TEST
DOI 10.1007/s11749-014-0392-4

ORIGINAL PAPER

A simultaneous confidence corridor for varying coefficient regression with sparse functional data

Lijie Gu · Li Wang · Wolfgang K. Härdle · Lijian Yang

Received: 19 January 2014 / Accepted: 27 June 2014

Abstract We consider a varying coefficient regression model for sparse functional data, with time varying response variable depending linearly on some time-independent covariates with coefficients as functions of time-dependent covariates. Based on spline smoothing, we propose data-driven simultaneous confidence corridors for the coefficient functions with asymptotically correct confidence level. Such confidence corridors are useful benchmarks for statistical inference on the global shapes of coefficient functions under any hypotheses. Simulation experiments corroborate with the theoretical results. An example in CD4/HIV study is used to illustrate how inference is made with computable p values on the effects of smoking, pre-infection CD4 cell percentage and age on the CD4 cell percentage of HIV infected patients under treatment.

Keywords B spline · Confidence corridor · Karhunen–Loève L^2 representation · Knots · Functional data · Varying coefficient

Mathematics Subject Classification (2000) 62G08 · 62G15 · 62G32

L. Gu · L. Yang (✉)
Center for Advanced Statistics and Econometrics Research,
Soochow University, Suzhou 215006, China
e-mail: yanglijian@suda.edu.cn

L. Wang
Department of Statistics, Iowa State University, Ames, IA 50011, USA

W. K. Härdle
Center for Applied Statistics and Economics (C.A.S.E.),
Humboldt-Universität zu Berlin, 610099 Berlin, Germany

W. K. Härdle
Lee Kong Chian School of Business, Singapore Management University, Singapore, Singapore

1 Introduction

Functional data, known also as "curve data", are commonly encountered in biomedical studies, epidemiology and social science, where information is collected over a time period for each subject. Conceptually, such data can be viewed as a simple random sample from the abstract space of functions, see for instance, Ferraty and Vieu (2006), Manteiga and Vieu (2007). For functional data analysis (FDA) approach without nonparametric smoothing, see Gabrys et al. (2010), and the recent comprehensive review in Horváth and Kokoszka (2012). In this, paper we have taken from Ramsay and Silverman (2005) the more convenient view of functional data as discretely recorded observations of independent stochastic processes contaminated with measurement errors.

In many longitudinal studies, repeated measurements are often collected at finite number of time points. If the time points of observation for every subjects are dense and regular, the data are termed dense functional data, see Cao et al. (2012a,b), and Zhu et al. (2012) for theoretical development and real examples of dense functional data. If, on the other hand, data collection is made at few and irregular time points for each subject, the data are frequently referred to as sparse longitudinal or sparse functional data, see James et al. (2000), James and Sugar (2003), Yao et al. (2005a), Hall et al. (2006), Zhou et al. (2008), Ma et al. (2012) for works on sparse functional data. It should be emphasized especially that by "sparse" we mean that the covariate is observed sparsely over a compact interval, not having anything to do with sparsity used in variable selection context such as the popular LASSO method. A crucial condition for sparse FDA is that the time points from all subjects are dense in the data range despite sparsity for any individual subject, see Assumption (A3) in Appendix A that the design density $f(t)$ has a positive lower bound c_f, which implies that the sampling frequency is almost uniform for the time covariate.

In longitudinal study, often, interest lies in studying the association between the covariates and the response variable. In recent years, there has been an increasing interest in nonparametric analysis of longitudinal data to enhance flexibility, see e.g., Yao and Li (2013). The varying coefficient model (VCM) proposed by Hastie and Tibshirani (1993) strikes a delicate balance between the simplicity of linear regression and the flexibility of multivariate nonparametric regression and has been widely applied in various settings, for instance, the Cobb–Douglas model for GDP growth in Liu and Yang (2010), and the longitudinal model for CD4 cell percentages in AIDS patients in Wu and Chiang (2000), Fan and Zhang (2000) and Wang et al. (2008). See Fan and Zhang (2008) for an extensive literature review of VCM.

To examine whether the association changes over time, Hoover et al. (1998) proposed the following VCM

$$Y(t) = \beta_0(t) + \mathbf{X}(t)^{\mathsf{T}}\boldsymbol{\beta}(t) + \varepsilon(t), \quad t \in \mathcal{T}, \tag{1}$$

where $\mathbf{X}(t) = (X_1(t), \ldots, X_d(t))^{\mathsf{T}}$ are covariates at time t, $\varepsilon(t)$ is a mean zero process, and $\boldsymbol{\beta}(t) = (\beta_1(t), \ldots, \beta_d(t))^{\mathsf{T}}$ are functions of t. Model (1) is a special case of functional linear models, see Ramsay and Silverman (2005) and Wu et al. (2010).

The coefficient functions $\beta_l(t)$s in model (1) can be estimated by, for example, kernel method in Hoover et al. (1998), basis function approximation method in Huang et al. (2002), polynomial spline method in Huang et al. (2004) and smoothing spline method in Brumback and Rice (1998). Fan and Zhang (2000) proposed a two-step method to overcome the computational burden of the smoothing spline method.

For some longitudinal studies, the covariates are independent of time, and their observations are cross-sectional. Take for instance the longitudinal CD4 cell percentage data among HIV seroconverters. This dataset contains 1,817 observations of CD4 cell percentages on 283 homosexual men infected with the HIV virus. Three of the covariates are observed at the time of HIV infection and hence by nature independent of the measurement time and frequency: X_{i1}, the ith patient's smoking status; X_{i2}, the ith patient's centered pre-infection CD4 percentage; and X_{i3} the ith patient's centered age at the time of HIV infection. A fourth predictor, however, is time dependent: T_{ij}, the time (in years) of the jth measurement of CD4 cell on the ith patient after HIV infection; while the response Y_{ij} is also time dependent: the jth measurement of the ith patient's CD4 cell percentage at time T_{ij}. Wu and Chiang (2000), Fan and Zhang (2000) and Wang et al. (2008) all contain detailed descriptions and analysis of this dataset.

A feasible VCM for multivariate functional data such as the above takes the form

$$Y_{ij} = \sum_{l=1}^{d} \eta_{il}\left(T_{ij}\right) X_{il} + \sigma\left(T_{ij}\right) \varepsilon_{ij}, \quad 1 \leq i \leq n, \ 1 \leq j \leq N_i, \tag{2}$$

where the measurement errors $\left(\varepsilon_{ij}\right)_{i=1,j=1}^{n,N_i}$ satisfy $\mathsf{E}\left(\varepsilon_{ij}\right) = 0$, $\mathsf{E}(\varepsilon_{ij}^2) = 1$, and $\{\eta_{il}(t), t \in \mathcal{T}\}$ are i.i.d copies of a L^2 process $\{\eta_l(t), t \in \mathcal{T}\}$, i.e., $\mathsf{E}\int_{\mathcal{T}} \eta_l^2(t)\mathrm{d}t < +\infty$, $l = 1, \ldots, d$. The common mean function of processes $\{\eta_l(t), t \in \mathcal{T}\}$ is denoted as $m_l(t) = \mathsf{E}\{\eta_l(t)\}$, $l = 1, \ldots, d$. The actual data set consists of $\left\{\mathbf{X}_i, T_{ij}, Y_{ij}\right\}$, $1 \leq i \leq n$, $1 \leq j \leq N_i$, in which the ith subject is observed N_i times, the time-independent covariates for the ith subject are $\mathbf{X}_i = (X_{il})_{l=1}^{d}$, $1 \leq i \leq n$, and the random measurement time $T_{ij} \in \mathcal{T} = [a, b]$. The aforementioned data example is called sparse functional as the number of measurements N_i for the ith subject is relatively low. (In the above CD4 example actually at most 14). In contrast, for a dense functional data $\lim_{n\to\infty} \min_{1\leq i\leq n} N_i = \infty$.

For the CD4 cell percentage data, we introduce a fourth time-independent covariate, the baseline $X_{i0} \equiv 1$, and denote by $m_l(t)$, $l = 0, 1, 2, 3$, the coefficient functions for baseline CD4 percentage, smoking status, centered pre-infection CD4 percentage and centered age, respectively. Figures 2, 3, 4, 5 contain spline estimates of the $m_l(t)$, $0 \leq l \leq 3$, and simultaneous confidence corridors (SCC) at various confidence levels.

In previous works the theoretical focus has mainly been on consistency and asymptotic normality of the estimators of the coefficient functions of interest, and the construction of pointwise confidence intervals. However, as demonstrated in Fan and Zhang (2000), this is unsatisfactory as investigators are often interested in testing whether some coefficient functions are significantly nonzero or varying, for which a SCC is needed. Take for instance, Fig. 3, which shows both the 95 and 20.277 % SCC of $m_1(t)$ contain the zero line completely, thus with a very high p value of 0.79723 the null hypothesis of $m_1(t) \equiv 0, t \in \mathcal{T}$ is not rejected. More details are in Sect. 6.

Construction of computationally simple SCCs with exact coverage probability is known to be difficult even with independent cross-sectional data; see, Wang and Yang (2009) and related earlier work Härdle and Luckhaus (1984) on uniform consistency. Most earlier methods proposed in the literature restrict to asymptotic conservative SCCs. Wu et al. (1998) developed asymptotic SCCs for the unknown coefficients based on local polynomial methods, which are computationally intensive, as the kernel estimator requires solving an optimization problem at every point. Huang et al. (2004) proposed approximating each coefficient function by a polynomial spline and developed spline SCCs, which are simpler to construct, while Xue and Zhu (2007) proposed maximum empirical likelihood estimators and constructed SCCs for the coefficient functions. All these SCCs are Bonferroni-type variability bands according to Hall and Titterington (1988). The idea is to invoke pointwise confidence intervals on a very fine grid of $[a, b]$, then adjust the level of these confidence intervals by the Bonferroni method to obtain uniform confidence bands, and finally bridge the gaps between the grid points via smoothness conditions on the coefficient curve. However, to use these bands in practice, one must have a priori bounds on the magnitude of the bias on each subinterval as well as a choice for the number of grid points. Chiang et al. (2001) proposed a bootstrap procedure to construct confidence intervals. However, theoretical properties of their procedures have not yet been developed.

In this paper, we derive SCCs with exact coverage probability for the coefficient functions $m_l(t)$, $l = 1, \ldots, d$, in (3) via extreme value theory of Gaussian processes and approximating coefficient functions by piecewise-constant splines. The results represent the first attempt at developing exact SCCs for the coefficient functions in VCM for sparse functional data. Our simulation studies indicate the proposed SCCs are computationally efficient and have the right coverage probability for finite samples. Our work parallels Zhu et al. (2012) which established asymptotic theory of SCC in the case of VCM for dense functional data. It is important to mention as well that the linear covariates in Zhu et al. (2012) are time dependent, which does not complicate the problem as they work with dense data instead of the sparse data we concentrate on. Our work can also be viewed as an extension of the univariate longitudinal regression in Ma et al. (2012) to varying coefficient regression, the latter corresponds exactly to the special case of $d = 1$, $X_{i1} \equiv 1$. Theorem 1 of Ma et al. (2012) follows from Theorems 1 and 2 in this paper with some slight modifications.

We organize our paper as follows. Section 2 describes spline estimators, and establish their asymptotic properties for sparse longitudinal data. Section 3.1 proposes asymptotic pointwise confidence intervals and SCCs constructed from piecewise constant splines. Section 3.2 describes actual steps to implement the proposed SCCs. In Sect. 4, we provide further insights into the estimation error structure of spline estimators. Section 5 reports findings from a simulation study. A real data example appears in Sect. 6. Proofs of technical lemmas are in Appendix A.

2 Spline estimation and asymptotic properties

For a functional data $\left\{\mathbf{X}_i, T_{ij}, Y_{ij}\right\}$, $1 \leq i \leq n$, $1 \leq j \leq N_i$, denote the eigenvalues and eigenfunctions sequences of its covariance operator $G_l\left(s, t\right) = \mathrm{cov}\left\{\eta_l(s), \eta_l(t)\right\}$ as $\left\{\lambda_{k,l}\right\}_{k=1}^{\infty}$, $\left\{\psi_{k,l}(t)\right\}_{k=1}^{\infty}$, in which $\lambda_{1,l} \geq \lambda_{2,l} \geq \cdots \geq 0$, $\sum_{k=1}^{\infty} \lambda_{k,l} < \infty$, and $\left\{\psi_{k,l}\right\}_{k=1}^{\infty}$ form an orthonormal basis of $L^2\left(\mathcal{T}\right)$. It follows from spectral theory that $G_l\left(s, t\right) = \sum_{k=1}^{\infty} \lambda_{k,l}\psi_{k,l}(s)\psi_{k,l}\left(t\right)$. For any $l = 1, \ldots, d$, the ith trajectory $\left\{\eta_{il}(t), t \in \mathcal{T}\right\}$ allows the Karhunen–Loève L^2 representation (Yao et al. 2005b): $\eta_{il}(t) = m_l(t) + \sum_{k=1}^{\infty} \xi_{ik,l}\phi_{k,l}(t)$, where the random coefficients $\xi_{ik,l}$ are uncorrelated with mean 0 and variances 1, and the functions $\phi_{k,l} = \sqrt{\lambda_{k,l}}\psi_{k,l}$, thus $G_l(s, t) = \sum_{k=1}^{\infty} \phi_{k,l}(s)\phi_{k,l}\left(t\right)$, and the response measurements (2) can be represented as follows:

$$Y_{ij} = \sum_{l=1}^{d} m_l\left(T_{ij}\right)X_{il} + \sum_{l=1}^{d}\sum_{k=1}^{\infty} \xi_{ik,l}\phi_{k,l}\left(T_{ij}\right)X_{il} + \sigma\left(T_{ij}\right)\varepsilon_{ij}. \tag{3}$$

Without loss of generality, we take $\mathcal{T} = [a, b]$ to be $[0, 1]$. Following Xue and Yang (2006), we approximate each coefficient function by the spline smoothing method. To describe the spline functions, one can divide the finite interval $[0, 1]$ into $(N_{\mathrm{s}} + 1)$ equal subintervals $\chi_J = \left[\upsilon_J, \upsilon_{J+1}\right)$, $J = 0, \ldots, N_{\mathrm{s}} - 1$, $\chi_{N_{\mathrm{s}}} = \left[\upsilon_{N_{\mathrm{s}}}, 1\right]$. A sequence of equally spaced points $\left\{\upsilon_J\right\}_{J=1}^{N_{\mathrm{s}}}$, called interior knots, are given as $\upsilon_0 = 0 < \upsilon_1 < \cdots < \upsilon_{N_{\mathrm{s}}} < 1 = \upsilon_{N_{\mathrm{s}}+1}$. Let $\upsilon_J = Jh_{\mathrm{s}}$ for $0 \leq J \leq N_{\mathrm{s}}+1$, where $h_{\mathrm{s}} = 1/\left(N_{\mathrm{s}} + 1\right)$ is the distance between neighboring knots. We denote by $G^{(-1)} = G^{(-1)}\left[0, 1\right]$ the space of functions that are constant on each subinterval χ_J, and the B-spline basis of $G^{(-1)}$, as $\left\{b_J(t)\right\}_{J=0}^{N_{\mathrm{s}}}$, which are simply indicator functions of intervals χ_J, $b_J(t) = I_{\chi_J}(t)$, $J = 0, 1, \ldots, N_{\mathrm{s}}$. For any $t \in [0, 1]$, define its location index as $J(t) = J_n(t) = \min\left\{\left[t/h_{\mathrm{s}}\right], N_{\mathrm{s}}\right\}$ so that $t \in \chi_{J(t)}$.

Next we define the space of spline coefficient functions on $\mathcal{T} \times \mathbb{R}^d$ as

$$\mathcal{M} = \left\{g\left(t, \mathbf{x}\right) = \sum_{l=1}^{d} g_l(t)x_l : g_l(t) \in G^{(-1)}, t \in \mathcal{T}, \mathbf{x} = \left(x_1, \ldots, x_d\right)^{\mathsf{T}} \in \mathbb{R}^d\right\},$$

and propose estimating the multivariate function $\sum_{l=1}^{d} m_l(t)x_l$ by

$$\hat{m}\left(t, \mathbf{x}\right) = \sum_{l=1}^{d} \hat{m}_l(t)x_l = \underset{g \in \mathcal{M}}{\mathrm{argmin}} \sum_{i=1}^{n}\sum_{j=1}^{N_i}\left\{Y_{ij} - g\left(T_{ij}, \mathbf{X}_i\right)\right\}^2. \tag{4}$$

Let $\sigma_Y^2(t, \mathbf{x})$ be the conditional variance of $\mathbf{Y}$ given $T = t$ and $\mathbf{X} = \mathbf{x} = \left(x_1, \ldots, x_d\right)^{\mathsf{T}} \in \mathbb{R}^d$

$$\sigma_Y^2(t, \mathbf{x}) = \mathsf{Var}(Y \mid T = t, \mathbf{X} = \mathbf{x}) = \sum_{l=1}^{d} G_l\left(t, t\right)x_l^2 + \sigma^2(t).$$

Next for any $t \in [0, 1]$, let

$$\Gamma_n(t) = c_{J(t),n}^{-2}\{n\mathsf{E}(N_1)\}^{-1}\mathsf{E}\mathbf{X}\mathbf{X}^{\mathsf{T}}\left[\int_{\chi_{J(t)}} \sigma_Y^2(u, \mathbf{X}) f(u)\,\mathrm{d}u \right.$$
$$\left. + \frac{\mathsf{E}\{N_1(N_1-1)\}}{\mathsf{E}N_1}\sum_{l=1}^{d} X_l^2 \int_{\chi_{J(t)}\times\chi_{J(t)}} G_l(u, v) f(u) f(v)\,\mathrm{d}u\mathrm{d}v\right], \quad (5)$$

where

$$c_{J,n} = \mathsf{E}b_J^2(T) = \int_0^1 b_J^2(t) f(t)\mathrm{d}t, \quad J = 0, \ldots, N_{\mathrm{s}}. \quad (6)$$

Further denote

$$\Sigma_n(t) = \mathbf{H}^{-1}\Gamma_n(t)\mathbf{H}^{-1} = \left\{\sigma_{n,ll'}^2(t)\right\}_{l,l'=1}^{d}, \quad (7)$$

where $\sigma_{n,ll'}^2(t)$ are later shown to be the asymptotic covariances between $\hat{m}_l(t)$ and $\hat{m}_{l'}(t)$.

Theorem 1 *Under Assumptions* (A1)–(A6) *in Appendix A, for any* $t \in [0, 1]$*, as* $n \to \infty$*,*

$$\Sigma_n^{-1/2}(t)\{\hat{\mathbf{m}}(t) - \mathbf{m}(t)\} \xrightarrow{\mathcal{L}} N(\mathbf{0}, \mathbf{I}_{d\times d}),$$

where $\hat{\mathbf{m}}(t) = \left(\hat{m}_1(t), \ldots, \hat{m}_d(t)\right)^{\mathsf{T}}$ *is the estimate of* $\mathbf{m}(t) = (m_1(t), \ldots, m_d(t))^{\mathsf{T}}$. *Furthermore, for any* $l = 1, \ldots, d$ *and* $\alpha \in (0, 1)$,

$$\lim_{n\to\infty} P\left\{\sigma_{n,ll}^{-1}(t)\left|\hat{m}_l(t) - m_l(t)\right| \le Z_{1-\alpha/2}\right\} = 1 - \alpha.$$

Remark 1 Note that $\Sigma_n(t) = \left\{\sigma_{n,ll'}^2(t)\right\}_{l,l'=1}^{d}$ in (7) is complicated to compute in practice. The next proposition suggests that, for any $t \in [0, 1]$, $\Gamma_n(t)$ in (5) can be simplified by

$$\tilde{\Gamma}_n(t) \equiv \mathsf{E}\left[\mathbf{X}\mathbf{X}^{\mathsf{T}}\frac{\sigma_Y^2(t, \mathbf{X})}{f(t)h_{\mathrm{s}}n\mathsf{E}(N_1)}\left\{1 + \frac{\mathsf{E}N_1(N_1-1)}{\mathsf{E}N_1}\frac{\sum_{l=1}^{d} X_l^2 G_l(t, t) f(t)h_{\mathrm{s}}}{\sigma_Y^2(t, \mathbf{X})}\right\}\right]. \quad (8)$$

Denote the supremum of a function ϕ on $[a, b]$ by $\|\phi\|_\infty = \sup_{t\in[a,b]}|\phi(t)|$. For any matrix $\mathbf{A} = \left(a_{ij}\right)$, define $\|\mathbf{A}\|_\infty = \max\left|a_{ij}\right|$, where the maximum is taken over all the elements of $\mathbf{A}$, while for a matrix function $\mathbf{A}(t) = \left(a_{ij}(t)\right)$, $\|\mathbf{A}\|_\infty = \sup_{t\in[a,b]}\|\mathbf{A}(t)\|_\infty$.

Proposition 1 *Under Assumptions* (A2)–(A6) *in Appendix A, there exists a constant* $c > 0$ *such that as* $n \to \infty$, $\|\Gamma_n(t) - \tilde{\Gamma}_n(t)\|_\infty = \mathcal{O}\left(n^{-1}h_{\mathrm{s}}^{r-1}\right) = \mathcal{O}\left(n^{-c}\right)$.

To derive the maximal deviation distribution of estimators $\hat{m}_l(t), l = 1, \ldots, d$, let

$$Q_{N_s+1}(\alpha) = b_{N_s+1} - a_{N_s+1}^{-1} \log\left\{-\frac{1}{2}\log(1-\alpha)\right\}, \quad \alpha \in (0,1) \tag{9}$$

$$a_{N_s+1} = \{2\log(N_s+1)\}^{1/2}, \quad b_{N_s+1} = a_{N_s+1} - \frac{\log\left(2\pi a_{N_s+1}^2\right)}{2a_{N_s+1}}. \tag{10}$$

Theorem 2 *Under Assumptions* (A1)–(A6) *in Appendix A, for $l = 1, \ldots, d$ and any $\alpha \in (0,1)$,*

$$\lim_{n\to\infty} P\left\{\sup_{t\in[0,1]} \sigma_{n,ll}^{-1}(t)\left|\hat{m}_l(t) - m_l(t)\right| \le Q_{N_s+1}(\alpha)\right\} = 1 - \alpha,$$

where $\sigma_{n,ll}(t)$ and $Q_{N_s+1}(\alpha)$ are given in (7) *and* (9)*, respectively.*

One reviewer has pointed out that the use of constant instead of higher order spline is not optimal, which we completely agree. Further research involving sophisticated nonstationary Gaussian process extreme value theory is needed to extend our present work to splines of any order, such as the popular cubic spline. To be precise, analog of Proposition 4 for higher order spline concerns the maximum of a standardized continuous Gaussian process over interval [0, 1], whereas for constant spline, the Gaussian process breaks down to N_s+1 weakly correlated standard Gaussian variables.

3 Asymptotic confidence regions

In this section, we construct the confidence regions for functions $m_l(t), l = 1, \ldots, d$.

3.1 Asymptotic confidence intervals and SCCs

Theorems 1 and 2 allow one to construct pointwise confidence intervals and SCCs for components $\hat{m}_l(t), l = 1, \ldots, d$. The next corollary provides the theoretical underpinning upon which SCCs can be actually implemented, see Sect. 3.2.

Corollary 1 *Under Assumptions* (A1)–(A6) *in Appendix A, for any $l = 1, \ldots, d$ and $\alpha \in (0,1)$, as $n \to \infty$,*

- (i) *an asymptotic* $100(1-\alpha)\%$ *pointwise confidence interval for $m_l(t)$, $t \in [0,1]$, is $\hat{m}_l(t) \pm \sigma_{n,ll}(t) Z_{1-\alpha/2}$, with $\sigma_{n,ll}(t)$ given in* (7)*, while $Z_{1-\alpha/2}$ is the* $100(1-\alpha/2)$*th percentile of the standard normal distribution.*
- (ii) *an asymptotic* $100(1-\alpha)\%$ *SCC for $m_l(t)$, with $Q_{N_s+1}(\alpha)$ given in* (9)*, is $\hat{m}_l(t) \pm \sigma_{n,ll}(t) Q_{N_s+1}(\alpha)$, $t \in [0,1]$.*

One reviewer has raised the interesting question whether our SCC would significantly improve by some form of bootstrapping. The answer is negative for now due to the lack of convincing procedures that simultaneously resample from the unknown distributions of both the unobserved error ε_{ij}s and the unobserved functional principal components $\xi_{ik,l}$s. On the other hand, further investigation in FDA will lead to theoretically sound resampling methods analogous to the smoothed bootstrap for nonparametric regression in Claeskens and Van Keilegom (2003).

3.2 Implementation

In the following, we describe procedures to construct the SCCs and the pointwise intervals given in Corollary 1. For any data set $\left(T_{ij}, Y_{ij}, X_{il}\right)_{i=1,j=1,l=1}^{n,N_i,d}$ from model (3), the spline estimators $\hat{m}_l(t)$, $l = 1, \ldots, d$, are obtained by (4), and the number of interior knots is taken to be $N_{\mathrm{s}} = [cN_{\mathrm{T}}^{1/3}(\log(n))]$, in which $N_{\mathrm{T}} = \sum_{i=1}^{n} N_i$ is the total sample size, $[a]$ denotes the integer part of a, and c is a positive constant.

To construct the SCCs, one needs to evaluate the functions $\sigma_{n,ll}^2(t)$, $l = 1, \ldots, d$, which are the diagonal elements of matrix $\Sigma_n(t)$ in (7). Based on Proposition 1, one can estimate each unknowns $f(t)$, $\sigma_Y^2(t, \mathbf{x})$, $G_l\,(t, t)$ and matrix $\mathbf{H}$ and then plug these estimators into the formula of the SCCs; see Wang and Yang (2009).

The number of interior knots for pilot estimation of $f(t)$, $\sigma_Y^2(t, \mathbf{x})$, and $G_l\,(t, t)$ is taken to be $N_{\mathrm{s}}^* = \left[0.5n^{1/3}\right]$, and $h_{\mathrm{s}}^* = 1/\left(1 + N_{\mathrm{s}}^*\right)$. The histogram estimator of the density function $f(t)$ is $\hat{f}(t) = N_{\mathrm{T}}^{-1} h_{\mathrm{s}}^{*-1} \sum_{i=1}^{n} \sum_{j=1}^{N_i} b_{J(t)}(T_{ij})$.

To estimate the covariance matrix $\Gamma_n\,(t)$ in (5), define the raw covariance term $R_{ij} = \left(Y_{ij} - \sum_{l=1}^{d} \hat{m}(T_{ij}) X_{il}\right)^2$, $1 \le j \le N_i$, $1 \le i \le n$, the estimator of $\sigma_Y^2(t, \mathbf{x})$ is

$$\hat{\sigma}_Y^2(t, \mathbf{x}) = \sum_{l=1}^{d} \sum_{J=0}^{N_{\mathrm{s}}^*} \hat{\rho}_{J,l} b_J(t) x_l^2 + \sum_{J=0}^{N_{\mathrm{s}}^*} \hat{\mu}_J b_J(t) = \sum_{l=1}^{d} \hat{G}_l\,(t, t)\, x_l^2 + \hat{\sigma}^2(t),$$

where $\{\hat{\rho}_{0,1}, \ldots, \hat{\rho}_{N_{\mathrm{s}}^*,d}, \hat{\mu}_0, \ldots, \hat{\mu}_{N_{\mathrm{s}}^*}\}^{\mathsf{T}}$ are solutions of the following least squares problem:

$$\begin{aligned}
&\left(\hat{\rho}_{0,1}, \ldots, \hat{\rho}_{N_{\mathrm{s}}^*,d}, \hat{\mu}_0, \ldots, \hat{\mu}_{N_{\mathrm{s}}^*}\right)^{\mathsf{T}} \\
&\quad = \underset{\left(\rho_{0,1}, \ldots, \mu_{N_{\mathrm{s}}^*}\right)^{\mathsf{T}} \in \mathbb{R}^{(N_{\mathrm{s}}^*+1)(d+1)}}{\operatorname{argmin}} \sum_{i=1}^{n} \sum_{j=1}^{N_i} \left\{ R_{ij} - \sum_{l=1}^{d} \sum_{J=0}^{N_{\mathrm{s}}^*} \rho_{J,l} b_J(T_{ij}) X_{il}^2 - \sum_{J=0}^{N_{\mathrm{s}}^*} \mu_J b_J(T_{ij}) \right\}^2.
\end{aligned}$$

The matrix $\Gamma_n\,(t)$ is estimated by substituting $f(t)$, $G_l\,(t, t)$ and $\sigma_Y^2(t, \mathbf{x})$ with $\hat{f}(t)$, $\hat{G}_l\,(t, t)$ and $\hat{\sigma}_Y^2(t, \mathbf{x})$. Define

$$\begin{aligned}
\hat{\Gamma}_n\,(t) \equiv \Bigg[& n^{-1} \sum_{i=1}^{n} X_{il} X_{il'} \hat{\sigma}_Y^2(t, \mathbf{X}_i) \left\{ \hat{f}(t) h_{\mathrm{s}} N_{\mathrm{T}} \right\}^{-1} \\
& \times \left\{ 1 + \left(\frac{\sum_{i=1}^{n} N_i^2}{N_{\mathrm{T}}} - 1 \right) \frac{\sum_{l=1}^{d} \hat{G}_l\,(t, t)\, X_{il}^2}{\hat{\sigma}_Y^2(t, \mathbf{X}_i)} \hat{f}(t) h_{\mathrm{s}} \right\} \Bigg]_{l,l'=1}^{d}.
\end{aligned}$$

The following proposition provides the consistent rate of $\hat{\Gamma}_n(t)$ to $\Gamma_n(t)$.

Proposition 2 *Under Assumptions* (A1)–(A6) *in Appendix A, there exists a constant* $c > 0$ *such that as* $n \to \infty$, $\left\| \hat{\Gamma}_n(t) - \Gamma_n(t) \right\|_\infty = \mathcal{O}_p\left(n^{-c}\right)$.

Proposition 2 implies that $\Gamma_n(t)$ can be replaced by $\hat{\Gamma}_n(t)$ with a negligible error. Define a $d \times d$ matrix $\hat{\mathbf{H}} = \left\{n^{-1} \sum_{i=1}^n X_{il} X_{il'}\right\}_{l,l'=1}^d$, then $\Sigma_n(t)$ can be estimated well by $\hat{\Sigma}_n(t) = \left\{\hat{\sigma}^2_{n,ll'}(t)\right\}_{l,l'=1}^d = \hat{\mathbf{H}}^{-1} \hat{\Gamma}_n(t) \hat{\mathbf{H}}^{-1}$. Therefore, as $n \to \infty$, $l = 1, \ldots, d$, the SCCs

$$\hat{m}_l(t) \pm \hat{\sigma}_{n,ll}(t) Q_{N_\mathrm{s}+1}(\alpha), \tag{11}$$

with $Q_{N_\mathrm{s}+1}(\alpha)$ given in (9), and the pointwise intervals $\hat{m}_l(t) \pm \hat{\sigma}_{n,ll}(t) Z_{1-\alpha/2}$ have asymptotic confidence level $1 - \alpha$.

4 Decomposition

In this section, we describe the representation of the spline estimators $\hat{m}_l(t)$, $l = 1, \ldots, d$, in (4), then break the estimation error $\hat{m}_l(t) - m_l(t)$ into three terms by the decomposition of Y_{ij} in model (3). Although such representation is not needed for applying the procedure described in Sect. 3.2 to analyze data, it provides insights into the proof of the main theoretical results in Sect. 2.

We consider the following rescaled B-spline basis $\{B_J(t)\}_{J=0}^{N_\mathrm{s}}$ for $G^{(-1)}$:

$$B_J(t) \equiv b_J(t) \left(c_{J,n}\right)^{-1/2}, \quad J = 0, \ldots, N_\mathrm{s}. \tag{12}$$

It is easily verified that $\mathsf{E}\{B_J(T)\}^2 = 1$ for $J = 0, 1, \ldots, N_\mathrm{s}$, and $B_J(t) B_{J'}(t) \equiv 0$ for $J \neq J'$. By simple linear algebra, the spline estimator $\hat{m}_l(t)$ defined in (4) equals

$$\hat{m}_l(t) = \sum_{J=0}^{N_\mathrm{s}} \hat{\gamma}_{J,l} B_J(t) = c_{J(t),n}^{-1/2} \hat{\gamma}_{J(t),l}, \quad l = 1, \ldots, d, \tag{13}$$

where the coefficients $\hat{\gamma} = \left(\hat{\gamma}_0^\mathsf{T}, \ldots, \hat{\gamma}_{N_\mathrm{s}}^\mathsf{T}\right)^\mathsf{T}$ with $\hat{\gamma}_J = \left(\hat{\gamma}_{J,1}, \ldots, \hat{\gamma}_{J,d}\right)^\mathsf{T}$ being the solution of the following least squares problem

$$\hat{\gamma} = \underset{\gamma = \left(\gamma_{0,1}, \ldots, \gamma_{N_\mathrm{s},d}\right)^\mathsf{T} \in \mathbb{R}^{d(N_\mathrm{s}+1)}}{\operatorname{argmin}} \sum_{i=1}^n \sum_{j=1}^{N_i} \left\{ Y_{ij} - \sum_{l=1}^d \sum_{J=0}^{N_\mathrm{s}} \gamma_{J,l} B_J\left(T_{ij}\right) X_{il} \right\}^2. \tag{14}$$

In the following, let $\mathbf{Y} = \left(Y_{11}, \ldots, Y_{1N_1}, \ldots, Y_{n1}, \ldots, Y_{nN_n}\right)^\mathsf{T}$ be the collection of all the Y_{ij}s. Let $\mathbf{B}(t) = \left(B_0(t), \ldots, B_{N_\mathrm{s}}(t)\right)^\mathsf{T}$ and $\mathbf{X}_i = (X_{i1}, \ldots, X_{id})^\mathsf{T}$ be two vectors of dimension $(N_s + 1)$ and d, respectively. Denote

$$\mathbf{D} = \left(\mathbf{B}(T_{11}) \otimes \mathbf{X}_1, \ldots, \mathbf{B}(T_{1N_1}) \otimes \mathbf{X}_1, \ldots, \mathbf{B}(T_{n1}) \otimes \mathbf{X}_n, \ldots, \mathbf{B}(T_{nN_n}) \otimes \mathbf{X}_n\right)^\top, \tag{15}$$

a $N_T \times ((N_s + 1)\, d)$ matrix, where " $\otimes$" denotes the Kronecker product. Solving the least squares problem in (14), we obtain

$$\hat{\gamma} = \left(\mathbf{D}^\top\mathbf{D}\right)^{-1}\left(\mathbf{D}^\top\mathbf{Y}\right). \tag{16}$$

Denote $\mathbf{x} = (x_1, \ldots, x_d)^\top$, thus Eq. (4) can be rewritten as

$$\sum_{l=1}^{d} \hat{m}_l(t)x_l = (\mathbf{B}(t) \otimes \mathbf{x})^\top \left(\mathbf{D}^\top\mathbf{D}\right)^{-1}\left(\mathbf{D}^\top\mathbf{Y}\right). \tag{17}$$

According to (15), one has $\mathbf{D}^\top\mathbf{D} = \sum_{i=1}^{n}\sum_{j=1}^{N_i}\left\{\mathbf{B}(T_{ij})\mathbf{B}(T_{ij})^\top \otimes \mathbf{X}_i\mathbf{X}_i^\top\right\}$, in which matrix $\mathbf{B}(T_{ij})\mathbf{B}(T_{ij})^\top = \mathrm{diag}\left\{B_0^2(T_{ij}), \ldots, B_{N_s}^2(T_{ij})\right\}$. So matrix $\mathbf{D}^\top\mathbf{D}$ should be a block diagonal matrix, and $N_T^{-1}\mathbf{D}^\top\mathbf{D} = \mathrm{diag}\hat{\mathbf{V}}_0, \ldots, \hat{\mathbf{V}}_{N_s}\}$, where

$$\hat{\mathbf{V}}_J = \left\{N_T^{-1}\sum_{i=1}^{n}\sum_{j=1}^{N_i} B_J^2(T_{ij})X_{il}X_{il'}\right\}_{l,l'=1}^{d}. \tag{18}$$

On the other hand, we have $\mathbf{D}^\top\mathbf{Y} = \sum_{i=1}^{n}\sum_{j=1}^{N_i}\left\{\mathbf{B}(T_{ij}) \otimes \mathbf{X}_i\right\}Y_{ij}$. Thus, $\hat{\gamma} = \left(\hat{\gamma}_0^\top, \ldots, \hat{\gamma}_{N_s}^\top\right)^\top$ can be easily calculated using

$$\hat{\gamma}_J = \hat{\mathbf{V}}_J^{-1}\left\{N_T^{-1}\sum_{i=1}^{n}\sum_{j=1}^{N_i} B_J(T_{ij})X_{il}Y_{ij}\right\}_{l=1}^{d}, \quad J = 0, \ldots, N_s. \tag{19}$$

Then the functions $\mathbf{m}(t) = (m_1(t), \ldots, m_d(t))^\top$ can be simply estimated by

$$\hat{\mathbf{m}}(t) = \left(\hat{m}_1(t), \ldots, \hat{m}_d(t)\right)^\top = c_{J(t),n}^{-1/2}\left(\hat{\gamma}_{J(t),1}, \ldots, \hat{\gamma}_{J(t),d}\right)^\top = c_{J(t),n}^{-1/2}\hat{\gamma}_{J(t)}. \tag{20}$$

Projecting the relationship in model (3) onto the space of spline coefficient functions on $\mathcal{T} \times \mathbb{R}^d$ as $\mathcal{M}$, we obtain the following important decomposition:

$$\sum_{l=1}^{d} \hat{m}_l(t)x_l = \sum_{l=1}^{d} \tilde{m}_l(t)x_l + \sum_{l=1}^{d} \tilde{\xi}_l(t)x_l + \sum_{l=1}^{d} \tilde{\varepsilon}_l(t)x_l, \tag{21}$$

where for any $l = 1, \ldots, d$,

$$\tilde{m}_l(t) = \sum_{J=0}^{N_s} \tilde{\gamma}_{J,l} B_J(t) = c_{J(t),n}^{-1/2} \tilde{\gamma}_{J(t),l}, \tag{22}$$

$$\tilde{\xi}_l(t) = \sum_{J=0}^{N_s} \tilde{\alpha}_{J,l} B_J(t) = c_{J(t),n}^{-1/2} \tilde{\alpha}_{J(t),l}, \quad \tilde{\varepsilon}_l(t) = \sum_{J=0}^{N_s} \tilde{\theta}_{J,l} B_J(t) = c_{J(t),n}^{-1/2} \tilde{\theta}_{J(t),l}, \tag{23}$$

where $(\tilde{\gamma}_{J,l}, J = 0, \ldots, N_s, l = 1, \ldots, d)^{\mathsf{T}}$, $\left(\tilde{\alpha}_{J,l}, J = 0, \ldots, N_s, l = 1, \ldots, d\right)^{\mathsf{T}}$, and $(\tilde{\theta}_{J,l}, J = 0, \ldots, N_s, l = 1, \ldots, d)^{\mathsf{T}}$ are solutions to (14) with Y_{ij} replaced by $\sum_{l=1}^{d} m_l \left(T_{ij}\right) X_{il}$, $\sum_{l=1}^{d} \sum_{k=1}^{\infty} \xi_{ik,l} \phi_{k,l} \left(T_{ij}\right) X_{il}$, and $\sigma \left(T_{ij}\right) \varepsilon_{ij}$, respectively.

Furthermore, under Assumption (A5) we can decompose $\hat{m}_l(t)$ as

$$\hat{m}_l(t) = \tilde{m}_l(t) + \tilde{\xi}_l(t) + \tilde{\varepsilon}_l(t), \qquad l = 1, \ldots, d. \tag{24}$$

The next two propositions concern the functions $\tilde{m}_l(t)$, $\tilde{\xi}_l(t)$, $\tilde{\varepsilon}_l(t)$, $l = 1, \ldots, d$, given in (22) and (23). Proposition 3 gives the uniform convergence rate of $\tilde{m}_l(t)$ to $m_l(t)$. Proposition 4 provides the asymptotic distribution for the maximum of the normalized error terms.

Proposition 3 *Under Assumptions* (A1), (A2) *and* (A4)–(A6) *in Appendix A, the functions $\tilde{m}_l(t)$, $l = 1, \ldots, d$ satisfy* $\sup_{t \in [0,1]} \sup_{1 \le l \le d} |\tilde{m}_l(t) - m_l(t)| = \mathcal{O}_p(h_s)$.

Proposition 4 *Under Assumptions* (A2)–(A6) *in Appendix A, for $\tau \in \mathbb{R}$, and $\sigma_{n,ll}(t)$, a_{N_s+1}, b_{N_s+1} as given in* (7) *and* (9),

$$\lim_{n \to \infty} P \left\{ \sup_{t \in [0,1]} \sigma_{n,ll}^{-1}(t) \left| \tilde{\xi}_l(t) + \tilde{\varepsilon}_l(t) \right| \le \tau / a_{N_s+1} + b_{N_s+1} \right\} = \exp\left(-2\mathrm{e}^{-\tau}\right).$$

5 Simulation

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

$$Y_{ij} = \left\{ m_1\left(T_{ij}\right) + \sum_{k=1}^{2} \xi_{ik,1} \phi_{k,1}\left(T_{ij}\right) \right\} X_{i1} + \left\{ m_2\left(T_{ij}\right) + \sum_{k=1}^{3} \xi_{ik,2} \phi_{k,2}\left(T_{ij}\right) \right\} X_{i2}$$
$$+ \sigma\left(T_{ij}\right) \varepsilon_{ij}, \quad 1 \le i \le n, 1 \le j \le N_i,$$

where $T \sim U[0, 1]$, $X_1 \sim N(0, 1)$, $X_2 \sim$ Binomial$[1, 0.5]$, $\xi_{k,1} \sim N(0, 1)$, $k = 1, 2$, $\xi_{k,2} \sim N(0, 1)$, $k = 1, 2, 3$, $\varepsilon \sim N(0, 1)$, and N_i is generated from a discrete uniform distribution from $2, \ldots, 14$, for $1 \le i \le n$. For the first component, we take $m_1(t) = \sin\{2\pi (t - 1/2)\}$, $\phi_{1,1}(t) = -2\cos\{\pi (t - 1/2)\}/\sqrt{5}$, $\phi_{2,1}(t) = \sin\{\pi (t - 1/2)\}/\sqrt{5}$, thus $\lambda_{1,1} = 2/5$, $\lambda_{2,1} = 1/10$. For the second

Table 1 Coverage percentages of the SCCs for functions m_1 (*left*) and m_2 (*right*), based on 500 replications

σ	n	$1-\alpha$	$c=0.3$	$c=0.5$	$c=0.8$	$c=1$
1.0	200	0.950	0.950, 0.952	0.944, 0.948	0.920, 0.904	0.886, 0.884
		0.990	0.990, 0.998	0.990, 0.990	0.976, 0.984	0.968, 0.974
	400	0.950	0.944, 0.948	0.950, 0.930	0.922, 0.912	0.908, 0.904
		0.990	0.996, 0.984	0.990, 0.988	0.984, 0.988	0.974, 0.966
	600	0.950	0.934, 0.962	0.954, 0.946	0.930, 0.952	0.930, 0.924
		0.990	0.992, 0.996	0.992, 0.986	0.988, 0.990	0.984, 0.990
	800	0.950	0.936, 0.934	0.960, 0.966	0.950, 0.964	0.956, 0.934
		0.990	0.998, 0.996	0.994, 0.994	0.986, 0.992	0.988, 0.988
0.5	200	0.950	0.936, 0.948	0.952, 0.942	0.916, 0.900	0.912, 0.890
		0.990	0.988, 0.994	0.992, 0.990	0.972, 0.974	0.972, 0.972
	400	0.950	0.916, 0.930	0.936, 0.932	0.928, 0.916	0.904, 0.898
		0.990	0.994, 0.984	0.992, 0.988	0.996, 0.988	0.978, 0.976
	600	0.950	0.924, 0.948	0.952, 0.954	0.926, 0.958	0.936, 0.938
		0.990	0.996, 0.994	0.994, 0.986	0.984, 0.990	0.990, 0.994
	800	0.950	0.942, 0.900	0.950, 0.960	0.942, 0.962	0.960, 0.938
		0.990	0.996, 0.998	0.996, 0.994	0.990, 0.996	0.992, 0.988

component, we take $m_2(t) = 5\,(t-0.6)^2$, $\phi_{1,2}(t) = 1$, $\phi_{2,2}(t) = \sqrt{2}\sin(2\pi t)$, $\phi_{3,2}(t) = \sqrt{2}\cos(2\pi t)$, thus $\lambda_{1,2} = \lambda_{2,2} = \lambda_{3,2} = 1$. The noise level is chosen to be $\sigma = 0.5, 1.0$, and the number of subjects n is taken to be 200, 400, 600, 800.

We consider the confidence levels $1-\alpha = 0.95$ and 0.99. Table 1 reports the coverage of the SCCs as the percentage out of the total 500 replications for which the true curve was covered by (11) at the 101 points $\{k/100, k = 0, \ldots, 100\}$.

In the above SCC construction, the number of interior knots N_s is determined by the sample size n and a tuning constant c as described in Sect. 3.2. We have experimented with $c = 0.3, 0.5, 0.8, 1.0$ in this simulation study. The simulation results in Table 1 reflect that the coverage percentages depend on the choice of c; however, the dependency becomes weaker when sample sizes increase. For large sample sizes $n = 600, 800$, the effect of the choice of c on the coverage percentages is insignificant. Because N_s varies with N_i, for $1 \leq i \leq n$, the data-driven selection of an "optimal" N_s remains an open problem. At all noise levels, the coverage percentages for the SCC (11) are very close to the nominal confidence levels 0.95 and 0.99 for $c = 0.5$. Note that since $\mathsf{E}N_1 = 8$, the total sample size $N_T \approx 8 \times 200, 8 \times 400, 8 \times 600, 8 \times 800$ which explains the closeness of coverage percentages in Table 1 to the nominal levels. These large N_Ts are realistic as we believe they are common for real data. For instance, the CD4 cell percentage data in Sect. 6 has $N_T = 1{,}817$.

For visualization of actual function estimates, Fig. 1 shows the true curve, the estimated curve, the asymptotic 95 % SCC and the pointwise confidence intervals at $\sigma = 0.5$ with $n = 200$. The same plot for $n = 600$ has shown significantly

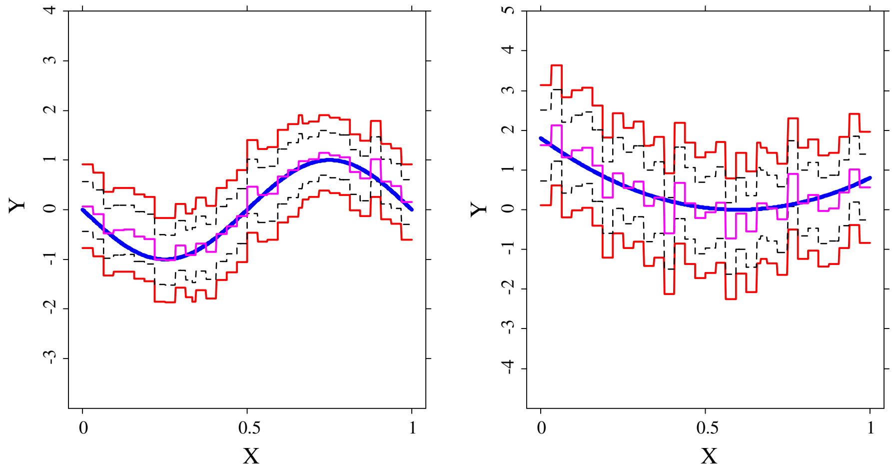

Fig. 1 Plots of 95 % SCC (11) (*upper and lower solid*), pointwise confidence intervals (*dashed*), the spline estimator (*thin*), and the true function (*middle thick*) at $\sigma = 0.5$, $n = 200$ for m_1 (*left*) and m_2 (*right*)

narrower SCC and pointwise confidence intervals as expected, but is not included to save space.

6 Real data analysis

To illustrate our method, we return to the CD4 cell percentage data discussed in Sect. 1 for further analysis. Since the actual visit times T_{ij} are irregularly spaced and vary from year 0 to year 6, we first transform the times by $Z_{ij} = F_{N_T}\left(T_{ij}\right)$, where F_{N_T} is the empirical cdf of times $\left\{T_{ij}\right\}_{i=1,j=1}^{n,N_i}$. Then the Z_{ij} values are distributed fairly uniformly. We have set a slightly smaller number of interior knots $N_{\mathrm{s}} = [0.3N_{\mathrm{T}}^{1/3}(\log(n))]$ to avoid singularity in solving the least squares problem.

The left plots of Figs. 2, 3, 4 and 5 depict the spline estimates, the asymptotic 95 % SCCs, the pointwise confidence intervals for $m_l\,(t)$, $l = 0, 1, 2, 3$, respectively. The horizontal solid line represents zero. Based on the shape of the SCCs, we are interested in testing the following hypotheses:

$H_{00} : m_0\,(t) \equiv a + bt$, for some $a, b \in \mathbb{R}$ v.s. $H_{10} : m_0\,(t) \neq a + bt$, for any $a, b \in \mathbb{R}$;

$H_{01} : m_1\,(t) \equiv 0$ v.s. $H_{11} : m_1\,(t) \neq 0$, for some $t \in [0, 6]$;

$H_{02} : m_2\,(t) \equiv c$ for some $c > 0$ v.s. $H_{12} : m_2\,(t) \neq c$, for any $c > 0$;

$H_{03} : m_3\,(t) \equiv 0$ v.s. $H_{13} : m_3\,(t) \neq 0$, for some $t \in [0, 6]$.

Asymptotic p values are calculated for each pair of hypotheses as $\hat{\alpha}_0 = 0.99072$, $\hat{\alpha}_1 = 0.79723$, $\hat{\alpha}_2 = 0.25404$, $\hat{\alpha}_3 = 0.10775$. Apparently, none of the null hypothesis is rejected. The p values are calculated as, for example

$$\hat{\alpha}_0 = 1 - \exp\left[-2\exp\left(-a_{N_{\mathrm{s}}+1}\left\{\max_{k=0}^{400}\left|\frac{\hat{m}_0\,(t_k) - \left(\hat{a} + \hat{b}t_k\right)}{\hat{\sigma}_{n,ll}\,(t_k)}\right| - b_{N_{\mathrm{s}}+1}\right\}\right)\right],$$

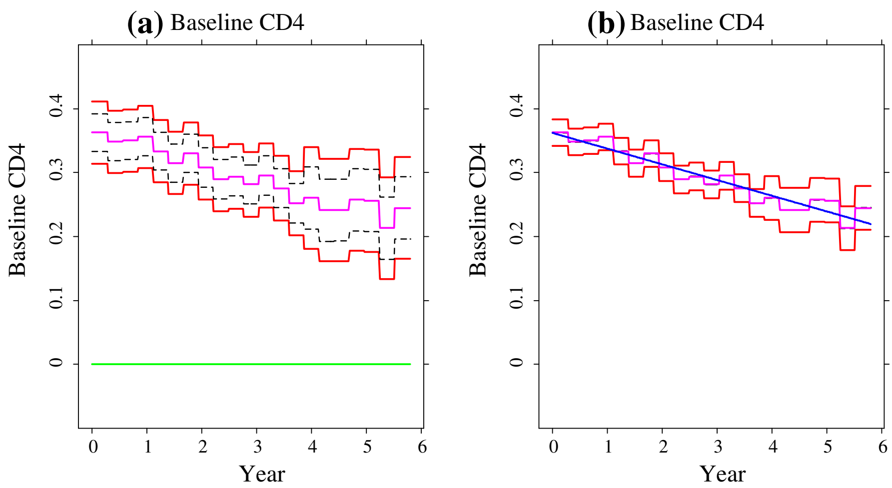

Fig. 2 Plots of **a** 95 % SCC (*upper and lower solid*), pointwise confidence intervals (*dashed*) and the spline estimator $\hat{m}_0$ (*middle solid*) for baseline effect; and **b** the same except with confidence level $1 - \hat{\alpha}_0$ and the estimated m_0 under H_{00} (*solid linear*)

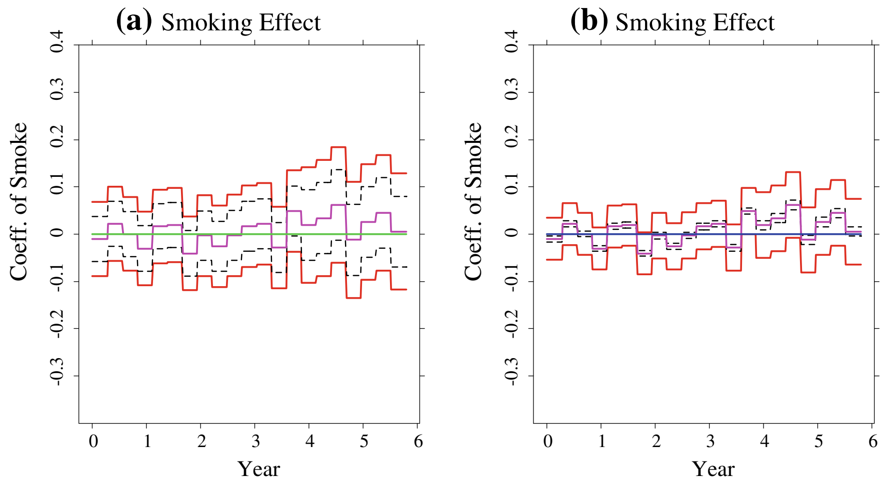

Fig. 3 Plots of **a** 95 % SCC (*upper and lower solid*), pointwise confidence intervals (*dashed*) and the spline estimator $\hat{m}_1$ (*middle solid*) for smoking effect; and **b** the same except with confidence level $1 - \hat{\alpha}_1$ and the estimated m_1 under H_{01} (*solid linear*)

where t_k, $k = 0, \ldots, 400$ are equally spaced grid points over the range of the actual visit times, while $\hat{a} + \hat{b}t$ is a least squares linear approximation to $\hat{m}_0(t)$. In other words, the p value $\hat{\alpha}_0$ is a solution of

$$\max_{k=0}^{400} \left| \frac{\hat{m}_0(t_k) - \left(\hat{a} + \hat{b}t_k\right)}{\hat{\sigma}_{n,ll}(t_k)} \right| = b_{N_s+1} - a_{N_s+1}^{-1} \log\left\{-\frac{1}{2}\log(1 - \hat{\alpha}_0)\right\}.$$

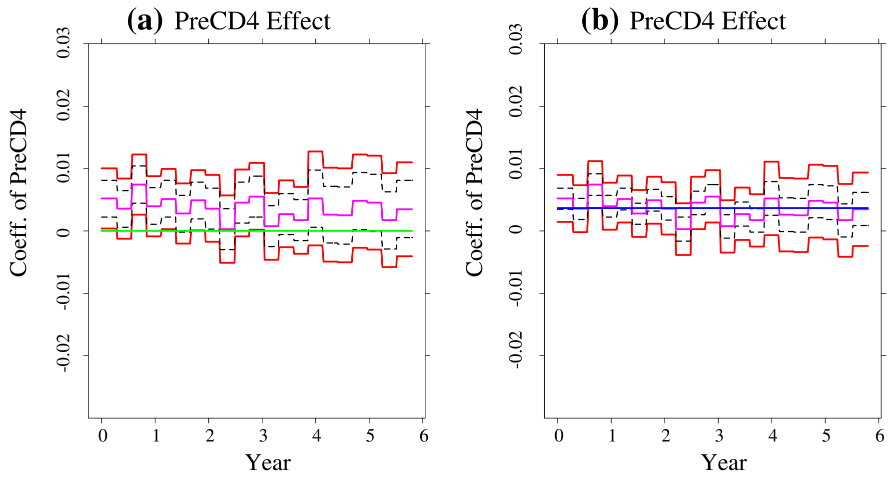

Fig. 4 Plots of **a** 95 % SCC (*upper and lower solid*), pointwise confidence intervals (*dashed*) and the spline estimator $\hat{m}_2$ (*middle solid*) for pre-infection CD4 effect; and **b** the same except with confidence level $1-\hat{\alpha}_2$ and the estimated m_2 under H_{02} (*solid linear*)

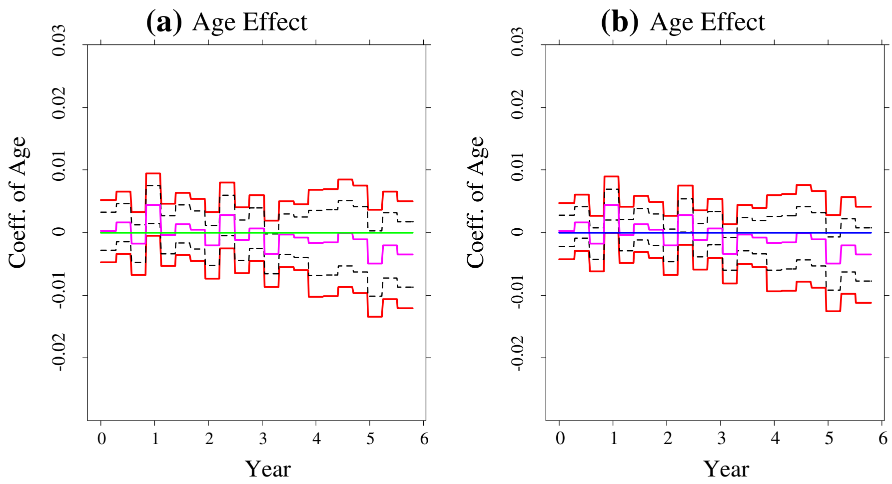

Fig. 5 Plots of **a** 95 % SCC (*upper and lower solid*), pointwise confidence intervals (*dashed*) and the spline estimator $\hat{m}_3$ (*middle solid*) for age effect; and **b** the same except with confidence level $1-\hat{\alpha}_3$ and the estimated m_3 under H_{03} (*solid linear*)

The right plots of Figs. 2, 3, 4 and 5 show the spline estimates, the $100(1-\hat{\alpha}_l)\%$ SCCs and the pointwise confidence intervals, and estimates of $m_l(t)$ under H_{0l}, $l = 0, 1, 2, 3$. From these figures, one can see the baseline CD4 percentage of the population is a decreasing linear function of time and greater than zero over the range of time. The effects of smoking status and age at HIV infection are insignificant, while the pre-infection CD4 percentage is positively proportional to the post-infection CD4 percentage. These findings are consistent with the observations in Wu and Chiang

(2000), Fan and Zhang (2000) and Wang et al. (2008), but are put on rigorous standing due to the quantification of type I errors by computing asymptotic p values relative to the SCCs.

Acknowledgments This work is part of Lijie Gu's dissertation and has been supported in part by the Deutsche Forschungsgemeinschaft through the CRC 649 "Economic Risk", the US National Science Foundation awards DMS 0905730, 1007594, 1106816, 1309800, Jiangsu Specially Appointed Professor Program SR10700111, Jiangsu Province Key-Discipline Program (Statistics) ZY107002, National Natural Science Foundation of China award 11371272, and Research Fund for the Doctoral Program of Higher Education of China award 20133201110002.

Appendix A

Throughout this section, $a_n \sim b_n$ means $\lim_{n\to\infty} b_n/a_n = c$, where c is some nonzero constant. For functions $a_n(t)$, $b_n(t)$, $a_n(t) = \mathcal{U}\{b_n(t)\}$ means $a_n(t)/b_n(t) \to 0$ as $n \to \infty$ uniformly for $t \in [0, 1]$, and $a_n(t) = \mathcal{U}\{b_n(t)\}$ means $a_n(t)/b_n(t) = \mathcal{O}(1)$ as $n \to \infty$ uniformly for $t \in [0, 1]$. We use $\mathcal{U}_p(\cdot)$ and $\mathcal{U}_p(\cdot)$ if the convergence is in the sense of uniform convergence in probability.

A.1 Technical assumptions

We define the modulus of continuity of a continuous function ϕ on $[a, b]$ by $\omega(\phi, \delta) = \max_{t,t'\in[a,b],|t-t'|\le\delta} \left|\phi(t) - \phi\left(t'\right)\right|$. For any $r \in (0, 1]$, denote the collection of order r Hõlder continuous function on $[0, 1]$ by

$$C^{0,r}[0,1] = \left\{\phi : \|\phi\|_{0,r} = \sup_{t\ne t',t,t'\in[0,1]} \frac{\left|\phi(t) - \phi\left(t'\right)\right|}{|t-t'|^r} < +\infty\right\},$$

in which $\|\phi\|_{0,r}$ is the $C^{0,r}$-seminorm of ϕ. Let $C[0, 1]$ be the collection of continuous function on $[0, 1]$. Clearly, $C^{0,r}[0, 1] \subset C[0, 1]$ and, if $\phi \in C^{0,r}[0, 1]$, then $\omega(\phi, \delta) \le \|\phi\|_{0,r}\,\delta^r$.

The following regularity assumptions are needed for the main results.

- (A1) The regression functions $m_l(t) \in C^{0,1}[0, 1]$, $l = 1, \ldots, d$.
- (A2) The set of random variables $\left(T_{ij}, \varepsilon_{ij}, N_i, \xi_{ik,l}, X_{il}\right)_{i=1,j=1,k=1,l=1}^{n,N_i,\infty,d}$ is a subset of variables $\left(T_{ij}, \varepsilon_{ij}, N_i, \xi_{ik,l}, X_{il}\right)_{i=1,j=1,k=1,l=1}^{\infty,\infty,\infty,d}$ consisting of independent random variables, in which all T_{ij}'s i.i.d with $T_{ij} \sim T$, where T is a random variable with probability density function $f(t)$; X_{il}'s i.i.d for each $l = 1, \ldots, d$; N_i's i.i.d with $N_i \sim N$, where $N > 0$ is a positive integer-valued random variable with $\mathsf{E}\{N^{2r}\} \le r!c_N^r$, $r = 2, 3, \ldots,$ for some constant $c_N > 0$. Variables$\left(\xi_{ik,l}\right)_{i=1,k=1,l=1}^{\infty,\infty,d}$ and $\left(\varepsilon_{ij}\right)_{i=1,j=1}^{\infty,\infty}$ are i.i.d $N(0, 1)$.
- (A3) The functions $f(t)$, $\sigma(t)$ and $\phi_{k,l}(t) \in C^{0,r}[0, 1]$ for some $r \in (0, 1]$ with $f(t) \in \left[c_f, C_f\right]$, $\sigma(t) \in [c_\sigma, C_\sigma]$, $t \in [0, 1]$, for constants $0 < c_f \le C_f < +\infty$, $0 < c_\sigma \le C_\sigma < +\infty$.

(A4) For $l = 1, \dots, d, \sum_{k=1}^{\infty} \left\| \phi_{k,l} \right\|_{\infty} < +\infty$, and $G_l(t,t) \in \left[c_{G,l}, C_{G,l}\right], t \in [0,1]$, for constants $0 < c_{G,l} \le C_{G,l} < +\infty$.

(A5) There exist constants $0 < c_{\mathbf{H}} \le C_{\mathbf{H}} < +\infty$ and $0 < c_\eta \le C_\eta < +\infty$, such that $c_{\mathbf{H}} I_{d\times d} \le \mathbf{H} = \{H_{ll'}\}_{l,l'=1}^{d} = \mathsf{E}\left(\mathbf{X}\mathbf{X}^{\mathsf{T}}\right) \le C_{\mathbf{H}} I_{d\times d}$. For some $\eta > 4$, $l = 1, \dots, d, c_\eta \le \mathsf{E}\left|X_l\right|^{8+\eta} \le C_\eta$.

(A6) As $n \to \infty$, the number of interior knots $N_{\mathrm{s}} = \mathcal{O}\left(n^{\vartheta}\right)$ for some $\vartheta \in (1/3, 1/2)$ while $N_{\mathrm{s}}^{-1} = \mathcal{O}\left\{n^{-1/3}\left(\log(n)\right)^{-1/3}\right\}$. The subinterval length $h_{\mathrm{s}} \sim N_{\mathrm{s}}^{-1}$.

Assumptions (A1)–(A3) are common conditions used in the literature; see for example, Ma et al. (2012). Assumption (A1) controls the rate of convergence of the spline approximation $\hat{m}_l, l = 1, \dots, d$. The requirement of N_i in Assumption (A2) ensures that the observation times are randomly scattered, reflecting sparse and irregular designs. Assumption (A4) guarantees that the random variable $\sum_{k=1}^{\infty} \xi_{ik,l}\phi_{k,l}(t)$ absolutely uniformly converges. Assumption (A5) is analog to Assumption (A2) in Liu and Yang (2010), ensuring that the X_{il}s are not multicollinear. Assumption (A6) describes the requirement of the growth rate of the dimension of the spline spaces relative to the sample size.

A.2 Preliminaries

Lemma 1 (Bosq (1998), Theorem 1.2). *Suppose that* $\{\xi_i\}_{i=1}^n$ *are i.i.d with* $\mathsf{E}(\xi_1) = 0, \sigma^2 = \mathsf{E}\xi_1^2$, *and there exists* $c > 0$ *such that for* $r = 3, 4, \dots$, $\mathsf{E}\left|\xi_1\right|^r \le c^{r-2} r! \mathsf{E}\xi_1^2 < +\infty$. *Then for each* $n > 1$, $t > 0$, $P(|S_n| \ge \sqrt{n}\sigma t) \le 2\exp\left(-t^2\left(4 + 2ct/\sqrt{n}\sigma\right)^{-1}\right)$, *in which* $S_n = \sum_{i=1}^n \xi_i$.

Lemma 2 *Under Assumptions* (A2)–(A6), *we have*

$$A_{n,1} = \sup_{0\le J\le N_{\mathrm{s}}, 1\le l,l'\le d} \frac{\left|\langle B_J X_l, B_J X_{l'}\rangle_{N_{\mathrm{T}}} - \langle B_J X_l, B_J X_{l'}\rangle\right|}{\sqrt{\langle B_J X_l, B_J X_l\rangle}\sqrt{\langle B_J X_{l'}, B_J X_{l'}\rangle}} = \mathcal{O}_p\left(\sqrt{\frac{\log(n)}{nh_{\mathrm{s}}}}\right),$$

where for any $J = 0, \dots, N_{\mathrm{s}}$ *and* $l, l' = 1, \dots, d$,

$$\langle B_J X_l, B_J X_{l'}\rangle_{N_{\mathrm{T}}} = N_{\mathrm{T}}^{-1} \sum_{i=1}^{n} \sum_{j=1}^{N_i} B_J^2(T_{ij}) X_{il} X_{il'},$$
$$\langle B_J X_l, B_J X_{l'}\rangle = \mathsf{E}\left\{B_J^2(T_{ij}) X_{il} X_{il'}\right\} = H_{ll'}.$$

Proof Let $\omega_{i,J} = \omega_{i,J,l,l'} = \sum_{j=1}^{N_i} B_J^2(T_{ij}) X_{il} X_{il'}$, then $\mathsf{E}\omega_{i,J} = \mathsf{E}N_1 H_{ll'} \sim 1$ and $\mathsf{E}\left(\omega_{ij,J}\right)^2 = \mathsf{E}\left\{\sum_{j=1}^{N_i} B_J^2(T_{ij})\right\}^2 \mathsf{E}\left(X_{il}X_{il'}\right)^2 \sim h_{\mathrm{s}}^{-1}$. Next define a sequence $D_n = n^{\alpha}$ with $\alpha(4+\eta/2) > 1$ and $\sqrt{\log(n)} D_n n^{-1/2} h_{\mathrm{s}}^{-1/2} \to 0$, $n^{1/2} h_{\mathrm{s}}^{1/2} D_n^{-(3+\eta/2)} \to 0$, which necessitates $\eta > 2$ according to Assumption (A5). We make use of the following truncated and tail decomposition

$$X_{ill'} = X_{il} X_{il'} = X_{ill',1}^{D_n} + X_{ill',2}^{D_n},$$

where $X_{ill',1}^{D_n} = X_{il}X_{il'}I\left\{|X_{il}X_{il'}| > D_n\right\}$, $X_{ill',2}^{D_n} = X_{il}X_{il'}I\left\{|X_{il}X_{il'}| \le D_n\right\}$. Correspondingly, the truncated and tail parts of $\omega_{i,J}$ are $\omega_{i,J,m} = B_J^2(T_{ij})X_{ill',m}^{D_n}$, $m = 1, 2$. According to Assumption (A5), for any $l, l' = 1, \ldots, d$,

$$\sum_{n=1}^{\infty} P\left\{|X_{nl}X_{nl'}| > D_n\right\} \le \sum_{n=1}^{\infty} \frac{\mathsf{E}\,|X_{nl}X_{nl'}|^{4+\eta/2}}{D_n^{4+\eta/2}} \le C_\eta \sum_{n=1}^{\infty} D_n^{-(4+\eta/2)} < \infty.$$

By Borel–Cantelli Lemma, one has $\sum_{j=1}^{N_i} B_J^2(T_{ij})X_{ill',1}^{D_n} = 0$, $a.s.$. So we obtain

$$\sup_{J,l,l'} \left| n^{-1} \sum_{i=1}^{n} \omega_{i,J,1} \right| = \mathcal{O}_{a.s.}\left(n^{-k}\right), \quad k \ge 1,$$

and

$$\begin{aligned}
\mathsf{E}\omega_{i,J,1} &= \mathsf{E}\left(X_{ill',1}^{D_n}\right) \mathsf{E}\left\{\sum_{j=1}^{N_i} B_J^2(T_{ij})\right\} \\
&\le D_n^{-(3+\eta/2)} \mathsf{E}\,|X_{il}X_{il'}|^{4+\eta/2}\, \mathsf{E}N_1 \mathsf{E}B_J^2(T_{ij}) \le cD_n^{-(3+\eta/2)}.
\end{aligned}$$

Next we considerate the truncated part $\omega_{i,J,2}$. For large n, $\mathsf{E}\left(\omega_{i,J,2}\right) = \mathsf{E}\left(\omega_{i,J}\right) - \mathsf{E}\left(\omega_{i,J,1}\right) \sim 1$, $\mathsf{E}\left(\omega_{i,J,2}\right)^2 = \mathsf{E}\left(\omega_{i,J}\right)^2 - \mathsf{E}\left(\omega_{i,J,1}\right)^2 \sim h_{\mathrm{s}}^{-1}$. Define $\omega_{i,J,2}^{*} = \omega_{i,J,2} - \mathsf{E}\left(\omega_{i,J,2}\right)$, then $\mathsf{E}\omega_{i,J,2}^{*} = 0$, and

$$\begin{aligned}
\mathsf{E}\left(\omega_{i,J,2}^{*}\right)^2 = \mathsf{E}\left(\omega_{i,J,2}\right)^2 - \left(\mathsf{E}\omega_{i,J,2}\right)^2 &= \mathsf{E}\left\{\sum_{j=1}^{N_i} B_J^2(T_{ij})X_{ill',2}^{D_n}\right\}^2 - \mathcal{U}(1) \\
&= \mathsf{E}\left(X_{ill',2}^{D_n}\right)^2 \mathsf{E}\left\{\sum_{j=1}^{N_i} B_J^2(T_{ij})\right\}^2 - \mathcal{U}(1).
\end{aligned}$$

Note that

$$\begin{aligned}
\mathsf{E}\left(X_{ill',2}^{D_n}\right)^2 \mathsf{E}\left\{\sum_{j=1}^{N_i} B_J^2(T_{ij})\right\}^2 &\ge \left\{\mathsf{E}\left(X_{ill'}\right)^2 - \mathsf{E}\left(X_{ill',1}^{D_n}\right)^2\right\} \mathsf{E}\left\{\sum_{j=1}^{N_i} B_J^4(T_{ij})\right\} \\
&\ge \left\{\mathsf{E}\left(X_{ill'}\right)^2 - \mathcal{U}(1)\right\} \mathsf{E}N_1 \mathsf{E}B_J^4(T_{ij}).
\end{aligned}$$

Thus, there exists c_ω such that for large n, $\mathsf{E}\left(\omega_{i,J,2}^{*}\right)^2 \ge c_\omega \mathsf{E}\left(X_{ill'}\right)^2 h_{\mathrm{s}}^{-1}$. Next for any $r > 2$

$$\begin{aligned}
\mathsf{E}\left|\omega_{i,J,2}^{*}\right|^{r} &= \mathsf{E}\left|\omega_{i,J,2}-\mathsf{E}\left(\omega_{i,J,2}\right)\right|^{r} \leq 2^{r-1}\left(\mathsf{E}\left|\omega_{i,J,2}\right|^{r}+\left|\mathsf{E}\left(\omega_{i,J,2}\right)\right|^{r}\right) \\
&= 2^{r-1}\left\{\mathsf{E}\left|X_{ill',2}^{D_n}\right|^{r}\mathsf{E}\left|\sum_{j=1}^{N_i}B_J^2(T_{ij})\right|^{r}+\mathcal{U}(1)\right\} \\
&= 2^{r-1}\left[\mathsf{E}\left|X_{ill',2}^{D_n}\right|^{r}\mathsf{E}\left\{\sum_{0\leq r_1,\ldots,r_{N_i}\leq r}^{r_1+\cdots+r_{N_i}=r}\binom{r}{r_1\cdots r_{N_i}}\prod_{j=1}^{N_i}\mathsf{E}B_J^{2r_j}\left(T_{ij}\right)\right\}+\mathcal{U}(1)\right],
\end{aligned}$$

then there exists $C_\omega > 0$ such that for any $r > 2$ and large n,

$$\begin{aligned}
\mathsf{E}\left|\omega_{i,J,2}^{*}\right|^{r} &\leq 2^{r-1}\left[D_n^{r-2}\mathsf{E}\left(X_{ill'}\right)^2\mathsf{E}\left\{N_1^r\max\prod_{j=1}^{N_i}\mathsf{E}B_J^{2r_j}\left(T_{ij}\right)\right\}+\mathcal{U}(1)\right] \\
&\leq 2^{r-1}\left[D_n^{r-2}\mathsf{E}\left(X_{ill'}\right)^2\left(\mathsf{E}N_1^r\right)C_B h_{\mathrm{s}}^{1-r}+\mathcal{U}(1)\right] \\
&\leq 2^r D_n^{r-2}\left(c_N^r r!\right)^{1/2}C_B h_{\mathrm{s}}^{2-r}c_\omega^{-1}\mathsf{E}\left(\omega_{i,J,2}^{*}\right)^2 \\
&\leq \left(C_\omega D_n h_{\mathrm{s}}^{-1}\right)^{r-2} r!\,\mathsf{E}\left(\omega_{i,J,2}^{*}\right)^2,
\end{aligned}$$

which implies that $\left\{\omega_{i,J,2}^{*}\right\}_{i=1}^{n}$ satisfies Cramér's condition with constant $C_\omega D_n h_{\mathrm{s}}^{-1}$. Applying Lemma 1 to $\sum_{i=1}^{n}\omega_{i,J,2}^{*}$, for $r > 2$ and any large enough $\delta > 0$, $P\left\{\left|n^{-1}\sum_{i=1}^{n}\omega_{i,J,2}^{*}\right| \geq \delta\left(nh_{\mathrm{s}}\right)^{-1/2}\left(\log(n)\right)^{1/2}\right\}$ is bounded by

$$2\exp\left\{\frac{-\delta^2\left(\log(n)\right)}{4+2C_\omega D_n h_{\mathrm{s}}^{-1}\delta\left(\log(n)\right)^{1/2}n^{-1/2}h_{\mathrm{s}}^{1/2}}\right\} \leq 2n^{-8}.$$

Hence

$$\sum_{n=1}^{\infty}P\left\{\sup_{0\leq J\leq N_{\mathrm{s}},1\leq l,l'\leq d}\left|n^{-1}\sum_{i=1}^{n}\omega_{i,J,2}^{*}\right| \geq \delta\left(nh_{\mathrm{s}}\right)^{-1/2}\left(\log(n)\right)^{1/2}\right\} < \infty.$$

Thus, $\sup_{J,l,l'}\left|n^{-1}\sum_{i=1}^{n}\omega_{i,J,2}^{*}\right| = \mathcal{O}_{a.s.}\left\{\left(nh_{\mathrm{s}}\right)^{-1/2}\left(\log(n)\right)^{1/2}\right\}$ as $n \to \infty$ by Borel–Cantelli Lemma. Furthermore,

$$\begin{aligned}
&\sup_{J,l,l'}\left|n^{-1}\sum_{i=1}^{n}\omega_{i,J}-\mathsf{E}\omega_{i,J}\right| \\
&\quad\leq \sup_{J,l,l'}\left|n^{-1}\sum_{i=1}^{n}\omega_{i,J,1}\right|+\sup_{J,l,l'}\left|n^{-1}\sum_{i=1}^{n}\omega_{i,J,2}^{*}\right|+\sup_{J,l,l'}\left|\mathsf{E}\omega_{i,J,1}\right|
\end{aligned}$$

$$= \mathcal{U}_{a.s.}\left(n^{-k}\right) + \mathcal{O}_{a.s.}\left\{(nh_{\mathrm{s}})^{-1/2}(\log(n))^{1/2}\right\} + \mathcal{U}\left(D_n^{-(3+\eta/2)}\right)$$
$$= \mathcal{O}_{a.s.}\left\{(nh_{\mathrm{s}})^{-1/2}(\log(n))^{1/2}\right\}.$$

Finally, we notice that

$$\sup_{J,l,l'}\left|\langle B_J X_l, B_J X_{l'}\rangle_{N_{\mathrm{T}}} - \langle B_J X_l, B_J X_{l'}\rangle\right| = \sup_{J,l,l'}\left|\left(nN_{\mathrm{T}}^{-1}\right)n^{-1}\sum_{i=1}^{n}\omega_{i,J} - (\mathsf{E}N_1)^{-1}\mathsf{E}\omega_{i,J}\right|$$
$$\leq \sup_{J,l,l'}(\mathsf{E}N_1)^{-1}\left|(n\mathsf{E}N_1)\,N_{\mathrm{T}}^{-1} - 1\right|\left|n^{-1}\sum_{i=1}^{n}\omega_{i,J}\right| + \sup_{J,l,l'}(\mathsf{E}N_1)^{-1}\left|n^{-1}\sum_{i=1}^{n}\omega_{i,J} - \mathsf{E}\omega_{i,J}\right|$$
$$= \mathcal{O}_p\left(n^{-1/2}\right) + \mathcal{O}_{a.s.}\left\{(nh_{\mathrm{s}})^{-1/2}(\log(n))^{1/2}\right\} = \mathcal{O}_p\left\{(nh_{\mathrm{s}})^{-1/2}(\log(n))^{1/2}\right\},$$

and $\langle B_J X_l, B_J X_l\rangle = H_{ll} = \mathcal{U}(1)$. Hence, $A_{n,1} = \mathcal{O}_p\left\{(nh_{\mathrm{s}})^{-1/2}(\log(n))^{1/2}\right\}$. □

For the random matrix $\hat{\mathbf{V}}_J$ defined in (18), the lemma below shows that its inverse can be approximated by the inverse of a deterministic matrix $\mathbf{H} = \mathsf{E}(\mathbf{X}\mathbf{X}^{\mathsf{T}})$.

Lemma 3 *Under Assumptions* (A2) *and* (A4)–(A6), *for any* $J = 0, \ldots, N_{\mathrm{s}}$, *we have*

$$\hat{\mathbf{V}}_J^{-1} = \mathbf{H}^{-1} + \mathcal{O}_p\left\{(nh_{\mathrm{s}})^{-1/2}(\log(n))^{1/2}\right\}. \tag{25}$$

Proof By Lemma 2, we have

$$\left\|\hat{\mathbf{V}}_J - \mathbf{H}\right\|_{\infty} = \mathcal{O}_p\left\{(nh_{\mathrm{s}})^{-1/2}(\log(n))^{1/2}\right\}.$$

Using the fact that for any matrices $\mathbf{A}$ and $\mathbf{B}$,

$$(\mathbf{A} + h\mathbf{B})^{-1} = \mathbf{A}^{-1} - h\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1} + \mathcal{O}(h^2),$$

we obtain (25). □

The next lemma implies that the difference between $\tilde{\boldsymbol{\xi}}(t)$ and $\hat{\boldsymbol{\xi}}(t)$ and the difference between $\tilde{\boldsymbol{\varepsilon}}(t)$ and $\hat{\boldsymbol{\varepsilon}}(t)$ are both negligible uniformly over $t \in [0, 1]$.

Lemma 4 *Under Assumption* (A2)–(A6), *for* $\tilde{\boldsymbol{\xi}}(t)$, $\tilde{\boldsymbol{\varepsilon}}(t)$ *given in* (36), (37) *and* $\hat{\boldsymbol{\xi}}(t)$, $\hat{\boldsymbol{\varepsilon}}(t)$ *given in* (38), (39), *as* $n \to \infty$, *we have*

$$\sup_{t\in[0,1]}\left\|\tilde{\boldsymbol{\xi}}(t) - \hat{\boldsymbol{\xi}}(t)\right\|_{\infty} = \mathcal{O}_p\left\{n^{-1}h_{\mathrm{s}}^{-3/2}\log(n)\right\}, \tag{26}$$
$$\sup_{t\in[0,1]}\left\|\tilde{\boldsymbol{\varepsilon}}(t) - \hat{\boldsymbol{\varepsilon}}(t)\right\|_{\infty} = \mathcal{O}_p\left\{n^{-1}h_{\mathrm{s}}^{-3/2}\log(n)\right\}. \tag{27}$$

Proof Comparing the equations of $\tilde{\boldsymbol{\xi}}(t)$ and $\hat{\boldsymbol{\xi}}(t)$ given in (A.2) and (A.4), we let

$$\frac{1}{N_{\mathrm{T}}}\sum_{i=1}^{n}\sum_{j=1}^{N_i} B_J(T_{ij})X_{il}\sum_{l''=1}^{d}\sum_{k=1}^{\infty}\xi_{ik,l''}\phi_{k,l''}\left(T_{ij}\right)X_{il''} = \frac{n}{N_{\mathrm{T}}}\sum_{l''=1}^{d}\sum_{i=1}^{n}\Omega_{i,J,l'',l}.$$

where $\Omega_{i,J,l'',l} = \Omega_i = n^{-1}\left[X_{il}X_{il''}\sum_{k=1}^{\infty}\left\{\sum_{j=1}^{N_i} B_J(T_{ij})\phi_{k,l''}\left(T_{ij}\right)\right\}\xi_{ik,l''}\right]$. Note that $\mathsf{E}\Omega_i = 0$ and

$$\begin{aligned}
\sigma^2_{\Omega_i,n} &= \mathsf{E}\left(\Omega_i^2\,\middle|\left(T_{ij}, N_i, X_{il}\right)_{i=1,j=1,l=1}^{n,N_i,d}\right)\\
&= n^{-2}\left[X_{il}X_{il''}\sum_{k=1}^{\infty}\left\{\sum_{j=1}^{N_i} B_J(T_{ij})\phi_{k,l''}\left(T_{ij}\right)\right\}^2\right]\\
&\le n^{-2}\left\{X_{il}^2X_{il''}^2\sum_{k=1}^{\infty}N_i\sum_{j=1}^{N_i} B_J^2(T_{ij})\phi_{k,l''}^2\left(T_{ij}\right)\right\}\\
&= n^{-2}\left\{X_{il}^2X_{il''}^2N_i\sum_{j=1}^{N_i} B_J^2(T_{ij})G_{l''}\left(T_{ij}, T_{ij}\right)\right\}\\
&\le Cn^{-2}h_{\mathrm{s}}^{-1}X_{il}^2X_{il''}^2N_i^2.
\end{aligned}$$

Given $\left(T_{ij}, N_i, X_{il}\right)_{i=1,j=1,l=1}^{n,N_i,d}$, $\left\{\sigma^{-1}_{\Omega_i,n}\Omega_i\right\}_{i=1}^{n}$ are i.i.d $N\,(0,1)$. It is easy to show that for any large enough $\delta > 0$,

$$\begin{aligned}
&P\left\{\frac{\left|\sum_{i=1}^{n}\Omega_i\right|}{\sqrt{\sum_{i=1}^{n}\sigma^2_{\Omega_i,n}}} \ge \delta\sqrt{\log(n)}\,\middle|\,\left(T_{ij}, N_i, X_{il}\right)_{i=1,j=1,l=1}^{n,N_i,d}\right\}\\
&\le 2\exp\left\{-\frac{1}{2}\delta^2\log(n)\right\} \le 2n^{-8},
\end{aligned}$$

$$P\left[\left|\sum_{i=1}^{n}\Omega_i\right| \ge \delta\left\{\frac{C\log(n)}{nh_{\mathrm{s}}}n^{-1}\sum_{i=1}^{n}X_{il}^2X_{il''}^2N_i^2\right\}^{1/2}\middle|\left(T_{ij}, N_i, X_{il}\right)_{i=1,j=1,l=1}^{n,N_i,d}\right] \le 2n^{-8}.$$

Note that $n^{-1}\sum_{i=1}^{n}X_{il}^2X_{il''}^2N_i^2 = \mathcal{O}_p\,(1)$, hence

$$\sum_{n=1}^{\infty}P\left\{\sup_{0\le J\le N_{\mathrm{s}},1\le l,l''\le d}\left|\sum_{i=1}^{n}\Omega_{i,J,l'',l}\right| \ge \delta\,(nh_{\mathrm{s}})^{-1/2}\,(\log(n))^{1/2}\right\} < \infty.$$

Thus, $\sup_{J,l,l''}\left|\sum_{i=1}^{n}\Omega_{i,J,l'',l}\right| = \mathcal{O}_p\left\{(nh_{\mathrm{s}})^{-1/2}\,(\log(n))^{1/2}\right\}$ as $n \to \infty$ by Borel–Cantelli Lemma. It follows that $\sup_{J,l}\left|nN_{\mathrm{T}}^{-1}\sum_{l''=1}^{d}\sum_{i=1}^{n}\Omega_{i,J,l'',l}\right| =$

$\mathcal{O}_p\left\{ (nh_\mathrm{s})^{-1/2} (\log(n))^{1/2}\right\}$. Finally, according to Lemma 25, we obtain (26). (27) is proved similarly. □

Denote the inverse matrix of $\mathbf{H}$ by $\mathbf{H}^{-1} = \{z_{ll'}\}_{l,l'=1}^{d}$. For any $l = 1, \ldots, d$, we rewrite the lth element of $\hat{\xi}_l(t)$ and $\hat{\varepsilon}_l(t)$ in (38) and (39) as the following

$$\hat{\xi}_l(t) = c_{J(t),n}^{-1/2} N_\mathrm{T}^{-1} \sum_{l''=1}^{d} \sum_{i=1}^{n} \sum_{k=1}^{\infty} R_{ik,\xi,J(t),l'',l}\xi_{ik,l''}, \tag{28}$$

$$\hat{\varepsilon}_l(t) = c_{J(t),n}^{-1/2} N_\mathrm{T}^{-1} \sum_{i=1}^{n} \sum_{j=1}^{N} R_{ij,\varepsilon,J(t),l}\varepsilon_{ij}, \tag{29}$$

where for any $0 \le J \le N_\mathrm{s}$,

$$R_{ik,\xi,J,l'',l} = \left(\sum_{l'=1}^{d} z_{ll'} X_{il'} X_{il''}\right) \left\{\sum_{j=1}^{N_i} B_J\left(T_{ij}\right) \phi_{k,l''}\left(T_{ij}\right)\right\}, \tag{30}$$

$$R_{ij,\varepsilon,J,l} = \left(\sum_{l'=1}^{d} z_{ll'} X_{il'}\right) B_J\left(T_{ij}\right) \sigma\left(T_{ij}\right). \tag{31}$$

Further denote

$$S_{ill''} = \left(\sum_{l'=1}^{d} z_{ll'} X_{il'} X_{il''}\right)^2, \quad s_{ll''} = \mathsf{E}\left(S_{ill''}\right), \quad 1 \le l, l'' \le d. \tag{32}$$

Lemma 5 *Under Assumptions* (A2)–(A6), *for* $R_{ik,\xi,J,l'',l}$, $R_{ij,\varepsilon,J,l}$ *in* (30), (31),

$$\mathsf{E}\left(\sum_{k=1}^{\infty} R_{ik,\xi,J,l'',l}^2\right) = c_{J,n}^{-1} s_{ll''}\left[(\mathsf{E}N_1)\int b_J(u)\, G_{l''}(u,u)\, f(u)\,\mathrm{d}u \right.$$
$$\left. + \mathsf{E}\left\{N_1(N_1-1)\right\}\int b_J(u)\, b_J(v)\, G_{l''}(u,v)\, f(u)\, f(v)\,\mathrm{d}u\mathrm{d}v\right],$$

$$\mathsf{E}R_{ij,\varepsilon,J,l}^2 = c_{J,n}^{-1} z_{ll} \int b_J(u)\, \sigma^2(u)\, f(u)\,\mathrm{d}u,$$

for $0 \le J \le N_\mathrm{s}$ *and* $0 \le l, l'' \le d$. *In addition, there exist* $0 < c_R < C_R < \infty$, *such that*

$$c_R s_{ll''} \le \mathsf{E}\left(\sum_{k=1}^{\infty} R_{ik,\xi,J,l'',l}^2\right) \le C_R s_{ll''}, \quad c_R \le \mathsf{E}R_{ij,\varepsilon,J,l}^2 \le C_R,$$

for $0 \leq J \leq N_{\mathrm{s}}$, $0 \leq l, l'' \leq d$, *and as* $n \to \infty$

$$A_{n,\xi} = \sup_{J,l'',l} \left| n^{-1} \sum_{i=1}^{n} \sum_{k=1}^{\infty} R_{ik,\xi,J,l'',l}^2 - \mathsf{E}\left(\sum_{k=1}^{\infty} R_{ik,\xi,J,l'',l}^2 \right) \right|$$
$$= \mathcal{O}_{a.s.} \left\{ (nh_{\mathrm{s}})^{-1/2} (\log(n))^{1/2} \right\},$$
$$A_{n,\varepsilon} = \sup_{J,l} \left| N_{\mathrm{T}}^{-1} \sum_{i=1}^{n} \sum_{j=1}^{N_i} R_{ij,\varepsilon,J,l}^2 - \mathsf{E} R_{ij,\varepsilon,J,l}^2 \right| = \mathcal{O}_{a.s.} \left\{ (nh_{\mathrm{s}})^{-1/2} (\log(n))^{1/2} \right\}.$$

Proof By independence of $\left\{T_{ij}\right\}_{j=1}^{\infty}$, $\{X_{il}\}_{l=1}^{d}$, N_i, the definition of B_J and (32),

$$\mathsf{E}\left(\sum_{k=1}^{\infty} R_{ik,\xi,J,l'',l}^2 \right) = \mathsf{E}\left(S_{ill''}\right) \mathsf{E} \sum_{k=1}^{\infty} \left\{ \sum_{j=1}^{N_i} B_J\left(T_{ij}\right) \phi_{k,l''}\left(T_{ij}\right) \right\}^2$$
$$= s_{ll''} \mathsf{E} \sum_{j=1}^{N_i} \sum_{j'=1}^{N_i} B_J\left(T_{ij}\right) B_J\left(T_{ij'}\right) G_{l''}\left(T_{ij}, T_{ij'}\right)$$
$$= s_{ll''} c_{J,n}^{-1} \left\{ (\mathsf{E} N_1) \int b_J(u) G_{l''}(u,u) f(u) \,\mathrm{d}u \right.$$
$$\left. + \mathsf{E}\left\{N_1(N_1 - 1)\right\} \int b_J(u) b_J(v) G_{l''}(u,v) f(u) f(v) \,\mathrm{d}u\mathrm{d}v \right\},$$

thus there exist constants $0 < c_R < C_R < \infty$ such that $c_R s_{ll''} \leq \mathsf{E}\left(\sum_{k=1}^{\infty} R_{1k,\xi,J,l'',l}^2\right) \leq C_R s_{ll''}$, $0 \leq J \leq N_{\mathrm{s}}$, $0 \leq l, l'' \leq d$.

If $s_{ll''} = 0$, one has $S_{ill''} = 0$, almost surely. Hence $n^{-1} \sum_{i=1}^{n} \sum_{k=1}^{\infty} R_{ik,\xi,J,l'',l}^2 = 0$, almost surely. In the case of $s_{ll''} > 0$, let $\zeta_{i,J} = \zeta_{i,J,l'',l} = \sum_{k=1}^{\infty} R_{ik,\xi,J,l'',l}^2$ for brevity. Under Assumption (A5), it is easy to verify that

$$0 < s_{ll''}^2 \leq \mathsf{E}\left(S_{ill''}\right)^2 \leq d^3 \sum_{l'=1}^{d} \mathsf{E}\left|z_{ll'} X_{il'} X_{il''}\right|^4 \leq d^3 \sum_{l'=1}^{d} z_{ll'} \left\{ \mathsf{E}\left|X_{il'}\right|^8 \mathsf{E}\left|X_{il''}\right|^8 \right\}^{1/2} < \infty.$$

So for large n,

$$\mathsf{E}\left(\zeta_{i,J}\right)^2 = \mathsf{E}\left\{ \left(S_{ill''}\right)^2 \left(\sum_{j=1}^{N_i} \sum_{j'=1}^{N_i} B_J\left(T_{ij}\right) B_J\left(T_{ij'}\right) G_{l''}\left(T_{ij}, T_{ij'}\right) \right)^2 \right\}$$
$$\geq \mathsf{E}\left(S_{ill''}\right)^2 \frac{1}{4} c_{G,l''}^2 \mathsf{E}\left\{ \sum_{j=1}^{N_i} B_J\left(T_{ij}\right) \right\}^4 \geq c \mathsf{E} \sum_{j=1}^{N_i} B_J^4\left(T_{ij}\right) \geq c h_{\mathrm{s}}^{-1},$$

and

$$\mathsf{E}\left(\zeta_{i,J}\right)^2 \leq \mathsf{E}\left(S_{ill''}\right)^2 4C_{G,l''}^2 \mathsf{E}\left\{\sum_{j=1}^{N_i} B_J\left(T_{ij}\right)\right\}^4$$
$$\leq c\mathsf{E}\left[N_1^3 \sum_{j=1}^{N_i} \mathsf{E}B_J^4\left(T_{ij}\right) \middle| N_1\right] \leq c\mathsf{E}N_1^4 \mathsf{E}B_J^4\left(T_{ij}\right) \leq ch_{\mathrm{s}}^{-1}.$$

Define a sequence $D_n = n^\alpha$ that satisfies $\alpha\left(2+\eta/4\right) > 1$, $D_n n^{-1/2} h_{\mathrm{s}}^{-1/2} (\log(n))^{1/2} \to 0$, $n^{1/2} h_{\mathrm{s}}^{1/2} D_n^{-(1+\eta/4)} \to 0$, which requires $\eta > 4$ provided by Assumption (A5). We make use of the following truncated and tail decomposition

$$S_{ill''} = \sum_{l'=1}^{d} \sum_{l'''=1}^{d} z_{ll'} z_{ll'''} X_{il'} X_{il'''} X_{il''}^2 = S_{ill'',1}^{D_n} + S_{ill'',2}^{D_n},$$

where

$$S_{ill'',1}^{D_n} = \sum_{l'=1}^{d} \sum_{l'''=1}^{d} z_{ll'} z_{ll'''} X_{il'} X_{il'''} X_{il''}^2 I\left\{\left|X_{il'} X_{il'''} X_{il''}^2\right| > D_n\right\},$$
$$S_{ill'',2}^{D_n} = \sum_{l'=1}^{d} \sum_{l'''=1}^{d} z_{ll'} z_{ll'''} X_{il'} X_{il'''} X_{il''}^2 I\left\{\left|X_{il'} X_{il'''} X_{il''}^2\right| \leq D_n\right\}.$$

Define correspondingly the truncated and tail parts of $\zeta_{i,J}$ as

$$\zeta_{i,J,m} = S_{ill'',m}^{D_n} \sum_{j=1}^{N_i} \sum_{j'=1}^{N_i} B_J\left(T_{ij}\right) B_J\left(T_{ij'}\right) G_{l''}\left(T_{ij}, T_{ij'}\right), \quad m = 1, 2.$$

According to Assumption (A5), for any $l', l'', l''' = 1, \ldots, d$,

$$\sum_{n=1}^{\infty} P\left\{\left|X_{nl'} X_{nl'''} X_{nl''}^2\right| > D_n\right\} \leq \sum_{n=1}^{\infty} \frac{\mathsf{E}\left|X_{nl'} X_{nl'''} X_{nl''}^2\right|^{2+\eta/4}}{D_n^{2+\eta/4}} \leq C_\eta \sum_{n=1}^{\infty} D_n^{-(2+\eta/4)} < \infty.$$

Borel–Cantelli Lemma implies that

$$P\left\{\omega \middle| \exists N(\omega), \left|X_{nl'} X_{nl'''} X_{nl''}^2(\omega)\right| \leq D_n \text{ for } n > N(\omega)\right\} = 1,$$
$$P\left\{\omega \middle| \exists N(\omega), \left|X_{il'} X_{il'''} X_{il''}^2(\omega)\right| \leq D_n, i = 1, \ldots, n \text{ for } n > N(\omega)\right\} = 1,$$
$$P\left\{\omega \middle| \exists N(\omega), I\left\{\left|X_{il'} X_{il'''} X_{il''}^2(\omega)\right| > D_n\right\} = 0, i = 1, \ldots, n \text{ for } n > N(\omega)\right\} = 1.$$

Furthermore, one has

$$n^{-1}\sum_{i=1}^{n}\left\{S_{ill'',1}^{D_n}\sum_{j=1}^{N_i}\sum_{j'=1}^{N_i}B_J\left(T_{ij}\right)B_J\left(T_{ij'}\right)G_{l''}\left(T_{ij},T_{ij'}\right)\right\}=0,\quad a.s.$$

Therefore, one has

$$\sup_{J,l,l''}\left|n^{-1}\sum_{i=1}^{n}\zeta_{i,J,1}\right|=\mathcal{O}_{a.s.}\left(n^{-k}\right),\quad k\geq 1.$$

Notice that

$$\begin{aligned}\mathsf{E}\left(S_{ill'',1}^{D_n}\right)&=\mathsf{E}\left[\sum_{l'=1}^{d}\sum_{l'''=1}^{d}z_{ll'}z_{ll'''}X_{il'}X_{il'''}X_{il''}^2 I\left\{\left|X_{il'}X_{il'''}X_{il''}^2\right|>D_n\right\}\right]\\&\leq D_n^{-(1+\eta/4)}\sum_{l'=1}^{d}\sum_{l'''=1}^{d}z_{ll'}z_{ll'''}\mathsf{E}\left|X_{il'}X_{il'''}X_{il''}^2\right|^{2+\eta/4}\\&\leq cD_n^{-(1+\eta/4)}.\end{aligned}$$

So for large n,

$$\begin{aligned}\mathsf{E}\left(\zeta_{i,J,1}\right)&=\mathsf{E}\left(S_{ill'',1}^{D_n}\right)\mathsf{E}\left\{\sum_{j=1}^{N_i}\sum_{j'=1}^{N_i}B_J\left(T_{ij}\right)B_J\left(T_{ij'}\right)G_{l''}\left(T_{ij},T_{ij'}\right)\right\}\\&\leq cD_n^{-(1+\eta/4)}2C_{G,l''}\mathsf{E}\left\{\sum_{j=1}^{N_i}B_J\left(T_{ij}\right)\right\}^2\\&\leq cD_n^{-(1+\eta/4)}\mathsf{E}\left(N_1^2\right)\mathsf{E}B_J^2\left(T_{ij}\right)\\&\leq cD_n^{-(1+\eta/4)}.\end{aligned}$$

Next we considerate the truncated part $\zeta_{i,J,2}$. For large n, $\mathsf{E}\left(\zeta_{i,J,2}\right)=\mathsf{E}\left(\zeta_{i,J}\right)-\mathsf{E}\left(\zeta_{i,J,1}\right)\sim 1$, $\mathsf{E}\left(\zeta_{i,J,2}\right)^2=\mathsf{E}\left(\zeta_{i,J}\right)^2-\mathsf{E}\left(\zeta_{i,J,1}\right)^2\sim h_{\mathrm{s}}^{-1}$. Define $\zeta_{i,J,2}^{*}=\zeta_{i,J,2}-\mathsf{E}\left(\zeta_{i,J,2}\right)$, then $\mathsf{E}\zeta_{i,J,2}^{*}=0$, and there exist c_ζ, $C_\zeta>0$ such that for $r>2$ and large n,

$$\begin{aligned}\mathsf{E}\left(\zeta_{i,J,2}^{*}\right)^2&=\mathsf{E}\left|S_{ill'',2}^{D_n}\right|^2\mathsf{E}\left|\sum_{j=1}^{N_i}\sum_{j'=1}^{N_i}B_J\left(T_{ij}\right)B_J\left(T_{ij'}\right)G_{l''}\left(T_{ij},T_{ij'}\right)\right|^2-\left(\mathsf{E}\zeta_{i,J,2}\right)^2\\&\geq\left\{\mathsf{E}\left|S_{ill''}\right|^2-\mathsf{E}\left|S_{ill'',1}^{D_n}\right|^2\right\}\frac{1}{4}c_{G,l''}^2\mathsf{E}\left\{\sum_{j=1}^{N_i}B_J\left(T_{ij}\right)\right\}^4-\mathcal{U}(1)\end{aligned}$$

$$
\begin{aligned}
&\geq \left\{\mathsf{E}\left|S_{ill''}\right|^2-\mathcal{U}(1)\right\}\frac{1}{4}c_{G,l''}^2\mathsf{E}\left\{\sum_{j=1}^{N_i}B_J^4\left(T_{ij}\right)\right\}-\mathcal{U}(1)\\
&\geq \frac{1}{2}\mathsf{E}\left|S_{ill''}\right|^2\frac{1}{4}c_{G,l''}^2\mathsf{E}N_1\mathsf{E}B_J^4\left(T_{ij}\right)-\mathcal{U}(1)\\
&\geq c_\zeta\mathsf{E}\left|S_{ill''}\right|^2h_{\mathrm{s}}^{-1},
\end{aligned}
$$

and

$$
\begin{aligned}
\mathsf{E}\left|\zeta_{i,J,2}^*\right|^r&=\mathsf{E}\left|\zeta_{i,J,2}-\mathsf{E}\left(\zeta_{i,J,2}\right)\right|^r\leq 2^{r-1}\left(\mathsf{E}\left|\zeta_{i,J,2}\right|^r+\left|\mathsf{E}\left(\zeta_{i,J,2}\right)\right|^r\right)\\
&=2^{r-1}\left\{\mathsf{E}\left|S_{ill'',2}^{D_n}\right|^r\mathsf{E}\left|\sum_{j=1}^{N_i}\sum_{j'=1}^{N_i}B_J\left(T_{ij}\right)B_J\left(T_{ij'}\right)G_{l''}\left(T_{ij},T_{ij'}\right)\right|^r+\mathcal{U}(1)\right\}\\
&\leq 2^{r-1}\left[\left(cD_n\right)^{r-2}\mathsf{E}\left|S_{ill''}\right|^2\left(2C_{G,l''}\right)^r\mathsf{E}\left\{\sum_{j=1}^{N_i}B_J(T_{ij})\right\}^{2r}+\mathcal{U}(1)\right]\\
&\leq 2^{r-1}\left[\left(cD_n\right)^{r-2}\mathsf{E}\left|S_{ill''}\right|^2\left(2C_{G,l''}\right)^r\left(\mathsf{E}N_1^{2r}\right)C_Bh_{\mathrm{s}}^{1-r}+\mathcal{U}(1)\right]\\
&\leq 2^r\left(cD_n\right)^{r-2}\left(2C_{G,l''}\right)^rc_N^rr!C_Bh_{\mathrm{s}}^{2-r}c_\zeta^{-1}\mathsf{E}\left(\zeta_{i,J,2}^*\right)^2\\
&\leq \left(C_\zeta D_nh_{\mathrm{s}}^{-1}\right)^{r-2}r!\mathsf{E}\left(\zeta_{i,J,2}^*\right)^2,
\end{aligned}
$$

which implies that $\left\{\zeta_{i,J,2}^*\right\}_{i=1}^n$ satisfies Cramér's condition. Applying Lemma 1 to $\sum_{i=1}^n\zeta_{i,J,2}^*$, for $r>2$ and any large enough $\delta>0$,

$$
\begin{aligned}
&P\left\{\left|n^{-1}\sum_{i=1}^n\zeta_{i,J,2}^*\right|\geq\delta(nh_{\mathrm{s}})^{-1/2}(\log(n))^{1/2}\right\}\\
&\quad\leq 2\exp\left\{\frac{-\delta^2\log(n)}{4+2C_\zeta D_nh_{\mathrm{s}}^{-1}\delta\left(\log(n)\right)^{1/2}n^{-1/2}h_{\mathrm{s}}^{1/2}}\right\}\leq 2n^{-8}.
\end{aligned}
$$

Hence

$$
\sum_{n=1}^\infty P\left\{\sup_{J,l'',l}\left|n^{-1}\sum_{i=1}^n\zeta_{i,J,2}^*\right|\geq\delta\left(nh_{\mathrm{s}}\right)^{-1/2}\left(\log(n)\right)^{1/2}\right\}<\infty.
$$

Thus, $\sup_{J,l'',l}\left|n^{-1}\sum_{i=1}^n\zeta_{i,J,2}^*\right|=\mathcal{O}_{a.s.}\left\{(nh_{\mathrm{s}})^{-1/2}\left(\log(n)\right)^{1/2}\right\}$ as $n\to\infty$ by the Borel–Cantelli lemma. Furthermore, we have

$$
A_{n,\xi}\leq\sup_{J,l,l''}\left|n^{-1}\sum_{i=1}^n\zeta_{i,J,1}\right|+\sup_{J,l'',l}\left|n^{-1}\sum_{i=1}^n\zeta_{i,J,2}^*\right|+\sup_{J,l'',l}\left|\mathsf{E}\left(\zeta_{i,J,1}\right)\right|
$$

$$= \mathcal{U}_{a.s.}\left(n^{-k}\right) + \mathcal{O}_{a.s.}\left\{(nh_{\mathrm{s}})^{-1/2}(\log(n))^{1/2}\right\} + \mathcal{U}\left(D_n^{-(1+\eta/4)}\right)$$
$$= \mathcal{O}_{a.s.}\left\{(nh_{\mathrm{s}})^{-1/2}(\log(n))^{1/2}\right\}.$$

The properties of $R_{ij,\varepsilon,J,l}$ are obtained similarly. □

Next define two $d \times d$ matrices

$$\Gamma_{\xi,n}(t) = c_{J(t),n}^{-1} N_{\mathrm{T}}^{-2} \sum_{l''=1}^{d} \sum_{i=1}^{n} \sum_{k=1}^{\infty} \left\{ \sum_{j=1}^{N_i} B_{J(t)}(T_{ij}) \phi_{k,l''}\left(T_{ij}\right) \right\}^2 X_{il''}^2 \mathbf{X}_i \mathbf{X}_i^{\mathsf{T}},$$
$$\Gamma_{\varepsilon,n}(t) = c_{J(t),n}^{-1} N_{\mathrm{T}}^{-2} \sum_{i=1}^{n} \sum_{j=1}^{N_i} B_{J(t)}^2(T_{ij}) \sigma^2\left(T_{ij}\right) \mathbf{X}_i \mathbf{X}_i^{\mathsf{T}}.$$

Lemma 6 *For any $t \in \mathbb{R}$, the conditional covariance matrices of $\hat{\boldsymbol{\xi}}(t)$ and $\hat{\boldsymbol{\varepsilon}}(t)$ on $\left(T_{ij}, N_i, X_{il}\right)_{i=1,j=1,l=1}^{n,N_i,d}$ are*

$$\Sigma_{\xi,n}(t) = \mathsf{E}\left\{\hat{\boldsymbol{\xi}}(t)\hat{\boldsymbol{\xi}}^{\mathsf{T}}(t) \,\middle|\, \left(T_{ij}, N_i, X_{il}\right)_{i=1,j=1,l=1}^{n,N_i,d}\right\} = \mathbf{H}^{-1}\Gamma_{\xi,n}(t)\mathbf{H}^{-1},$$
$$\Sigma_{\varepsilon,n}(t) = \mathsf{E}\left\{\hat{\boldsymbol{\varepsilon}}(t)\hat{\boldsymbol{\varepsilon}}^{\mathsf{T}}(t) \,\middle|\, \left(T_{ij}, N_i, X_{il}\right)_{i=1,j=1,l=1}^{n,N_i,d}\right\} = \mathbf{H}^{-1}\Gamma_{\varepsilon,n}(t)\mathbf{H}^{-1},$$

and with $\Sigma_n(t)$ defined in (7),

$$\sup_{t\in[0,1]} \left\| \left\{\Sigma_{\xi,n}(t) + \Sigma_{\varepsilon,n}(t)\right\} - \Sigma_n(t) \right\|_\infty = \mathcal{O}_{a.s.}\left\{n^{-3/2} h_{\mathrm{s}}^{-3/2} (\log(n))^{1/2}\right\}. \quad (33)$$

Proof Note that

$$\hat{\boldsymbol{\xi}}(t)\hat{\boldsymbol{\xi}}^{\mathsf{T}}(t) = c_{J(t),n}^{-1}\mathbf{H}^{-1}\left\{\frac{1}{N_{\mathrm{T}}^2}\sum_{i=1}^{n}\sum_{j=1}^{N_i} B_{J(t)}(T_{ij}) X_{il} \sum_{l''=1}^{d}\sum_{k=1}^{\infty} \xi_{ik,l''}\phi_{k,l''}\left(T_{ij}\right) X_{il''}\right.$$
$$\left.\times \sum_{i=1}^{n}\sum_{j=1}^{N_i} B_{J(t)}(T_{ij}) X_{il'} \sum_{l''=1}^{d}\sum_{k=1}^{\infty} \xi_{ik,l''}\phi_{k,l''}\left(T_{ij}\right) X_{il''}\right\}_{l,l'=1}^{d} \mathbf{H}^{-1}.$$

Thus,

$$\Sigma_{\xi,n}(t) = \mathsf{E}\left\{\hat{\boldsymbol{\xi}}(t)\hat{\boldsymbol{\xi}}^{\mathsf{T}}(t) \,\middle|\, \left(T_{ij}, N_i, X_{il}\right)_{i=1,j=1,l=1}^{n,N_i,d}\right\} = c_{J(t),n}^{-1}\mathbf{H}^{-1}$$
$$\times \left[N_{\mathrm{T}}^{-2} \sum_{l''=1}^{d}\sum_{i=1}^{n}\sum_{k=1}^{\infty} \left\{\sum_{j=1}^{N_i} B_{J(t)}(T_{ij})\phi_{k,l''}\left(T_{ij}\right)\right\}^2 X_{il''}^2 \mathbf{X}_i\mathbf{X}_i^{\mathsf{T}} \right]\mathbf{H}^{-1}$$
$$= \mathbf{H}^{-1}\Gamma_{\xi,n}(t)\mathbf{H}^{-1}.$$

Similarly, we can derive the conditional covariance matrix of $\hat{\boldsymbol{\varepsilon}}(t)$. Next let

$$\Psi_{ik,\xi,J,l,l',l''} = \left\{\sum_{j=1}^{N_i} B_J(T_{ij})\phi_{k,l''}\left(T_{ij}\right)\right\}^2 X_{il''}^2 X_{il} X_{il'},$$
$$\Psi_{ij,\varepsilon,J,l,l'} = B_J^2(T_{ij})\sigma^2\left(T_{ij}\right) X_{il} X_{il'}.$$

Similar to the proof of Lemma 5,

$$\mathsf{E}\left(\sum_{k=1}^{\infty}\Psi_{ik,\xi,J,l,l',l''}\right) = c_{J,n}^{-1}\mathsf{E}\left(X_{il''}^2 X_{il} X_{il'}\right)\left[(\mathsf{E}N_1)\int_{\chi_J} G_{l''}(u,u)\, f(u)\,\mathrm{d}u\right.$$
$$\left. + \mathsf{E}\left\{N_1(N_1-1)\right\}\int_{\chi_J\times\chi_J} G_{l''}(u,v)\, f(u)\, f(v)\,\mathrm{d}u\mathrm{d}v\right],$$
$$\mathsf{E}\Psi_{ij,\varepsilon,J,l,l'} = c_{J,n}^{-1}\mathsf{E}\left(X_{il}X_{il'}\right)\int_{\chi_J}\sigma^2(u)\, f(u)\,\mathrm{d}u,$$

and as $n\to\infty$,

$$\sup_{J,l,l',l''}\left| n^{-1}\sum_{i=1}^{n}\sum_{k=1}^{\infty}\Psi_{ik,\xi,J,l,l',l''} - \mathsf{E}\left(\sum_{k=1}^{\infty}\Psi_{ik,\xi,J,l,l',l''}\right)\right|$$
$$= \mathcal{O}_{a.s.}\left\{(nh_{\mathrm{s}})^{-1/2}(\log(n))^{1/2}\right\},$$

$$\sup_{J,l,l'}\left| N_{\mathrm{T}}^{-1}\sum_{i=1}^{n}\sum_{j=1}^{N_i}\Psi_{ij,\varepsilon,J,l,l'} - \mathsf{E}\Psi_{ij,\varepsilon,J,l,l'}\right| = \mathcal{O}_{a.s.}\left\{(nh_{\mathrm{s}})^{-1/2}(\log(n))^{1/2}\right\}.$$

Furthermore,

$$\sup_{J,l,l',l''}\left| N_T^{-2}\sum_{i=1}^{n}\sum_{k=1}^{\infty}\Psi_{ik,\xi,J,l,l',l''} - n^{-1}(\mathsf{E}N_1)^{-2}\mathsf{E}\left(\sum_{k=1}^{\infty}\Psi_{ik,\xi,J,l,l',l''}\right)\right|$$
$$\leq \sup_{J,l,l',l''} n^{-1}(\mathsf{E}N_1)^{-2}\left\{\left|\left(\frac{n\mathsf{E}N_1}{N_T}\right)^2 - 1\right|\left| n^{-1}\sum_{i=1}^{n}\sum_{k=1}^{\infty}\Psi_{ik,\xi,J,l,l',l''}\right|\right.$$
$$\left. + \left| n^{-1}\sum_{i=1}^{n}\sum_{k=1}^{\infty}\Psi_{ik,\xi,J,l,l',l''} - \mathsf{E}\left(\sum_{k=1}^{\infty}\Psi_{ik,\xi,J,l,l',l''}\right)\right|\right\}$$
$$= \mathcal{O}_{a.s.}\left\{n^{-3/2}h_{\mathrm{s}}^{-1/2}(\log(n))^{1/2}\right\},$$

and

$$\sup_{J,l,l'}\left|N_T^{-2}\sum_{i=1}^{n}\sum_{j=1}^{N_i}\Psi_{ik,\varepsilon,J,l,l'}-(n\mathsf{E}N_1)^{-1}\mathsf{E}\Psi_{ik,\varepsilon,J,l,l'}\right|$$
$$\leq \sup_{J,l,l'}(n\mathsf{E}N_1)^{-1}\left\{\left|\frac{n\mathsf{E}N_1}{N_T}-1\right|\left|N_T^{-1}\sum_{i=1}^{n}\sum_{j=1}^{N_i}\Psi_{ik,\varepsilon,J,l,l'}\right|\right.$$
$$\left.+\left|N_T^{-1}\sum_{i=1}^{n}\sum_{j=1}^{N_i}\Psi_{ik,\varepsilon,J,l,l'}-\mathsf{E}\Psi_{ik,\varepsilon,J,l,l'}\right|\right\}$$
$$=\mathcal{O}_{a.s.}\left\{n^{-3/2}h_{\mathrm{s}}^{-1/2}(\log(n))^{1/2}\right\}.$$

Notice that

$$\Sigma_n(t)=\mathbf{H}^{-1}c_{J(t),n}^{-1}(n\mathsf{E}N_1)^{-1}$$
$$\times\left\{(\mathsf{E}N_1)^{-1}\mathsf{E}\left(\sum_{l''=1}^{d}\sum_{k=1}^{\infty}\Psi_{ik,\xi,J(t),l,l',l''}\right)+\mathsf{E}\Psi_{ij,\varepsilon,J(t),l,l'}\right\}_{l,l'=1}^{d}\mathbf{H}^{-1},$$

$$\Sigma_{\xi,n}(t)+\Sigma_{\varepsilon,n}(t)$$
$$=\mathbf{H}^{-1}c_{J(t),n}^{-1}N_{\mathrm{T}}^{-2}\left\{\sum_{l''=1}^{d}\sum_{i=1}^{n}\sum_{k=1}^{\infty}\Psi_{ik,\xi,J(t),l,l',l''}+\sum_{i=1}^{n}\sum_{j=1}^{N_i}\Psi_{ij,\varepsilon,J(t),l,l'}\right\}_{l,l'=1}^{d}\mathbf{H}^{-1},$$

and (35) implies $\sup_{t\in[0,1]}\left|c_{J(t),n}\right|=\mathcal{O}(h_{\mathrm{s}})$. Hence (33) holds. □

Given $\left(T_{ij},N_i,X_{il}\right)_{i=1,j=1,l=1}^{n,N_i,d}$, let $\sigma_{\xi_l,n}^2(t)$ and $\sigma_{\varepsilon_l,n}^2(t)$ be the conditional variances of $\hat{\xi}_l(t)$ and $\hat{\varepsilon}_l(t)$ defined in (38) and (39), respectively. Lemma 6 implies that

$$\sup_{t\in[0,1]}\left|\sigma_{\xi_l,n}^2(t)+\sigma_{\varepsilon_l,n}^2(t)-\sigma_{n,ll}^2(t)\right|=\mathcal{O}_{a.s.}\left\{n^{-3/2}h_{\mathrm{s}}^{-3/2}(\log(n))^{1/2}\right\}.\quad(34)$$

Lemma 7 *Under Assumptions* (A2)–(A6)*, for* $l=1,\ldots,d$*,* $\eta_l(t)$ *defined in* (40) *is a Gaussian process consisting of* $(N_{\mathrm{s}}+1)$ *standard normal variables* $\left\{\eta_{J,l}\right\}_{J=0}^{N_{\mathrm{s}}}$ *such that* $\eta_l(t)=\eta_{J(t),l}$ *for* $t\in[0,1]$*, and there exists a constant* $C>0$ *such that for large n,* $\sup_{0\leq J\neq J'\leq N_{\mathrm{s}}}\left|\mathsf{E}\eta_{J,l}\eta_{J',l}\right|\leq Ch_{\mathrm{s}}$.

Proof For any fixed $l=1,\ldots,d$ and $0\leq J\leq N_{\mathrm{s}}$, $\mathcal{L}\left\{\eta_{J,l}\left|\left(T_{ij},N_i,X_{il}\right)_{i=1,j=1,l=1}^{n,N_i,d}\right.\right\}$ $=N(0,1)$ by Assumption (A2), so $\mathcal{L}\left\{\eta_{J,l}\right\}=N(0,1)$, for $0\leq J\leq N_{\mathrm{s}}$.

Next we derive the upper bound for $\sup_{0\leq J\neq J'\leq N_{\mathrm{s}}}\left|\mathsf{E}\eta_{J,l}\eta_{J',l}\right|$. Let

$$\bar{R}_{\xi,J(t),l}=N_{\mathrm{T}}^{-1}\sum_{l''=1}^{d}\sum_{i=1}^{n}\sum_{k=1}^{\infty}R_{ik,\xi,J(t),l'',l}^2,\quad \bar{R}_{\varepsilon,J(t),l}=N_{\mathrm{T}}^{-1}\sum_{i=1}^{n}\sum_{j=1}^{N_i}R_{ij,\varepsilon,J(t),l}^2,$$

then we have

$$\sigma_{\xi_l,n}(t)=\left\{c_{J(t),n}^{-1}N_{\mathrm{T}}^{-2}\sum_{l''=1}^{d}\sum_{i=1}^{n}\sum_{k=1}^{\infty}R_{ik,\xi,J(t),l'',l}^{2}\right\}^{1/2}=\left\{c_{J(t),n}^{-1}N_{\mathrm{T}}^{-1}\bar{R}_{\xi,J(t),l}\right\}^{1/2},$$

$$\sigma_{\varepsilon_l,n}(t)=\left\{c_{J(t),n}^{-1}N_{\mathrm{T}}^{-2}\sum_{i=1}^{n}\sum_{j=1}^{N_i}R_{ij,\varepsilon,J(t),l}^{2}\right\}^{1/2}=\left\{c_{J(t),n}^{-1}N_{\mathrm{T}}^{-1}\bar{R}_{\varepsilon,J(t),l}\right\}^{1/2}.$$

For $J\neq J'$, by (31) and the definition of B_J,

$$R_{ij,\varepsilon,J,l}R_{ij,\varepsilon,J',l}=\left(\sum_{l'=1}^{d}z_{ll'}X_{il'}\right)^{2}B_J\left(T_{ij}\right)B_{J'}\left(T_{ij}\right)\sigma^2\left(T_{ij}\right)=0,$$

along with the conditional independence of $\hat{\xi}_l(t),\hat{\varepsilon}_l(t)$ on $\left(T_{ij},N_i,X_{il}\right)_{i=1,j=1,l=1}^{n,N_i,d}$, and independence of $\xi_{ik,l}$, T_{ij}, N_i, $\{X_{il}\}_{l=1}^{d}$, $1\leq j\leq N_i$, $1\leq i\leq n$, $k=1,2,\ldots,$

$$\begin{aligned}\mathsf{E}\left(\eta_{J,l}\eta_{J',l}\right)&=\mathsf{E}\Big[(\bar{R}_{\xi,J,l}+\bar{R}_{\varepsilon,J,l})^{-1/2}(\bar{R}_{\xi,J',l}+\bar{R}_{\varepsilon,J',l})^{-1/2}\\&\quad\times N_{\mathrm{T}}^{-1}\mathsf{E}\Big\{\Big(\sum_{l''=1}^{d}\sum_{i=1}^{n}\sum_{k=1}^{\infty}R_{ik,\xi,J,l'',l}\xi_{ik,l''}\Big)\Big(\sum_{l''=1}^{d}\sum_{i=1}^{n}\sum_{k=1}^{\infty}R_{ik,\xi,J',l'',l}\xi_{ik,l''}\Big)\\&\quad+\left(\sum_{i=1}^{n}\sum_{j=1}^{N_i}R_{ij,\varepsilon,J,l}\varepsilon_{ij}\right)\left(\sum_{i=1}^{n}\sum_{j=1}^{N_i}R_{ij,\varepsilon,J',l}\varepsilon_{ij}\right)\Big|\left(T_{ij},N_i,X_{il}\right)_{i=1,j=1,l=1}^{n,N_i,d}\Big\}\Big]\\&=\mathsf{E}C_{n,J,J',l},\end{aligned}$$

in which

$$\begin{aligned}C_{n,J,J',l}&=(\bar{R}_{\xi,J,l}+\bar{R}_{\varepsilon,J,l})^{-1/2}(\bar{R}_{\xi,J',l}+\bar{R}_{\varepsilon,J',l})^{-1/2}\\&\quad\times\left\{N_{\mathrm{T}}^{-1}\sum_{l''=1}^{d}\sum_{i=1}^{n}\sum_{k=1}^{\infty}R_{ik,\xi,J,l'',l}R_{ik,\xi,J',l'',l}\right\}.\end{aligned}$$

Note that according to definitions of $R_{ik,\xi,J,l'',l}$, $R_{ij,\varepsilon,J,l}$, and Lemma 5, for $0\leq J\leq N_{\mathrm{s}}$

$$\bar{R}_{\xi,J(t),l}+\bar{R}_{\varepsilon,J(t),l}\geq\bar{R}_{\varepsilon,J(t),l}\geq ER_{ij,\varepsilon,J,l}^{2}-A_{n,\varepsilon}\geq c_R-A_{n,\varepsilon},$$

$$P\left[\inf_{0\leq J\neq J'\leq N_{\mathrm{s}}}\left\{(\bar{R}_{\xi,J,l}+\bar{R}_{\varepsilon,J,l})(\bar{R}_{\xi,J',l}+\bar{R}_{\varepsilon,J',l})\right\}\geq\left(c_R-\delta\sqrt{\frac{\log(n)}{nh_{\mathrm{s}}}}\right)^{2}\right]\geq 1-2n^{-8}.$$

Thus for large n, with probability $\geq 1-2n^{-8}$, the denominator of $C_{n,J,J',l}$ is uniformly greater than $c_R^2/4$. On the other hand, we consider the numerator of $C_{n,J,J',l}$.

$$\mathsf{E}\left(N_{\mathrm{T}}^{-1}\sum_{l''=1}^{d}\sum_{i=1}^{n}\sum_{k=1}^{\infty}R_{ik,\xi,J,l'',l}R_{ik,\xi,J',l'',l}\right)=\mathsf{E}\left\{N_{\mathrm{T}}^{-1}\sum_{l''=1}^{d}\sum_{i=1}^{n}\left(\sum_{l'=1}^{d}z_{ll'}X_{il'}X_{il''}\right)^2\right.$$
$$\left.\times\left(\sum_{j=1}^{N_i}\sum_{j'=1}^{N_i}B_J\left(T_{ij}\right)B_{J'}\left(T_{ij'}\right)G_{l''}\left(T_{ij},T_{ij'}\right)\right)\right\}\sim h_{\mathrm{s}}.$$

Applying Bernstein's inequality, there exists $C_0>0$ such that, for large n,

$$P\left(\sup_{0\leq J\neq J'\leq N_{\mathrm{s}}}\left|N_{\mathrm{T}}^{-1}\sum_{l''=1}^{d}\sum_{i=1}^{n}\sum_{k=1}^{\infty}R_{ik,\xi,J,l'',l}R_{ik,\xi,J',l'',l}\right|\leq C_0h_{\mathrm{s}}\right)\geq 1-2n^{-8}.$$

Putting the above together, for large n, $C_1=C_0\left(c_R^2/4\right)^{-1}$,

$$P\left(\sup_{0\leq J\neq J'\leq N_{\mathrm{s}}}\left|C_{n,J,J',l}\right|\leq C_1h_{\mathrm{s}}\right)\geq 1-4n^{-8}.$$

Note that as a continuous random variable, $\sup_{0\leq J\neq J'\leq N_{\mathrm{s}}}\left|C_{n,J,J',l}\right|\in[0,1]$, thus

$$\mathsf{E}\left(\sup_{0\leq J\neq J'\leq N_{\mathrm{s}}}\left|C_{n,J,J',l}\right|\right)=\int_0^1P\left(\sup_{0\leq J\neq J'\leq N_{\mathrm{s}}}\left|C_{n,J,J',l}\right|>u\right)\mathrm{d}u.$$

For large n, $C_1h_{\mathrm{s}}<1$ and then $\mathsf{E}\left(\sup_{0\leq J\neq J'\leq N_{\mathrm{s}},l}\left|C_{n,J,J'}\right|\right)$ is

$$\int_0^{C_1h_{\mathrm{s}}}P\left\{\sup_{0\leq J\neq J'\leq N_{\mathrm{s}},l}\left|C_{n,J,J',l}\right|>u\right\}\mathrm{d}u+\int_{C_1h_{\mathrm{s}}}^1P\left\{\sup_{0\leq J\neq J'\leq N_{\mathrm{s}},l}\left|C_{n,J,J',l}\right|>u\right\}\mathrm{d}u$$
$$\leq\int_0^{C_1h_{\mathrm{s}}}1\mathrm{d}u+\int_{C_1h_{\mathrm{s}}}^1 4n^{-8}\mathrm{d}u\leq C_1h_{\mathrm{s}}+4n^{-8}\leq Ch_{\mathrm{s}}$$

for some $C>0$ and large enough n. The lemma now follows from

$$\sup_{0\leq J\neq J'\leq N_{\mathrm{s}}}\left|\mathsf{E}\left(C_{n,J,J',l}\right)\right|\leq\mathsf{E}\left(\sup_{0\leq J\neq J'\leq N_{\mathrm{s}}}\left|C_{n,J,J',l}\right|\right)\leq Ch_{\mathrm{s}}.$$

This completes the proof of the lemma. □

Lemma 8 *Under Assumptions* (A2)–(A6)*, for* $\eta_l(t)$, $\sigma_{n,ll}(t)$, $l = 1, \ldots, d$*, defined in* (40) *and* (7)*, one has* $\left|\sigma_{n,ll}(t)^{-1}\left\{\hat{\xi}_l(t) + \hat{\varepsilon}_l(t)\right\} - \eta_l(t)\right| = \left|r_{n,l}(t) - 1\right| |\eta_l(t)|$*, where* $r_{n,l}(t) = \sigma_{n,ll}^{-1}(t)\left\{\sigma_{\xi_l,n}^2(t) + \sigma_{\varepsilon_l,n}^2(t)\right\}^{1/2}$*, and as* $n \to \infty$*,*

$$\sup_{t\in[0,1]}\left\{a_{N_s+1}\left|r_{n,l}(t) - 1\right|\right\} = \mathcal{O}_{a.s.}\left\{(nh_s)^{-1/2}\left(\log\left(N_s+1\right)\log(n)\right)^{1/2}\right\}.$$

Proof By Lemma 5, $\sigma_{n,ll}^2(t)$ in (7) can be rewritten as

$$\begin{aligned}\sigma_{n,ll}^2(t) &= c_{J(t),n}^{-1}\left(n\mathsf{E}N_1\right)^{-1}\left\{(\mathsf{E}N_1)^{-1}\sum_{l''=1}^{d}\mathsf{E}\left(\sum_{k=1}^{\infty}R_{ik,\xi,J(t),l'',l}^2\right) + \mathsf{E}R_{ij,\varepsilon,J(t),l}^2\right\}\\ &\sim n^{-1}h_s^{-1}.\end{aligned}$$

Hence, according to (34) and (10),

$$\begin{aligned}\sup_{t\in[0,1]}\left\{a_{N_s+1}\left|r_{n,l}(t) - 1\right|\right\} &= \sup_{t\in[0,1]}\left\{a_{N_s+1}\left|\sigma_{n,ll}^{-1}(t)\left\{\sigma_{\xi_l,n}^2(t) + \sigma_{\varepsilon_l,n}^2(t)\right\}^{1/2} - 1\right|\right\}\\ &\le \sup_{t\in[0,1]}\left\{a_{N_s+1}\left|\sigma_{n,ll}^{-2}(t)\left\{\sigma_{\xi_l,n}^2(t) + \sigma_{\varepsilon_l,n}^2(t)\right\} - 1\right|\right\}\\ &= \sup_{t\in[0,1]}\left\{a_{N_s+1}\sigma_{n,ll}^{-2}(t)\left|\sigma_{\xi_l,n}^2(t) + \sigma_{\varepsilon_l,n}^2(t) - \sigma_{n,ll}^2(t)\right|\right\}\\ &= \mathcal{O}_{a.s.}\left\{(nh_s)^{-1/2}\left(\log\left(N_s+1\right)\log(n)\right)^{1/2}\right\}.\end{aligned}$$

This completes the proof. □

A.3 Proofs of Propositions 1– 4

Proof of Proposition 1 By Assumption (A3) on the continuity of functions $\phi_{k,l}(t)$, $\sigma^2(t)$ and $f(t)$ on [0, 1] and Assumption (A4), for any $t, u \in [0, 1]$ satisfying $|t - u| \le h_s$,

$$|G_l(t,t) - G_l(u,u)| \le \sum_{k=1}^{\infty}\left|\phi_{k,l}^2(t) - \phi_{k,l}^2(u)\right| \le 2\sum_{k=1}^{\infty}\left\|\phi_{k,l}\right\|_{\infty}\omega\left(\phi_{k,l}, h_s\right) \le Ch_s^r.$$

Furthermore,

$$\left|\int_{\chi_{J(t)}}\{G_l(t,t)f(t) - G_l(u,u)f(u)\}\,\mathrm{d}u\right| \le Ch_s^{1+r} = \mathcal{O}\left(h_s^{1+r}\right),$$

$$\left|\int_{\chi_{J(t)}\times\chi_{J(t)}}\left\{G_l(t,t)f^2(t) - G_l(u,v)f(u)f(v)\right\}\mathrm{d}u\mathrm{d}v\right| \le Ch_s^{2+r} = \mathcal{O}\left(h_s^{2+r}\right),$$

$$\left|\int_{\chi_{J(t)}}\left\{\sigma^2(t)f(t)-\sigma^2(u)\,f(u)\right\}\mathrm{d}u\right|\le Ch_{\mathrm{s}}^{1+r}=\mathcal{O}\left(h_{\mathrm{s}}^{1+r}\right).$$

According to the definition of $C_{J,n}$ in (6),

$$C_{J,n}=\int_{[\upsilon_J,\upsilon_{J+1}]}f(x)\mathrm{d}x=f(\upsilon_J)h_{\mathrm{s}}+\int_{[\upsilon_J,\upsilon_{J+1}]}\{f(x)-f(\upsilon_J)\}\mathrm{d}x, \tag{35}$$

thus, $|C_{J,n}-f(\upsilon_J)h_{\mathrm{s}}|\le w(f,h_{\mathrm{s}})h_{\mathrm{s}}$ for all $J=0,\ldots,N_{\mathrm{s}}$. Therefore,

$$\begin{aligned}\Gamma_n(t)&=\left\{f(t)h_{\mathrm{s}}+\mathcal{U}\left(h_{\mathrm{s}}^{1+r}\right)\right\}^{-2}(n\mathsf{E}N_1)^{-1}\,\mathsf{E}\left[\left\{\sigma_Y^2(t,\mathbf{X})\,f(t)\,h_{\mathrm{s}}+\mathcal{U}_p\left(h_{\mathrm{s}}^{1+r}\right)\right.\right.\\&\quad\left.\left.+\frac{\mathsf{E}\{N_1(N_1-1)\}}{\mathsf{E}N_1}\sum_{l=1}^{d}X_l^2G_l(t,t)\,f^2(t)\,h_{\mathrm{s}}^2+\mathcal{U}_p\left(h_{\mathrm{s}}^{2+r}\right)\right\}\mathbf{XX}^{\mathsf{T}}\right]\\&=\mathsf{E}\left[\mathbf{XX}^{\mathsf{T}}\sigma_Y^2(t,\mathbf{X})\left\{f(t)h_{\mathrm{s}}n\mathsf{E}N_1\right\}^{-1}\left\{1+\frac{\mathsf{E}\{N_1(N_1-1)\}}{\mathsf{E}N_1}\right.\right.\\&\quad\left.\left.\times\frac{\sum_{l=1}^{d}X_l^2G_l(t,t)\,f(t)\,h_{\mathrm{s}}}{\sigma_Y^2(t,\mathbf{X})}\right\}\left\{1+\mathcal{U}_p\left(h_{\mathrm{s}}^r\right)\right\}\right]=\tilde{\Gamma}_n(t)+\mathcal{U}\left(n^{-1}h_{\mathrm{s}}^{r-1}\right),\end{aligned}$$

establishing the proposition. □

Proof of Proposition 2 The result follows from standard theory of kernel and spline smoothing, as in Wang and Yang (2009), thus omitted. □

Proof of Proposition 3 According to the result on page 149 of de Boor (2001), there exist functions $g_l\in G^{(-1)}[0,1]$ that satisfies $\|m_l-g_l\|_\infty=\mathcal{O}(h_{\mathrm{s}})$ for $l=1,\ldots,d$. By the definition of $\tilde{m}_l(t)$ in (22),

$$\tilde{\boldsymbol{m}}(t)=(\tilde{m}_1(t),\ldots,\tilde{m}_d(t))^{\mathsf{T}}=c_{J(t),n}^{-1/2}\left(\tilde{\gamma}_{J(t),1},\ldots,\tilde{\gamma}_{J(t),d}\right)^{\mathsf{T}}=c_{J(t),n}^{-1/2}\tilde{\gamma}_{J(t)},$$

where $\tilde{\gamma}_J=\hat{\mathbf{V}}_J^{-1}\left\{N_{\mathrm{T}}^{-1}\sum_{i=1}^{n}\sum_{j=1}^{N_i}B_J(T_{ij})X_{il}\sum_{l'=1}^{d}m_{l'}(T_{ij})X_{il'}\right\}_{l=1}^{d}$ for $\hat{\mathbf{V}}_J$ defined in (18).

Let $\tilde{\boldsymbol{g}}(t)=(\tilde{g}_1(t),\ldots,\tilde{g}_d(t))^{\mathsf{T}}$, then $\tilde{\mathbf{m}}_l(t)-\tilde{\mathbf{g}}_l(t)$ equals to

$$c_{J(t),n}^{-1/2}\hat{\mathbf{V}}_{J(t)}^{-1}\left[\frac{1}{N_{\mathrm{T}}}\sum_{i=1}^{n}\sum_{j=1}^{N_i}B_{J(t)}(T_{ij})X_{il}\sum_{l'=1}^{d}\left\{m_{l'}(T_{ij})-g_{l'}(T_{ij})\right\}X_{il'}\right]_{l=1}^{d}.$$

Observing that $\tilde{g}_l\equiv g_l$ as $g_l\in G^{(-1)}[0,1]$, there is a decomposition similar to (24), $\tilde{m}_l(t)=\tilde{m}_l(t)-\tilde{g}_l(t)+g_l(t)$, $l=1,\ldots,d$.

By (35), $\sup_{t\in[0,1]}\left|c_{J(t),n}\right| = \mathcal{O}\left(h_{\mathrm{s}}\right)$. Next $\mathsf{E}|B_J(T_{ij})| = c_{J,n}^{-1/2}\int b_J(x)f(x)\mathrm{d}x \sim h_{\mathrm{s}}^{1/2}$, thus $\sup_{t\in[0,1]}\left|B_{J(t)}(T_{ij})\right| = \mathcal{O}_p(h_{\mathrm{s}}^{1/2})$. Then it is easy to show that $\|\tilde{m}_l - \tilde{g}_l\|_\infty = \mathcal{O}_p\left(h_{\mathrm{s}}^{-1/2}h_{\mathrm{s}}^{1/2}h_{\mathrm{s}}\right) = \mathcal{O}_p\left(h_{\mathrm{s}}\right)$. Hence, for $l = 1,\ldots,d$,

$$\|\tilde{m}_l - m_l\|_\infty \le \|\tilde{m}_l - \tilde{g}_l\|_\infty + \|m_l - g_l\|_\infty = \mathcal{O}_p\left(h_{\mathrm{s}}\right),$$

which completes the proof. □

Note that $B_J(t) \equiv b_J c_{J,n}^{-1/2}$, $t \in [0,1]$, so the terms $\tilde{\xi}_l(t)$ and $\tilde{\varepsilon}_l(t)$, $l = 1,\ldots,d$, defined in (23) are

$$\tilde{\boldsymbol{\xi}}(t) = \left(\tilde{\xi}_1(t),\ldots,\tilde{\xi}_d(t)\right)^{\mathsf{T}} = c_{J(t),n}^{-1/2}\left(\tilde{\alpha}_{J(t),1},\ldots,\tilde{\alpha}_{J(t),d}\right)^{\mathsf{T}} = c_{J(t),n}^{-1/2}\tilde{\boldsymbol{\alpha}}_{J(t)}, \tag{36}$$

$$\tilde{\boldsymbol{\varepsilon}}(t) = \left(\tilde{\varepsilon}_1(t),\ldots,\tilde{\varepsilon}_d(t)\right)^{\mathsf{T}} = c_{J(t),n}^{-1/2}\left(\tilde{\theta}_{J(t),1},\ldots,\tilde{\theta}_{J(t),d}\right)^{\mathsf{T}} = c_{J(t),n}^{-1/2}\tilde{\boldsymbol{\theta}}_{J(t)}, \tag{37}$$

where

$$\tilde{\alpha}_J = \hat{\mathbf{V}}_J^{-1}\left\{N_{\mathrm{T}}^{-1}\sum_{i=1}^{n}\sum_{j=1}^{N_i}B_J(T_{ij})X_{il}\sum_{l''=1}^{d}\sum_{k=1}^{\infty}\xi_{ik,l''}\phi_{k,l''}\left(T_{ij}\right)X_{il''}\right\}_{l=1}^{d},$$

$$\tilde{\theta}_J = \hat{\mathbf{V}}_J^{-1}\left\{N_{\mathrm{T}}^{-1}\sum_{i=1}^{n}\sum_{j=1}^{N_i}B_J(T_{ij})X_{il}\sigma\left(T_{ij}\right)\varepsilon_{ij}\right\}_{l=1}^{d}.$$

According to Lemma 3, the inverse of the random matrix $\hat{\mathbf{V}}_J$ can be approximated by that of a deterministic matrix $\mathbf{H} = \mathsf{E}(\mathbf{X}\mathbf{X}^{\mathsf{T}})$. Substituting $\hat{\mathbf{V}}_J$ with $\mathbf{H}$ in (36) and (37), we define the random vectors

$$\hat{\boldsymbol{\xi}}(t) = c_{J(t),n}^{-1/2}\mathbf{H}^{-1}\left\{\frac{1}{N_{\mathrm{T}}}\sum_{i=1}^{n}\sum_{j=1}^{N_i}B_{J(t)}(T_{ij})X_{il}\sum_{l''=1}^{d}\sum_{k=1}^{\infty}\xi_{ik,l''}\phi_{k,l''}\left(T_{ij}\right)X_{il''}\right\}_{l=1}^{d}, \tag{38}$$

$$\hat{\boldsymbol{\varepsilon}}(t) = c_{J(t),n}^{-1/2}\mathbf{H}^{-1}\left\{\frac{1}{N_{\mathrm{T}}}\sum_{i=1}^{n}\sum_{j=1}^{N_i}B_{J(t)}(T_{ij})X_{il}\sigma\left(T_{ij}\right)\varepsilon_{ij}\right\}_{l=1}^{d}. \tag{39}$$

Proof of Proposition 4 Given $\left(T_{ij}, N_i, X_{il}\right)_{i=1,j=1,l=1}^{n,N_i,d}$, let $\sigma_{\xi_l,n}^2(t)$ and $\sigma_{\varepsilon_l,n}^2(t)$ be the conditional variances of $\hat{\xi}_l(t)$ and $\hat{\varepsilon}_l(t)$ defined in (38) and (39), respectively. Define

$$\eta_l(t) = \left\{\sigma_{\xi_l,n}^2(t) + \sigma_{\varepsilon_l,n}^2(t)\right\}^{-1/2}\left\{\hat{\xi}_l(t) + \hat{\varepsilon}_l(t)\right\}. \tag{40}$$

By Lemma 7, $\eta_l(t)$ is a Gaussian process consisting of (N_s+1) standard normal variables $\left\{\eta_{J,l}\right\}_{J=0}^{N_s}$ such that $\eta_l(t) = \eta_{J(t),l}$ for $t \in [0,1]$. Thus, for any $\tau \in \mathbb{R}$,

$$P\left(\sup_{t\in[0,1]} |\eta_l(t)| \le \tau/a_{N_s+1} + b_{N_s+1}\right) \\ = P\left(|\max\{\eta_{0,l}, \ldots, \eta_{N_s,l}\}| \le \tau/a_{N_s+1} + b_{N_s+1}\right).$$

By Theorem 1.5.3 in Leadbetter et al. (1983), if $\xi_0, \ldots, \xi_{N_s}$ are i.i.d. standard normal r.v.'s, then for $\tau \in \mathbb{R}$

$$P\left(|\max\{\xi_0, \ldots, \xi_{N_s}\}| \le \tau/a_{N_s} + b_{N_s}\right) \to \exp(-2e^{-\tau}).$$

Next by Lemma 11.1.2 in Leadbetter et al. (1983),

$$P\left(|\max\{\eta_{0,l}, \ldots, \eta_{N_s,l}\}| \le \tau/a_{N_s+1} + b_{N_s+1}\right) \\ -P\left(|\max\{\xi_0, \ldots, \xi_{N_s}\}| \le \tau/a_{N_s+1} + b_{N_s+1}\right) \\ \le \frac{4}{2\pi} \sum_{0\le J<J'\le N_s} |\mathsf{E}\eta_{J,l}\eta_{J',l}|(1-|\mathsf{E}\eta_{J,l}\eta_{J',l}|^2)^{-1/2} \exp\left\{\frac{-(\tau/a_{N_s+1}+b_{N_s+1})^2}{1+\mathsf{E}\eta_{J,l}\eta_{J',l}}\right\}.$$

According to Lemma 7, there exists a constant $C > 0$ such that $\sup_{0\le J\ne J'\le N_s} \left|\mathsf{E}\eta_{J,l}\eta_{J',l}\right| \le Ch_s$ for large n. Thus, as $n\to\infty$,

$$P\left(|\max\{\eta_{0,l}, \ldots, \eta_{N_s,l}\}| \le \tau/a_{N_s+1} + b_{N_s+1}\right) \\ -P\left(|\max\{\xi_0, \ldots, \xi_{N_s}\}| \le \tau/a_{N_s+1} + b_{N_s+1}\right) \to 0.$$

Therefore, for any $\tau \in \mathbb{R}$,

$$\lim_{n\to\infty} P\left(\sup_{t\in[0,1]} |\eta_l(t)| \le \tau/a_{N_s+1} + b_{N_s+1}\right) = \exp\left(-2e^{-\tau}\right). \tag{41}$$

By Lemma 8, we have

$$a_{N_s+1}\left(\sup_{t\in[0,1]} \sigma_{n,ll}^{-1}(t)\left|\hat{\xi}_l(t)+\hat{\varepsilon}_l(t)\right| - \sup_{t\in[0,1]}|\eta_l(t)|\right) \\ = \mathcal{O}_p\left\{\log(N_s+1)(nh_s)^{-1/2}(\log(n))^{1/2}\right\} = \mathcal{O}_p(1).$$

On the other hand, Lemma 4 ensures that

$$
\begin{aligned}
& a_{N_s+1}\left(\sup_{t\in[0,1]} \sigma_{n,ll}^{-1}(t)\left|\tilde{\xi}_l(t)+\tilde{\varepsilon}_l(t)\right| - \sup_{t\in[0,1]} \sigma_{n,ll}^{-1}(t)\left|\hat{\xi}_l(t)+\hat{\varepsilon}_l(t)\right|\right) \\
&= \mathcal{O}_p\left\{(\log(N_s+1)\, nh_s)^{1/2} n^{-1} h_s^{-3/2}\log(n)\right\} \\
&= \mathcal{O}_p\left\{n^{-1/2} h_s^{-1}(\log(N_s+1))^{1/2}\log(n)\right\} = \mathcal{O}_p(1).
\end{aligned}
$$

Then the proof follows from (41) and Slutsky's Theorem. □

A.4 Proof of Theorem 1

For any vector $\mathbf{a}=(a_1,\ldots,a_d)^\mathsf{T}\in\mathbb{R}^d$, $\mathsf{E}\left[\sum_{l=1}^d a_l\left\{\hat{\xi}_l(t)+\hat{\varepsilon}_l(t)\right\}\right]=0$. Using the conditional independence of $\hat{\xi}_l(t)$, $\hat{\varepsilon}_l(t)$ on $\left(T_{ij},N_i,X_{il}\right)_{i=1,j=1,l=1}^{n,N_i,d}$, we have

$$
\begin{aligned}
& \mathsf{Var}\left[\sum_{l=1}^d a_l\left\{\hat{\xi}_l(t)+\hat{\varepsilon}_l(t)\right\}\middle|\left(T_{ij},N_i,X_{il}\right)_{j=1,i=1,l=1}^{N_i,n,d}\right] \\
&= \sum_{l=1}^d\sum_{l'=1}^d a_l a_{l'}\mathsf{E}\left\{\hat{\xi}_l(t)\hat{\xi}_{l'}(t)+\hat{\varepsilon}_l(t)\hat{\varepsilon}_{l'}(t)\middle|\left(T_{ij},N_i,X_{il}\right)_{j=1,i=1,l=1}^{N_i,n,d}\right\} \\
&= \mathbf{a}^\mathsf{T}\left\{\Sigma_{\xi,n}(t)+\Sigma_{\varepsilon,n}(t)\right\}\mathbf{a}.
\end{aligned}
$$

Meanwhile, Assumption (A2) entails that for any $t\in[0,1]$, given $\left(T_{ij},N_i,X_{il}\right)_{j=1,i=1,l=1}^{N_i,n,d}$, the conditional distribution of $\left[\mathbf{a}^\mathsf{T}\left\{\Sigma_{\xi,n}(t)+\Sigma_{\varepsilon,n}(t)\right\}\mathbf{a}\right]^{-1/2}$ $\sum_{l=1}^d a_l\left\{\hat{\xi}_l(t)+\hat{\varepsilon}_l(t)\right\}$ is a standard normal distribution. So we have

$$
\left[\mathbf{a}^\mathsf{T}\left\{\Sigma_{\xi,n}(t)+\Sigma_{\varepsilon,n}(t)\right\}\mathbf{a}\right]^{-1/2}\sum_{l=1}^d a_l\left\{\hat{\xi}_l(t)+\hat{\varepsilon}_l(t)\right\}\sim N(0,1).
$$

Using (33), we have as $n\to\infty$

$$
\left[\mathbf{a}^\mathsf{T}\Sigma_n(t)\mathbf{a}\right]^{-1/2}\sum_{l=1}^d a_l\left\{\hat{\xi}_l(t)+\hat{\varepsilon}_l(t)\right\}\xrightarrow{\mathcal{L}} N(0,1).
$$

Therefore, $\left[\mathbf{a}^\mathsf{T}\Sigma_n(t)\mathbf{a}\right]^{-1/2}\sum_{l=1}^d a_l\left\{\hat{m}_l(t)-m_l(t)\right\}\xrightarrow{\mathcal{L}} N(0,1)$ follows from (24), Proposition 3, Lemma 4 and Slutsky's Theorem. Applying Cramér–Wold's device, we obtain $\Sigma_n^{-1/2}(t)\left\{\hat{m}_l(t)-m_l(t)\right\}_{l=1}^d\xrightarrow{\mathcal{L}} N(\mathbf{0},\mathbf{I}_{d\times d})$, and consequently, $\sigma_{n,ll}^{-1}(t)\left\{\hat{m}_l(t)-m_l(t)\right\}\xrightarrow{\mathcal{L}} N(0,1)$ for any $t\in[0,1]$ and $l=1,\ldots,d$. □

A.5 Proof of Theorem 2

By Proposition 3, $\|\tilde{m}_l - m_l\|_\infty = \mathcal{O}_p(h_\mathrm{s})$, $l = 1, \ldots, d$, so

$$a_{N_\mathrm{s}+1}\left\{\sup_{t\in[0,1]} \sigma_{n,ll}^{-1}(t)\,|\tilde{m}_l(t) - m_l(t)|\right\} = \mathcal{O}_p\left\{(nh_\mathrm{s})^{1/2}\left(\log(N_\mathrm{s}+1)\right)^{1/2} h_\mathrm{s}\right\} = \mathcal{O}_p(1).$$

According to (24), it is easy to show that

$$a_{N_\mathrm{s}+1}\left\{\sup_{t\in[0,1]} \sigma_{n,ll}^{-1}(t)\left|\hat{m}_l(t) - m_l(t)\right| - \sup_{t\in[0,1]} \sigma_{n,ll}^{-1}(t)\left|\tilde{\xi}_l(t) + \tilde{\varepsilon}_l(t)\right|\right\} = \mathcal{O}_p(1).$$

Meanwhile, Proposition 4 entails that, for any $\tau \in \mathbb{R}$,

$$\lim_{n\to\infty} P\left\{a_{N_\mathrm{s}+1}\left(\sup_{t\in[0,1]} \sigma_{n,ll}^{-1}(t)\left|\tilde{\xi}_l(t) + \tilde{\varepsilon}_l(t)\right| - b_{N_\mathrm{s}+1}\right) \le \tau\right\} = \exp\left(-2\mathrm{e}^{-\tau}\right).$$

Thus Slutsky's Theorem implies that

$$\lim_{n\to\infty} P\left\{a_{N_\mathrm{s}+1}\left(\sup_{t\in[0,1]} \sigma_{n,ll}^{-1}(t)\left|\hat{m}_l(t) - m_l(t)\right| - b_{N_\mathrm{s}+1}\right) \le \tau\right\} = \exp\left(-2\mathrm{e}^{-\tau}\right).$$

Let $\tau = -\log\left\{-\frac{1}{2}\log(1-\alpha)\right\}$, the definition of $Q_{N_\mathrm{s}+1}(\alpha)$ in (9) entails

$$\begin{aligned}
&\lim_{n\to\infty} P\left\{m_l(t) \in \hat{m}_l(t) \pm \sigma_{n,ll}(t) Q_{N_\mathrm{s}+1}(\alpha),\ \forall t \in [0,1]\right\} \\
&= \lim_{n\to\infty} P\left\{\sup_{t\in[0,1]} \sigma_{n,ll}^{-1}(t)\left|\hat{m}_l(t) - m_l(t)\right| \le Q_{N_\mathrm{s}+1}(\alpha)\right\} = 1-\alpha.
\end{aligned}$$

Theorem 2 is proved. □

References

Bosq D (1998) Nonparametric statistics for stochastic processes. Springer, New York

Brumback B, Rice JA (1998) Smoothing spline models for the analysis of nested and crossed samples of curves (with Discussion). J Am Stat Assoc 93:961–994

Cao G, Yang L, Todem D (2012) Simultaneous inference for the mean function based on dense functional data. J Nonparametr Stat 24:359–377

Cao G, Wang J, Wang L, Todem D (2012) Spline confidence bands for functional derivatives. J Stat Plan Inference 142:1557–1570

Chiang CT, Rice JA, Wu CO (2001) Smoothing spline estimation for varying coefficient models with repeatedly measured dependent variables. J Am Stat Assoc 96:605–619

Claeskens G, Van Keilegom I (2003) Bootstrap confidence bands for regression curves and their derivatives. Ann Stat 31:1852–1884

de Boor C (2001) A practical guide to splines. Springer, New York

Fan J, Zhang JT (2000) Functional linear models for longitudinal data. J R Stat Soc Ser B 62:303–322

Fan J, Zhang WY (2000) Simultaneous confidence bands and hypothesis testing in varying-coefficient models. Scand J Stat 27:715–731
Fan J, Zhang WY (2008) Statistical methods with varying coefficient models. Stat Interface 1:179–195
Ferraty F, Vieu P (2006) Nonparametric functional data analysis: theory and practice. Springer, New York
Gabrys R, Horváth L, Kokoszka P (2010) Tests for error correlation in the functional linear model. J Am Stat Assoc 105:1113–1125
Hall P, Müller HG, Wang JL (2006) Properties of principal component methods for functional and longitudinal data analysis. Ann Stat 34:1493–1517
Hall P, Titterington DM (1988) On confidence bands in nonparametric density estimation and regression. J Mult Anal 27:228–254
Härdle W, Luckhaus S (1984) Uniform consistency of a class of regression function estimators. Ann Stat 12:612–623
Hastie T, Tibshirani R (1993) Varying-coefficient models. J R Stat Soc Ser B 55:757–796
Hoover DR, Rice JA, Wu CO, Yang LP (1998) Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data. Biometrika 85:809–822
Horváth L, Kokoszka P (2012) Inference for functional data with applications. Springer, New York
Huang JZ, Wu CO, Zhou L (2002) Varying-coefficient models and basis function approximations for the analysis of repeated measurements. Biometrika 89:111–128
Huang JZ, Wu CO, Zhou L (2004) Polynomial spline estimation and inference for varying coefficient models with longitudinal data. Stat Sin 14:763–788
James GM, Hastie T, Sugar C (2000) Principal component models for sparse functional data. Biometrika 87:587–602
James GM, Sugar CA (2003) Clustering for sparsely sampled functional data. J Am Stat Assoc 98:397–408
Leadbetter MR, Lindgren G, Rootzén H (1983) Extremes and related properties of random sequences and processes. Springer, New York
Liu R, Yang L (2010) Spline-backfitted kernel smoothing of additive coefficient model. Econ Theory 26:29–59
Ma S, Yang L, Carroll RJ (2012) A simultaneous confidence band for sparse longitudinal regression. Stat Sin 22:95–122
Manteiga W, Vieu P (2007) Statistics for functional data. Comput Stat Data Anal 51:4788–4792
Ramsay JO, Silverman BW (2005) Functional data analysis. Springer, New York
Wang L, Li H, Huang JZ (2008) Variable selection in nonparametric varying-coefficient models for analysis of repeated measurements. J Am Stat Assoc 103:1556–1569
Wang L, Yang L (2009) Polynomial spline confidence bands for regression curves. Stat Sin 19:325–342
Wu CO, Chiang CT (2000) Kernel smoothing on varying coefficient models with longitudinal dependent variable. Stat Sin 10:433–456
Wu CO, Chiang CT, Hoover DR (1998) Asymptotic confidence regions for kernel smoothing of a varying-coefficient model with longitudinal data. J Am Stat Assoc 93:1388–1402
Wu Y, Fan J, Müller HG (2010) Varying-coefficient functional linear regression. Bernoulli 16:730–758
Xue L, Yang L (2006) Additive coefficient modelling via polynomial spline. Stat Sin 16:1423–1446
Xue L, Zhu L (2007) Empirical likelihood for a varying coefficient model with longitudinal data. J Am Stat Assoc 102:642–654
Yao W, Li R (2013) New local estimation procedure for a non-parametric regression function for longitudinal data. J R Stat Soc Ser B 75:123–138
Yao F, Müller HG, Wang JL (2005a) Functional linear regression analysis for longitudinal data. Ann Stat 33:2873–2903
Yao F, Müller HG, Wang JL (2005b) Functional data analysis for sparse longitudinal data. J Am Stat Assoc 100:577–590
Zhou L, Huang J, Carroll RJ (2008) Joint modelling of paired sparse functional data using principal components. Biometrika 95:601–619
Zhu H, Li R, Kong L (2012) Multivariate varying coefficient model for functional responses. Ann Stat 40:2634–2666