Web-based supporting materials for "FMEM: Functional Mixed effects Models for Longitudinal Functional Responses" by Hongtu Zhu, Kehui Chen, X. Luo, Ying Yuan, and Jane-Ling Wang

A Proofs of Theorems

Recall that $y_i = (y_{i1}, \ldots, y_{iT_i})^{\mathrm{T}}$ and $X_i = [x_{i1} \cdots x_{iT_i}]$. Furthermore, we define $Z_i = [z_{i1} \cdots z_{iT_i}]$, $e_{i,G} = (e_{i1,G}, \ldots, e_{iT_i,G})^{\mathrm{T}}$, and $e_{i,L} = (e_{i1,L}, \ldots, e_{iT_i,L})^{\mathrm{T}}$. For notational simplicity, we omit the subscription of a bandwidth h when there is no confusion. We present three lemmas that will be used in the proof of main theorems.

Lemma 1: Under assumptions (A.1) (or (A.1b)), (A.2), (A.3), and (A.6)-(A.7), if $\log M = o(Mh)$ *and there exists* $\gamma_n \to \infty$, *with* $n^{1/2}\gamma_n^{1-q} = o(1)$ *and* $n^{-1/2}\gamma_n \log M = o(1)$ *for some* $q > 2$, *that satisfies (A.7), we have*

$$\sup_s n^{1/2} |(nM)^{-1} \sum_{i,m} X_i W_i^{-1}(s_m) e_{i,L}(s_m) K_h(s_m - s)| = o_p(1),$$

where $W_i(s)$ *is a* $T_i \times T_i$ *working covariance, and it is Lipchitz continuous along* s.

Lemma 2: Under assumptions (A.1) (or (A.1b)), (A.2), (A.3), and (A.6)-(A.8), we have

$$\sup_s |(nM)^{-1} \sum_{i,m} X_i W_i^{-1}(s_m)(Z_i^{\mathrm{T}} b_i(s_m) + e_{i,G}(s_m)) K_h(s_m - s)| = O_p((\log n/n)^{1/2}),$$

where $W_i(s)$ *is a* $T_i \times T_i$ *working covariance, and it is Lipchitz continuous along* s.

Lemma 3: Under assumptions (A.1), (A.3), and (A.6), for any $r \geq 0$, *we have*

$$\sup_{s \in [0,1]} | \int K_h(u - s) \frac{(u - s)^r}{h^r} d[F_M(u) - F(u)]| = O_p((Mh)^{-1/2}). \tag{34}$$

If (A.1) is replaced with (A.1b), then we have

$$\sup_{s \in [0,1]} | \int K_h(u - s) \frac{(u - s)^r}{h^r} d[F_M(u) - F(u)]| = O_p((Mh)^{-1}). \tag{35}$$

Proof of Theorem 1. For each s, consider a $T_i \times T_i$ working variance matrix $W_i(s)$, and define

$$Q_1(s) = E(X_i W_i(s)^{-1} X_i^{\mathrm{T}}),$$

and

$$Q_2(s, s') = E(X_i W_i(s)^{-1} \Sigma_{y_i,G}(s, s') W_i(s')^{-1} X_i^{\mathrm{T}}).$$

We will derive asymptotic results for any $W_i(s)$ that is Lipschitz continuous along s and for which Q_1, Q_2 exist. A special case when $W_i \equiv I_{T_i \times T_i}$ leads to the asymptotic results for $\hat{\beta}(s)$ obtained in Step 1 with $Q_1(s) = E(X_i X_i^{\mathrm{T}})$ and $Q_2(s, s') = E(X_i \Sigma_{y_i,G}(s, s') X_i^{\mathrm{T}})$. Consequently, it can be shown (similar to the proof of Theorem 2) that $\hat{\Sigma}_{y_i,G}(s, s)$ is a sup-norm consistent estimator of $\Sigma_{y_i,G}$. Theorem 1 then follows by observing that the refined estimator $\tilde{\beta}(s)$ obtained in Step III used $W_i(s) = \hat{\Sigma}_{y_i,G}(s, s)$, which is a sup-norm consistent estimator of $\Sigma_{y_i,G}$. Note that (A.5) implies that $\Sigma_{y_i,G}(s, s)$ are Lipschitz continuous along s. In this case,

$$Q_1(s) = E(X_i \Sigma_{y_i,G}(s, s)^{-1} X_i^{\mathrm{T}})$$

and

$$Q_2(s, s') = E(X_i \Sigma_{y_i,G}(s, s)^{-1} \Sigma_{y_i,G}(s, s') \Sigma_{y_i,G}(s', s')^{-1} X_i^{\mathrm{T}}).$$

For $\gamma = 0$, 1, we introduce some $p_x \times 1$ matrix notations:

$$\begin{aligned}
R_\gamma &= 1/(nM) \sum_{i,m} X_i W_i^{-1}(s_m) y_i(s_m) K_h(s_m - s)[(s_m - s)/h]^\gamma, \\
R_\gamma^{(1)} &= 1/(nM) \sum_{i,m} X_i W_i^{-1}(s_m)(Z_i^{\mathrm{T}} b_i(s_m) + e_{i,G}(s_m)) K_h(s_m - s)[(s_m - s)/h]^\gamma, \\
R_\gamma^{(2)} &= 1/(nM) \sum_{i,m} X_i W_i^{-1}(s_m) e_{i,L}(s_m) K_h(s_m - s)[(s_m - s)/h]^\gamma, \qquad (36) \\
R_\gamma^{*} &= 1/(nM) \sum_{i,m} X_i W_i^{-1}(s_m)[y_i(s_m) - X_i^{\mathrm{T}} \beta(s) - X_i^{\mathrm{T}} \dot{\beta}(s)(s - s_m)] K_h(s_m - s)[(s_m - s)/h]^\gamma.
\end{aligned}$$

For $\gamma = 0$, 1, 2, we introduce a $p_x \times p_x$ matrix notation:

$$L_\gamma = 1/(nM) \sum_{i,m} X_i W_i^{-1}(s_m) X_i^{\mathrm{T}} K_h(s_m - s)[(s_m - s)/h]^\gamma, \qquad (37)$$

We define

$$L = \begin{pmatrix} L_0 & L_1 \\ L_1 & L_2 \end{pmatrix}, R = \begin{pmatrix} R_0 \\ R_1 \end{pmatrix}, R^{(1)} = \begin{pmatrix} R_0^{(1)} \\ R_1^{(1)} \end{pmatrix}, R^{(2)} = \begin{pmatrix} R_0^{(2)} \\ R_1^{(2)} \end{pmatrix}, R^* = \begin{pmatrix} R_0^* \\ R_1^* \end{pmatrix}. \tag{38}$$

Direct calculations lead to

$$\tilde{\beta}(s) = [(1,0) \otimes I_{p_x}]L^{-1}R \quad \text{and} \quad \tilde{\beta}(s) - \beta(s) = [(1,0) \otimes I_{p_x}]L^{-1}R^*.$$

Proof of Theorem 1 (i): We first consider the bias part. Based on Lemma 3, under a random design (A.1), we have

$$L = \begin{pmatrix} L_0 & L_1 \\ L_1 & L_2 \end{pmatrix} = \begin{pmatrix} Q_1(s)f(s) & 0 \\ 0 & Q_1(s)f(s)u_2 \end{pmatrix} + O_p(n^{-1/2} + (Mh)^{-1/2} + h);$$

or under a prefixed design (A.1b), we have

$$L = \begin{pmatrix} L_0 & L_1 \\ L_1 & L_2 \end{pmatrix} = \begin{pmatrix} Q_1(s)f(s) & 0 \\ 0 & Q_1(s)f(s)u_2 \end{pmatrix} + O_p(n^{-1/2} + (Mh)^{-1} + h).$$

Using $W_i(s_m) = W_i(s) + O(h)$ and Lemma 3, we have

$$E(R^*|\mathcal{S}) = \begin{pmatrix} Q_1(s)f(s)0.5h^2\ddot{\beta}(s)u_2 \\ Q_1(s)f(s)0.5h^2\ddot{\beta}(s)u_3 \end{pmatrix} + o(h^2),$$

$$Bias(\tilde{\beta}(s)|\mathcal{S}) = E(\tilde{\beta}(s)|\mathcal{S}) - \beta(s) = \frac{1}{2}\ddot{\beta}(s)h^2u_2(1 + o(1)). \tag{39}$$

Next, we consider the variance part. Simple calculation shows that

$$\begin{aligned} \text{var}(R_0|\mathcal{S}) &= \{o(1) + 1\}\frac{1}{(Mn)^2}\text{var}(\sum_{i,m} K_h(s_m - s)X_iW_i^{-1}(s)\left(Z_i^{\mathrm{T}}b_i(s_m) + e_i(s_m)\right)|\mathcal{S}) \qquad (40) \\ &= [\frac{1}{nMh}[\tilde{Q}(s) - Q_2(s,s)]f(s)v_2 + \frac{1}{n}Q_2(s,s)f(s)^2](1 + o(1)) \\ &= \frac{1}{n}Q_2(s,s)f(s)^2(1 + o(1)), \end{aligned}$$

where $\tilde{Q}(s) = E(X_i W_i^{-1}(s)\Sigma_{y,i}(s,s)W_i^{-1}(s)X_i^{\mathrm{T}})$. Note that

$$\mathrm{var}(R|\mathcal{S}) = \begin{pmatrix} 1 & o(1) \\ o(1) & o(1) \end{pmatrix} \otimes \mathrm{var}(R_0|\mathcal{S}) = \begin{pmatrix} n^{-1}Q_2(s,s)f(s)^2 & 0 \\ 0 & 0 \end{pmatrix} + o(n^{-1}),$$

we have $\mathrm{var}(\tilde{\beta}(s)|\mathcal{S}) = n^{-1}Q_1^{-1}(s)Q_2(s,s)Q_1^{-1}(s)\{1+o(1)\}$. This achieves the minimum when $W_i(s) = \Sigma_{y_i,G}(s,s)$, and then $\mathrm{var}(\tilde{\beta}(s)|\mathcal{S}) = n^{-1}\{E(X_i\{\Sigma_{y_i,G}(s,s)\}^{-1}X_i^{\mathrm{T}})\}^{-1}\{1+o(1)\}$.

Proof of Theorem 1 (ii): $\sqrt{n}[\tilde{\beta}(s) - E(\tilde{\beta}|\mathcal{S})]$ can be decomposed as

$$\sqrt{n}[(1,0)\otimes I_{p_x}]L^{-1}(s)[R^{(1)}(s) + R^{(2)}(s)],$$

where $L(s)$, $R^{(1)}(s)$ and $R^{(2)}(s)$ are defined in (38). Consider

$$L(s) = \begin{pmatrix} L_0 & L_1 \\ L_1 & L_2 \end{pmatrix} = \begin{pmatrix} Q_1(s)f(s) & 0 \\ 0 & Q_1(s)f(s)u_2 \end{pmatrix} + R_n(s),$$

in which it follows from Lemma 3 that $\sup_s R_n(s) = O_p(n^{-1/2} + (Mh)^{-1/2} + h)$. Lemma 1 implies that

$$\sup_s \sqrt{n}[(1,0)\otimes I_{p_x}]L^{-1}(s)R^{(2)}(s) = o(1), \quad a.s.,$$

which is negligible relative to $[(1,0)\otimes I_{p_x}]L^{-1}(s)R^{(1)}(s)$.

For any vector $a \in R^{p_x}$, we define $H_n(s) = \sqrt{n}a^{\mathrm{T}}[(1,0)\otimes I_{p_x}]L^{-1}(s)R^{(1)}(s)$ and show the following two results.

- (1.a) For any set of finite points $s_1, \ldots, s_K$, the sequence $(H_n(s_1), \ldots, H_n(s_K))$ converges in distribution to a multivariate Gaussian vector $G(0, a^{\mathrm{T}}R_K a)$, where

  $$R_{K;i,j} = Q_1(s_i)^{-1}Q_2(s_i,s_j)Q_1(s_j)^{-1}.$$

  This can be easily proved by using the standard central limit theorem, so we omit the details.

- (1.b) For every ϵ, $\eta > 0$, there exists a partition of $[0,1]$ into to finitely many sets

$S_1, \ldots, S_K$ such that

$$\limsup_{n\to\infty} P(\sup_l \sup_{s,s'\in S_l} |H_n(s) - H_n(s')| \geq \epsilon) \leq \eta.$$

This condition is equivalent to the asymptotic tightness, which will be verified below.

Let $M_i(s_m) = X_i W_i^{-1}(s_m)[Z_i^{\mathrm{T}} b_i(s_m) + e_i(s_m)]$ and

$$\phi_i(s) = M^{-1} \sum_{m=1}^{M} f(s)^{-1} a^{\mathrm{T}} Q_1(s)^{-1} M_i(s_m) K_h(s_m - s).$$

With some algebraic calculations, we can obtain

$$H_n(s) = n^{-1/2} \sum_{i=1}^{n} \phi_i(s) + \tilde{R}_n(s),$$

where $\tilde{R}_n(s) = O(R_n(s))$ and $\sup_s \tilde{R}_n(s) = O_p(n^{-1/2} + (Mh)^{-1/2} + h)$.

For $\delta > 0$ and $h < \delta$, it follows from (A.1)-(A.3) and (A.5) that there is a $C > 0$ such that with probability 1, we have

$$|\phi_i(s) - M^{-1} \sum_{m=1}^{M} f(s)^{-1} a^{\mathrm{T}} Q_1(s)^{-1} M_i(s) K_h(s_m - s)| \leq Ch,$$

and $\sup_{|s_1 - s_2| < \delta} |n^{-1/2} \sum_{i=1}^{n} \{\phi_i(s_1) - \phi_i(s_2)\}| \leq C(\delta + h)$. Furthermore, for any $\epsilon > 0$ and η, one can find large enough n, M, and small enough h such that

$$P(\sup_s \tilde{R}_n(s) > \epsilon/4) < \eta \quad \text{and} \quad Ch \leq \epsilon/4.$$

So with a partition $S_1, S_2, \ldots, S_K$ that $\sup_l \sup_{s,s'\in S_l} |s - s'| \leq \delta$ and $\delta = \epsilon/(4C)$, we have

$$\limsup_{n\to\infty} P(\sup_l \sup_{s,s'\in S_l} |H_n(s) - H_n(s')| \geq \epsilon) \leq \eta.$$

It follows from Theorem 18.14 in Van der Vaart (2000) (1.a), and (1.b) that $H_n(s)$ converges to a centered Gaussian process with covariance function $a^{\mathrm{T}} R(s, s') a$, with $R(s, s') = Q_1(s)^{-1} Q_2(s, s') Q_1(s')^{-1}$. For $W_i(s) = \Sigma_{y_i,G}(s, s)$, we have $R(s, s') = Q^*(s, s)^{-1} Q^*(s, s') Q^{-1}(s', s')$, with $Q^*(s, s') = E(X_i \Sigma_{y_i,G}(s, s)^{-1} \Sigma_{y_i,G}(s, s') \Sigma_{y_i,G}(s', s')^{-1} X_i^{\mathrm{T}})$. This completes the proof of part (ii). □

Proof of Theorem 2:

Proof of Theorem 2 (i): For any fixed s and s', $\hat{\Sigma}_{e,G}(s,s')$ is given by

$$\begin{aligned}\hat{\Sigma}_{e,G}(s,s')) &= [\sum_{m\neq m'} K_h(s_m-s)K_h(s_{m'}-s')]^{-1} \\ &\quad \times[\sum_{m\neq m'} K_h(s_m-s)K_h(s_{m'}-s')\hat{\Sigma}_e^{LS}(s_m,s_{m'})],\end{aligned}$$

where $\hat{\Sigma}_e^{LS}(s_m,s_{m'}) = (1-a_2g^{\mathrm{T}}g)^{-1}(v(s,s')-a_2g^{\mathrm{T}}Gu(s,s'))$, and details are given in Section 2 Step (II) (S1). For each fixed pair (m,m'), $\Sigma_e^{LS}(s_m,s_{m'})$ is the least square estimate based on $\hat{u}_{ij}(s_m)\hat{u}_{ij}(s_{m'})$, and $\hat{u}_{ij}(s_m) = y_{ij}(s_m) - x_{ij}^{\mathrm{T}}\hat{\beta}(s_m)$ is an estimate of $u_{ij}(s_m) = y_{ij}(s_m) - x_{ij}^{\mathrm{T}}\beta(s_m)$.

Here $\hat{\beta}(s)$ is the initial estimator in step (I), which can be viewed as one of the general weighted estimator using $W_i(s) \equiv I_{T_i\times T_i}$. As shown in the proof of Theorem 1,

$$\hat{\beta}(s) - \beta(s) = [(1,0)\otimes I_{p_x}]L^{-1}R^*,$$

and

$$\sup_s |\hat{\beta}(s) - \beta(s)| = \sup_s [(1,0)\otimes I_{p_x}]L^{-1}(R^{(1)}+R^{(2)}) + O(h_1^2).$$

It is easy to see that lemma 2 holds for $W_i(s) \equiv I_{T_i\times T_i}$. From Lemma 2,

$$\sup_s [(1,0)\otimes I_{p_x}]L^{-1}R^{(1)} = O_p((\log n/n)^{1/2}),$$

and following the same argument as lemma 2, we have

$$\sup_s [(1,0)\otimes I_{p_x}]L^{-1}R^{(2)} = O_p((\log n/n)^{1/2}).$$

Together we have $\sup_s |\hat{\beta}(s) - \beta(s)| = O_p(h_1^2 + (\log n/n)^{1/2})$ and then $\sup_s |\hat{u}_{ij}(s) - u_{ij}(s)| = O_p(h_1^2 + (\log n/n)^{1/2}) = O_p((\log n/n)^{1/2})$.

Therefore direct calculation leads to

$$\sup_{(s,s')\in[0,1]^2} |E[\hat{\Sigma}_{e,G}(s,s')] - \Sigma_{e,G}(s,s')| = O_p(h_{\Sigma_e}^2 + h_1^2 + (\log n/n)^{1/2}) = O_p((\log n/n)^{1/2}). \quad (41)$$

Now we show that

$$\sup_{(s,s')\in[0,1]^2} |\hat{\Sigma}_{e,G}(s,s') - E[\hat{\Sigma}_{e,G}(s,s')]| = O_p(\log n/n)^{1/2}). \quad (42)$$

Lemma 3 implies

$$\sup_{s,s'} |\frac{1}{M(M-1)}[\sum_{m\neq m'} K_h(s_m - s)K_h(s_{m'} - s')]| = O_p(1).$$

By (A.2) and Law of large numbers, $(1 - a_2 g^{\mathrm{T}} g)^{-1} = O_p(1)$ and $a_2 g^{\mathrm{T}} G \mathbf{1}_{p_z^2\times 1} = O_p(1)$, where $\mathbf{1}_{p_z^2\times 1}$ denotes a $p_z^2 \times 1$ all-ones vector. Further combining with the result, $\sup_s |\hat{u}_{ij}(s) - u_{ij}(s)| = O_p((\log n/n)^{1/2})$, we have

$$\begin{aligned}&\sup_{(s,s')\in[0,1]^2} |\hat{\Sigma}_{e,G}(s,s') - E[\hat{\Sigma}_{e,G}(s,s')]| \\ =&O_p(\sup_{s,s'} |D_1(s,s') - ED_1(s,s')|) + O_p(\sup_{s,s'} |D_2(s,s') - ED_2(s,s')|) + O_p((\log n/n)^{1/2}),\end{aligned}$$

where

$$D_1(s,s') = \frac{1}{M(M-1)\sum_{i=1}^n T_i} \sum_{i,j,m\neq m'} u_{ij}(s_m)u_{ij}(s_{m'})K_{m,m'}(s,s'),$$

and

$$D_2(s,s') = \frac{1}{M(M-1)\sum_{i=1}^n T_i^2} \sum_{i,j_1,j_2,m\neq m'} u_{ij_1}(s_m)u_{ij_2}(s_{m'})\mathbf{1}^{\mathrm{T}}_{p_z^2\times 1} z_{ij_1} \otimes z_{ij_2} K_{m,m'}(s,s'),$$

with $K_{m,m'}(s,s') = K_h(s_m - s)K_h(s_{m'} - s')$.

For D_1, let $\alpha_n = (\log n/n)^{1/2}$, $F_{imm'} = \frac{1}{T_i}\sum_j u_{ij}(s_m)u_{ij}(s_{m'})$, and $Q_n = \alpha_n^{-1}$. We define

$$G_n(s,s') = \frac{1}{nM^2h^2} \sum_{i,m,m'} F_{imm'}\mathbf{1}(-h \le s - s_m < h)\mathbf{1}(-h \le s' - s_{m'} < h),$$

$$G_n^{Q_n}(s,s') = \frac{1}{nM^2h^2} \sum_{i,m,m'} F_{imm'}\mathbf{1}(|F_{imm'}| < Q_n)\mathbf{1}(-h \le s - s_m < h)\mathbf{1}(-h \le s' - s_{m'} < h),$$

where $\mathbf{1}(x)$ is the indicator function. By (A.3) and lemma 4 of Li and Hsing (2010), we have $\sup_{s,s'} |D_1(s,s') - ED_1(s,s')| \le C \sup_{s,s'} |G_n(s,s') - EG_n(s,s')|$. To show $\sup_{s,s'} |D_1(s,s') - ED_1(s,s')| = O_p(\alpha_n)$, we will prove that

$$\sup_{s,s'} |G_n^{Q_n}(s,s') - G_n(s,s')| = o_p(\alpha_n), \tag{43}$$

$$\sup_{s,s'} |E(G_n^{Q_n}(s,s')) = o_p(\alpha_n), \tag{44}$$

$$\sup_{s,s'} |G_n^{Q_n}(s,s') - E(G_n^{Q_n}(s,s'))| = O_p(\alpha_n). \tag{45}$$

The proof of Eq. (43), (44) and (45) mirrors the one-dimensional correspondences, Eq. 43, 44, and 45 given in the proof of lemma 2. See also lemma 3 and lemma 4 in Li and Hsing (2010). The same arguments apply for D_2.

Combining (41) and (42), part (i) follows.

Proof of Theorem 2 (ii) and (iii): It follows from Theorem 2 (i) that we have

$$||\hat{\Sigma}_{e,G} - \Sigma_{e,G}||_2 = \{\int_0^1 \int_0^1 |\hat{\Sigma}_{e,G}(s,s') - \Sigma_{e,G}(s,s')|^2 ds ds'\}^{1/2}$$
$$= O_p((\log n/n)^{1/2}).$$

It follows from Lemma 4.3 of Bosq (2000) that

$$|\hat{\lambda}_l^e - \lambda_l^e| \leq ||\hat{\Sigma}_{e,G} - \Sigma_{e,G}||_2 \quad \text{and} \quad ||\hat{\phi}_l^e - \phi_l^e||_2 \leq 2\sqrt{2}\delta_l^{-1}||\hat{\Sigma}_{e,G} - \Sigma_{e,G}||_2,$$

for $l = 1 \ldots, E$, where δ_l is the minimum eigenvalue gap before l. It follows from assumption (A.10) that δ_l is a positive constant. Therefore, it yields Theorem 2 (ii) and (iii). □

Almost identical arguments could be used to prove the following results regarding $\Sigma_{b,kk'}$.

Theorem 2b. *Under (A.1) (or (A.1b)) and (A.2)-(A.8), (A.10), if $h_1 = O((\log n/n)^{1/4})$ and $h_2 = O(\log n/n)^{1/4}$, then we have the following results hold for $1 \leq k, k' \leq p_z$:*

(i) $\sup_{s,s'} |\hat{\Sigma}_{b,kk'}(s,s') - \Sigma_{b,kk'}(s,s')| = O_p((\log n/n)^{1/2})$;

(ii) For $1 \leq l \leq E$, $\{\int_0^1 |\hat{\psi}_l^{bkk'}(s) - \psi_l^{bkk'}(s)|^2 ds\}^{1/2} = O_p((\log n/n)^{1/2})$;

(iii) For $1 \leq l \leq E$, $|\hat{\lambda}_l^{bkk'} - \lambda_l^{bkk'}| = O_p((\log n/n)^{1/2})$.

Proof of Theorem 3: Let $d(s) = C\tilde{\beta}(s) - \text{bias}(C\tilde{\beta}(s)) - \beta_0(s)$ and

$$\boldsymbol{\omega}(s) = \{C\sum_{i=1}^n X_i \hat{\Sigma}_{y_i,G}(s,s)^{-1} X_i^{\mathrm{T}}]^{-1} C^{\mathrm{T}}\}^{-1/2} d(s).$$

Using similar argument as in Theorem 1 (ii), we can show that $\boldsymbol{\omega}(s)$ converges to a centered Gaussian process $G_C(\cdot)$ with covariance function

$$\boldsymbol{\gamma}_\omega(s,s') = \{CQ^*(s,s)C^{\mathrm{T}}\}^{-1/2} R(s,s') \{CQ^*(s',s')C^{\mathrm{T}}\}^{-1/2},$$

which yields Theorem 3 (i).

We prove Theorem 3 (ii) as follows. Under H_{1n}, we have

$$\boldsymbol{\omega}(s) \stackrel{asymp}{\sim} GP(\boldsymbol{\eta}_\omega, \boldsymbol{\gamma}_\omega),$$

where $\boldsymbol{\eta}_\omega(s) = \{C[\sum_{i=1}^n X_i \hat{\Sigma}_{y_i,G}(s,s)^{-1} X_i^{\mathrm{T}}]^{-1} C^{\mathrm{T}}\}^{-1/2} n^{-\tau/2} \boldsymbol{d}(s)$, $\boldsymbol{\gamma}_\omega(s,t) = (\gamma_{\omega,kl}(s,t))_{1\leq k\leq l\leq q}$, and $\gamma_{\omega,kl}(s,t) = \mathrm{cov}(\omega_k(s), \omega_l(t))$. We consider a Hilbert space of q-dimensional vectors of functions in $L_2(\mathcal{S})$, denoted by $\mathbb{H}$. Define the inner product of two q-dimensional functions as $< \boldsymbol{f}, \boldsymbol{g} >_{\mathbb{H}} = \sum_{l=1}^q < f_l, g_l >$. By the multivariate version of Mercer's theorem, there exists a set of orthonormal basis functions $\boldsymbol{\phi}_r = (\phi_{r1}, \ldots, \phi_{rq})$ in $\mathbb{H}$ such that

$$\boldsymbol{\gamma}_\omega(s,t) = \sum_{r=1}^\infty \lambda_r \boldsymbol{\phi}_r(s) \boldsymbol{\phi}_r^{\mathrm{T}}(t),$$

in which $\gamma_{\omega,kl}(s,t) = \sum_{r=1}^\infty \lambda_r \phi_{rk}(s) \phi_{rl} t)$ and $< \boldsymbol{\phi}_r, \boldsymbol{\phi}_r' >_{\mathbb{H}} = \delta_{rr'}$.

Let $\xi_r = < \boldsymbol{\omega}, \boldsymbol{\phi}_r >_{\mathbb{H}}$, we have $\xi_r \sim N(\pi_r, \lambda_r)$, where $\pi_r = < \boldsymbol{\eta}_\omega, \boldsymbol{\phi}_r >_{\mathbb{H}}$. It is assumed that the eigenvalues are ordered in decreasing values. Without loss of generality, the first m eigenvalues are assumed to be positive. If all eigenvalues are positive, we set $m = \infty$. It is easy to see that

$$S_n = \int_{\mathcal{S}} \boldsymbol{\omega}(s)^{\mathrm{T}} \boldsymbol{\omega}(s) ds = \sum_{r=1}^\infty \xi_r^2 \stackrel{d}{=} \sum_{r=1}^m \lambda_r A_r + \sum_{r=m+1}^\infty \pi_r^2,$$

where $A_r \sim \chi^2(\lambda_r^{-1} \pi_r^2)$. Similar results have been obtained and discussed in (Zhang, 2011; Zhang and Chen, 2007).

Under the null hypothesis H_0, $\pi_r = 0$ holds for all r. So the null distribution of S_n is a mixture of χ^2. Under H_{1n}, we have $\pi_r^2 = n^{1-r} \delta_r^2$, where δ_r is given by

$$\delta_r = \sum_{l=1}^q \int_{\mathcal{S}} \{C[\frac{1}{n} \sum_{i=1}^n X_i \hat{\Sigma}_{y_i,G}(s,s)^{-1} X_i^{\mathrm{T}}]^{-1} C^{\mathrm{T}}\}^{-1/2} d_l(s) \phi_{rl}(s) ds.$$

Note that $A_r \stackrel{d}{=} [Z_r + n^{(1-\tau)/2} \lambda_r^{-1/2} \delta_r]^2$, where $Z_r \sim N(0,1)$. Thus, we have

$$S_n \stackrel{d}{=} \sum_{r=1}^m \lambda_r Z_r^2 + 2n^{(1-\tau)/2} \delta_\lambda Z_r + n^{1-\tau} \delta^2,$$

where $\delta_\lambda = \sum_{r=1}^m \lambda_r^{1/2} \delta_r$ and $\delta^2 = \sum_{r=1}^m \delta_r^2 > 0$.

We consider two cases of δ_λ. For $\delta_\lambda = 0$, it is easy to show that the power tends to 1 as $n \to \infty$, since the second term is zero and the third term dominates the first term. For $\delta_\lambda^2 > 0$, the second term and the third term dominate the first term as $n \to \infty$. Following similar arguments as in Theorem 3 in (Zhang, 2011), we can show that S_n is asymptotically normally distributed under H_{1n} with mean $n^{1-\tau}\delta^2$ and variance $4n^{1-\tau}\delta_\lambda^2$. Therefore, we have

$$P(S_n > S_{n,\alpha} \mid H_{1n}) = 1 - \Phi(\frac{S_{n,\alpha}^* - n^{1-\tau}\delta^2}{\sqrt{4n^{1-\tau}\delta_\lambda^2}}) + o(1) = \Phi(\frac{n^{(1-\tau)/2}\delta^2}{2\delta_\lambda}) + o(1),$$

which tends to 1 as $n \to \infty$ This completes the proof of Theorem 3. □

Proof of Theorem 4: The proof of Theorem 4 is basically the same as the proof of Theorem 5 in Zhu et al. (2012). We omit the details here.

B Proofs of Lemmas

Proof of Lemma 1: The proof of this Lemma is similar to Lemma 8 of Zhu et al. (2012). We give a unified proof for both fixed and random designs and allow short-range correlation of error processes. Without loss of generality, we only prove the case that $p_x = 1$, since the extension to $p_x > 1$ is straightforward.

It follows from (A.3) and Lemma 2 of Li and Hsing (2010) that,

$$\begin{aligned}
&\sup_s n^{1/2} |(nM)^{-1} \sum_{i,m} \mathbf{X}_i W_i^{-1}(s_m) \boldsymbol{e}_{i,L}(s_m) K_h(s_m - s)| \\
\leq & C \sup_s n^{1/2} |(nMh)^{-1} \sum_{i,m} \mathbf{X}_i W_i^{-1}(s_m) \boldsymbol{e}_{i,L}(s_m) \mathbf{1}(-h \leq s - s_m < h)|. \qquad (46)
\end{aligned}$$

We introduce some notation and let

$$G_n(s) = n^{1/2}(hMn)^{-1} \sum_{i,m} \mathbf{X}_i W_i^{-1}(s_m) \boldsymbol{e}_{i,L}(s_m) \mathbf{1}(-h \leq s - s_m < h).$$

For some γ_n satisfing $n^{1/2}\gamma_n^{1-q} = o(1)$ and $n^{-1/2}\gamma_n \log M = o(1)$, where q satisfies (A.7), define

$$G_n^{\gamma_n}(s) = n^{1/2}(nMh)^{-1}\sum_{i,m}\mathbf{X}_i W_i^{-1}(s_m)\boldsymbol{e}_{i,L}(s_m)\mathbf{1}(\|\boldsymbol{e}_{i,L}(s_m)\|_2 < \gamma_n)\mathbf{1}(-h \le s - s_m < h).$$

By (46), we only need to prove $\sup_s |G_n(s)| = o_p(1)$, so we will complete the proof by showing the following three results.

$$\sup_s |G_n^{\gamma_n}(s) - G_n(s)| = o_p(1), \tag{47}$$

$$\sup_s |E(G_n^{\gamma_n}(s))| = o_p(1), \tag{48}$$

$$\sup_s |G_n^{\gamma_n}(s) - E(G_n^{\gamma_n}(s))| = o_p(1). \tag{49}$$

For (47), we have

$$\sup_s |G_n^{\gamma_n}(s) - G_n(s)| \le \tag{50}$$
$$n^{1/2}\gamma_n^{1-q}\frac{1}{n}\sum_i \sup_s \|\mathbf{X}_i W_i^{-1}(s)\|_2 \frac{1}{hM}\sum_m \|\boldsymbol{e}_{i,L}(s_m)\|_2^q \mathbf{1}(\|\boldsymbol{e}_{i,L}(s_m)\|_2 > \gamma_n)\mathbf{1}(-h \le s - s_m < h),$$

which is bounded by $n^{1/2}\gamma_n^{1-q} = o_p(1)$ from above. This leads to (47). Similarly, we can prove (48).

For (49), let

$$Z_i(s) = (hM)^{-1}\sum_m \mathbf{X}_i W_i^{-1}(s_m)\boldsymbol{e}_{i,L}(s_m)\mathbf{1}(\|\boldsymbol{e}_{i,L}(s_m)\|_2 < \gamma_n)\mathbf{1}(-h \le s - s_m < h).$$

Then $\sup_s |G_n^{\gamma_n}(s) - E(G_n^{\gamma_n}(s))| = \sup_s n^{-1/2}|\sum_{i=1}^n \{Z_i(s) - E(Z_i(s))\}|$. It follows from (A.2) and the Lipchitz continuity of $W_i(s)$ that $\sup_{s\in[0,1]} \|\mathbf{X}_i W_i^{-1}(s)\|_2$ is bounded and

$$2f_l \le E[(hM)^{-1}\sum_{m=1}^M \mathbf{1}(-h \le s_m - s \le h)] \le 2f_u.$$

Therefore, if $P(\Omega_S) = 1 - p_M \to 1$ in a set Ω_S, then we have $|Z_i(s) - E(Z_i(s))| < C_1\gamma_n$ and

$$\sum_{i=1}^n \mathrm{var}(Z_i(s)) \le \sum_{i=1}^n E(Z_i(s))^2 \le n(Mh)^{-2}O(Mh\sum_{k=0}^{\lceil Mh\rceil} k^{-\delta}) = O(n(Mh)^{-1}),$$

where C_1 does not depend on s. If (A.1) is replaced by (A.1b), then $|Z_i(s) - E(Z_i(s))| = O(\gamma_n)$ and $\sum_i \mathrm{var}(Z_i(s)) = O(n(Mh)^{-1})$ still hold.

For any design S, as s varies in [0, 1], $(s-h, s+h)$ covers some consecutive points of s_m. The pair of starting point and ending point have $M(M-1)/2$ possibilities, which does not depend on the design S. So we have

$$P(\sup_{s\in[0,1]} n^{-1/2}|\sum_{i=1}^{n}\{Z_i(s)-E(Z_i(s))\}|>x \mid S\in\Omega_S) \tag{51}$$

$$\leq M^2\exp(-\frac{x^2}{2C_2(Mh)^{-1}+2/3C_1\gamma_n x/\sqrt{n}}). \tag{52}$$

Let

$$x_n=\frac{1}{3}CC_1n^{-1/2}\gamma_n\log M+0.5\sqrt{\frac{4}{9}C^2C_1^2\log M^2n^{-1}\gamma_n^2+16CC_2\log M(Mh)^{-1}}.$$

It follows from (A.6) and $\log M=o(Mh)$ that $x_n\to 0$ as $n\to\infty$. Then, we have

$$P(\sup_{s\in[0,1]} n^{-1/2}|\sum_{i=1}^{n}Z_i(s)-E(Z_i(s))|>x_n \mid S\in\Omega_S)\leq M^2\exp(-C\log M)=M^{2-C}.$$

For some $C>2$, we have

$$P(\sup_{s\in[0,1]}|G_n^{\gamma_n}(s)-E(G_n^{\gamma_n}(s))|>x_n)$$

$$\leq P(\sup_{s\in[0,1]} n^{-1/2}|\sum_{i=1}^{n}Z_i(s)-E(Z_i(s))|>x_n|S\in\Omega_S)+p_M\leq M^{2-C}+p_M=o(1). \tag{53}$$

The proof of Lemma 1 is now complete. □

Remark: This result is similar to Lemma 2 in Li and Hsing (2010), but the proof is much simpler. We completely remove the constrains on h and n (C5 in their paper). The result is almost surely for pre-fixed design as in (A.1b). For random design (A.1), we first prove the result for relatively good (uniform) designs and add back the very small probability of bad designs. So the result is in probability rather than almost surely.

Proof of Lemma 2: Without loss of generality, we prove the case with $p_x=1$, since it is easy to prove the case with $p_x>1$. It follows from (A.3) and lemma 2 of Li and Hsing

(2010) that

$$\begin{aligned}&\sup_s |(nM)^{-1}\sum_{i,m}\mathbf{X}_i W_i^{-1}(s_m)(\mathbf{Z}_i^T\boldsymbol{b}_i(s_m)+\boldsymbol{e}_{i,G}(s_m))K_h(s_m-s)|\\ \le & C\sup_s |(nMh)^{-1}\sum_{i,m}\mathbf{X}_i W_i^{-1}(s_m)(\mathbf{Z}_i^T\boldsymbol{b}_i(s_m)+\boldsymbol{e}_{i,G}(s_m))\mathbf{1}(-h\le s-s_m<h)|. \qquad (54)\end{aligned}$$

Moreover, it follows from (A.2) and the Lipchitz continuity of $W_i(s)$ that $\sup_{s\in[0,1]}\|\mathbf{X}_i W_i^{-1}(s)\|_2$ is bounded almost surely. Thus, we only need to show that

$$\sup_s |(nMh)^{-1}\sum_{i,m}(\mathbf{Z}_i^T\boldsymbol{b}_i(s_m)+\boldsymbol{e}_{i,G}(s_m))\mathbf{1}(-h\le s-s_m<h)| = O_p((\log n/n)^{1/2}).$$

Let $\alpha_n = (\log n/n)^{1/2}$, $Q_n = \alpha_n^{-1} = (n/\log n)^{1/2}$, and $F_i(s_m) = Z_i^T\boldsymbol{b}_i(s_m)+\boldsymbol{e}_{i,G}(s_m)$. Define $G_n(s) = (nMh)^{-1}\sum_{i,m}F_i(s_m)\mathbf{1}(-h\le s-s_m<h)$ and

$$G_n^{Q_n}(s) = (nMh)^{-1}\sum_{i,m}F_i(s_m)\mathbf{1}(\|F_i(s_m)\|_2<Q_n)\mathbf{1}(-h\le s-s_m<h).$$

We will show that

$$\sup_s |G_n^{Q_n}(s)-G_n(s)| = o_p(\alpha_n), \qquad (55)$$

$$\sup_s |E(G_n^{Q_n}(s)) = o_p(\alpha_n)|, \qquad (56)$$

$$\sup_s |G_n^{Q_n}(s)-E(G_n^{Q_n}(s))| = O_p(\alpha_n). \qquad (57)$$

The proof of Lemma 2 will be complete by combining (55)-(57).

For (55), we have

$$\begin{aligned}&\alpha_n^{-1}\sup_s |G_n^{Q_n}(s)-G_n(s)|\\ \le&\alpha_n^{-1}Q_n^{1-q}\frac{1}{nMh}\sum_{i,m}\|F_i(s_m)\|^q\mathbf{1}(\|F_i(s_m)\|>Q_n)\mathbf{1}(-h\le s-s_m<h)\\ \le&\alpha_n^{-1}Q_n^{1-q}O_p(1)=o_p(1), \qquad (58)\end{aligned}$$

in which (A,7) and (A.8) are used for some $q>2$. Similarly, we can prove (56).

For (57), we define an equally-spaced grid $\mathcal{G}=\{\nu_k\}$ with $\nu_k = kh\alpha_n^2$ for $k=0,\ldots,[1/(h\alpha_n^2)]$, where [.] denotes the greatest integer part. For any $\nu_k \le s \le \nu_{k+1}$, we have

$$
\begin{aligned}
&|G_n^{Q_n}(s) - EG_n^{Q_n}(s)| \\
=&|G_n^{Q_n}(\nu_k) - EG_n^{Q_n}(\nu_k)| + |G_n^{Q_n}(s) - G_n^{Q_n}(\nu_k)| + |EG_n^{Q_n}(s) - EG_n^{Q_n}(\nu_k)|,
\end{aligned}
$$

where $|G_n^{Q_n}(s) - G_n^{Q_n}(\nu_k)|$ is bounded from above by

$$
\begin{aligned}
&|(Mh)^{-1}\sum_{m=1}^{M} Q_n \mathbf{1}(s_m \in [\nu_k - h, s-h] \cup [\nu_k + h, s+h])| \\
\le&|(Mh)^{-1}\sum_{m=1}^{M} Q_n \mathbf{1}(s_m \in [\nu_k - h, \nu_{k+1} - h] \cup [\nu_k + h, \nu_{k+1} + h])|.
\end{aligned}
$$

Define $I_k = [\nu_k - h, \nu_{k+1} - h] \cup [\nu_k + h, \nu_{k+1} + h]$ for $k = 0, \ldots, [1/(h\alpha_n^2)]$. It follows from (A.1) that

$$
2f_l M\alpha_n^2 h \le E\sum_{m=1}^{M} \mathbf{1}(s_m \in I_k) \le 2f_u M\alpha_n^2 h.
$$

Bernstein inequality implies that $P(\sup_k \sum_{m=1}^{M} \mathbf{1}(s_m \in I_k) > 2Cf_u M\alpha_n^2 h) \to 0$ as $M \to \infty$. Given Ω_{S_1}, we have

$$
\sup_k \sum_{m=1}^{M} \mathbf{1}(s_m \in I_k) < 2Cf_u M\alpha_n^2 h \quad \text{and} \quad P(\Omega_{S_1}) = 1 - p_M^{(1)} \to 1. \tag{59}
$$

Using (59), we have

$$
\begin{aligned}
&\sup_s |G_n^{Q_n}(s) - EG_n^{Q_n}(s)| \\
\le&\sup_k \{|G_n^{Q_n}(\nu_k) - EG_n^{Q_n}(\nu_k)| + |(Mh)^{-1}\sum_{m=1}^{M} Q_n \mathbf{1}(s_m \in I_k)| + |(Mh)^{-1}\sum_{m=1}^{M} Q_n E\mathbf{1}(s_m \in I_k)|\} \\
\le&\sup_k |G_n^{Q_n}(\nu_k) - G^{Q_n}(\nu_k)| + (Mh)^{-1}Q_n 2Cf_u M\alpha_n^2 h + (Mh)^{-1}Q_n 2f_u M\alpha_n^2 h \\
\le&\sup_k |G_n^{Q_n}(\nu_k) - G^{Q_n}(\nu_k)| + 2Cf_u\alpha_n + 2f_u\alpha_n, \\
=&\sup_k ||G_n^{Q_n}(\nu_k) - EG_n^{Q_n}(\nu_k)|| + O(\alpha_n).
\end{aligned}
$$

Let $Z_i(s) = (hM)^{-1} \sum_m F_i(s_m) \mathbf{1}(\|F_i(s_m)\| < Q_n) \mathbf{1}(-h \le s - s_m < h)$. Then, we have

$$\sup_k |G_n^{Q_n}(\nu_k) - E(G_n^{Q_n}(\nu_k))| = \sup_k n^{-1} |\sum_{i=1}^n Z_i(\nu_k) - E(Z_i(\nu_k))|.$$

It also follows from (A.1) that

$$2f_l \le E[(hM)^{-1} \sum_{m=1}^M \mathbf{1}(-h \le s_m - s \le h)] \le 2f_u.$$

So in a set Ω_{S_2} with $P(\Omega_{S_2}) = 1 - p_M^{(2)} \to 1$, we have $\sup_k |Z_i(\nu_k) - E(Z_i(\nu_k))| < C_1 Q_n$ and

$$\sum_{i=1}^n \mathrm{var}(Z_i) \le \sum_{i=1}^n E(Z_i)^2 \le C_2 n.$$

If (A.1) is replaced by (A.1b), $\sup_k |Z_i - E(Z_i)| = O_p(Q_n)$ and $\sum_i \mathrm{var}(Z_i) = O(n)$ still hold.

Finally, we have

$$\begin{aligned} &P(\sup_k n^{-1} |\sum_{i=1}^n Z_i(\nu_k) - E(Z_i(\nu_k))| > C\alpha_n | S \in (\Omega_{S_1} \cap \Omega_{S_2})) \\ \le& (\alpha_n^2 h)^{-1} \exp(-\frac{n^2 C^2 \alpha_n^2}{2C_2 n + 2/3 C_1 Q_n n \alpha_n}) = (\alpha_n^2 h)^{-1} \exp(-C^* \log n), \end{aligned}$$

where $C^* = C^2/(2C_2 + 2/3C_1)$. For large enough C^*, we have

$$P(\sup_k |G_n^{Q_n}(\nu_k) - E(G_n^{Q_n}(\nu_k))| > C\alpha_n) \le h^{-1} n^{1-C^*} + p_M^{(1)} + P_M^{(2)} = o(1),$$

which finishes the proof of Lemma 2. □

Proof of Lemma 3: Similar to Lemma 2 in Zhu et al. (2012), Eq. (22) can be proved by using empirical process techniques. Specifically, it follows from (A.1) and (A.3) that

$$\{K\left(\frac{. - s}{h}\right) \frac{(. - s)^r}{h^r}, s \in [0, 1]\} \quad \text{is a Donsker class.}$$

Eq. (23) is the same as (60) in Lemma 8 of Zhu et al. (2012). It can be proved by using Taylor expansion and (A.1b). □

C Additional Simulation Results

Simulation 4. The fourth simulation is to evaluate the accuracy of the estimators of the

Table 1: ASD data analysis: Distributions of scan availability.

Available scans	1	2	3	5	6
Number of subjects	137	78	34	5	1

eigenvalues and eigenfucntions of the covariance functions $\Sigma_b(.,.)$, $\Sigma_{e,G}(.,.)$ and $\Sigma_{e,L}$. We used the same parameter values as those in Simulation 1. We set $c = 0.1$ and $n = 50$ and 100, and generated 100 datasets for each combination. The accuracy of all kinds of estimators improves with the sample size. The estimated eigenfunctions were plotted in Figures 1 and 2, in which the mean and the pointwise $5th$ and $95th$ percentiles of the estimated functions were plotted along with the true eigenfunctions. Figures 3 and 4 show the boxplots for the estimates of the eigenvalues and σ^2, which are quite close to their true values.

ASD Data Analysis. Table 1 presents the distribution of scan availability. We compared FMEM with WFMM and PFFR. Figures 5 and 6 present the estimated coefficient functions with their 95% confidence intervals corresponding to WFMM and PFFR. Inspecting Figure 5 reveals that the estimated coefficient functions from WFMM are very bumpy since wavelet is a poor choice for intrinsically smooth functions.

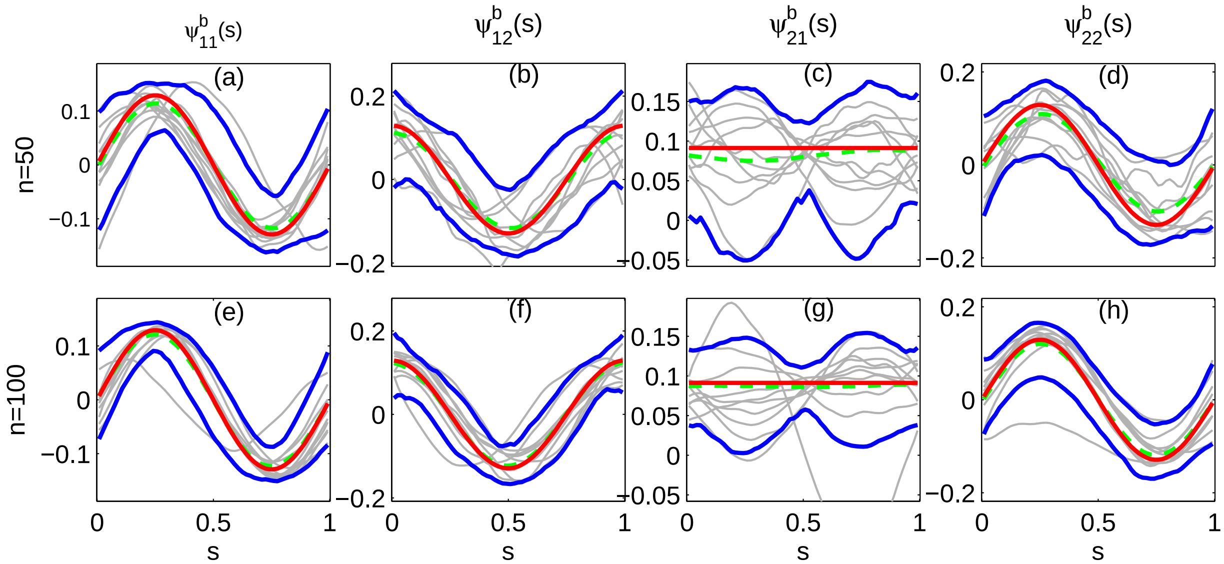

Figure 1: Simulations 4: the estimates of the first two eigenfunctions $\psi_{l,k}^{b}(\cdot)$ for $l, k = 1, 2$ and their pointwise confidence intervals. The red solid, green dashed and blue solid, curves are, respectively, the true eigenfunctions, the pointwise means, and their pointwise $5th$ and $95th$ percentiles of estimated eigenfunctions based on 100 replications.

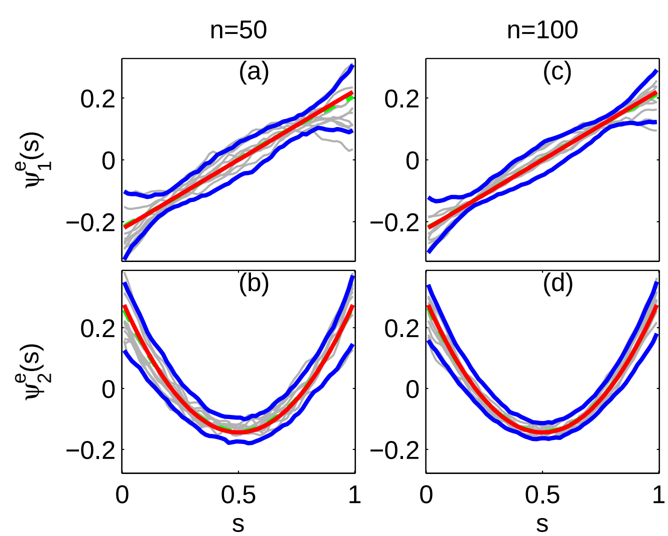

Figure 2: Simulations 4: the estimates of the first two eigenfunctions $\psi_{k}^{e}, k = 1, 2$ and their pointwise confidence interval. The red solid, green dashed and blue solid, curves are, respectively, the true eigenfunctions, the pointwise means and their pointwise $5th$ and $95th$ percentiles of estimated eigenfunctions based on 100 replications.

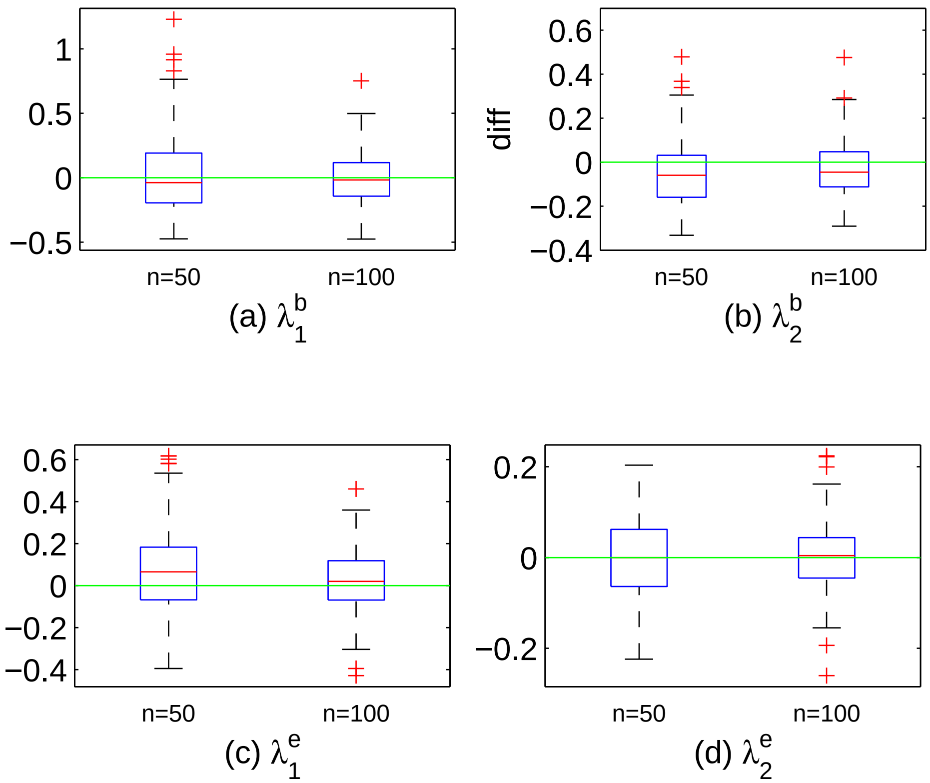

Figure 3: Simulation 4: boxplots of the differences between the estimated eigenvalues $\hat{\lambda}_k^b$ and $\hat{\lambda}_k^e, k = 1, 2$ and their true values based on 100 replications.

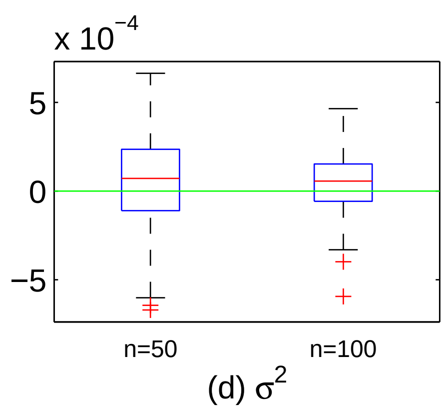

Figure 4: Simulation 4: boxplots of the differences between the estimated σ^2 and its true values based on 100 replications.

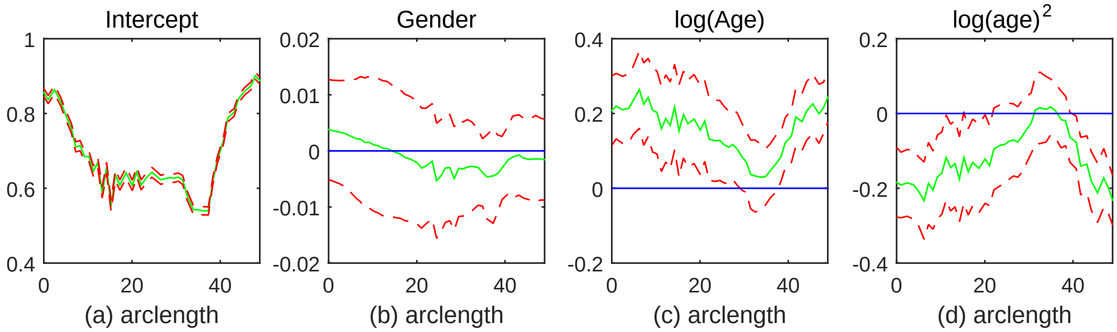

Figure 5: WFMM: 95% pointwise posterior credible intervals for coefficient functions. The solid curves are the estimated coefficient functions and the dashed curves are the 95% credible intervals. The thin horizontal line is the line crossing the origin (0, 0). Computational time is 7.9 seconds.

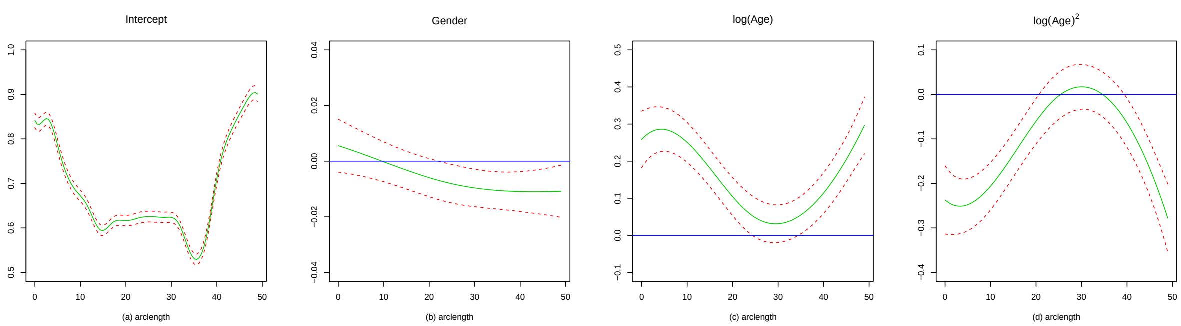

Figure 6: PFFR: 95% confidence bands for coefficient functions. The solid curves are the estimated coefficient functions and the dashed curves are the 95% confidence bands. The thin horizontal line is the line crossing the origin (0, 0). Computational time is 6.078 hours.

References

Bosq, D. (2000), *Linear Processes in Function Spaces: Theory and Applications*, New York: Springer-Verlag.

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

Van der Vaart, A. W. (2000), *Asymptotic Statistics*, vol. 3, Cambridge university press.

Zhang, J. (2011), "Statistical inferences for linear models with functional responses," *Statistica Sinica*, 21, 1431–1451.

Zhang, J. and Chen, J. (2007), "Statistical inference for functional data," *The Annals of Statistics*, 35, 1052–1079.

Zhu, H. T., Li, R., and Kong, L. (2012), "Multivariate varying coefficient model for functional responses," *Annals of Statistics*, 40, 2634–2666.