Supplement to "Kernel-based Method for Detecting Structural Break in Distribution of Functional Data"

Peijun Sang and Bing Li

In this Supplement we provide the proofs of the theorems, lemmas, and corollaries, extra details for numerical implementations and additional simulation results in the manuscript. The equation labels such as (1) and (2) are for the equations in the manuscript; equation labels such as (A1) and (A2) are for the equations in this supplement.

S.1 Proofs

S.1.1 Proof of Theorems 3.1 and 3.2

Proof of Theorem 3.1. For any $\tau \in [0,1]$, define a partial sum of i.i.d $\mathfrak{M}_X$-valued random variables:

$$S_n(\tau) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor n\tau \rfloor} \kappa(X_i - \mu_X, \cdot). \tag{A1}$$

By Theorem 1 of Kuelbs (1973),

$$S_n(\tau) \xrightarrow{\mathcal{L}} \mathcal{N}(\tau),$$

where $\mathcal{N}(\tau)$ is a centered $\mathfrak{M}_X$-valued Gaussian process defined on $[0,1]$ such that $\mathrm{E}\{\mathcal{N}(\tau)\} = 0 \in \mathfrak{M}_X$ and $\mathrm{E}\{\langle \mathcal{N}(\tau_1), f_1\rangle_{\mathfrak{M}_X} \langle \mathcal{N}(\tau_2), f_2\rangle_{\mathfrak{M}_X}\} = \min\{\tau_1,\tau_2\}\langle \Sigma_{XX} f_1, f_2\rangle_{\mathfrak{M}_X}$. Moreover, one has

$$\sup_{0\le\tau\le1} \left\| n^{-1/2} \sum_{1\le i\le \lfloor n\tau\rfloor} \kappa(X_i - \mu_X, \cdot) \right\|_{\mathfrak{M}_X} = O_p(1).$$

Therefore,

$$\begin{aligned}\frac{d_n(\tau)}{n} =& \frac{1}{n} \left\| \sum_{i=1}^{\lfloor n\tau \rfloor} \kappa(X_i,\cdot) - \frac{\lfloor n\tau\rfloor}{n} \sum_{i=1}^{n} \kappa(X_i,\cdot) \right\|^2_{\mathfrak{M}_X} \\ =& \|S_n(\tau) - \tau S_n(1)\|^2_{\mathfrak{M}_X} + o_p(1).\end{aligned}$$

Using a similar argument, we show that $S_n(\tau) - \tau S_n(1) \xrightarrow{\mathcal{L}} \mathcal{N}^0(\tau)$, where $\mathcal{N}^0(\tau)$ is a centered $\mathfrak{M}_X$-valued Gaussian process defined on $[0,1]$ such that $\mathrm{E}\{\mathcal{N}^0(\tau)\} = 0 \in \mathfrak{M}_X$ and $\mathrm{E}\{\langle \mathcal{N}^0(\tau_1), f_1\rangle_{\mathfrak{M}_X} \langle \mathcal{N}^0(\tau_2), f_2\rangle_{\mathfrak{M}_X}\} = (\min\{\tau_1,\tau_2\} - \tau_1\tau_2)\langle \Sigma_{XX} f_1, f_2\rangle_{\mathfrak{M}_X}$. In other words, the limiting process has the same distribution as

$$\sum_{\nu=1}^{\infty} \sqrt{\theta_\nu} B_\nu(\tau) \varphi_\nu,$$

where $\{(\theta_\nu, \varphi_\nu) : \nu = 1, \ldots\}$ are eigenvalues and the corresponding eigenfunctions of Σ_{XX}, and B_ν's are iid standard Brownian bridges. By the continuous mapping theorem, one has

$$\frac{d_n(\tau)}{n} \xrightarrow{\mathcal{L}} \sum_{\nu=1}^{\infty} \theta_\nu B_\nu^2$$

in the Skorohod topology of $D[0,1]$. This completes the proof. □

Proof of Theorem 3.2. Let

$$S^0_{n,k^*} = \frac{1}{\sqrt{n}} \sum_{i=1}^{k^*} \epsilon_i \quad \text{and} \quad S^1_{n,k^*} = \frac{1}{\sqrt{n}} \sum_{i=k^*+1}^{n} \epsilon_i.$$

It follows that

$$T_n = \sup_{0 \le \tau \le 1} \frac{d_n(\tau)}{n} \ge \frac{S_{k^*}}{n} = \left\| \left(1 - \frac{k^*}{n}\right) S^0_{n,k^*} - \frac{k^*}{n} S^1_{n,k^*} - \frac{k^*}{\sqrt{n}} \frac{n-k^*}{n} \delta \right\|^2_{\mathfrak{M}_X}.$$

By using a similar argument for $S_n(\tau)$ in (A1), we have

$$\left(1 - \frac{k^*}{n}\right) S^0_{n,k^*} - \frac{k^*}{n} S^1_{n,k^*} = O_p(1),$$

since $k^* = \lfloor n\tau^* \rfloor$ for some $\tau^* \in (0,1)$. Therefore,

$$\begin{aligned}
&\left\| \left(1 - \frac{k^*}{n}\right) S^0_{n,k^*} - \frac{k^*}{n} S^1_{n,k^*} - \frac{k^*}{\sqrt{n}} \frac{n-k^*}{n} \delta \right\|_{\mathfrak{M}_X} \\
&\ge \left\| \frac{k^*}{\sqrt{n}} \frac{n-k^*}{n} \delta \right\|_{\mathfrak{M}_X} - O_p(1) \\
&\ge c_{1,n} \sqrt{n} \tau^* (1 - \tau^*) \|\delta\|_{\mathfrak{M}_X} \to \infty.
\end{aligned}$$

as $n \to \infty$, where $c_{1,n}$ can be taken as a positive sequence converging to 1. □

S.1.2 Proof of Theorem 3.3

The proof of Theorem 3.3 is similar to that of Lemma A.1 in Aue *et al.* (2018). But our theoretical results do not require a specific assumption on the data generation process as in model (2.1) in Aue *et al.* (2018).

Proof of Theorem 3.3. We first show part (i) of Theorem 3.3. To prove $\hat{\tau}_n \xrightarrow{\mathbb{P}} \tau^*$, it suffices to show that $|\hat{k}_n - k^*|$ is bounded in probability.

By the definition of $\hat{k}_n$, it follows that

$$\hat{k}_n = \min\left\{k : R_{n,k} = \max_{1 \le k' \le n} R_{n,k'}\right\},$$

where

$$R_{n,k} = S_k - S_{k^*}.$$

Under Assumption 3, after some algebra, for $1 \le k < k^*$, we can rewrite $R_{n,k}$ as

$$R_{n,k} = \frac{1}{n}\langle E_k^{(1)} + D_k^{(1)}, E_k^{(2)} + D_k^{(2)} \rangle_{\mathfrak{M}_X}, \tag{A2}$$

where

$$D_k^{(1)} = -(k-k^*)\frac{n-k^*}{n}\delta \quad \text{and} \quad D_k^{(2)} = -(k+k^*)\frac{n-k^*}{n}\delta$$

are deterministic functions and

$$E_k^{(1)} = -\sum_{i=k+1}^{k^*} \epsilon_i - \frac{k-k^*}{n}\sum_{i=1}^{n}\epsilon_i \text{ and } E_k^{(2)} = \sum_{i=1}^{k}\epsilon_i + \sum_{i=1}^{k^*}\epsilon_i - \frac{k+k^*}{n}\sum_{i=1}^{n}\epsilon_i$$

are random functions. Similarly, for $k^* < k \le n$, one has

$$R_{n,k} = \frac{1}{n}\langle E_k^{(3)} + D_k^{(3)}, E_k^{(4)} + D_k^{(4)} \rangle_{\mathfrak{M}_X}, \tag{A3}$$

with deterministic functions

$$D_k^{(3)} = -(k-k^*)\frac{k^*}{n}\delta \quad \text{and} \quad D_k^{(4)} = -(2n-k-k^*)\frac{k^*}{n}\delta$$

and random functions

$$E_k^{(3)} = -\sum_{i=k^*+1}^{k} \varepsilon_i - \frac{k-k^*}{n}\sum_{i=1}^{n}\epsilon_i \text{ and } E_k^{(4)} = \sum_{i=1}^{k}\epsilon_i + \sum_{i=1}^{k^*}\epsilon_i - \frac{k+k^*}{n}\sum_{i=1}^{n}\epsilon_i.$$

We will focus on the analysis of $R_{n,k}$ for $1 \le k < k^*$ below, as the case of $k^* < k \le n$ can be analyzed in a similar manner by (A3).

For any $1 \le N < k^*$, since $k^* = \lfloor n\tau^* \rfloor$ for some $\tau^* \in (0,1)$, it follows that

$$\max_{1\le k\le k^*-N} \frac{k+k^*}{n}\left(\frac{n-k^*}{n}\right)^2 = \frac{2k^*-N}{n}\left(\frac{n-k^*}{n}\right)^2 \to 2\tau^*(1-\tau^*)^2,$$

as $n \to \infty$. By the definition of $D_k^{(1)}$ and $D_k^{(2)}$ in (A2), we have

$$\begin{aligned}\max_{1\le k\le k^*-N} \frac{1}{n}\langle D_k^{(1)}, D_k^{(2)}\rangle_{\mathfrak{M}_X} &= \max_{1\le k\le k^*-N}(k-k^*)\frac{k+k^*}{n}\left(\frac{n-k^*}{n}\right)^2 \|\delta\|^2_{\mathfrak{M}_X} \\ &= -2N\tau^*(1-\tau^*)\|\delta\|^2_{\mathfrak{M}_X} + o(1).\end{aligned}$$

Hence,

$$\lim_{N\to\infty}\limsup_{n\to\infty}\max_{1\le k\le k^*-N}\frac{1}{n}\langle D_k^{(1)}, D_k^{(2)}\rangle_{\mathfrak{M}_X} = -\infty. \tag{A4}$$

Next it will be shown that this deterministic term with the limit established in (A4) is the dominant term on the right-hand side of (A2). To this end, we first show that, for any $a > 0$,

$$\lim_{N\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(\max_{1\le k\le k^*-N}\left|\frac{\langle E_k^{(1)}, E_k^{(2)}\rangle_{\mathfrak{M}_X}}{\langle D_k^{(1)}, D_k^{(2)}\rangle_{\mathfrak{M}_X}}\right| > a\right) = 0. \tag{A5}$$

In fact, for any $n \geq 1$ and $1 \leq k < k^*$,

$$|\langle D_k^{(1)}, D_k^{(2)} \rangle_{\mathfrak{M}_X}| = (n - k^*)(k^* - k)\frac{k + k^*}{n}\frac{n - k^*}{n}\|\delta\|_{\mathfrak{M}_X}^2 \\ \geq c_0 n(k^* - k)$$

for some constant $c_0 > 0$. Then by the Cauchy-Schwarz inequality, we have

$$\left|\frac{\langle E_k^{(1)}, E_k^{(2)} \rangle_{\mathfrak{M}_X}}{\langle D_k^{(1)}, D_k^{(2)} \rangle_{\mathfrak{M}_X}}\right| \leq c_0^{-1}\frac{\langle E_k^{(1)}, E_k^{(2)} \rangle_{\mathfrak{M}_X}}{n(k^* - k)} \leq c_0^{-1}\frac{\|E_k^{(1)}\|_{\mathfrak{M}_X}\|E_k^{(2)}\|_{\mathfrak{M}_X}}{n(k^* - k)}.$$

Taking the maximum of the right-hand side of the above inequality over $1 \leq k \leq k^* - N$ yields

$$\max_{1 \leq k \leq k^* - N}\left|\frac{\langle E_k^{(1)}, E_k^{(2)} \rangle_{\mathfrak{M}_X}}{\langle D_k^{(1)}, D_k^{(2)} \rangle_{\mathfrak{M}_X}}\right| \leq c_1 \max_{1 \leq k \leq k^* - N}\frac{\|E_k^{(1)}\|_{\mathfrak{M}_X}}{k^* - k} \cdot \max_{1 \leq k \leq k^* - N}\frac{\|E_k^{(2)}\|_{\mathfrak{M}_X}}{n}.$$

By the definition of $E_k^{(1)}$, applying the triangle inequality leads to

$$\max_{1 \leq k \leq k^* - N}\frac{\|E_k^{(1)}\|_{\mathfrak{M}_X}}{k^* - k} \leq \max_{1 \leq k \leq k^* - N}\frac{1}{k^* - k}\left\|\sum_{i=k+1}^{k^*}\epsilon_i\right\|_{\mathfrak{M}_X} + \frac{1}{n}\left(\left\|\sum_{i=1}^{k^*}\epsilon_i\right\|_{\mathfrak{M}_X} + \left\|\sum_{i=k^*+1}^{n}\epsilon_i\right\|_{\mathfrak{M}_X}\right). \tag{A6}$$

Note that for any $p \in (1/2, 1)$,

$$\max_{1 \leq k \leq k^* - N}\frac{1}{k^* - k}\left\|\sum_{i=k+1}^{k^*}\epsilon_i\right\|_{\mathfrak{M}_X} \overset{d}{=} \max_{N \leq k \leq k^* - 1}\frac{1}{k}\left\|\sum_{i=1}^{k}\epsilon_i\right\|_{\mathfrak{M}_X} \leq \frac{1}{N^{1-p}}\sup_{k \geq 1}\frac{1}{k^p}\left\|\sum_{i=1}^{k}\epsilon_i\right\|_{\mathfrak{M}_X}.$$

By the Markov inequality, $\sup_{k \geq 1}\frac{1}{k^p}\left\|\sum_{i=1}^{k}\epsilon_i\right\|_{\mathfrak{M}_X} = O_p(1)$. Moreover, by applying the strong law of large numbers in $\mathfrak{M}_X$ (see Theorem 2.4 in Bosq 2000) to the right-hand side of (A6), we have

$$\lim_{N \to \infty}\limsup_{n \to \infty}\mathbb{P}\left(\max_{1 \leq k \leq k^* - N}\frac{\|E_k^{(1)}\|_{\mathfrak{M}_X}}{k^* - k} > a\right) = 0$$

for any $a > 0$. We may apply similar arguments to show that

$$\lim_{N \to \infty}\limsup_{n \to \infty}\mathbb{P}\left(\max_{1 \leq k \leq k^* - N}\frac{\|E_k^{(2)}\|_{\mathfrak{M}_X}}{n} > a\right) = 0$$

Thus (A5) holds.

Similar arguments lead to

$$\lim_{N \to \infty}\limsup_{n \to \infty}\mathbb{P}\left(\max_{1 \leq k \leq k^* - N}\left|\frac{\langle E_k^{(1)}, D_k^{(2)} \rangle_{\mathfrak{M}_X}}{\langle D_k^{(1)}, D_k^{(2)} \rangle_{\mathfrak{M}_X}}\right| > a\right) = 0. \tag{A7}$$

and

$$\lim_{N \to \infty}\limsup_{n \to \infty}\mathbb{P}\left(\max_{1 \leq k \leq k^* - N}\left|\frac{\langle E_k^{(2)}, D_k^{(1)} \rangle_{\mathfrak{M}_X}}{\langle D_k^{(1)}, D_k^{(2)} \rangle_{\mathfrak{M}_X}}\right| > a\right) = 0. \tag{A8}$$

Combining (A5), (A7) and (A8) with (A4) implies that, for any $M < 0$,

$$\lim_{N \to \infty}\limsup_{n \to \infty}\mathbb{P}\left(\max_{1 \leq k \leq k^* - N}R_{n,k} > M\right) = 0. \tag{A9}$$

Using a similar argument, we can show that

$$\lim_{N\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\max_{k^*+N\le k\le n} R_{n,k} > M\right) = 0. \tag{A10}$$

Note that $\hat{k}_n$ maximizes $R_{n,k}$ and $R_{n,k^*} = 0$. It implies that, for any $M < 0$,

$$\mathbb{P}(|\hat{k}_n - k^*| > N) = \mathbb{P}(\max_{1\le k\le k^*-N} R_{n,k} > M) + \mathbb{P}\left(\max_{k^*+N\le k\le n} R_{n,k} > M\right)$$

Thus, by (A9) and (A10), we have

$$\lim_{N\to\infty} \limsup_{n\to\infty} \mathbb{P}(|\hat{k}_n - k^*| > N) = 0,$$

which implies that $|\hat{k}_n - k^*|$ is bounded in probability.

Next, we show part (ii) of Theorem 3.3. To establish the limiting distribution of $\hat{k}_n - k^*$ in (3.8), it suffices to show that as $n \to \infty$,

$$(R_{n,k^*+k} : k \in [-N, N]) \xrightarrow{d} (2\tau^*(1-\tau^*)G(k) : k \in [-N, N]) \tag{A11}$$

for any $N \ge 1$. Actually, by part (i) of Theorem 3.3 and the continuous mapping theorem, (A11) implies (3.8).

By the definition of $D_k^{(1)}$ and $D_k^{(2)}$, we have

$$\begin{aligned}
&\max_{k^*-N\le k\le k^*} \left|\frac{1}{n}\langle D_k^{(1)}, D_k^{(2)}\rangle_{\mathfrak{M}_X} - 2\tau^*(1-\tau^*)^2\|\delta\|_{\mathfrak{M}_X}^2(k-k^*)\right| \\
&\le \max_{k^*-N\le k\le k^*} \left|(k-k^*)\left\{\frac{k+k^*}{n}\left(\frac{n-k^*}{n}\right)^2 - 2\tau^*(1-\tau^*)^2\right\}\right| \|\delta\|_{\mathfrak{M}_X}^2 \\
&\le N\|\delta\|_{\mathfrak{M}_X}^2 \max_{k^*-N\le k\le k^*} \left|\frac{k+k^*}{n}\left(\frac{n-k^*}{n}\right)^2 - 2\tau^*(1-\tau^*)^2\right| \\
&= o(1)
\end{aligned} \tag{A12}$$

as $n \to \infty$. Using a similar arguments, one can obtain

$$\max_{k^*\le k\le k^*+N} \left|\frac{1}{n}\langle D_k^{(3)}, D_k^{(4)}\rangle_{\mathfrak{M}_X} + 2(\tau^*)^2(1-\tau^*)\|\delta\|_{\mathfrak{M}_X}^2(k-k^*)\right| = o(1). \tag{A13}$$

Equations (A12) and (A13) establish the limit of the constant term in $R_{n,k}$ for $k < k^*$ and $k \ge k^*$, respectively.

To find the limit of the random terms in $R_{n,k}$, we first note that

$$\frac{1}{n}\langle E_k^{(1)}, D_k^{(2)}\rangle_{\mathfrak{M}_X} = \frac{n-k^*}{n}\frac{k+k^*}{n}\sum_{i=k+1}^{k^*}\langle \epsilon_i, \delta\rangle_{\mathfrak{M}_X} + \frac{n-k^*}{n}\frac{k+k^*}{n}\frac{k-k^*}{n}\sum_{i=1}^{n}\langle \epsilon_i, \delta\rangle_{\mathfrak{M}_X}. \tag{A14}$$

The first term on the right-hand side of (A14) can be bounded by

$$\begin{aligned}
&\max_{k^*-N\le k\le k^*} \left|\frac{n-k^*}{n}\frac{k+k^*}{n}\sum_{i=k+1}^{k^*}\langle \epsilon_i, \delta\rangle_{\mathfrak{M}_X} - 2\tau^*(1-\tau^*)\sum_{i=k+1}^{k^*}\langle \epsilon_i, \delta\rangle_{\mathfrak{M}_X}\right| \\
&= \max_{k^*-N\le k\le k^*} \left|\frac{n-k^*}{n}\frac{k+k^*}{n} - 2\tau^*(1-\tau^*)\right| \cdot \max_{k^*-N\le k\le k^*} \left|\sum_{i=k+1}^{k^*}\langle \epsilon_i, \delta\rangle_{\mathfrak{M}_X}\right| \\
&= o_p(1).
\end{aligned} \tag{A15}$$

For the second term on the right-hand side of (A14), $1/n\sum_{i=1}^{k^*}\langle\epsilon_i,\delta\rangle_{\mathfrak{M}_X}$ and $\sum_{i=k^*+1}^{n}\langle\varepsilon_i,\delta\rangle_{\mathfrak{M}_X}$ converge to 0 by the strong law of large numbers. Therefore,

$$\max_{k^*-N\le k\le k^*}\left|\frac{n-k^*}{n}\frac{k+k^*}{n}\frac{k-k^*}{n}\sum_{i=1}^{n}\langle\epsilon_i,\delta\rangle_{\mathfrak{M}_X}\right| = o_p(1). \tag{A16}$$

Combining (A14), (A15) and (A16), we have

$$\max_{k^*-N\le k\le k^*}\left|\frac{1}{n}\langle E_k^{(1)},D_k^{(2)}\rangle_{\mathfrak{M}_X} - 2\tau^*(1-\tau^*)\sum_{i=k+1}^{k^*}\langle\epsilon_i,\delta\rangle_{\mathfrak{M}_X}\right| = o_p(1). \tag{A17}$$

Similar arguments imply

$$\max_{k^*\le k\le k^*+N}\left|\frac{1}{n}\langle E_k^{(3)},D_k^{(4)}\rangle_{\mathfrak{M}_X} - 2\tau^*(1-\tau^*)\sum_{i=k+1}^{k^*}\langle\epsilon_i,\delta\rangle_{\mathfrak{M}_X}\right| = o_p(1). \tag{A18}$$

It remains to show that the other four terms on the right-hand side of $R_{n,k}$ in (A2) and (A3) are asymptotically negligible. Applying the Cauchy-Schwarz inequality yields

$$\begin{aligned}\max_{k^*-N\le k\le k^*}\left|\frac{1}{n}\langle E_k^{(1)},E_k^{(2)}\rangle_{\mathfrak{M}_X}\right| &\le \max_{k^*-N\le k\le k^*}\frac{1}{n}\|E_k^{(1)}\|_{\mathfrak{M}_X}\|E_k^{(2)}\|_{\mathfrak{M}_X}\\ &\le \max_{k^*-N\le k\le k^*}\|E_k^{(1)}\|_{\mathfrak{M}_X}\max_{k^*-N\le k\le k^*}\frac{1}{n}\|E_k^{(2)}\|_{\mathfrak{M}_X}.\end{aligned} \tag{A19}$$

By the triangle inequality, we have

$$\max_{k^*-N\le k\le k^*}\|E_k^{(1)}\|_{\mathfrak{M}_X} \le \max_{k^*-N\le k\le k^*}\left\|\sum_{i=k+1}^{k^*}\epsilon_i\right\|_{\mathfrak{M}_X} + \frac{N}{n}\left\|\sum_{i=1}^{n}\epsilon_i\right\|_{\mathfrak{M}_X} = O_p(1)+o_p(1), \tag{A20}$$

where the last equality holds since $\sum_{i=k+1}^{k^*}\epsilon_i$ contains at most N terms and $1/n\sum_{i=1}^{n}\epsilon_i = o_p(1)$ by the strong law of large numbers. Moreover, applying the triangle inequality again yields

$$\begin{aligned}\max_{k^*-N\le k\le k^*}\frac{1}{n}\|E_k^{(2)}\|_{\mathfrak{M}_X} &\le \max_{k^*-N\le k\le k^*}\frac{1}{n}\left\|\sum_{i=}^{k}\epsilon_i\right\|_{\mathfrak{M}_X} + \frac{1}{n}\left\|\sum_{i=1}^{k^*}\epsilon_i\right\|_{\mathfrak{M}_X} + \max_{k^*-N\le k\le k^*}\frac{k+k^*}{n^2}\left\|\sum_{i=1}^{n}\epsilon_i\right\|_{\mathfrak{M}_X}\\ &= o_p(1),\end{aligned} \tag{A21}$$

where the last equality holds by the strong law of large numbers. Equations (A19), (A20) and (A21) imply

$$\max_{k^*-N\le k\le k^*}\left|\frac{1}{n}\langle E_k^{(1)},E_k^{(2)}\rangle_{\mathfrak{M}_X}\right| = o_p(1). \tag{A22}$$

Lastly, using a similar argument for (A19), one obtains

$$\begin{aligned}\max_{k^*-N\le k\le k^*}\left|\frac{1}{n}\langle D_k^{(1)},E_k^{(2)}\rangle_{\mathfrak{M}_X}\right| &\le \max_{k^*-N\le k\le k^*}\|D_k^{(1)}\|_{\mathfrak{M}_X}\max_{k^*-N\le k\le k^*}\frac{1}{n}\|E_k^{(2)}\|_{\mathfrak{M}_X}\\ &\le N\|\delta\|_{\mathfrak{M}_X}\frac{n-k^*}{n}\max_{k^*-N\le k\le k^*}\frac{1}{n}\|E_k^{(2)}\|_{\mathfrak{M}_X}\\ &= o_p(1),\end{aligned} \tag{A23}$$

according to (A21). Equations (A12), (A17), (A22) and (A23) imply that (A11) holds for $k^*-N\le k\le k^*$. Equations (A13), (A18) plus similar arguments applied to the remaining two terms in (A3) imply (A11) holds for $k^*\le k\le k^*+N$ as well.

□

S.1.3 Proof of Theorem 3.4

The proof of Theorem 3.4 is similar to that of Theorem 3.4. We first establish the following lemma to show part (i) of Theorem 3.4.

Lemma S1. *Under the assumptions of Theorem 3.4, $\|\delta_n\|^2_{\mathfrak{M}_X}|\hat{k}_n - k^*|$ is bounded in probability.*

Proof. Given any $N \geq 1$, let $N_\delta = \|\delta_n\|^{-2}_{\mathfrak{M}_X} N$. As $n\|\delta_n\|^2_{\mathfrak{M}_X} \to \infty$, $N_\delta/n = o(1)$ and

$$\max_{1\leq k\leq k^*-N_\delta} \left(\frac{k+k^*}{n}\right)\left(\frac{n-k^*}{n}\right)^2 = \left(\frac{2k^*-N_\delta}{n}\right)\left(\frac{n-k^*}{n}\right)^2 \to 2\tau^*(1-\tau^*)^2$$

as $n \to \infty$. Therefore, as $n \to \infty$,

$$\begin{aligned}
\max_{1\leq k\leq k^*-N_\delta} \frac{1}{n}\langle D_k^{(1)}, D_k^{(2)}\rangle_{\mathfrak{M}_X} &= \max_{1\leq k\leq k^*-N_\delta} \|\delta_n\|^2_{\mathfrak{M}_X}(k-k^*)\left(\frac{k+k^*}{n}\right)\left(\frac{n-k^*}{n}\right)^2 \\
&= \|\delta_n\|^2_{\mathfrak{M}_X}(-N_\delta)\left(\frac{2k^*-N_\delta}{n}\right)\left(\frac{n-k^*}{n}\right)^2 \\
&\to -2N\tau^*(1-\tau^*)^2.
\end{aligned}$$

Using the similar arguments to establish equations (A4), (A5), (A7) and (A8), we show that it is the dominant term for the right-hand side of (A2). Furthermore, for any $M < 0$,

$$\lim_{N\to\infty}\limsup_{n\to\infty} \mathbb{P}\left(\max_{1\leq k\leq k^*-N_\delta} R_{n,k} > M\right) = 0.$$

Similarly,

$$\lim_{N\to\infty}\limsup_{n\to\infty} \mathbb{P}\left(\max_{k^*+N_\delta\leq k\leq n} R_{n,k} > M\right) = 0.$$

These two equations imply that $\|\delta_n\|^2_{\mathfrak{M}_X}|\hat{k}_n - k^*|$ is bounded in probability. $\square$

To prove part (ii) of Theorem 3.4, with the aid of Lemma S1, it suffices to establish the following lemma.

Lemma S2. *Under the assumptions of Theorem 3.4,*

$$(R_{n,k^*+k(x)} : x \in [-N, N]) \xrightarrow{\mathcal{L}} (2\tau^*(1-\tau^*)U(x) : x \in [-N, N]),$$

for any $N \geq 1$ as $n \to \infty$, where $k(x) = \left\lfloor \|\delta_n\|^{-2}_{\mathfrak{M}_X} x\right\rfloor$.

Proof. By the definition of $D_k^{(1)}$ and $D_k^{(2)}$, we have

$$\begin{aligned}
&\sup_{x\in[-N,0]}\left|\frac{1}{n}\langle D^{(1)}_{k^*+k(x)}, D^{(2)}_{k^*+k(x)}\rangle_{\mathfrak{M}_X} - 2\tau^*(1-\tau^*)^2 x\right| \\
&= \sup_{x\in[-N,0]}\left|k(x)\left\{\frac{2k^*+k(x)}{n}\right\}\left(\frac{n-k^*}{n}\right)^2\|\delta_n\|^2_{\mathfrak{M}_X} - 2\tau^*(1-\tau^*)^2 x\right| \\
&= o(1),
\end{aligned}$$

where the last equality holds since $k^* = \lfloor n\tau^* \rfloor$, $k(x) = o(n)$ and $|k(x) \cdot \|\delta_n\|^2_{\mathfrak{M}_X} - x| = o(1)$ uniformly over $x \in [-N, 0]$. Applying similar arguments leads to

$$\sup_{x\in[0,N]} \left| \frac{1}{n} \langle D^{(3)}_{k^*+k(x)}, D^{(4)}_{k^*+k(x)} \rangle_{\mathfrak{M}_X} - \{-2(\tau^*)^2(1-\tau^*)x\} \right| = o(1).$$

The preceding two equations imply that the limit of the drift term in $R_{n,k^*+k(x)}$ is identical to that of $2\tau^*(1-\tau^*)U(x)$.

By Theorem 1 and equation (1.9) in Philipp (1980), there exist two independent sequences of one-parameter $\mathfrak{M}_X$-valued Gaussian processes $(Z^{(1)}_n(x) : x \in \mathbb{R})_{n\in\mathbb{N}}$ and $(Z^{(2)}_n(x) : x \in \mathbb{R})_{n\in\mathbb{N}}$ such that

$$\sup_{x\in[-N,0]} \left\| \|\delta_n\|_{\mathfrak{M}_X} \sum_{i=k^*+k(x)+1}^{k^*} \epsilon_i - Z^{(1)}_n(-x) \right\|^2_{\mathfrak{M}_X} = o_p(1) \tag{A24}$$

and

$$\sup_{x\in[0,N]} \left\| \|\delta_n\|_{\mathfrak{M}_X} \sum_{i=k^*}^{k^*+k(x)+1} \epsilon_i - Z^{(2)}_n(x) \right\|^2_{\mathfrak{M}_X} = o_p(1). \tag{A25}$$

Moreover, these two processes satisfy that $\mathrm{E}\{Z^{(j)}_n(x)\} = 0 \in \mathfrak{M}_X$ and

$$\mathrm{E}\left\{\langle Z^{(j)}_n(x_1), f_1\rangle_{\mathfrak{M}_X} \langle Z^{(j)}_n(x_2), f_2\rangle_{\mathfrak{M}_X}\right\} = \min\{x_1, x_2\} \langle \Sigma_{XX} f_1, f_2\rangle_{\mathfrak{M}_X}$$

for any $x_1, x_2 \in [0, N]$, $f_1, f_2 \in \mathfrak{M}_X$, $n \in \mathbb{N}$ and $j = 1, 2$. Next we will show that

$$\sup_{x\in[-N,0]} \left| \frac{1}{n} \langle E^{(1)}_{k^*+k(x)}, D^{(2)}_{k^*+k(x)} \rangle_{\mathfrak{M}_X} - 2\tau^*(1-\tau^*) \frac{\langle Z^{(1)}_n(-x), \delta_n\rangle}{\|\delta_n\|_{\mathfrak{M}_X}} \right| = o_p(1). \tag{A26}$$

To this end, note that

$$\begin{aligned}
&\sup_{x\in[-N,0]} \left| \frac{1}{n} \langle E^{(1)}_{k^*+k(x)}, D^{(2)}_{k^*+k(x)} \rangle_{\mathfrak{M}_X} - 2\tau^*(1-\tau^*) \frac{\langle Z^{(1)}_n(-x), \delta_n\rangle}{\|\delta_n\|_{\mathfrak{M}_X}} \right| \\
&= \sup_{x\in[-N,0]} \left| \frac{\{2k^* + k(x)\}(n-k^*)}{n^2} \|\delta_n\|_{\mathfrak{M}_X} \sum_{i=k^*}^{k^*+k(x)+1} \frac{\langle \epsilon_i, \delta_n\rangle_{\mathfrak{M}_X}}{\|\delta_n\|_{\mathfrak{M}_X}} \right. \\
&\quad \left. -2\tau^*(1-\tau^*) \frac{\langle Z^{(1)}_n(-x), \delta_n\rangle}{\|\delta_n\|_{\mathfrak{M}_X}} + \frac{k(x)\{2k^* + k(x)\}(n-k^*)}{n^3} \sum_{i=1}^{n} \langle \epsilon_i, \delta_n\rangle_{\mathfrak{M}_X} \right| \\
&\leq \sup_{x\in[-N,0]} H_1(x) + \sup_{x\in[-N,0]} H_2(x),
\end{aligned} \tag{A27}$$

where

$$\begin{aligned}
H_1(x) &= \left| \frac{\{2k^* + k(x)\}(n-k^*)}{n^2} \sum_{i=k^*}^{k^*+k(x)+1} \langle \epsilon_i, \delta_n\rangle_{\mathfrak{M}_X} - 2\tau^*(1-\tau^*) \frac{\langle Z^{(1)}_n(-x), \delta_n\rangle}{\|\delta_n\|_{\mathfrak{M}_X}} \right|, \\
H_2(x) &= \left| \frac{k(x)\{2k^* + k(x)\}(n-k^*)}{n^3} \sum_{i=1}^{n} \langle \epsilon_i, \delta_n\rangle_{\mathfrak{M}_X} \right|.
\end{aligned}$$

By (A24) and $k(x)/n = O(1/(n\|\delta_n\|^2_{\mathfrak{M}_X}))$ uniformly over $x \in [-N, 0]$, we have

$$\sup_{x\in[-N,0]} H_2(x) = O(1) \frac{1}{\sqrt{n}\|\delta_n\|_{\mathfrak{M}_X}} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\langle \epsilon_i, \delta_n\rangle_{\mathfrak{M}_X}}{\|\delta_n\|_{\mathfrak{M}_X}} = o_p(1). \tag{A28}$$

Moreover, since $\{2k^* + k(x)\}(n-k^*)/n^2 \to 2\tau^*(1-\tau*)$ as $n \to \infty$,

$$\sup_{x\in[-N,0]} H_1(x) = O(1) \sup_{x\in[-N,0]} \left| \left\langle \|\delta_n\|_{\mathfrak{M}_X} \sum_{i=k^*}^{k^*+k(x)+1} \epsilon_i - Z_n^{(1)}(-x), \frac{\delta_n}{\|\delta_n\|_{\mathfrak{M}_X}} \right\rangle_{\mathfrak{M}_X} \right|. \tag{A29}$$

Applying the Cauchy-Schwarz inequality yields that

$$\left| \left\langle \|\delta_n\|_{\mathfrak{M}_X} \sum_{i=k^*}^{k^*+k(x)+1} \epsilon_i - Z_n^{(1)}(-x), \frac{\delta_n}{\|\delta_n\|_{\mathfrak{M}_X}} \right\rangle_{\mathfrak{M}_X} \right| \le \left\| \|\delta_n\|_{\mathfrak{M}_X} \sum_{i=k^*}^{k^*+k(x)+1} \epsilon_i - Z_n^{(1)}(-x) \right\|_{\mathfrak{M}_X}.$$

This inequality together with equations (A24) and (A28) implies that

$$\sup_{x\in[-N,0]} H_1(x) = o_p(1). \tag{A30}$$

Combining (A27), (A28) and (A30), we establish (A26). Similar arguments lead to

$$\sup_{x\in[0,N]} \left| \frac{1}{n} \langle E^{(3)}_{k^*+k(x)}, D^{(4)}_{k^*+k(x)} \rangle_{\mathfrak{M}_X} - 2\tau^*(1-\tau^*) \frac{\langle Z_n^{(2)}(x), \delta_n \rangle}{\|\delta_n\|_{\mathfrak{M}_X}} \right| = o_p(1). \tag{A31}$$

For $x \in [-N, 0]$, define

$$\Psi_n(x) = \frac{\langle Z_n^{(1)}(-x), \delta_n \rangle_{\mathfrak{M}_X}}{\|\delta_n\|_{\mathfrak{M}_X}}.$$

Letting $\Psi(x) = \lim_{n\to\infty} \Psi_n(x)$, then $(\Psi(x) : x \in [-N, 0])$ is a one-parameter Gaussian process. By the definition of σ^2, $\mathrm{E}\{\Psi(x)\} = 0$ and $\mathrm{E}\{\Psi(x_1)\Psi(x_2)\} = \sigma^2 \min\{-x_1, -x_2\}$ for any $x_1, x_2 \in [-N, 0]$. Thus $\Psi(x) \overset{d}{=} \sigma B(-x)$ as a stochastic defined on $[-N, 0]$, where $B(x)$ denotes the Brownian motion defined on $[0, N]$. Using the same arguments for $Z_n^{(2)}(x)$, we define $\Psi(x)$ for $[0, N]$ and show that $\Psi(x) \overset{d}{=} \sigma B(x)$ for $x \in [0, N]$. Therefore, to establish the weak convergence of $(R_{n,k^*+k(x)} : x \in [-N, N])$, it suffices to the remaining terms on the right-hand side of (A2) and (A3) are asymptotically negligible in probability after we have established (A26) and (A31).

To study the limit of $1/n \cdot \langle D_k^{(1)}, E_k^{(2)} \rangle_{\mathfrak{M}_X}$, note that

$$\max_{k^*-N_\delta \le k \le k^*} \frac{1}{n} \langle D_k^{(1)}, E_k^{(2)} \rangle_{\mathfrak{M}_X} \le Q_1 + Q_2 + Q_3,$$

where

$$Q_1 = \max_{k^*-N_\delta \le k \le k^*} \frac{k-k^*}{n} \frac{n-k^*}{n} \sum_{i=1}^{k} \langle \epsilon_i, \delta_n \rangle_{\mathfrak{M}_X},$$

$$Q_2 = \max_{k^*-N_\delta \le k \le k^*} \frac{k-k^*}{n} \frac{n-k^*}{n} \sum_{i=1}^{k^*} \langle \epsilon_i, \delta_n \rangle_{\mathfrak{M}_X},$$

$$Q_3 = \max_{k^*-N_\delta \le k \le k^*} \frac{k-k^*}{n} \frac{k+k^*}{n} \frac{n-k^*}{n} \sum_{i=1}^{n} \langle \epsilon_i, \delta_n \rangle_{\mathfrak{M}_X}.$$

By the definition of N_δ,

$$\max_{k^*-N_\delta \le k \le k^*} \left| \frac{k-k^*}{n} \right| \le \frac{N}{n\|\delta_n\|^2_{\mathfrak{M}_X}}.$$

It follows that

$$
\begin{aligned}
|Q_1| &\le O(1)\frac{N}{n\|\delta_n\|^2_{\mathfrak{M}_X}}\left|\max_{k^*-N_\delta\le k\le k^*}\sum_{i=1}^{k}\langle \epsilon_i,\delta_n\rangle_{\mathfrak{M}_X}\right| \\
&= O(1)\frac{N}{\sqrt{n\|\delta_n\|^2_{\mathfrak{M}_X}}}\left|\max_{k^*-N_\delta\le k\le k^*}\frac{1}{\sqrt{n}}\sum_{i=1}^{k}\frac{\langle \epsilon_i,\delta_n\rangle_{\mathfrak{M}_X}}{\|\delta_n\|_{\mathfrak{M}_X}}\right| \\
&= o_p(1),
\end{aligned}
$$

where the last inequality holds by applying Theorem 1 and equation (1.9) in Philipp (1980) and the continuous mapping theorem. The same arguments imply that $Q_2 = o_p(1)$ and $Q_3 = o_p(1)$. Hence,

$$
\max_{k^*-N_\delta\le k\le k^*}\frac{1}{n}\langle D_k^{(1)}, E_k^{(2)}\rangle_{\mathfrak{M}_X} = o_p(1).
$$

Lastly, we determine the stochastic order of $1/n\cdot\langle E_k^{(1)}, E_k^{(2)}\rangle_{\mathfrak{M}_X}$. Applying the Cauchy-Schwarz inequality yields

$$
\begin{aligned}
&\left|\frac{1}{n}\langle E_k^{(1)}, E_k^{(2)}\rangle_{\mathfrak{M}_X}\right| \\
&\le \left|\frac{1}{n}\left\langle \sum_{i=k+1}^{k^*}\epsilon_i + \frac{k-k^*}{n}\sum_{i=1}^{n}\epsilon_i,\ \sum_{i=1}^{k}\epsilon_i + \sum_{i=1}^{k^*}\epsilon_i - \frac{k+k^*}{n}\sum_{i=1}^{n}\epsilon_i\right\rangle_{\mathfrak{M}_X}\right| \\
&\le \frac{1}{n}\left\|\sum_{i=k+1}^{k^*}\epsilon_i + \frac{k-k^*}{n}\sum_{i=1}^{n}\epsilon_i\right\|_{\mathfrak{M}_X}\cdot\left\|\sum_{i=1}^{k}\epsilon_i + \sum_{i=1}^{k^*}\epsilon_i - \frac{k+k^*}{n}\sum_{i=1}^{n}\epsilon_i\right\|_{\mathfrak{M}_X} \\
&\le \frac{1}{n}\left(\left\|\sum_{i=k+1}^{k^*}\epsilon_i\right\|_{\mathfrak{M}_X} + \left\|\frac{k-k^*}{n}\sum_{i=1}^{n}\epsilon_i\right\|_{\mathfrak{M}_X}\right)\left(\left\|\sum_{i=1}^{k}\epsilon_i\right\|_{\mathfrak{M}_X} + \left\|\sum_{i=1}^{k^*}\epsilon_i\right\|_{\mathfrak{M}_X} + \left\|\frac{k+k^*}{n}\sum_{i=1}^{n}\epsilon_i\right\|_{\mathfrak{M}_X}\right).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\max_{k^*-N_\delta\le k\le k^*}\left|\frac{1}{n}\langle E_k^{(1)}, E_k^{(2)}\rangle_{\mathfrak{M}_X}\right| \le& \frac{1}{n}\left(\max_{k^*-N_\delta\le k\le k^*}\left\|\sum_{i=k+1}^{k^*}\epsilon_i\right\|_{\mathfrak{M}_X} + \max_{k^*-N_\delta\le k\le k^*}\left\|\frac{k-k^*}{n}\sum_{i=1}^{n}\epsilon_i\right\|_{\mathfrak{M}_X}\right) \\
&\times\left(\max_{k^*-N_\delta\le k\le k^*}\left\|\sum_{i=1}^{k}\epsilon_i\right\|_{\mathfrak{M}_X} + \left\|\sum_{i=1}^{k^*}\epsilon_i\right\|_{\mathfrak{M}_X} + \max_{k^*-N_\delta\le k\le k^*}\left\|\frac{k+k^*}{n}\sum_{i=1}^{n}\epsilon_i\right\|_{\mathfrak{M}_X}\right) \\
&= \frac{1}{n}O_p\left(\sqrt{N_\delta} + \frac{N_\delta}{\sqrt{n}}\right)O_p(\sqrt{n}) = O_p\left(\sqrt{\frac{N_\delta}{n}} + \frac{N_\delta}{n}\right) \\
&= o_p(1),
\end{aligned}
$$

where the second equality follows from Theorem 1 and equation (1.9) in Philipp (1980) and the continuous mapping theorem. Using a similar argument we can show that the remaining two terms on the right-hand side of (A3) are asymptotically negligible in probability as $n\to\infty$. Thus the proof is completed.

□

Proof of Theorem 3.4. The proof of Theorem 3.4 follows from Lemmas 1 and 2 and the continuous mapping theorem. □

S.2 Additional details of sample-level implementations

In this section, we present more details regarding sample-level implementations for Section 4 in the main text.

To test the null hypothesis against the alternative one by using T_n defined in (3.7), an accurate estimate of the critical value of T_n under H_0, C_α, is needed. According to Theorem 3.1, estimating C_α amounts to finding the estimated eigenvalues of Σ_{XX}. Under H_0, we estimate μ_X and Σ_{XX} as $\hat{\mu}_X = 1/n \sum_{i=1}^n \kappa(X_i,)$ and the pooled covariance

$$\hat{\Sigma}_{XX} = \frac{1}{n}\sum_{i=1}^n \{\kappa(X_i,) - \hat{\mu}_X\} \otimes \{\kappa(X_i,) - \hat{\mu}_X\}.$$

Moreover, as pointed out by Wang (2012) and Li and Song (2017), when performing eigen-decomposition on $\hat{\Sigma}_{XX}$, one should restrict their attention to $\mathfrak{M}_X^0$, which is spanned by $\{\kappa(X_i, \cdot) - \hat{\mu}_X, i = 1, \ldots, n\}$, rather than the one spanned by $\{\kappa(X_i, \cdot), i = 1, \ldots, n\}$. To this end, we employ the coordinate representation method for operators introduced in Li and Song (2017) to conduct the eigen-decomposition on $\hat{\Sigma}_{XX}$.

Specifically, let $K_X = (\kappa(X_i, X_j))_{ij}$ denote the $n \times n$ Gram matrix. Let I_n and 1_n denote the $n \times n$ identity matrix and the n-dimensional column vector consisting of 1's, respectively. Then, by Proposition 3 of Li and Song (2017), the coordinate representation for $\hat{\Sigma}_{XX}$ with respect to $\mathfrak{M}_X^0$ is $G_X = n^{-1} Q K_X Q$, where $Q = I_n - 1_n 1_n^\intercal / n$. Any $\varphi \in \mathfrak{M}_X^0$ can be written as $\varphi = [\varphi]^\intercal Q b_X$, where $b_X = \{\kappa(X_1, \cdot), \ldots, \kappa(X_n, \cdot)\}^\intercal$ and $[\varphi]$ denotes the vector of the coordinates of φ with respect to Qb_X. Now we want to look for $\varphi \in \mathfrak{M}_X^0$ to maximize, subject to $\langle \varphi, \varphi \rangle_{\mathfrak{M}_X} = 1$,

$$\langle \varphi, \hat{\Sigma}_{XX}\varphi \rangle_{\mathfrak{M}_X} = n^{-1}[\varphi]^\intercal K_X Q K_X Q[\varphi] = n[\varphi] G_X^2 [\varphi],$$

where the last equality holds as $Q[\varphi] = [\varphi]$ by Proposition 2 of Li and Song (2017) and $Q^2 = Q$.

Since $\langle \varphi, \varphi \rangle_{\mathfrak{M}_X} = [\varphi]^\intercal G_X [\varphi]$, letting $z = G_X^{1/2}[\varphi]$, we have $\langle \varphi, \varphi \rangle_{\mathfrak{M}_X} = [\varphi]^\intercal G_X [\varphi] = 1$ if and only $z^\intercal z = 1$. We consider the Moore-Penrose inverse of G_X, denoted as $G_X^\dagger$, to obtain $[\varphi]$ from z: $[\varphi] = (G_X^\dagger)^{1/2} z$. This representation leads to the following standard eigen-decomposition problem: for $\nu = 1, \ldots, p$,

$$z_\nu = \arg\max_{z \in \mathbb{R}^n} z^\intercal G_X G_X^\dagger G_X z,$$

subject to $z^\intercal z = 1$ and $z^\intercal z_k = 0$ for any $k < \nu$. Namely, $z_1, \ldots, z_p$ are the first p eigenvectors of the matrix $G_X G_X^\dagger G_X$. Consequently, $\hat{\theta}_\nu$, the estimated νth largest eigenvalue of Σ_{XX}, is that of G_X, and the corresponding estimated eigenfunction is given by $\hat{\varphi}_\nu = z_\nu^\intercal (G_X^\dagger)^{1/2} Q b_X$ for $\nu = 1, \ldots, p$.

To select p, the number of eigenvalues, to approximate the limiting distribution of T_n under the null, we use the fraction of variance explained. That is, we choose p as

$$\hat{p} = \min\left\{ p : \frac{\sum_{\nu=1}^p \hat{\theta}_\nu}{\sum_{\nu \geq 1} \hat{\theta}_\nu} \geq 0.9 \right\}.$$

This strategy is widely adopted in the literature of FDA: see Zhu *et al.* (2014), Aue *et al.* (2018) and Hörmann *et al.* (2022), for example. Alternatively, we can adopt the leave-one-curve out cross-validation method proposed by Yao *et al.* (2005). Let $\hat{\kappa}^{(-i)}(X_i,)$ denote the predicted value for $\kappa(X_i,)$ based on (2.5) when the ith curve is removed from the preceding eigen-analysis. Then we choose optimal p by minimizing the leave-one-out sum of residual squares:

$$\sum_{i=1}^{n}\sum_{j\neq i}[\kappa(X_i, X_j) - \langle \hat{\kappa}^{(-i)}(X_i, \cdot), \kappa(X_j, \cdot)\rangle_{\mathfrak{M}_X}]^2.$$

S.3 Additional results of simulation studies

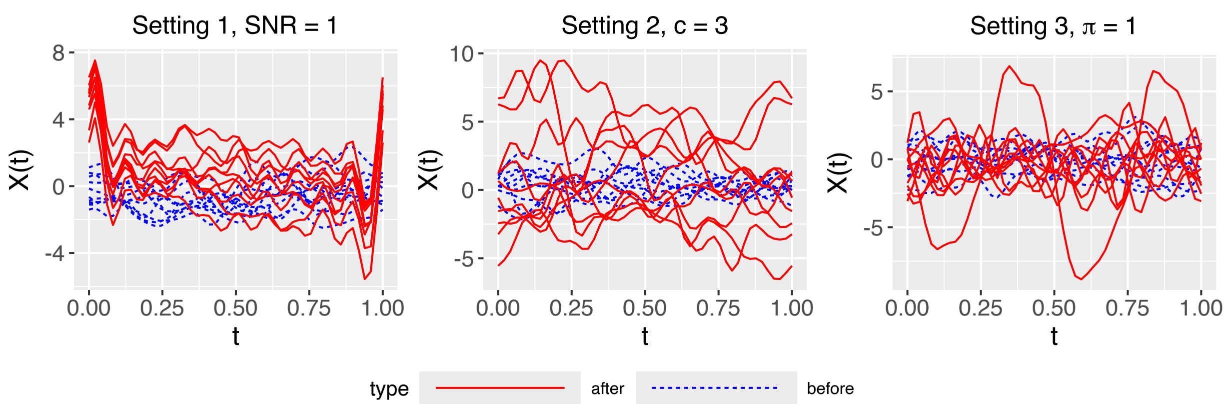

Figure S1: Profiles of functional data generated from the three settings. In each panel, red solid and blue dotted lines represent the trajectories of 10 randomly selected subjects after and before the structural break, respectively.

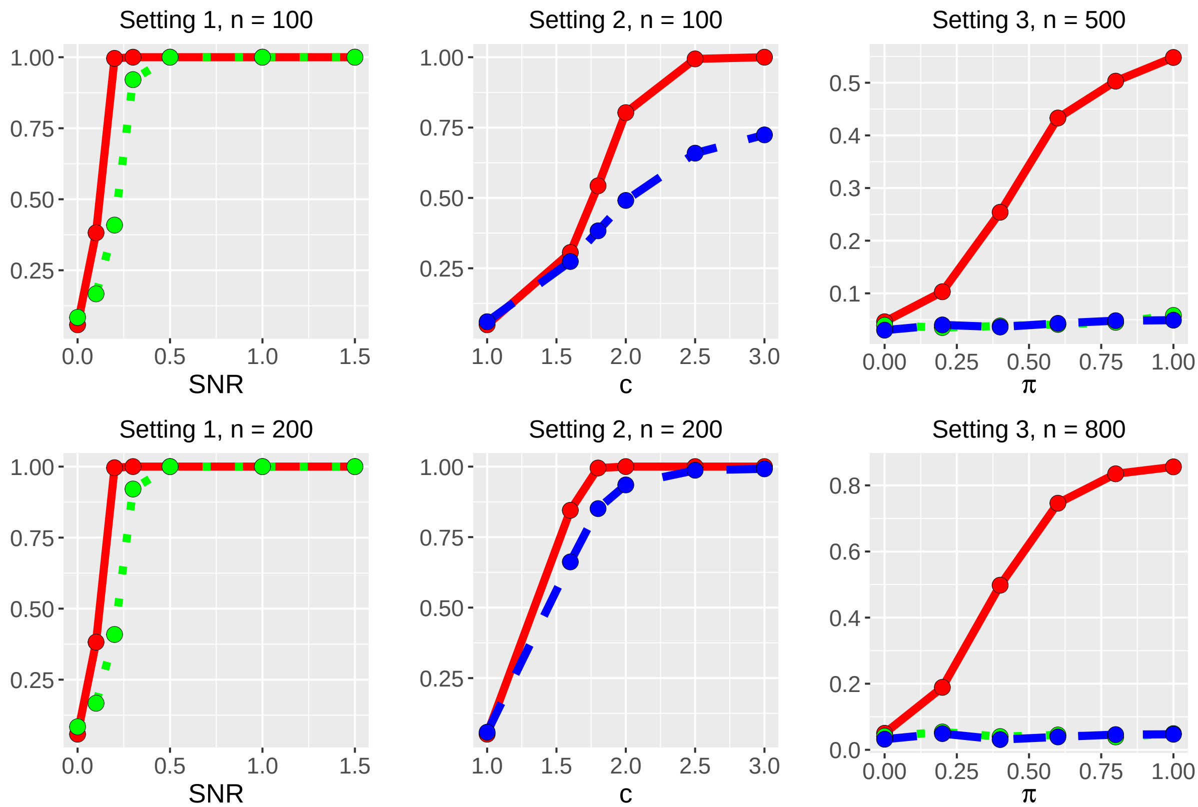

Figure S2: Power curves of the three detection methods under the three simulation settings with $\tau^* = 0.25$ for various sample sizes. In these panels, the red solid, green dotted and blue dashed lines represent the power curves of FF-dist, FF-mean and FF-cov, respectively.

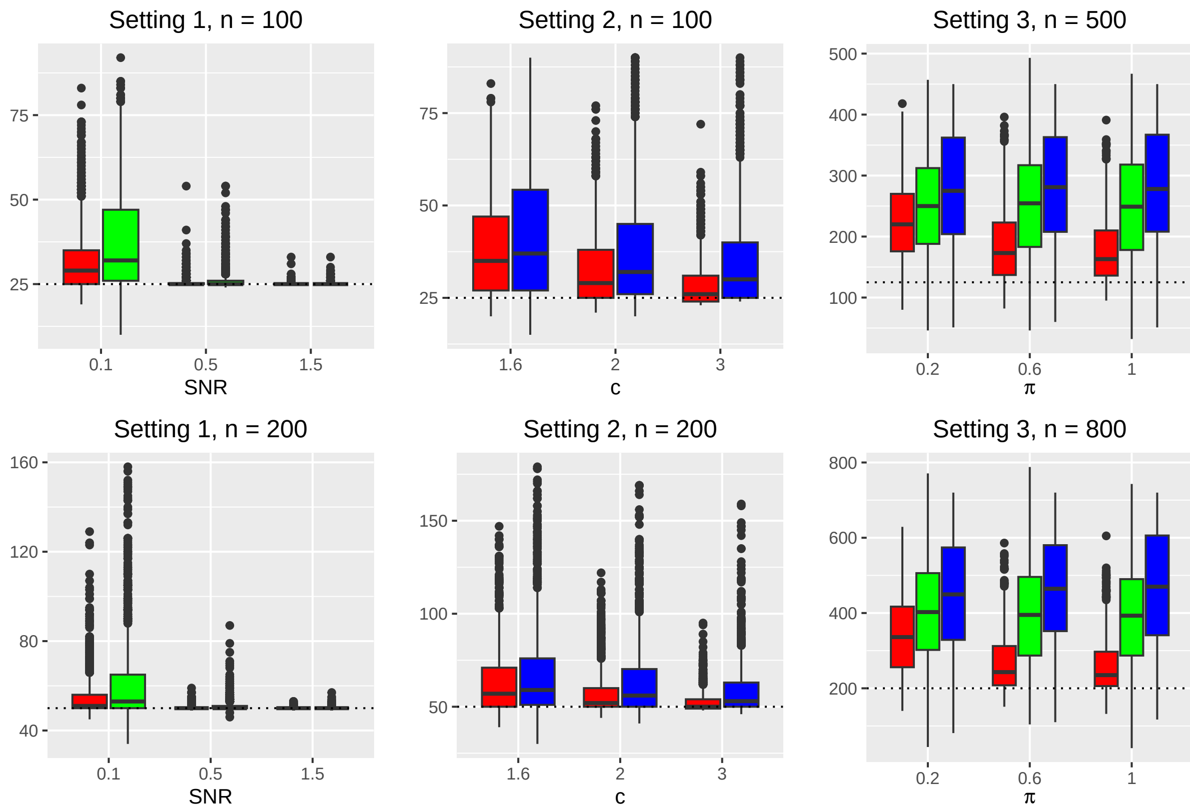

Figure S3: Boxplots of the estimated structural breaks for the three detection methods under the three simulation settings with $\tau^* = 0.25$ across 1000 Monte Carlo simulations. In these panels, the red, green and blue boxplots are for FF-dist, FF-mean and FF-cov, respectively.

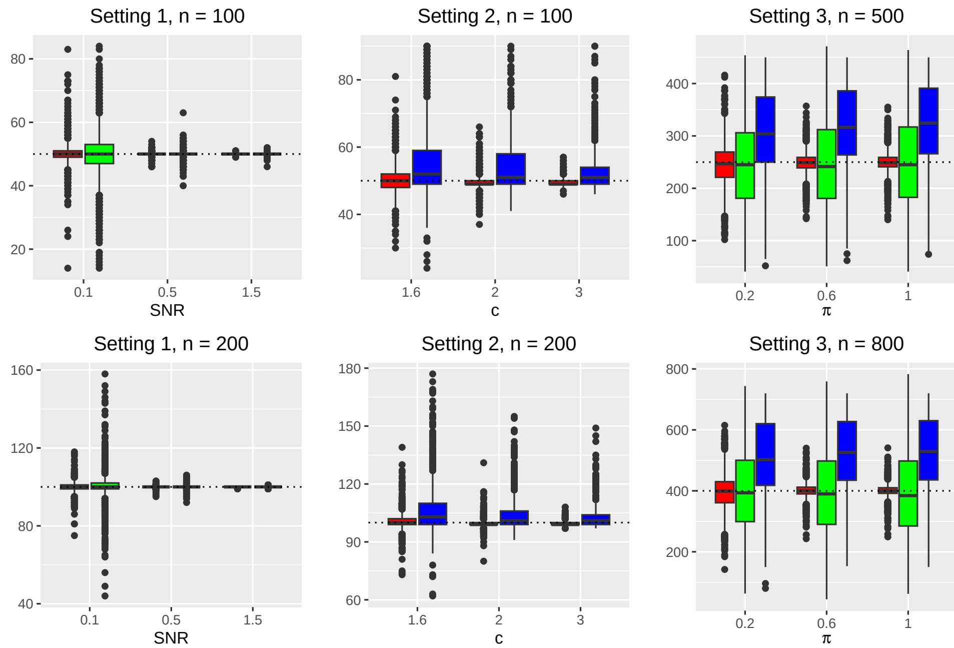

Figure S4: Boxplots of the estimated structural breaks for the three detection methods under the three simulation settings with $\tau^* = 0.50$ across 1000 Monte Carlo simulations. In these panels, the red, green and blue boxplots are for FF-dist, FF-mean and FF-cov, respectively. The dotted horizontal line represents the true structural break time.

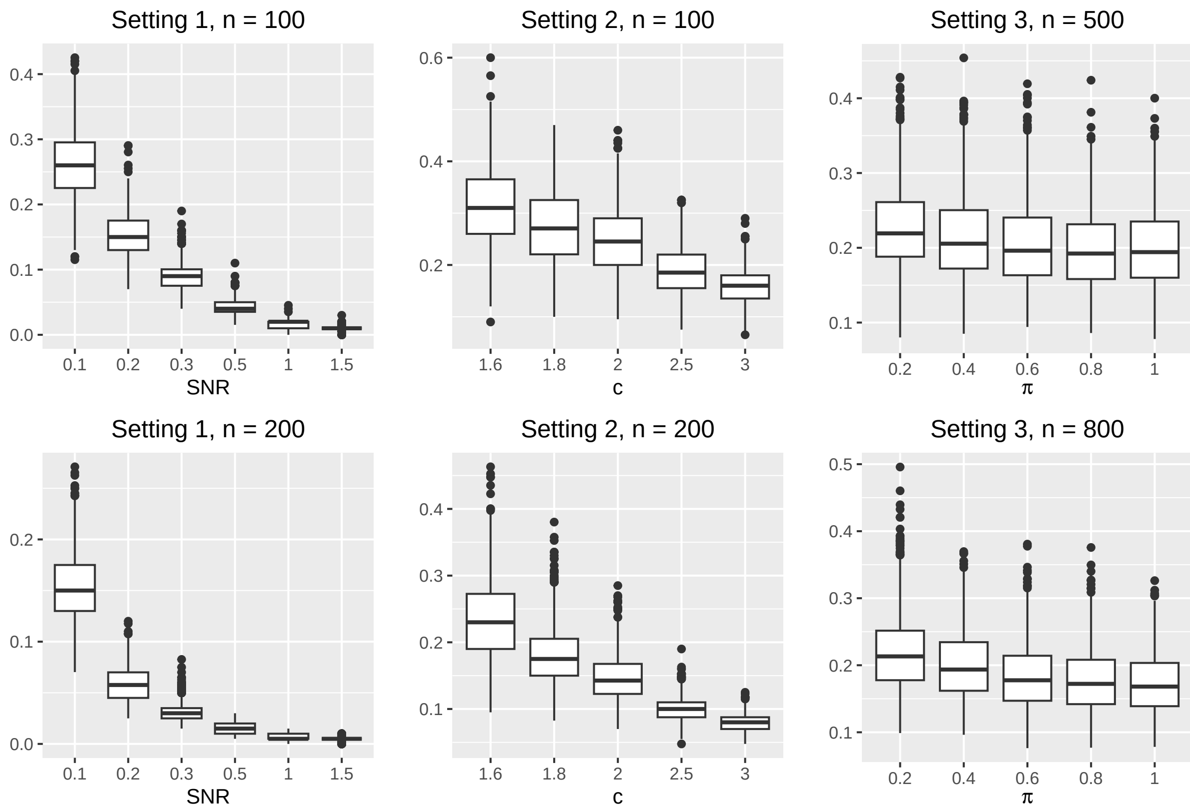

Figure S5: Boxplots of the lengths of the bootstrap confidence intervals for τ^* under the three simulation settings with $\tau^* = 0.25$ across 1000 Monte Carlo simulations.

S.3.1 Sensitivity analysis

In this subsection, we investigate the sensitivity of our method to the choice of γ and the reproducing kernel κ.

In addition to the median heuristic developed in Section 4.1, we further consider $\gamma = C_0 m^{-1}$, where $m = 50$ denotes the number of observations per curve in our simulations. This choice of γ is recommended by Wynne and Duncan (2022), where they show that this rate can ensure appealing theoretical properties. Here we consider $C_0 \in \{1, 3, 5\}$. Table S1 and Figure S6 summarize the size and power of our method in detecting the structure under different simulation settings with these candidate values of γ. They indicate that the performance of our method is sensitive to the choice of γ. More specifically, if γ becomes larger, the corresponding test becomes more conservative. Furthermore, comparing these numerical results with those obtained from the proposed median heuristic, we conclude that our proposed strategy of selecting the tuning parameter yields a powerful test while controlling the type I error reasonably well.

Setting	n	C_0		
		1	3	5
1	100	0.050	0.004	0.000
	200	0.072	0.012	0.000
2	100	0.028	0.006	0.000
	200	0.041	0.005	0.000
3	500	0.000	0.000	0.000
	800	0.003	0.000	0.000

Table S1: Empirical sizes of the proposed method with various tuning parameters under the three simulation settings with the nominal level $\alpha = 0.05$.

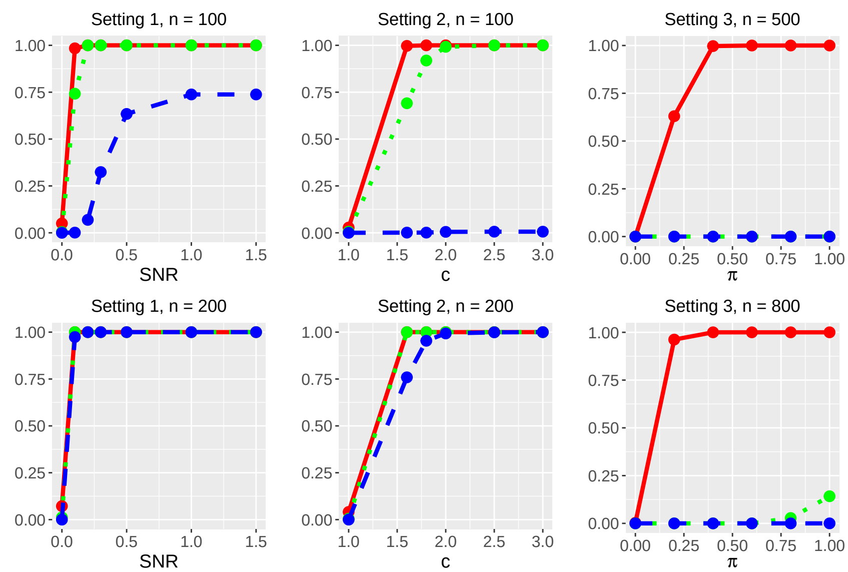

Figure S6: Power curves of our method with different choices of γ under the three simulation settings with $\tau^* = 0.50$ for various sample sizes. In these panels, the red solid, green dotted and blue dashed lines represent the power curves of $C_0 = 1$, 3, and 5, respectively.

Then we investigate the performance of our method with other kernels: the Laplace kernel and the quadratic kernel. Table S2 and Figure S7 display the size and power of our testing method with these two kernels under the three simulation settings. It turns out that the performance of our method with the Laplace kernel is similar to that with the Gaussian kernel in terms of detecting the structural break. In contrast, our method with the quadratic kernel cannot detect the structural break in any of these three settings. This study justifies the importance of selecting a proper kernel when implementing our method.

Setting	n	Kernel	
		Laplace	Quadratic
1	100	0.062	0.000
	200	0.080	0.000
2	100	0.043	0.000
	200	0.050	0.000
3	500	0.041	0.000
	800	0.046	0.000

Table S2: Empirical sizes of the proposed method with different kernels under the three simulation settings with the nominal level $\alpha = 0.05$.

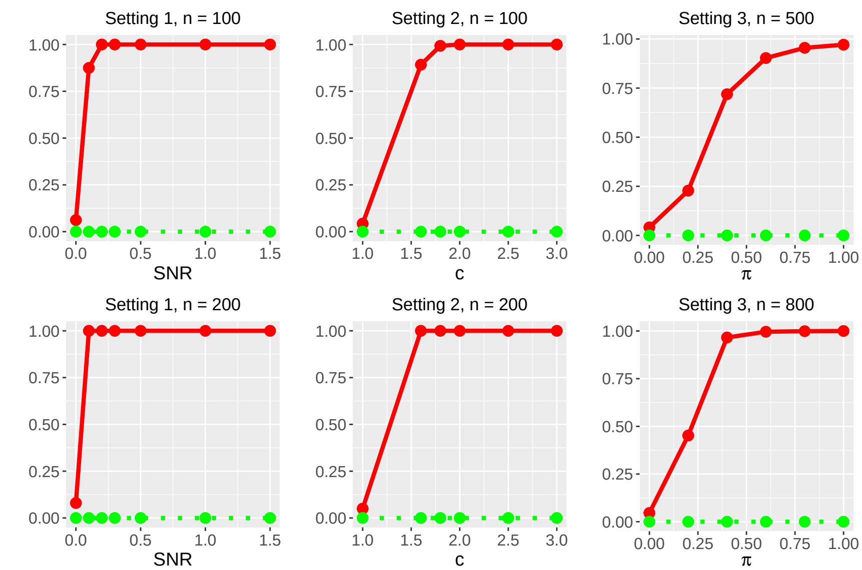

Figure S7: Power curves of our method with different choices of γ under the three simulation settings with $\tau^* = 0.50$ for various sample sizes. In these panels, the red solid, green dotted and blue dashed lines represent the power curves of $C_0 = 1$, 3, and 5, respectively.

S.4 Additional results of real data analysis

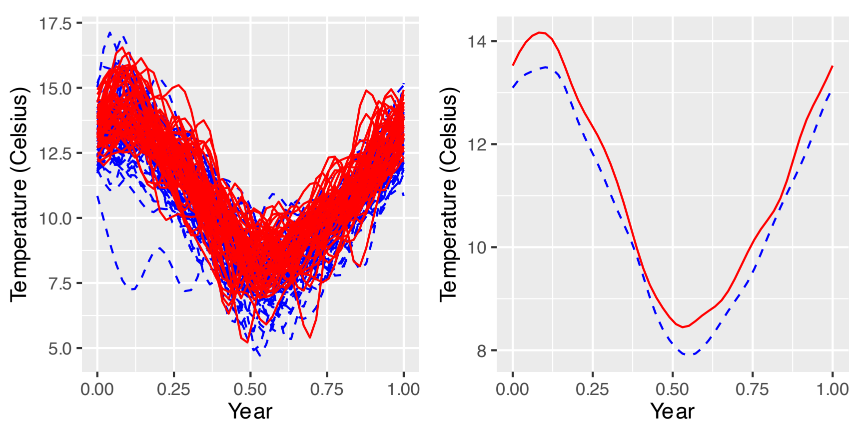

Figure S8: Left: profiles of minimum temperatures after and before the estimated structural break year based on FF-mean. Right: the mean functions of the temperature profiles after and before the estimated structural break. In both panels, the red solid and blue dashed lines represent the profiles from after and before the structural break, respectively.

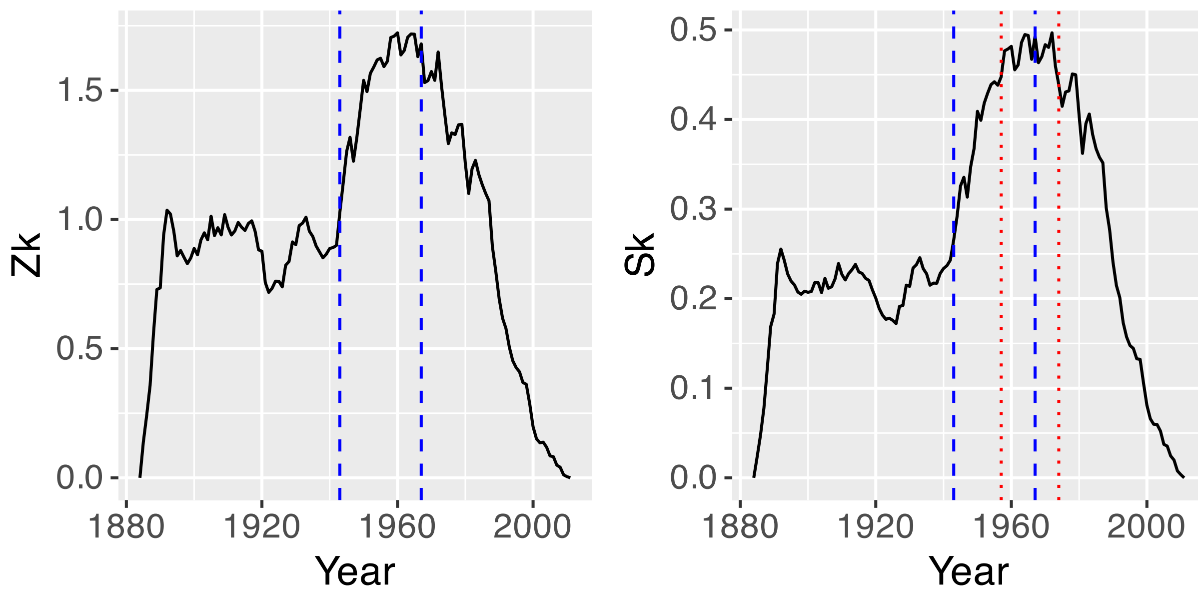

Figure S9: Trace of functional CUSUM statistics from FF-mean (left) and FF-dist (right).

References

Aue, A., Rice, G., and Sönmez, O. (2018). Detecting and dating structural breaks in functional data without dimension reduction. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **80**(3), 509–529.

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

Hörmann, S., Kuenzer, T., and Rice, G. (2022). Estimating the conditional distribution in functional regression problems. *Electronic Journal of Statistics*, **16**(2), 5751–5778.

Kuelbs, J. (1973). The invariance principle for Banach space valued random variables. *Journal of Multivariate Analysis*, **3**(2), 161–172.

Li, B. and Song, J. (2017). Nonlinear sufficient dimension reduction for functional data. *The Annals of Statistics*, **45**(3), 1059–1095.

Philipp, W. (1980). Weak and L^p-invariance principles for sums of B-valued random variables. *The Annals of Probability*, **8**(1), 68–82.

Wang, Q. (2012). Kernel principal component analysis and its applications in face recognition and active shape models. *arXiv preprint arXiv:1207.3538*.

Wynne, G. and Duncan, A. B. (2022). A kernel two-sample test for functional data. *Journal of Machine Learning Research*, **23**(73), 1–51.

Yao, F., Müller, H.-G., and Wang, J.-L. (2005). Functional data analysis for sparse longitudinal data. *Journal of the American Statistical Association*, **100**(470), 577–590.

Zhu, H., Yao, F., and Zhang, H. H. (2014). Structured functional additive regression in reproducing kernel Hilbert spaces. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **76**(3), 581–603.