evaluate the percentile $Q_{1-\alpha}$ and estimate the variance function $V(x,x')$. An estimator $\hat{V}(x,x')$ of $V(x,x')$ is

$$\hat{V}(x,x') = \hat{G}_{p_1,p_2}(x,x')^2 + \hat{G}_{p_1,p_2}(x,x)\hat{G}_{p_1,p_2}(x',x') + \sum\nolimits_{k=1}^{\kappa} \hat{\phi}_k^2(x)\,\hat{\phi}_k^2(x')\left(n^{-1}\sum\nolimits_{i=1}^{n}\hat{\xi}_{ik}^4 - 3\right),$$

where $\hat{\phi}_k$ and $\hat{\xi}_{ik}$ are the estimators of ϕ_k and ξ_{ik} respectively. According to [24], the estimates of eigenfunctions and eigenvalues correspond to the solutions $\hat{\phi}_k$ and $\hat{\lambda}_k$ of the eigen-equations,

$$\int_0^1 \hat{G}_{p_1,p_2}(x,x')\hat{\phi}_k(x)\,dx = \hat{\lambda}_k\hat{\phi}_k(x'), \tag{13}$$

where the $\hat{\phi}_k$ are subject to $\int_0^1 \hat{\phi}_k^2(t)\,dt = \hat{\lambda}_k$ and $\int_0^1 \hat{\phi}_k(t)\,\hat{\phi}_{k'}(t)\,dt = 0$ for $k' < k$. Since N is sufficiently large, (13) can be approximated by $N^{-1}\sum_{j=1}^{N}\hat{G}_{p_1,p_2}(j/N, j'/N)\hat{\phi}_k(j/N) = \hat{\lambda}_k\hat{\phi}_k(j'/N)$. For the same reason, the estimation of ξ_{ik} has the form of

$$\hat{\xi}_{ik} = N^{-1}\sum\nolimits_{j=1}^{N}\hat{\lambda}_k^{-1}\left(Y_{ij} - \hat{m}_{p_1}(j/N)\right)\hat{\phi}_k(j/N).$$

To choose the number of principal components, κ, [18] described two methods. The first method is the "pseudo-AIC" criterion proposed in [23]. The second is a simple "fraction of variation explained" method, i.e. select the number of eigenvalues that can explain, say, 95% of the variation in the data. From our experience in the numerical studies, the simple "fraction of variation explained" method often works well.

Finally, to evaluate $Q_{1-\alpha}$, we need to simulate the Gaussian random field $\zeta_Z(x,x')$ in (4). The definition of $\zeta_Z(x,x')$ involves $\phi_k(x)$ and $V(x,x')$, which are replaced by their estimators described above. The fourth moment of ξ_{1k} is replaced by the empirical moments of $\hat{\xi}_{ik}$. We simulate a large number of independent realizations of $\zeta_Z(x,x')$, and take the maximal absolute deviation for each copy of $\zeta_Z(x,x')$. Then $Q_{1-\alpha}$ is estimated by the empirical percentiles of these maximum values.

5 Simulation

To illustrate the finite-sample performance of the spline approach, we generated data from the model

$$Y_{ij} = m\left(j/N\right) + \sum\nolimits_{k=1}^{2} \xi_{ik}\phi_k\left(j/N\right) + \sigma\varepsilon_{ij}, 1 \leq j \leq N, 1 \leq i \leq n,$$

where $\xi_{ik} \sim N(0,1), k = 1,2$, $\varepsilon_{ij} \sim N(0,1)$, for $1 \leq i \leq n$, $1 \leq j \leq N$, $m(x) = 10 + \sin\left\{2\pi\left(x - 1/2\right)\right\}$, $\phi_1(x) = -2\cos\left\{\pi\left(x - 1/2\right)\right\}$ and $\phi_2(x) = \sin\left\{\pi\left(x - 1/2\right)\right\}$. This setting implies $\lambda_1 = 2$ and $\lambda_2 = 0.5$. The noise levels are set to be $\sigma = 0.5$ and 1.0. The number of subjects n is taken to be 50, 100, 200, 300 and 500, and under each sample size the number of observations per curve is assumed to be $N = 4[n^{0.3}\log(n)]$. This simulated process has a similar design as one of the simulation models in [23], except that each subject is densely observed. We consider both linear and cubic spline estimators, and use confidence levels $1 - \alpha = 0.95$ and 0.99 for our simultaneous confidence envelops. Each simulation is repeated 500 times.

Figure 1 depicts a simulated data set with $n = 200$ and $\sigma = 0.5$. Table 1 shows the empirical frequency that the true surface $G(x, x')$ is entirely covered by the confidence envelopes. At both noise levels, one observes that, as sample size increases, the true coverage probability of the confidence envelopes becomes closer to the nominal confidence level, which shows a positive confirmation of Theorem 2.2.

We present two estimation schemes: a) both mean and covariance functions are estimated by linear splines, i.e., $p_1 = p_2 = 2$; b) both are estimated by cubic splines, i.e. $p_1 = p_2 = 4$. Since the true covariance function is smooth in our our simulation, the cubic spline estimator provides better estimate of the covariance function. However, as can been seen from Table 1, the two spline estimators behave rather similarly in terms of coverage probability. We also did simulation studies for the cases $p_1 = 4$, $p_2 = 2$ and $p_1 = 2$, $p_2 = 4$, the coverage rates are not shown here because they are similar to the cases presented in Table 1.

We show in Figure 2 the spline covariance estimator and the 95% confidence envelops for $n = 200$ and $\sigma = 0.5$. The two panels of Figure 2 correspond to linear ($p_1 = p_2 = 2$) and cubic ($p_1 = p_2 = 4$) spline estimators respectively. In each panel, the true covariance function is overlayed by the two confidence envelopes.

Table 1: Uniform coverage rates from 500 replications using spline (5).

σ	n	$1-\alpha$	Coverage proportion ($p_1 = p_2 = 4$)	Coverage proportion ($p_1 = p_2 = 2$)
0.5	50	0.950	0.720	0.710
		0.990	0.824	0.828
	100	0.950	0.858	0.834
		0.990	0.946	0.930
	200	0.950	0.912	0.898
		0.990	0.962	0.956
	300	0.950	0.890	0.884
		0.990	0.960	0.958
	500	0.950	0.908	0.894
		0.990	0.976	0.964
1.0	50	0.950	0.626	0.690
		0.990	0.720	0.796
	100	0.950	0.752	0.796
		0.990	0.874	0.904
	200	0.950	0.798	0.852
		0.990	0.912	0.944
	300	0.950	0.822	0.828
		0.990	0.922	0.936
	500	0.950	0.864	0.858
		0.990	0.946	0.946

APPENDIX

Throughout this section, C and c mean some positive constant in this whole section. For any continuous function ϕ on $[0,1]$, let $\omega(\phi,\delta) = \max_{x,x'\in[0,1],|x-x'|\leq\delta} |\phi(x) - \phi(x')|$ be its modulus of continuity. For any vector $\mathbf{a} = (a_1,...,a_n) \in R^n$, denote the norm $\|\mathbf{a}\|_r = (|a_1|^r + \cdots + |a_n|^r)^{1/r}$, $1 \leq r < +\infty$, $\|\mathbf{a}\|_\infty = \max(|a_1|,...,|a_n|)$. For any $n \times n$ symmetric matrix $\mathbf{A}$, denote by $\lambda_{\min}(\mathbf{A})$ and $\lambda_{\max}(\mathbf{A})$ its smallest and largest eigenvalues, and its L_r norm as $\|\mathbf{A}\|_r = \max_{\mathbf{a}\in R^n,\mathbf{a}\neq\mathbf{0}} \|\mathbf{A}\mathbf{a}\|_r \|\mathbf{a}\|_r^{-1}$. In particular, $\|\mathbf{A}\|_2 = \lambda_{\max}(\mathbf{A})$, and if $\mathbf{A}$ is also nonsingular, $\|\mathbf{A}^{-1}\|_2 = \lambda_{\min}^{-1}(\mathbf{A})$. Also L_∞ norm for a matrix $\mathbf{A} = (a_{ij})_{i=1,j=1}^{m,n}$ is $\|\mathbf{A}\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^n |a_{ij}|$.

A.1 Preliminaries

For any positive integer p, denote the theoretical and empirical inner product matrices of $\{B_{J,p}(x)\}_{J=1-p}^{N_{\mathrm{s}}}$ as

$$\mathbf{V}_p = \left(\int_0^1 B_{J,p}\left(x\right) B_{J',p}\left(x\right) dx\right)_{J,J'=1-p}^{N_{\mathrm{s}}}, \quad \hat{\mathbf{V}}_p = \left(N^{-1}\sum_{j=1}^{N} B_{J,p}\left(j/N\right) B_{J',p}\left(j/N\right)\right)_{J,J'=1-p}^{N_{\mathrm{s}}}. \tag{A.1}$$

The next lemma is a special case of Theorem 13.4.3, Page 404 of [4]. Let p be a positive integer, a matrix $\mathbf{A} = (a_{ij})$ is said to have bandwidth p if $a_{ij} = 0$ when $|i-j| \geq p$, and p is the smallest integer with this property.

Lemma A.1. *If a matrix $\mathbf{A}$ with bandwidth p has an inverse $\mathbf{A}^{-1}$ and $d = \|\mathbf{A}\|_2 \|\mathbf{A}^{-1}\|_2$ is the condition number of $\mathbf{A}$, then $\|\mathbf{A}^{-1}\|_\infty \leq 2c_0\left(1-\nu\right)^{-1}$, with $c_0 = \nu^{-2p}\|\mathbf{A}^{-1}\|_2$, $\nu = \left(\frac{d^2-1}{d^2+1}\right)^{1/(4p)}$.*

We establish next that the theoretical inner product matrix $\mathbf{V}_p$ defined in (A.1) has an inverse with bounded l_∞ norm.

Lemma A.2. *For any positive integer p, there exists a constant $M_p > 0$ depending only on p, such that $\left\|\mathbf{V}_p^{-1}\right\|_\infty \leq M_p h_{\mathrm{s}}^{-1}$, for a large enough n, where $h_{\mathrm{s}} = (N_{\mathrm{s}}+1)^{-1}$.*

Proof. According to Lemma A.1 in [21], $\mathbf{V}_p$ is invertible since it is a symmetric matrix with all eigenvalues positive, i.e. $0 < c_p N_s^{-1} \leq \lambda_{\min}\left(\mathbf{V}_p\right) \leq \lambda_{\max}\left(\mathbf{V}_p\right) \leq C_p N_s^{-1} < \infty$, where c_p and C_p are positive real numbers. The compact support of B spline basis makes $\mathbf{V}_p$ of bandwidth p, hence one can apply Lemma A.1. Since $d_p = \lambda_{\max}\left(\mathbf{V}_p\right)/\lambda_{\min}\left(\mathbf{V}_p\right) \leq C_p/c_p$, hence

$$\nu_p = \left(d_p^2-1\right)^{1/4p}\left(d_p^2+1\right)^{-1/4p} \leq \left(C_p^2 c_p^{-2}-1\right)^{1/4p}\left(C_p^2 c_p^{-2}+1\right)^{-1/4p} < 1.$$

If $p=1$, then $\mathbf{V}_p^{-1} = h_{\mathrm{s}}^{-1}\mathbf{I}_{N_{\mathrm{s}}+p}$, the lemma holds with $M_p = 1$. If $p > 1$, let $\mathbf{u}_{1-p} = \left(1, \mathbf{0}_{N_{\mathrm{s}}+p-1}^{\mathrm{T}}\right)^{\mathrm{T}}$, $\mathbf{u}_0 = \left(\mathbf{0}_{p-1}^{\mathrm{T}}, 1, \mathbf{0}_{N_{\mathrm{s}}}^{\mathrm{T}}\right)^{\mathrm{T}}$, then $\|\mathbf{u}_{1-p}\|_2 = \|\mathbf{u}_0\|_2 = 1$. Also lemma A.1 in [21] implies that

$$\begin{aligned}
\lambda_{\min}\left(\mathbf{V}_p\right) &= \lambda_{\min}\left(\mathbf{V}_p\right)\|\mathbf{u}_{1-p}\|_2^2 \leq \mathbf{u}_{1-p}^{\mathrm{T}}\mathbf{V}_p\mathbf{u}_{1-p} = \|B_{1-p,p}\|_2^2, \\
\mathbf{u}_0^{\mathrm{T}}\mathbf{V}_p\mathbf{u}_0 &= \|B_{0,p}\|_2^2 \leq \lambda_{\max}\left(\mathbf{V}_p\right)\|\mathbf{u}_0\|_2^2 = \lambda_{\max}\left(\mathbf{V}_p\right),
\end{aligned}$$

hence $d_p = \lambda_{\max}\left(\mathbf{V}_p\right)/\lambda_{\min}\left(\mathbf{V}_p\right) \geq \|B_{0,p}\|_2^2\|B_{1-p,p}\|_2^{-2} = r_p > 1$, where r_p is an constant depending only on p. Thus $\nu_p = \left(d_p^2-1\right)^{1/4p}\left(d_p^2+1\right)^{-1/4p} \geq \left(r_p^2-1\right)^{1/4p}\left(r_p^2+1\right)^{-1/4p} > 0$.

Applying Lemma A.1 and putting the above bounds together, one obtains

$$
\begin{aligned}
& \left\|\mathbf{V}_p^{-1}\right\|_\infty h_{\mathrm{s}} \leq 2\nu_p^{-2p}\left\|\mathbf{V}_p^{-1}\right\|_2 (1-\nu_p)^{-1} h_{\mathrm{s}} \\
\leq\ & 2\left(r_p^2-1\right)^{-1/2}\left(r_p^2+1\right)^{1/2}\lambda_{\min}^{-1}(\mathbf{V}_p) \\
& \times\left[1-\left(C_p^2c_p^{-2}-1\right)^{1/4p}\times\left(C_p^2c_p^{-2}+1\right)^{-1/4p}\right]^{-1} h_{\mathrm{s}} \\
\leq\ & 2\left(r_p^2-1\right)^{-1/2}\left(r_p^2+1\right)^{1/2} c_p^{-1}\left[1-\left(C_p^2c_p^{-2}-1\right)^{1/4p}\left(C_p^2c_p^{-2}+1\right)^{-1/4p}\right]^{-1} \equiv M_p.
\end{aligned}
$$

The lemma is proved. □

Lemma A.3. *For any positive integer p, there exists a constant $M_p > 0$ depending only on p, such that* $\left\|(\mathbf{V}_p\otimes\mathbf{V}_p)^{-1}\right\|_\infty \leq M_p^2 h_{\mathrm{s}}^{-2}$.

Proof. By Lemma A.2, $\left\|(\mathbf{V}_p\otimes\mathbf{V}_p)^{-1}\right\|_\infty = \left\|\mathbf{V}_p^{-1}\otimes\mathbf{V}_p^{-1}\right\|_\infty \leq \left(\left\|\mathbf{V}_p^{-1}\right\|_\infty\right)^2 \leq M_p^2 h_{\mathrm{s}}^{-2}$. □

Lemma A.4. *Under Assumption (A3), for $\mathbf{V}_{p_2,2}$ and $\hat{\mathbf{V}}_{p_2,2}$ defined in (7),* $\left\|\mathbf{V}_{p_2,2}-\hat{\mathbf{V}}_{p_2,2}\right\|_\infty = O\left(N_{\mathrm{s}_2}N^{-1}\right)$ *and* $\left\|\hat{\mathbf{V}}_{p_2,2}^{-1}\right\|_\infty = O\left(h_{\mathrm{s}_2}^{-2}\right)$.

Proof. Note that

$$
\begin{aligned}
\hat{\mathbf{V}}_{p_2,2} &= \left\{N^{-2}\sum_{1\leq j\neq j'\leq N} B_{JJ',p_2}(j/N,j'/N)\,B_{J''J''',p_2}(j/N,j'/N)\right\}_{J,J',J'',J'''=1-p_2}^{N_{\mathrm{s}_2}} \\
&= \left\{N^{-2}\left[\sum_{j=1}^N B_{JJ',p_2}(j/N,j/N)\right]\left[\sum_{j=1}^N B_{J''J''',p_2}(j/N,j/N)\right]\right\}_{J,J',J'',J'''=1-p_2}^{N_{\mathrm{s}_2}} \\
&\quad -\left\{N^{-2}\sum_{j=1}^N B_{JJ',p_2}(j/N,j/N)\,B_{J''J''',p_2}(j/N,j/N)\right\}_{J,J',J'',J'''=1-p_2}^{N_{\mathrm{s}_2}} \\
&= \hat{\mathbf{V}}_{p_2}\otimes\hat{\mathbf{V}}_{p_2} \\
&\quad -\left\{N^{-1}\left[\int_0^1 B_{JJ',p_2}(x,x)\,B_{J''J''',p_2}(x,x)\,dx + O(N^{-1}h_{\mathrm{s}_2}^{-1})\right]\right\}_{J,J',J'',J'''=1-p_2}^{N_{\mathrm{s}_2}} \\
&= \hat{\mathbf{V}}_{p_2}\otimes\hat{\mathbf{V}}_{p_2} + O\left(N^{-1}h_{\mathrm{s}_2} + N^{-2}h_{\mathrm{s}_2}^{-1}\right).
\end{aligned}
$$

Hence, $\left\|\hat{\mathbf{V}}_{p_2,2}-\hat{\mathbf{V}}_{p_2}\otimes\hat{\mathbf{V}}_{p_2}\right\|_\infty = O\left(N^{-1}h_{\mathrm{s}_2}+N^{-2}h_{\mathrm{s}_2}^{-1}\right)\times N_{\mathrm{s}_2}^2 = O\left(N_{\mathrm{s}_2}N^{-1}\right)$. Since

$$\begin{aligned}
&\left\|\mathbf{V}_{p_2}\otimes\mathbf{V}_{p_2}-\hat{\mathbf{V}}_{p_2}\otimes\hat{\mathbf{V}}_{p_2}\right\|_\infty \\
= & \max_{1-p_2\le J',J'''\le N_{\mathrm{s}_2}} \sum_{J,J''=1-p_2}^{N_{\mathrm{s}_2}} \left| N^{-2}\sum_{j,j'=1}^{N} B_{JJ',p_2}\left(j/N,j'/N\right)B_{J''J''',p_2}\left(j/N,j'/N\right) \right. \\
& \left. -\int_0^1\int_0^1 B_{JJ',p_2}\left(x,x'\right)B_{J''J''',p_2}\left(x,x'\right)dxdx' \right| \\
\le & \max_{1-p_2\le J',J'''\le N_{\mathrm{s}_2}} \sum_{J,J''=1-p_2}^{N_{\mathrm{s}_2}} \sum_{j,j'=1}^{N} \int_{(j'-1)/N}^{j'/N}\int_{(j-1)/N}^{j/N} \left|B_{JJ',p_2}\left(j/N,j'/N\right)\right. \\
& \left. \times B_{J''J''',p_2}\left(j/N,j'/N\right)-B_{JJ',p_2}\left(x,x'\right)B_{J''J''',p_2}\left(x,x'\right)\right| dxdx' \\
\le & Ch_{\mathrm{s}_2}^{-2}\left(Nh_{\mathrm{s}_2}\right)^2\times N^{-2}\times N^{-2}h_{\mathrm{s}_2}^{-2} = CN^{-2}h_{\mathrm{s}_2}^{-2},
\end{aligned}$$

applying Assumption (A3) one has $\left\|\mathbf{V}_{p_2}\otimes\mathbf{V}_{p_2}-\hat{\mathbf{V}}_{p_2,2}\right\|_\infty = O\left(N_{\mathrm{s}_2}N^{-1}\right)$.

According to Lemma A.3, for any $\left(N_{\mathrm{s}_2}+p_2\right)^2$ vector $\boldsymbol{\tau}$, one has $\|(\mathbf{V}_{p_2}\otimes\mathbf{V}_{p_2})^{-1}\boldsymbol{\tau}\|_\infty \le h_{\mathrm{s}_2}^{-2}\|\boldsymbol{\tau}\|_\infty$. Hence, $\|(\mathbf{V}_{p_2}\otimes\mathbf{V}_{p_2})\boldsymbol{\tau}\|_\infty \ge h_{\mathrm{s}_2}^2\|\boldsymbol{\tau}\|_\infty$. Note that

$$\left\|\hat{\mathbf{V}}_{p_2,2}\boldsymbol{\tau}\right\|_\infty \ge \|(\mathbf{V}_{p_2}\otimes\mathbf{V}_{p_2})\boldsymbol{\tau}\|_\infty - \left\|(\mathbf{V}_{p_2}\otimes\mathbf{V}_{p_2})\boldsymbol{\tau}-\hat{\mathbf{V}}_{p_2,2}\boldsymbol{\tau}\right\|_\infty = O\left(h_{\mathrm{s}_2}^2\right)\|\boldsymbol{\tau}\|_\infty .$$

If τ satisfies that $\left\|\hat{\mathbf{V}}_{p_2,2}^{-1}\right\|_\infty = \left\|\hat{\mathbf{V}}_{p_2,2}^{-1}\boldsymbol{\tau}\right\|_\infty \le O\left(h_{\mathrm{s}_2}^{-2}\right)\|\boldsymbol{\tau}\|_\infty = O\left(h_{\mathrm{s}_2}^{-2}\right)$, the lemma is proved. □

LEMMA A.5. *For $\hat{\mathbf{V}}_{p_2,2}$ defined in (7) and any $N(N-1)$ vector $\boldsymbol{\rho}=\left(\rho_{jj'}\right)$, there exists a constant $C>0$, such that* $\sup_{(x,x')\in[0,1]^2}\left\|N^{-2}\boldsymbol{B}_{p_2}^{\mathrm{T}}(x,x')\hat{\mathbf{V}}_{p_2,2}^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{\rho}\right\|_\infty \le C\|\boldsymbol{\rho}\|_\infty$.

PROOF. Firstly, one has that

$$\begin{aligned}
\left\|N^{-2}\mathbf{X}^{\mathrm{T}}\boldsymbol{\rho}\right\|_2 &= \left\|N^{-2}\left\{\sum_{1\le j\ne j'\le N} B_{JJ',p_2}\left(j/N,j'/N\right)\rho_{jj'}\right\}_{J,J'=1-p_2}^{N_{\mathrm{s}_2}}\right\|_2 \\
&\le \|\boldsymbol{\rho}\|_\infty \max_{1-p_2\le J,J'\le N_{\mathrm{s}_2}}\left\|N^{-2}\sum_{1\le j\ne j'\le N} B_{JJ',p_2}\left(j/N,j'/N\right)\right\|_2 \le h_{\mathrm{s}_2}^2\|\boldsymbol{\rho}\|_\infty .
\end{aligned}$$

Note that for any matrix $\mathbf{A}=(a_{ij})_{i=1,j=1}^{m,n}$ and any n by 1 vector $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_n)^{\mathrm{T}}$, one has $\|\mathbf{A}\boldsymbol{\alpha}\|_\infty \le \|\mathbf{A}\|_\infty\|\boldsymbol{\alpha}\|_2$, the above together yield that

$$\begin{aligned}
&\sup_{(x,x')\in[0,1]^2}\left\|N^{-2}\boldsymbol{B}_{p_2}^{\mathrm{T}}(x,x')\hat{\mathbf{V}}_{p_2,2}^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{\rho}\right\|_\infty \\
\le & \sup_{(x,x')\in[0,1]^2}\left\|\boldsymbol{B}_{p_2}^{\mathrm{T}}(x,x')\right\|_\infty\left\|N^{-2}\mathbf{X}^{\mathrm{T}}\boldsymbol{\rho}\right\|_2\left\|\hat{\mathbf{V}}_{p_2,2}^{-1}\right\|_\infty \\
\le & C\sup_{(x,x')\in[0,1]^2}\left\|\boldsymbol{B}_{p_2}^{\mathrm{T}}(x,x')\right\|_\infty h_{\mathrm{s}_2}^{-2}h_{\mathrm{s}_2}^2\|\boldsymbol{\rho}\|_\infty \le C\|\boldsymbol{\rho}\|_\infty ,
\end{aligned}$$

which proves the lemma. $\square$

Denote $\widetilde{\boldsymbol{\phi}}_{kk'}(x,x') = \boldsymbol{B}_{p_2}^{\mathrm{T}}(x,x')\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{\phi}_{kk'}$, and

$$\boldsymbol{\phi}_{kk'} = \left(\phi_k(2/N)\,\phi_{k'}(1/N),\ldots,\phi_k(1)\,\phi_{k'}(1/N),\ldots,\phi_k(1/N)\,\phi_{k'}(1),\ldots,\phi_k(1-1/N)\,\phi_{k'}(1)\right)^{\mathrm{T}}.$$

We cite next an important result concerning the function $\widetilde{\boldsymbol{\phi}}_{kk'}(x,x')$. The first part is from [3], p. 149 and the second is from Theorem 5.1 of [9].

Theorem A.1. *There is an absolute constant $C_g > 0$ such that for every $g \in C[0,1]$, there exists a function $g^* \in \mathcal{H}^{(p-1)}[0,1]$ such that $\|g-g^*\|_\infty \leq C_g\omega(g,h_{\mathrm{s}})$ and in particular, if $g \in C^{p-1,\mu}[0,1]$, for some $\mu \in (0,1]$, then $\|g-g^*\|_\infty \leq C_g h_{\mathrm{s}}^{p-1+\mu}$. Therefore, under Assumption (A2), $\left\|\boldsymbol{\phi}_{kk'} - \widetilde{\boldsymbol{\phi}}_{kk'}\right\|_\infty = O\left(h_{\mathrm{s}_2}^{p_2}\right)$.*

A.2 Proof of Proposition 2.1

Recall that the errors defined in Section 2 are $U_{ij} = Y_{ij} - m(j/N)$ and $U_{ij,p_1} = Y_{ij} - \hat{m}_{p_1}(j/N)$. For simplicity, denote

$$\mathbf{X}_1 = \begin{pmatrix} B_{1-p_1,p_1}(1/N) & \cdots & B_{N_{\mathrm{s}_1},p_1}(1/N) \\ \cdots & \cdots & \cdots \\ B_{1-p_1,p_1}(N/N) & \cdots & B_{N_{\mathrm{s}_1},p_1}(N/N) \end{pmatrix}_{N\times\left(N_{\mathrm{s}_1}+p_1\right)}$$

for the positive integer p_1. We decompose $\hat{m}_{p_1}(j/N)$ into three terms $\tilde{m}_{p_1}(j/N)$, $\tilde{\xi}_{p_1}(j/N)$ and $\tilde{\varepsilon}_{p_1}(j/N)$ in the space $\mathcal{H}^{(p_1-2)}$ of spline functions: $\hat{m}_{p_1}(x) = \tilde{m}_{p_1}(x) + \tilde{\varepsilon}_{p_1}(x) + \tilde{\xi}_{p_1}(x)$, where

$$\tilde{m}_p(x) = \{B_{1-p_1,p_1}(x),\ldots,B_{N_{\mathrm{s}},p_1}(x)\}\left(\mathbf{X}_1^{\mathrm{T}}\mathbf{X}_1\right)^{-1}\mathbf{X}_1^{\mathrm{T}}\mathbf{m},$$

$$\tilde{\varepsilon}_{p_1}(x) = \{B_{1-p_1,p_1}(x),\ldots,B_{N_{\mathrm{s}},p_1}(x)\}\left(\mathbf{X}_1^{\mathrm{T}}\mathbf{X}_1\right)^{-1}\mathbf{X}_1^{\mathrm{T}}\mathbf{e},$$

$$\tilde{\xi}_{p_1}(x) = \{B_{1-p_1,p_1}(x),\ldots,B_{N_{\mathrm{s}},p_1}(x)\}\left(\mathbf{X}_1^{\mathrm{T}}\mathbf{X}_1\right)^{-1}\mathbf{X}_1^{\mathrm{T}}\sum_{k=1}^{\kappa}\overline{\xi}_{\cdot k}\boldsymbol{\phi}_k,$$

in which $\mathbf{m} = (m(1/N),\ldots,m(N/N))^{\mathrm{T}}$ is the signal vector, $\mathbf{e} = \left(\sigma(1/N)\,\overline{\varepsilon}_{\cdot 1},\ldots,\sigma(N/N)\,\overline{\varepsilon}_{\cdot N}\right)^{\mathrm{T}}$, $\overline{\varepsilon}_{\cdot j} = n^{-1}\sum_{i=1}^n \varepsilon_{ij}, 1 \leq j \leq N$ is the noise vector and $\boldsymbol{\phi}_k = (\phi_k(1/N),\ldots,\phi_k(N/N))^{\mathrm{T}}$ are the eigenfunction vectors, and $\overline{\xi}_{\cdot k} = n^{-1}\sum_{i=1}^n \xi_{ik}$, $1 \leq k \leq \kappa$. Thus, one can write $\hat{U}_{ij,p_1} = m(j/N) - \tilde{m}_{p_1}(j/N) - \tilde{\xi}_{p_1}(j/N) - \tilde{\varepsilon}_{p_1}(j/N) + U_{ij}$. We calculate the difference of $\hat{G}_{p_1,p_2}(x,x') - \tilde{G}_{p_2}(x,x')$ by checking the difference $\hat{\bar{U}}_{\cdot jj',p_1} - \bar{U}_{\cdot jj'}$ first. For any $1 \leq j \neq j' \leq N$, one has

$$
\begin{aligned}
& \hat{\hat{U}}_{\cdot jj',p_1} - \bar{U}_{\cdot jj'} \\
= \;& n^{-1}\sum_{i=1}^{n}\Big[U_{ij}\left(\tilde{m}_{p_1}-m\right)(j'/N) + U_{ij'}\left(\tilde{m}_{p_1}-m\right)(j/N) + U_{ij}\tilde{\xi}_{p_1}(j'/N) \\
+ \;& U_{ij'}\tilde{\xi}_{p_1}(j/N) + U_{ij}\tilde{\varepsilon}_{p_1}(j'/N) + U_{ij'}\tilde{\varepsilon}_{p_1}(j/N)\Big] + \left(\hat{m}_{p_1}-m\right)(j'/N)\left(\hat{m}_{p_1}-m\right)(j/N).
\end{aligned}
$$

Next, we calculate the super norm of each part of $\hat{G}_{p_1,p_2}(x,x') - \tilde{G}_{p_2}(x,x')$ respectively. One can write $\tilde{\varepsilon}_{p_1}(j'/N) = \sum_{J=1-p_1}^{N_{s_1}} B_{J,p_1}(j'/N)\, w_J$, where $\{w_J\}_{J=1-p_1}^{N_{s_1}} = N^{-1}\hat{\mathbf{V}}_{p_1}^{-1}\mathbf{X}_1^{\mathrm{T}}\mathbf{e}$ and

$$
\begin{aligned}
& \mathbf{X}^{\mathrm{T}}\left\{n^{-1}\sum_{i=1}^{n} U_{ij}\tilde{\varepsilon}_{p_1}(j'/N)\right\}_{1\le j\ne j'\le N} \\
= \;& \left\{n^{-1}\sum_{i=1}^{n}\sum_{1\le j\ne j'\le N} U_{ij}\tilde{\varepsilon}_{p_1}(j'/N)\, B_{JJ',p_2}(j/N, j'/N)\right\}_{J,J'=1-p_2}^{N_{s_2}}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
& \left\|\mathbf{X}^{\mathrm{T}}\left\{n^{-1}\sum_{i=1}^{n} U_{ij}\tilde{\varepsilon}_{p_1}(j'/N)\right\}_{1\le j\ne j'\le N}\right\|_2^2 \\
= \;& \sum_{J,J'=1-p_2}^{N_{s_2}}\left[n^{-1}\sum_{i=1}^{n}\sum_{1\le j\ne j'\le N} U_{ij}\left(\sum_{J''=1-p_1}^{N_{s_1}} B_{J'',p_1}(j'/N)\, w_{J''}\right) B_{JJ',p_2}(j/N, j'/N)\right]^2 \\
\le \;& \sum_{J,J'=1-p_2}^{N_{s_2}}\sum_{J''=1-p_1}^{N_{s_1}}\left[n^{-1}\sum_{i=1}^{n}\sum_{1\le j\ne j'\le N} U_{ij}B_{J'',p_1}(j'/N)\, B_{JJ',p_2}(j/N, j'/N)\right]^2 \times \sum_{J''=1-p_1}^{N_{s_1}} w_{J''}^2 \\
= \;& I \times II,
\end{aligned}
$$

where $I = \sum_{J,J'=1-p_2}^{N_{s_2}}\sum_{J''=1-p_1}^{N_{s_1}}\left\{n^{-1}\sum_{i=1}^{n}\sum_{1\le j\ne j'\le N} U_{ij}B_{JJ',p_2}(j/N, j'/N)\, B_{J'',p_1}(j'/N)\right\}^2$ and $II = \left\|N^{-1}\hat{\mathbf{V}}_{p_1}^{-1}\mathbf{X}_1^{\mathrm{T}}\mathbf{e}\right\|_2^2$.

Let $h_* = \min\{h_{s_1}, h_{s_2}\}$. The definition of spline function implies that

$$
\begin{aligned}
E[I] \;=\;& n^{-1}\sum_{J,J'=1-p_2}^{N_{s_2}}\sum_{J''=1-p_1}^{N_{s_1}}\left\{\sum_{j,j''=1}^{N} E\left(U_{1j}U_{1j''}\right) B_{J,p_2}(j/N)\, B_{J,p_2}(j''/N)\right. \\
& \left.\times\sum_{j'\ne j}^{N}\sum_{j'''\ne j''}^{N} B_{J',p_2}(j'/N)\, B_{J'',p_1}(j'/N)\, B_{J',p_2}(j'''/N)\, B_{J'',p_1}(j'''/N)\right\} \\
\le\;& C(G,\sigma^2)n^{-1}N^4 h_{s_2}^2 h_*^2 N_{s_2}\max\{N_{s_1}, N_{s_2}\} \le C(G,\sigma^2)n^{-1}N^4 h_{s_2} h_*.
\end{aligned}
$$

Hence, $\left\|N^{-2}\mathbf{X}^{\mathrm{T}}\left\{n^{-1}\sum_{i=1}^{n}U_{ij}\tilde{\varepsilon}_{p_1}\left(j'/N\right)\right\}_{1\leq j\neq j'\leq N}\right\|_2 = O\left(n^{-1/2}h_{\mathrm{s}_2}^{1/2}h_*^{1/2}\right)$. Meanwhile, one has that $II \leq C_{p_1}\left\|N^{-1}\mathbf{X}_1^{\mathrm{T}}\mathbf{e}\right\|_2^2 h_{\mathrm{s}_1}^{-2} = O_{\mathrm{a.s.}}\left\{(Nnh_{\mathrm{s}_1}^2)^{-1}\right\}$. By Lemma A.5, one has

$$\begin{aligned} & n^{1/2}\sup_{x,x'\in[0,1]}\left\|\boldsymbol{B}_{p_2}^{\mathrm{T}}(x,x')\hat{\mathbf{V}}_{p_2,2}^{-1}N^{-2}\mathbf{X}^{\mathrm{T}}\left\{n^{-1}\sum_{i=1}^{n}U_{ij}\tilde{\varepsilon}_{p_1}\left(j'/N\right)\right\}_{1\leq j\neq j'\leq N}\right\|_\infty \\ = \; & O\left\{n^{-1/2}N^{-1/2}h_{\mathrm{s}_1}^{-1}h_{\mathrm{s}_2}^{1/2}h_*^{1/2}h_{\mathrm{s}_2}^{-2}\right\} = O\left\{n^{-1/2}N^{-1/2}h_{\mathrm{s}_1}^{-1/2}h_{\mathrm{s}_2}^{-3/2}\right\} = o\left(1\right). \end{aligned}$$

Similarly, one has $\tilde{\xi}_{p_1}\left(j'/N\right) = \sum_{J=1-p_1}^{N_{s_1}} B_{J,p_1}\left(j'/N\right)s_J$, where $\{s_J\}_{J=1-p_1}^{N_{s_1}} = N^{-1}\hat{\mathbf{V}}_{p_1}^{-1}\mathbf{X}_1^{\mathrm{T}}\sum\limits_{k=1}^{\kappa}\bar{\xi}_{\cdot k}\boldsymbol{\phi}_k$. Assumption (A3) ensures that

$$\begin{aligned} & \left\|\mathbf{X}^{\mathrm{T}}\left\{n^{-1}\sum_{i=1}^{n}U_{ij}\tilde{\xi}_{p_1}\left(j'/N\right)\right\}_{1\leq j\neq j'\leq N}\right\|_2^2 \\ \leq \; & \sum_{J,J'=1-p_2}^{N_{s_2}}\sum_{J''=1-p_1}^{N_{s_1}}\left[n^{-1}\sum_{i=1}^{n}\sum_{1\leq j\neq j'\leq N}U_{ij}B_{J'',p_1}\left(j'/N\right)B_{JJ',p_2}\left(j/N,j'/N\right)\right]^2\times\sum_{J''=1-p_1}^{N_{s_1}}s_{J''}^2 \\ = \; & I\times III, \end{aligned}$$

where $III = \left\|N^{-1}\hat{\mathbf{V}}_{p_1}^{-1}\mathbf{X}_1^{\mathrm{T}}\sum_{k=1}^{\kappa}\bar{\xi}_{\cdot k}\boldsymbol{\phi}_k\right\|_2^2$. Note that

$$N^{-1}\mathbf{X}_1^{\mathrm{T}}\sum_{k=1}^{\kappa}\bar{\xi}_{\cdot k}\boldsymbol{\phi}_k = \left\{\sum_{k=1}^{\kappa}\bar{\xi}_{\cdot k}N^{-1}\sum_{j=1}^{N}B_{J,p_1}\left(j/N\right)\phi_k\left(j/N\right)\right\}_{J=1-p_1}^{N_{\mathrm{s}_1}}$$

and

$$E\left[\sum_{k=1}^{\kappa}\bar{\xi}_{\cdot k}N^{-1}\sum_{j=1}^{N}B_{J,p_1}\left(j/N\right)\phi_k\left(j/N\right)\right]^2 = O\left(n^{-1}h_{\mathrm{s}_1}^2\right),$$

hence $III \leq C_{p_1}\left\|N^{-1}\mathbf{X}_1^{\mathrm{T}}\sum_{k=1}^{\kappa}\bar{\xi}_{\cdot k}\boldsymbol{\phi}_k\right\|_2^2 h_{\mathrm{s}_1}^{-2} = O_p\left\{(nh_{\mathrm{s}_1})^{-1}\right\}$ and

$$\begin{aligned} & n^{1/2}\sup_{x,x'\in[0,1]}\left\|\boldsymbol{B}_{p_2}^{\mathrm{T}}(x,x')\hat{\mathbf{V}}_{p_2,2}^{-1}N^{-2}\mathbf{X}^{\mathrm{T}}\left\{n^{-1}\sum_{i=1}^{n}U_{ij}\tilde{\xi}_{p_1}\left(j'/N\right)\right\}_{1\leq j\neq j'\leq N}\right\|_\infty \\ = \; & O\left\{n^{-1/2}h_{\mathrm{s}_1}^{-1/2}h_{\mathrm{s}_2}^{1/2}h_*^{1/2}h_{\mathrm{s}_2}^{-2}\right\} = O(n^{-1/2}h_{\mathrm{s}_2}^{-3/2}) = o\left(1\right). \end{aligned}$$

Next one obtains that

$$
\begin{aligned}
& E\left\{n^{-1}N^{-2}\sum_{i=1}^{n}\sum_{1\le j\neq j'\le N}U_{ij}B_{J,p_2}\left(j/N\right)B_{J',p_2}\left(j'/N\right)\left(m-\tilde{m}_{p_1}\right)\left(j'N\right)\right\}^2 \\
= \; & n^{-1}N^{-4}\sum_{j,j''=1}^{N}E\left(U_{1j}U_{1j''}\right)B_{J,p_2}\left(j/N\right)B_{J,p_2}\left(j''/N\right) \\
& \times\sum_{j'\neq j}^{N}\sum_{j'''\neq j''}^{N}B_{J',p_2}\left(j'/N\right)B_{J',p_1}\left(j'''/N\right)\left(m-\tilde{m}_{p_1}\right)\left(j'/N\right)\left(m-\tilde{m}_{p_1}\right)\left(j'''/N\right) \\
\le \; & C(G,\sigma^2)h_{\mathrm{s}_1}^{2p_1}n^{-1}N^{-4}\left(Nh_{\mathrm{s}_2}\right)^4 = C(G,\sigma^2)h_{\mathrm{s}_1}^{2p_1}n^{-1}h_{\mathrm{s}_2}^4.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
& n^{1/2}\sup_{x,x'\in[0,1]}\left\|\boldsymbol{B}_{p_2}^{\mathrm{T}}(x,x')\hat{\mathbf{V}}_{p_2,2}^{-1}N^{-2}\mathbf{X}^{\mathrm{T}}\left\{n^{-1}\sum_{i=1}^{n}U_{ij}\left(m\left(j'/N\right)-\tilde{m}_{p_1}\left(j'/N\right)\right)\right\}_{1\le j\neq j'\le N}\right\|_\infty \\
= \; & O(h_{\mathrm{s}_2}^{-2}h_{\mathrm{s}_1}^{p_1}h_{\mathrm{s}_2}) = O(h_{\mathrm{s}_2}^{-1}h_{\mathrm{s}_1}^{p_1}) = o\left(1\right).
\end{aligned}
$$

Finally, we derive the upper bound of $\sup_{x,x'\in[0,1]}\left\|\boldsymbol{B}_{p_2}^{\mathrm{T}}(x,x')\hat{\mathbf{V}}_{p_2,2}^{-1}N^{-2}\mathbf{X}^{\mathrm{T}}\left(\mathbf{m}-\hat{\mathbf{m}}_{p_1}\right)^{\otimes 2}\right\|_\infty$, where $\left(\mathbf{m}-\hat{\mathbf{m}}_{p_1}\right)^{\otimes 2} = \left\{\left(m-\hat{m}_{p_1}\right)\left(j/N\right)\left(m-\hat{m}_{p_1}\right)\left(j'/N\right)\right\}_{1\le j\neq j'\le N}$. In order to apply Lemma A.5, one needs to find the upper bound of $\left\|\left(\mathbf{m}-\hat{\mathbf{m}}_{p_1}\right)^{\otimes 2}\right\|_\infty$. Using the similar proof as Lemma A.8 in [21] and Assumption (A3), one has

$$
\sup_{x\in[0,1]}\left|\tilde{\varepsilon}_{p_1}\left(x\right)\right| + \sup_{x\in[0,1]}\left|\tilde{\xi}_{p_1}\left(x\right)\right| = o\left(n^{-1/2}\right).
$$

Therefore,

$$
\begin{aligned}
& \sup_{(x,x')\in[0,1]^2}\left(m\left(x\right)-\hat{m}_{p_1}\left(x\right)\right)\left(m\left(x'\right)-\hat{m}_{p_1}\left(x'\right)\right) \\
\le \; & \left[\sup_{x\in[0,1]}\left(m\left(x\right)-\hat{m}_{p_1}\left(x\right)\right)\right]^2 \\
\le \; & \left(\sup_{x\in[0,1]}\left|m\left(x\right)-\tilde{m}_{p_1}\left(x\right)\right| + \sup_{x\in[0,1]}\left|\tilde{\varepsilon}_{p_1}\left(x\right)\right| + \sup_{x\in[0,1]}\left|\tilde{\xi}_{p_1}\left(x\right)\right|\right)^2 \\
\le \; & \left[O\left(h_{\mathrm{s}_1}^{p_1}+n^{-1/2}\right)\right]^2 = O\left(h_{\mathrm{s}_1}^{2p_1}+n^{-1}+h_{\mathrm{s}_1}^{p_1}n^{-1/2}\right) = o\left(n^{-1/2}\right).
\end{aligned}
$$

Hence $\left\|\left(\mathbf{m}-\hat{\mathbf{m}}_{p_1}\right)^{\otimes 2}\right\|_\infty = o\left(n^{-1/2}\right)$. Hence, the proposition has been proved. □

A.3 Proof of Proposition 3.1

By the definition of $\bar{\mathbf{U}}_1$, one has that

$$\begin{aligned}\tilde{\mathcal{U}}_{1p_2}(x,x') &= N^{-2}\boldsymbol{B}_{p_2}^{\mathrm{T}}(x,x')\hat{\mathbf{V}}_{p_2,2}^{-1}\mathbf{X}^{\mathrm{T}}\left\{n^{-1}\sum_{i=1}^{n}\sum_{k\neq k'}^{\kappa}\xi_{ik}\xi_{ik'}\phi_k(j/N)\,\phi_{k'}(j'/N)\right\} \\ &= \sum_{k\neq k'}^{\kappa}\left(n^{-1}\sum_{i=1}^{n}\xi_{ik}\xi_{ik'}\right)\widetilde{\boldsymbol{\phi}}_{kk'}(x,x').\end{aligned}$$

Theorem A.1 and Assumption (A3) imply that

$$\begin{aligned}\left\|\tilde{\mathcal{U}}_{1p_2}-\mathcal{U}_1\right\|_\infty &= \left\|\tilde{\mathcal{U}}_{1p_2}-\sum_{k\neq k'}^{\kappa}\left(n^{-1}\sum_{i=1}^{n}\xi_{ik}\xi_{ik'}\right)\widetilde{\boldsymbol{\phi}}_{kk'}\right\|_\infty \\ &\leq \kappa^2\max_{1\leq k\neq k'\leq\kappa}\left\|n^{-1}\sum_{i=1}^{n}\xi_{ik}\xi_{ik'}\right\|_\infty\left\|\boldsymbol{\phi}_{kk'}-\widetilde{\boldsymbol{\phi}}_{kk'}\right\|_\infty \\ &= O_p\left(h_{s_2}^{p_2}n^{-1/2}\right)=o_p\left(n^{-1/2}\right).\end{aligned}$$

Similarly,

$$\begin{aligned}\left\|\tilde{\mathcal{U}}_{2p_2}-\mathcal{U}_2\right\|_\infty &= \left\|\tilde{\bar{U}}_{2p_2}-\sum_{k=1}^{\kappa}\left(n^{-1}\sum_{i=1}^{n}\xi_{ik}^2\right)\boldsymbol{\phi}_{kk}\right\|_\infty \\ &\leq \kappa\max_{1\leq k\leq\kappa}\left|n^{-1}\sum_{i=1}^{n}\xi_{ik}^2\right|\left\|\boldsymbol{\phi}_{kk}-\widetilde{\boldsymbol{\phi}}_{kk}\right\|_\infty \\ &= O_p\left(h_{s_2}^{p_2}\right)=o_p\left(n^{-1/2}\right).\end{aligned}$$

Therefore, one has $\left\|\tilde{\mathcal{U}}_{1p_2}+\tilde{\mathcal{U}}_{2p_2}-\mathcal{U}_1-\mathcal{U}_2\right\|_\infty=o_p\left(n^{-1/2}\right)$. Denote that

$$N^{-2}\mathbf{X}^{\mathrm{T}}\bar{\mathbf{U}}_3=\left\{n^{-1}\sum_{i=1}^{n}A_{iJJ'}\right\}_{J,J'=1-p_2}^{N_{s_2}},$$

where $A_{iJJ'}=N^{-2}\sum_{1\leq j\neq j'\leq N}B_{J,p_2}(j/N)\,\sigma(j/N)\,B_{J',p_2}(j'/N)\,\sigma(j'/N)\,\varepsilon_{ij}\varepsilon_{ij'}$. It is easy to see that $EA_{iJJ'}=0$ and $EA_{iJJ'}^2=O\left(h_{s_2}^2N^{-2}\right)$. Using standard arguments in [21], one has $\left\|N^{-2}\mathbf{X}^{\mathrm{T}}\bar{\mathbf{U}}_3\right\|_\infty=o_{\text{a.s.}}\left\{N^{-1}n^{-1/2}h_{s_2}\log^{1/2}(n)\right\}$. Therefore, according to the definition of $\tilde{\bar{U}}_{3p_2}(x,x')$, one has

$$\begin{aligned}&\sup_{x,x'\in[0,1]^2}\left\|\boldsymbol{B}_{p_2}^{\mathrm{T}}(x,x')\hat{\mathbf{V}}_{p_2,2}^{-1}N^{-2}\mathbf{X}^{\mathrm{T}}\bar{\mathbf{U}}_3\right\|_\infty \\ \leq\ & C_{p_2}\sup_{x,x'\in[0,1]^2}\left\|\mathbf{B}_{p_2}(x,x')\right\|_\infty\left\|\hat{\mathbf{V}}_{p_2,2}^{-1}N^{-2}\mathbf{X}^{\mathrm{T}}\bar{\mathbf{U}}_3\right\|_\infty \\ =\ & o_{\text{a.s.}}\left\{N^{-1}n^{-1/2}h_{s_2}^{-1}\log^{1/2}n\right\}=o_{\text{a.s.}}\left(n^{-1/2}\right).\end{aligned}$$

Likewise, in order to get the upper bound of $\left|\tilde{\mathcal{U}}_{4p_2}(x,x')\right|$, one has

$$\begin{aligned} & N^{-2}\mathbf{X}^{\mathrm{T}}\bar{\mathbf{U}}_4 \\ = & \left\{2n^{-1}N^{-2}\sum\nolimits_{i=1}^{n}\sum\nolimits_{k=1}^{\kappa}\xi_{ik}\sum_{1\le j\ne j'\le N}\phi_k(j/N)\,\sigma(j'/N)\,B_{JJ',p_2}(j/N,j'/N)\,\varepsilon_{ij'}\right\}_{J,J'=1-p_2}^{N_{s_2}}. \end{aligned}$$

Let $D_{iJJ'}=N^{-2}\sum_{k=1}^{\kappa}\left(\xi_{ik}\sum_{1\le j\ne j'\le N}\phi_k(j/N)\,\sigma(j'/N)\,B_{JJ',p}(j/N,j'/N)\,\varepsilon_{ij'}\right)$,
then $ED_{iJJ'}=0$,

$$\begin{aligned} ED^2_{iJJ'} &= N^{-4}\sum_{k=1}^{\kappa}\sum_{1\le j\ne j'\le N}\phi_k^2(j/N)\,\sigma^2(j'/N)\,B^2_{JJ',p_2}(j/N,j'/N)\,E\xi_{ik}^2E\varepsilon_{ij'}^2 \\ &\le CN^{-4}\sum_{k=1}^{\kappa}\sum_{1\le j\ne j'\le N}\phi_k^2(j/N)\,B^2_{J,p_2}(j/N)\,\sigma^2(j'/N)\,B^2_{J',p_2}(j'/N)=O\left(h_{s_2}^2N^{-2}\right). \end{aligned}$$

Similar arguments in [21] leads to $\left\|\frac{\mathbf{X}^{\mathrm{T}}\bar{\mathbf{U}}_4}{N^2}\right\|_\infty=o_{\text{a.s.}}\left\{N^{-1}n^{-1/2}h_{s_2}\log^{1/2}(n)\right\}$. Thus,

$$\left\|\boldsymbol{B}^{\mathrm{T}}_{p_2}(x,x')\hat{\mathbf{V}}^{-1}_{p_2,2}\frac{\mathbf{X}^{\mathrm{T}}\bar{\mathbf{U}}_4}{N^2}\right\|_\infty=o_{\text{a.s.}}\left\{N^{-1}n^{-1/2}h_{s_2}^{-1}\log^{1/2}(n)\right\}=o_{\text{a.s.}}\left(n^{-1/2}\right). \quad \square$$

A.4 Proofs of Theorems 2.1 and 2.2

We next provide the proofs of Theorems 2.1 and 2.2.

Proof of Theorem 2.1. By Propositions 3.1,

$$E[\tilde{G}_{p_1,p_2}(x,x')-G(x,x')]^2=E\left[\mathcal{U}_1(x,x')+\mathcal{U}_2(x,x')-G(x,x')\right]^2+o(1).$$

Let $\bar{\xi}_{\cdot kk'}=n^{-1}\sum_{i=1}^{n}\xi_{ik}\xi_{ik'}$, $1\le k,k'\le\kappa$. According to (10) and (11), one has

$$\mathcal{U}_1(x,x')+\mathcal{U}_2(x,x')-G(x,x')=\sum_{k\ne k'}^{\kappa}\bar{\xi}_{\cdot kk'}\phi_k(x)\,\phi_{k'}(x')+\sum_{k=1}^{\kappa}\left(\bar{\xi}_{\cdot kk}-1\right)\phi_k(x)\,\phi_k(x').$$

Since

$$\begin{aligned} & nE\left[\mathcal{U}_1(x,x')+\mathcal{U}_2(x,x')-G(x,x')\right]^2 \\ = & \sum_{k,k'=1}^{\kappa}\phi_k^2(x)\,\phi_{k'}^2(x')+\sum_{k,k'=1}^{\kappa}\phi_k(x')\,\phi_k(x)\,\phi_{k'}(x)\,\phi_{k'}(x')+\sum_{k=1}^{\kappa}\phi_k^2(x)\,\phi_k^2(x')\left(E\xi_{1k}^4-3\right) \\ = & G(x,x')^2+G(x,x)\,G(x',x')+\sum_{k=1}^{\kappa}\phi_k^2(x)\,\phi_k^2(x')\left(E\xi_{1k}^4-3\right)\equiv V(x,x'), \end{aligned}$$

and the desired result follows from Proposition 2.1. $\square$

Next define $\zeta(x,x')=n^{1/2}V^{-1/2}(x,x')\left\{\mathcal{U}_1(x,x')+\mathcal{U}_2(x,x')-G(x,x')\right\}$.

LEMMA A.6. *Under Assumptions (A2)-(A4),* $\sup_{(x,x')\in[0,1]^2}|\zeta_Z(x,x')-\zeta(x,x')| = o_{a.s.}(1)$, *where* $\zeta_Z(x,x')$ *is given in (4).*

PROOF. According to Assumption (A4) and multivariate Central Limit Theorem

$$\sqrt{n}\left\{\bar{\xi}_{\cdot kk'}, \left(\bar{\xi}^2_{\cdot k}-1\right)\left(E\xi^4_{1k}-1\right)^{-1/2}\right\}_{1\le k<k'\le\kappa} \to_d N\left(\mathbf{0}_{\kappa(\kappa+1)/2}, \mathbf{I}_{\kappa(\kappa+1)/2}\right).$$

Applying Skorohod's Theorem, there exist i.i.d. variables $Z_{kk'} = Z_{k'k} \sim N(0,1), Z_k \sim N(0,1), 1\le k<k'\le\kappa$ such that as $n\to\infty$,

$$\max_{1\le k<k'\le\kappa}\left\{\left|\sqrt{n}\bar{\xi}_{\cdot kk'}-Z_{kk'}\right|, \left|\sqrt{n}\left(\bar{\xi}^2_{\cdot k}-1\right)-Z_k\left(E\xi^4_{1k}-1\right)^{1/2}\right|\right\} = o_{\text{a.s.}}(1). \tag{A.2}$$

The desired result follows from (A.2). □

PROOF OF THEOREM 2.2. According to Lemma A.6, Proposition 3.1 and Theorem 2.1, as $n\to\infty$,

$$\begin{aligned} & P\left\{\sup_{(x,x')\in[0,1]^2} n^{1/2}\left|\tilde{\mathcal{U}}_{1p_2}(x,x')+\tilde{\mathcal{U}}_{2p_2}(x,x')-G(x,x')\right|V(x,x')^{-1/2}\le Q_{1-\alpha}\right\} \\ = \ & P\left\{\sup_{(x,x')\in[0,1]^2} n^{1/2}\left|\mathcal{U}_1(x,x')+\mathcal{U}_2(x,x')-G(x,x')\right|V(x,x')^{-1/2}\le Q_{1-\alpha}\right\} \\ = \ & P\left\{\sup_{(x,x')\in[0,1]^2}|\zeta_Z(x,x')|\le Q_{1-\alpha}\right\}, \forall\alpha\in(0,1). \end{aligned}$$

The desired result follows from Proposition 2.1 and equation (12). □

References

[1] CAI, T. AND HALL, P. (2005). Prediction in functional linear regression. *Annals of Statistics*, **34**, 2159-2179.

[2] CLAESKENS, G. AND VAN KEILEGOM, I. (2003). Bootstrap confidence bands for regression curves and their derivatives. *Annals of Statistics* **31**, 1852–1884.

[3] DE BOOR, C. (2001). *A Practical Guide to Splines*. Springer-Verlag, New York.

[4] DEVORE,R. AND LORENTZ, G.(1993) *Constructive approximation : polynomials and splines approximation* Springer-Verlag, Berlin; New York.

[5] FAN, J., HUANG, T. AND LI, R. (2007). Analysis of longitudinal data with semiparametric estimation of covariance function. *Journal of the American Statistical Association* **102**, 632–642.

[6] Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice.* Springer Series in Statistics, Springer: Berlin.

[7] Hall, P., Müller, H. G. and Wang, J. L. (2006). Properties of principal component methods for functional and longitudinal data analysis. *Annals of Statistics* **34**, 1493–1517.

[8] Härdle, W. and Marron, J. S. (1991). Bootstrap simultaneous error bars for nonparametric regression. *Annals of Statistics* **19**, 778–796.

[9] Huang, J. (2003). Local asymptotics for polynomial spline regression. *Annals of Statistics* **31**, 1600–1635.

[10] Huang, J. and Yang, L. (2004). Identification of nonlinear additive autoregressive models. *Journal of the Royal Statistical Society Series B* **66**, 463–477.

[11] James, G. M., Hastie, T. and Sugar, C. (2000). Principal Component Models for Sparse Functional Data. *Biometrika* **87**, 587–602.

[12] James, G. M. and Silverman, B. W. (2005). Functional adaptive model estimation. *Journal of the American Statistical Association* **100**, 565–576.

[13] Li, Y. and Hsing, T. (2007). On rates of convergence in functional linear regression. *Journal of Multivariate Analysis* **98**, 1782–1804.

[14] Li, Y. and Hsing, T. (2010). Uniform convergence rates for nonparametric regression and principal component analysis in functional/longitudinal data. *Annals of Statistics* **38**, 3321-3351.

[15] Li, Y. and Hsing, T. (2010). Deciding the dimension of effective dimension reduction space for functional and high-dimensional data. *Annals of Statistics* **38**, 3028-3062.

[16] Li, Y., Wang, N. and Carroll, R. J. (2010). Generalized functional linear models with semiparametric single-index interactions, *Journal of the American Statistical Association*, **105**, 621-633.

[17] Morris, J. S. and Carroll, R. J. (2006). Wavelet-based functional mixed models. *Journal of the Royal Statistical Society, Series B* **68**, 179–199.

[18] Müller, H. G. (2009). Functional modeling of longitudinal data. *In: Longitudinal Data Analysis (Handbooks of Modern Statistical Methods)*, Ed. Fitzmaurice, G., Davidian, M., Verbeke, G., Molenberghs, G., Wiley, New York, 223–252.

[19] Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis.* Second Edition. Springer Series in Statistics. Springer: New York.

[20] Wang, N., Carroll, R. J. and Lin, X. (2005). Efficient semiparametric marginal estimation for longitudinal/clustered data. *Journal of the American Statistical Association* **100**, 147–157.

[21] Wang, L. and Yang, L. (2009) Spline estimation of single index model. *Statistica Sinica* **19 (2)**, 765-783.

[22] Yao, F. and Lee, T. C. M. (2006). Penalized spline models for functional principal component analysis. *Journal of the Royal Statistical Society, Series B* **68**, 3–25.

[23] Yao, F., Müller, H. G. and Wang, J. L. (2005a). Functional linear regression analysis for longitudinal data. *Annals of Statistics* **33**, 2873–2903.

[24] Yao, F., Müller, H. G. and Wang, J. L. (2005b). Functional data analysis for sparse longitudinal data. *Journal of the American Statistical Association* **100**, 577–590.

[25] Zhao, X., Marron, J. S. and Wells, M. T. (2004) The functional data analysis view of longitudinal data. *Statistica Sinica* **14**, 789–808.

[26] Zhao, Z. and Wu, W. (2008). Confidence bands in nonparametric time series regression. *Annals of Statistics* **36**, 1854–1878.

[27] Zhou, L., Huang, J. and Carroll, R. J. (2008). Joint modelling of paired sparse functional data using principal components. *Biometrika* **95**, 601–619.

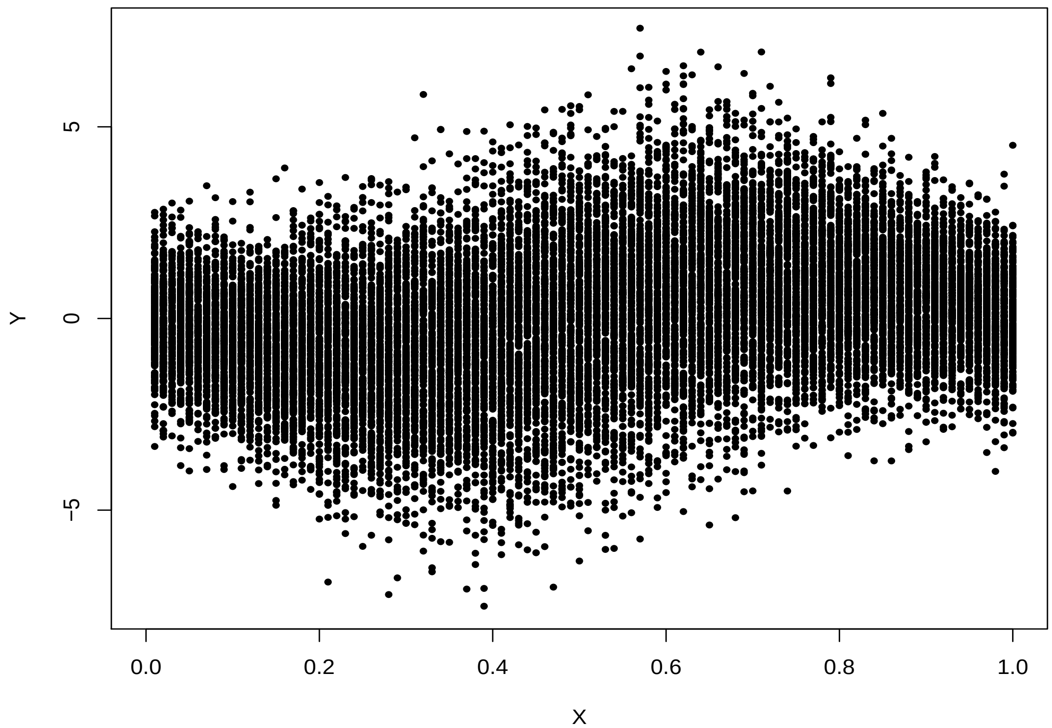
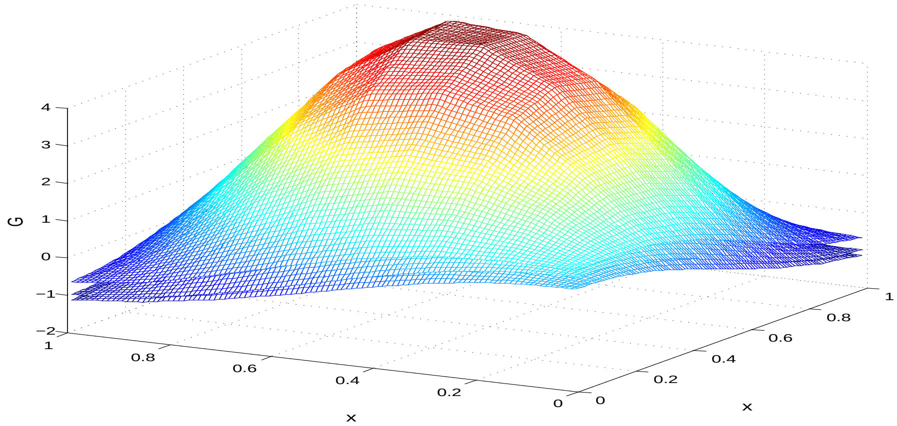


Figure 1: Plots of the simulated data: n=200, N=100 and σ=0.5

(a) 

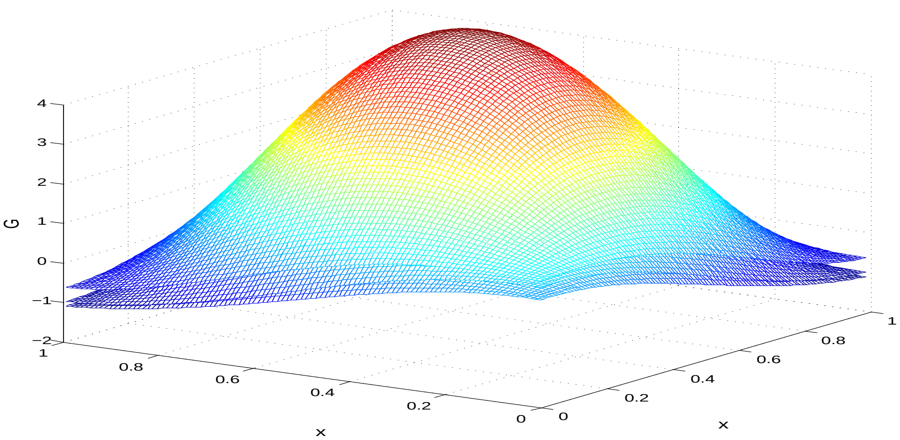
(b) 

Figure 2: Plots of the true covariance functions (middle surfaces) of the simulated data and their 95% confidence envelopes (5) (upper and lower surfaces): n=200, N=100, σ=0.5, (a) $p_1 = p_2 = 2$; (b) $p_1 = p_2 = 4$