

Supplementary Material for “Online Change Rate Learning for Functional Data”

This supplement provides the regularity assumptions, the corollary for the covariance function, the data-driven online bandwidth selection procedure, some extra simulations, an extended real data example and the technical results. To be more specific, Section S1 presents the regularity assumptions for the theoretical results, Section S2 derives a corollary for the covariance function, Section S3 details the data-driven online bandwidth selection procedure, Section S4 presents some extra simulation results, Section S5 analyzes an extended real data example, Section S6 introduces a lemma used in the proof of Theorem 2, and Sections S7–S10 present the respective proofs of the four theorems, respectively. For consistency, we also follow the same notations as those used in the main text.

S1. Assumptions

To derive the theoretical results for the proposed online estimators, we need some regularity assumptions as follows.

Assumption 1. (Time points and true functions)

(1.1) *The observed time points $\{T_{kij} : k = 1, \dots, K, i = 1, \dots, n_k, j = 1, \dots, m_{ki}\}$ are independent and identically distributed copies of a random variable $T \in [0, 1]$ with the probability density $f(\cdot)$. f is bounded from below and above,*

$$0 < m_f \leq \min_{t \in [0, 1]} f(t) \leq \max_{t \in [0, 1]} f(t) \leq M_f < \infty,$$

and the third-order derivative of $f(t)$, $f^{(3)}(t)$, is continuous and bounded.

(1.2) *The stochastic process X is independent of T , and the random noise ϵ is independent of T and also Φ which is the stochastic part of X .*

(1.3) *The third-order derivative of $\mu(t)$, $\mu^{(3)}(t)$, is continuous and bounded.*

(1.4) *The third-order partial derivative of $\gamma(s, t)$, $\gamma^{(u, v)}(s, t)$, $u + v = 3$, are continuous and bounded on $[0, 1]^2$.*

(1.5) *The first-order derivative of $\Phi(t)$, $\Phi^{(1)}(t)$, is continuous and bounded.*

Assumption 2. (Kernel for local quadratic smoother)

(2.1) *The one-dimensional kernel $W(\cdot)$ is a symmetric probability density function on $[-1, 1]$. $W(\cdot)$ is Lipschitz continuous, i.e. there exists a constant $0 < L_c < \infty$ such that*

$$|W(u) - W(v)| \leq L_c |u - v| \quad \text{for any } u, v \in [-1, 1].$$

Assumption 3. (Asymptotic normality for $\tilde{\mu}_{(K)}^{(1)}(t)$)

(3.1) The online bandwidth for $\tilde{\mu}_{(K)}^{(1)}(t)$, $\tilde{h}_{\mu(K)} \rightarrow 0$, $1/(S_{K,1}\tilde{h}_{\mu(K)}^3) \rightarrow 0$, and $S_{K,2}/S_{K,1}^2 \rightarrow 0$.

(3.2) $\max\{1/(S_{K,1}^2\tilde{h}_{\mu(K)}), S_{K,2}/S_{K,1}^3, S_{K,3}\tilde{h}_{\mu(K)}^3/S_{K,1}^3\}/\{1/(S_{K,1}\tilde{h}_{\mu(K)}^3) + S_{K,2}/S_{K,1}^2\}^{3/2} \rightarrow 0$.

(3.3) $\sup_{t \in [0,1]} E|\Phi(t)|^3 < \infty$, and $E|\epsilon|^3 < \infty$.

Assumption 4. (Asymptotic normality for $\tilde{\gamma}_{(K)}^{(1,0)}(s, t)$)

(4.1) The online bandwidth for $\tilde{\gamma}_{(K)}^{(1,0)}(s, t)$, $\tilde{h}_{\gamma(K)} \rightarrow 0$, $1/(S_{K,2}\tilde{h}_{\gamma(K)}^4) \rightarrow 0$, $S_{K,3}/(S_{K,2}^2\tilde{h}_{\gamma(K)}^3) \rightarrow 0$, and $S_{K,4}/S_{K,2}^2 \rightarrow 0$.

(4.2) $\max\{1/(S_{K,2}^2\tilde{h}_{\gamma(K)}^3), S_{K,3}/(S_{K,1}^2\tilde{h}_{\gamma(K)}^2), S_{K,4}/(S_{K,2}^3\tilde{h}_{\gamma(K)}), S_{K,5}/S_{K,2}^3, S_{K,6}\tilde{h}_{\gamma(K)}^3/S_{K,2}^3\}/\{1/(S_{K,2}\tilde{h}_{\gamma(K)}^4) + S_{K,3}/(S_{K,2}^2\tilde{h}_{\gamma(K)}^2) + S_{K,4}/S_{K,2}^2\}^{3/2} \rightarrow 0$.

(4.3) $\sup_{t \in [0,1]} E|\Phi(t)|^6 < \infty$, and $E|\epsilon|^6 < \infty$.

Assumption 5. (Data block)

(5.1) $\max_{k \leq K} s_{k,j}/S_{K,j} \rightarrow 0$ as $K \rightarrow \infty$ for $j = 1, 2$.

Remark: Assumptions (1.1)-(1.2) are common in FDA (Cao et al., 2016; Zhang and Wang, 2016; Zhang and Li, 2022). The independence assumption in (1.2) simplifies the theoretical derivations but does not capture certain practical scenarios, such as longitudinal or spatially clustered functional data, which often exhibit the dependence both within and across subjects (Mateu and Romano, 2017). If the observations satisfy the weak dependence (e.g., α -mixing), the consistency of our online estimators can be established via the nonparametric techniques analogous to the proofs used in the offline analysis (Horváth and Kokoszka, 2012). When the dependence is stronger, one may adjust the data-driven bandwidth selector by incorporating covariance-adjusted penalty terms (Brabanter et al., 2011), or introduce a forgetting factor into the recursive updates to down-weight older blocks (Chu and Chan, 2015). Extending these ideas to the offline derivative estimation for longitudinal or spatially clustered functional data remains undeveloped and considerably complex, while developing an online counterpart presents even greater challenges. As such, we leave a detailed theoretical analysis for future work. Assumptions (1.3)-(1.4) provide the necessary smoothness and boundedness conditions for the mean and covariance functions to derive the consistent and asymptotically normal estimators. Assumption (1.5) stipulates that Φ must be continuously differentiable, an additional condition necessary for estimating $\gamma^{(1,0)}(s, t)$ but not for estimating $\mu^{(1)}(t)$. As a result, when estimating $\gamma^{(1,0)}(s, t)$, the processes with high irregularities (e.g., infinite oscillations or jumps), such as Brownian motion and Poisson process, fall outside our current framework. While there is no established solution for handling the irregular processes, we outline two possible workarounds in the online framework. One may pre-smooth each trajectory via the kernel smoothing, penalized spline or wavelet shrinkage methods to obtain smooth and differentiable sample paths, thereby implying that Φ ,

the stochastic part of X , is differentiable (Wand and Jones, 1994; Antoniadis and Fan, 2001; Claeskens et al., 2009). Then OCRL can estimate $\gamma^{(1,0)}(s, t)$ directly and the induced pre-smoothing bias would need to be quantified and integrated into the theoretical analysis. Alternatively, one can target a generalized derivative by performing the local linear regression on the increments (Boker et al., 2011). In this variant, OCRL aggregates the weighted first-order difference quotients to yield a consistent estimator for the generalized derivative, thereby leading to distinct asymptotic behavior. A full theoretical development of these extensions lies beyond the scope of the present work but suggests promising directions for future research.

Assumption 2, which specifies the fundamental requirements for the kernel function, is widely used in the literature on the kernel-based method (Wand and Jones, 1994; Brabanter et al., 2011). It implies that $\alpha(W), \beta(W), R(W)$ and $U(W)$ in Theorems 1 and 2 are finite.

Assumptions 3 and 4 are introduced to establish the asymptotic properties of $\tilde{\mu}_{(K)}^{(1)}(t)$ and $\tilde{\gamma}^{(1,0)}(s, t)$ and are considered mild in nonparametric FDA (Li and Hsing, 2010; Zhang and Wang, 2016; Yang and Yao, 2023). Assumptions (3.1) and (4.1) ensure the consistency of our new estimators, serving as foundational conditions for local polynomial regression. Assumptions (3.2) and (4.2) ensure the applicability of the Lyapunov central limit theorem. Assumptions (3.3) and (4.3) provide the moment conditions necessary for the theoretical proofs.

Assumption 5 requires that the number of measurements per block should be comparable to, but significantly smaller than, the total number of measurements. This assumption is also reasonable and can be readily met in streaming data (Yang and Yao, 2023; Yang et al., 2024).

S2. Corollary for Covariance Function

The asymptotic normality of $\tilde{\gamma}^{(1,0)}(s, t)$ exhibits three distinct forms, determined by the relative order of the average number of measurements per subject $\bar{m}_{2,K}$ and N_K , where $\bar{m}_{2,K}$. Each form corresponds to a specific sampling scheme, reflecting the varying levels of the data density and their influence on the estimator's asymptotic behavior.

Corollary A.1. *Suppose that the conditions in Theorem 2 hold and $\limsup_K S_{K,2} S_{K,4} \tilde{h}_{\gamma(K)}^2 / S_{K,3}^2 = C_0 \in (0, \infty)$.*

(i) *When $\bar{m}_{2,K}/N_K \rightarrow 0$ and $\tilde{h}_{\gamma(K)} = O(S_{K,2}^{-1/8})$, we have*

$$\Gamma_\gamma(s, t) = \{1 + I(s = t)\} \left\{ \frac{R(W)U(W)}{\alpha^2(W)} \frac{E_1(s, t)}{f(s)f(t)} \right\}$$

and

$$\sqrt{\frac{S_{K,2}}{\Gamma_\gamma(s, t) \rho_{\gamma(K), -4}}} \left\{ \tilde{\gamma}_{(K)}^{(1,0)}(s, t) - \gamma^{(1,0)}(s, t) - \frac{1}{6} \frac{\beta(W)}{\alpha(W)} \gamma^{(3,0)}(s, t) \rho_{\gamma(K), 2} - \frac{1}{2} \alpha(W) \gamma^{(1,2)}(s, t) \rho_{\gamma(K), 2} \right\} \xrightarrow{d} N(0, 1).$$

(ii) When $\bar{m}_{2,K}/N_K \rightarrow C$ and $S_{K,4}\tilde{h}_{\gamma(K)}^4/S_{K,2} \rightarrow C_1^2$ with $0 < C, C_1 < \infty$, we have

$$\Gamma_\gamma(s, t) = \{1 + I(s = t)\} \left\{ \frac{R(W)U(W)}{\alpha^2(W)} \frac{E_1(s, t)}{f(s)f(t)} + \frac{C_1}{C_0^{\frac{1}{2}}} \frac{U(W)}{\alpha^2(W)} \frac{E_2(s, t)}{f(s)} \right\} \\ + C_1^2 \frac{E_\Phi \left[\{\Phi(s)\Phi(t)f^{(1)}(s) + \Phi^{(1)}(s)\Phi(t)f(s)\}^2 \right] - \{\gamma^{(1,0)}(s, t)f(s) + \gamma(s, t)f^{(1)}(s)\}^2}{f^2(s)}$$

and

$$\sqrt{\frac{S_{K,2}^2}{\Gamma_\gamma(s, t)S_{K,4}}} \left\{ \tilde{\gamma}_{(K)}^{(1,0)}(s, t) - \gamma^{(1,0)}(s, t) - \frac{1}{6} \frac{\beta(W)}{\alpha(W)} \gamma^{(3,0)}(s, t) \rho_{\gamma(K),2} - \frac{1}{2} \alpha(W) \gamma^{(1,2)}(s, t) \right. \\ \left. \rho_{\gamma(K),2} \right\} \xrightarrow{d} N(0, 1).$$

(iii) When $\bar{m}_{2,K}/N_K \rightarrow \infty$, $\tilde{h}_{\gamma(K)} = o(N_K^{-1})$, and $\bar{m}_{2,K}\tilde{h}_{\gamma(K)}^3 \rightarrow \infty$, we have

$$\Gamma_\gamma(s, t) = \frac{E_\Phi \left[\{\Phi(s)\Phi(t)f^{(1)}(s) + \Phi^{(1)}(s)\Phi(t)f(s)\}^2 \right] - \{\gamma^{(1,0)}(s, t)f(s) + \gamma(s, t)f^{(1)}(s)\}^2}{f^2(s)}$$

and

$$\sqrt{\frac{S_{K,2}^2}{\Gamma_\gamma(s, t)S_{K,4}}} \left\{ \tilde{\gamma}_{(K)}^{(1,0)}(s, t) - \gamma^{(1,0)}(s, t) \right\} \xrightarrow{d} N(0, 1).$$

In the online change rate estimation of the covariance function, functional data are categorized into three types based on their asymptotic distributions: non-dense, semi-dense and ultra-dense. Notably, Corollary A.1 reveals a discontinuity in the expression for the asymptotic variance, reflected as a jump in $\Gamma_\gamma(s, t)$ when $s = t$. However, this discontinuity vanishes when the data are ultra-dense.

S3. Bandwidth Selection

Based on Corollary 2 in the main text, the online learning bandwidths of $\mu_{(K)}^{(1)}(t)$ and $\gamma_{(K)}^{(1,0)}(s, t)$ in practical applications are

$$\tilde{h}_{\mu(K)}^{opt} = \left\{ 27 \frac{\alpha^2(W)}{\beta^2(W)} \frac{\tilde{\nu}_\mu}{\tilde{\theta}_\mu} \right\}^{1/7} S_{K,1}^{-1/7} \text{ and } \tilde{h}_{\gamma(K)}^{opt} = \left\{ 27 \frac{\alpha^2(W)}{\beta^2(W)} \frac{\tilde{\nu}_\gamma}{\tilde{\theta}_\gamma} \right\}^{1/8} S_{K,2}^{-1/8},$$

where

$$\tilde{\theta}_\mu = \int \left\{ \tilde{\mu}_{(K)}^{(3)}(t) \right\}^2 \tilde{f}_{(K)}(t) dt,$$

$$\begin{aligned}
\tilde{\nu}_\mu &= \int \tilde{\Gamma}_{\mu(K)}(t) \tilde{f}_{(K)}(t) dt, \\
\tilde{\theta}_\gamma &= \iint \left\{ \tilde{\gamma}_{(K)}^{(3,0)}(s, t) + 3 \frac{\alpha^2(W)}{\beta(W)} \tilde{\gamma}_{(K)}^{(1,2)}(s, t) \right\}^2 \tilde{f}_{(K)}(s) \tilde{f}_{(K)}(t) ds dt, \\
\tilde{\nu}_\gamma &= \iint \tilde{\Gamma}_{\gamma(K)}(s, t) \tilde{f}_{(K)}(s) \tilde{f}_{(K)}(t) ds dt.
\end{aligned}$$

For $\tilde{\theta}_\mu$, we estimate $\mu^{(3)}(t)$ using the online local quartic polynomial regression with the pilot bandwidth $h_{\tilde{\theta}_{\mu(K)}} = G_{\theta_\mu} S_{K,1}^{-1/9}$ and the candidate bandwidth sequence $\{\eta_{\theta_{\mu(K)}, \tilde{l}}\}_{\tilde{l}=1}^J$, where $\eta_{\theta_{\mu(K)}, \tilde{l}} = \{(J - \tilde{l} + 1)/J\}^{1/9}$ (Ruppert et al., 1995). We get $\tilde{\mu}_{(K)}^{(3)}(t) = \varepsilon_{5,4}^\top \{\tilde{P}_{\theta_{\mu(K)}, 1}(t)\}^{-1} \tilde{q}_{\theta_{\mu(K)}, 1}(t)$, where $\varepsilon_{q,v}$ is a q unit vector with 1 on the v th position. Meanwhile, we get $\tilde{f}_{(K)}(t) = \varepsilon_{5,1}^\top \{\tilde{P}_{\theta_{\mu(K)}, 1}(t)\}^{-1}$. Then, we integrate $\tilde{\mu}_{(K)}^{(3)}(t) \tilde{f}_{(K)}(t)$ over $[0, 1]$ to obtain $\tilde{\theta}_\mu$. Similarly, θ_γ can be estimated using the pilot online learning bandwidth $h_{\tilde{\theta}_{\gamma(K)}} = G_{\theta_\gamma} S_{K,2}^{-1/10}$ and the candidates $\{\eta_{\theta_{\gamma(K)}, \tilde{l}}\}_{\tilde{l}=1}^J$, where $\eta_{\theta_{\gamma(K)}, \tilde{l}} = \{(J - \tilde{l} + 1)/J\}^{1/10} h_{\tilde{\theta}_{\gamma(K)}}$. For $\tilde{\nu}_\mu$ and $\tilde{\nu}_\gamma$, Theorems 1 and 2 show different expressions under the non-dense, semi-dense and ultra-dense sampling schemes. We give the corresponding estimation method as follows.

- (a) $\tilde{\mu}_{(K)}^{(1)}(t)$ for the non-dense functional data. According to Case (i) in Corollary 1 in the main text, we have

$$\nu_\mu = \frac{U(W)}{\alpha^2(W)} \int \{\gamma(t, t) + \sigma^2\} dt.$$

Denote $r^2(t) = \gamma(t, t) + \sigma^2$. We use the online method to estimate $r^2(t)$.

- (a.1) Obtain $\tilde{\mu}_{(K)}(t)$ by applying the online local linear regression to Y_{kij} where the pilot bandwidth is $h_{\tilde{\mu}_{(K)}} = G_\mu S_{K,1}^{-1/5}$ and the candidate bandwidth sequence is $\{\eta_{\mu(K), \tilde{l}}\}_{\tilde{l}=1}^J$ with $\eta_{\mu(K), \tilde{l}} = \{(J - \tilde{l} + 1)/J\}^{1/5} h_{\tilde{\mu}_{(K)}}$.
- (a.2) Compute $\tilde{r}_{(K)}^2(T_{kij}) = \{Y_{kij} - \tilde{\mu}_{(K)}(T_{kij})\}^2$ and smooth $\tilde{r}_{(K)}^2(T_{kij})$ using the online local linear regression with $h_{\tilde{\mu}_{(K)}}$ and $\{\eta_{\mu(K), \tilde{l}}\}_{\tilde{l}=1}^J$.
- (b) $\tilde{\gamma}_{(K)}^{(1,0)}(s, t)$ for the non-dense functional data. Following the method of Yang and Yao (2023), we online estimate $\mu(t)$ by choosing the candidate bandwidth sequence of length 10. Then we get the raw covariance observations $\tilde{C}_{ki,j,l}$. According to Case (i) in Corollary A.1, we have

$$\nu_\gamma = \{1 + I(s = t)\} \frac{R(W)U(W)}{\alpha^2(W)} \iint E_1(s, t) ds dt.$$

Under model (1), $E_1(s, t) = E_\Phi \{\Phi^2(s) \Phi^2(t)\} + \sigma^2 \{\gamma(s, s) + \gamma(t, t)\} + \sigma^4$. We use the online method to estimate $E_1(s, t)$.

- (b.1) Obtain $\tilde{\gamma}_{(K)}(s, t)$ by applying the online local linear regression to $\tilde{C}_{ki,j,l}$ where the pilot bandwidth is $h_{\tilde{\gamma}_{(K)}} = G_{\gamma} S_{K,2}^{-1/6}$ and the candidate bandwidth sequence is $\{\eta_{\gamma_{(K)}, \tilde{l}}\}_{\tilde{l}=1}^J$ with $\eta_{\gamma_{(K)}, \tilde{l}} = \{(J - \tilde{l} + 1)/J\}^{1/6} h_{\tilde{\gamma}_{(K)}}$.
- (b.2) Compute $\tilde{\sigma}_{(K)}^2(t) = \tilde{r}_{(K)}^2(t) - \tilde{\gamma}_{(K)}(t, t)$, where $\tilde{r}_{(K)}^2(t)$ is obtained as in (a), and integrate $\tilde{\sigma}_{(K)}^2(t)$ to get $\tilde{\sigma}_{(K)}^2$.
- (b.3) Compute $\tilde{E}_{1(K)}(T_{kij}, T_{kil}) = \{\tilde{C}_{ki,j,l} - \tilde{\gamma}_{(K)}(T_{kij}, T_{kil})\}^2 + \tilde{\gamma}_{(K)}^2(T_{kij}, T_{kil})$ and smooth $\tilde{E}_{1(K)}(T_{kij}, T_{kil})$ using the online local linear regression with $h_{\tilde{\gamma}_{(K)}}$ and $\{\eta_{\gamma_{(K)}, \tilde{l}}\}_{\tilde{l}=1}^J$.
- (c) $\tilde{\mu}_{(K)}(t)$ for the semi-dense functional data. According to Case (ii) in Corollary 1 in the main text, we have

$$\begin{aligned} \nu_{\mu} = & \frac{U(W)}{\alpha^2(W)} \int \{\gamma(t, t) + \sigma^2\} dt + C_1 \int \gamma(t, t) \frac{\{f^{(1)}(t)\}^2}{f(t)} dt + C_1 \int \gamma^{(1,1)}(t, t) f(t) dt \\ & + 2C_1 \int \gamma^{(1,0)}(t, t) f^{(1)}(t) dt. \end{aligned}$$

It is the sum of the following four terms. We use the online method to estimate each term.

- (c.1) Obtain the first term with the same techniques as (a).
- (c.2) Estimate the second term. Specifically, obtain $\tilde{X}(t)$ by smoothing Y_{kij} using the online local linear regression with $h_{\tilde{\mu}_{(K)}}$ and $\{\eta_{\mu_{(K)}, \tilde{l}}\}_{\tilde{l}=1}^J$. Compute $\{Y_{kij} - \tilde{X}_{ki}(T_{kij})\}^2$ and integrate it to get $\tilde{\sigma}_{(K)}^2$. Compute $\tilde{\gamma}_{(K)}(t, t) = \tilde{r}_{(K)}^2(t) - \tilde{\sigma}_{(K)}^2$. Obtain $\tilde{f}_{(K)}^{(1)}(t)$ by applying the online local linear regression to $\tilde{f}_{(K)}(t)$ with $h_{\tilde{\mu}_{(K)}}$ and $\{\eta_{\mu_{(K)}, \tilde{l}}\}_{\tilde{l}=1}^J$. When estimating θ_{μ} online, we have $\tilde{f}_{(K)}(t)$.
- (c.3) Estimate the third term. Specifically, obtain $\tilde{\gamma}_{(K)}^{(1,1)}(t, t)$ by applying the online local cubic regression to $\tilde{r}_{(K)}^2(T_{kij})$ where the pilot bandwidth is $h_{\tilde{\mu}_{\nu_1(K)}} = G_{\mu_{\nu_1}} S_{K,1}^{-1/9}$ and the candidate bandwidth sequence is $\{\eta_{\mu_{\nu_1(K)}, \tilde{l}}\}_{\tilde{l}=1}^J$ with $\eta_{\mu_{\nu_1(K)}, \tilde{l}} = \{(J - \tilde{l} + 1)/J\}^{1/9} h_{\tilde{\mu}_{\nu_1(K)}}$.
- (c.4) Estimate the fourth term. Specifically, obtain $\tilde{\gamma}_{(K)}^{(1,0)}(t, t)$ by applying the online local quadratic regression to $\tilde{r}_{(K)}^2(T_{kij})$ where the pilot bandwidth is $h_{\tilde{\mu}_{\nu_2(K)}} = G_{\mu_{\nu_2}} S_{K,1}^{-1/7}$ and the candidate bandwidth sequence is $\{\eta_{\mu_{\nu_2(K)}, \tilde{l}}\}_{\tilde{l}=1}^J$ with $\eta_{\mu_{\nu_2(K)}, \tilde{l}} = \{(J - \tilde{l} + 1)/J\}^{1/7} h_{\tilde{\mu}_{\nu_2(K)}}$.
- (c.5) For C_1 , we provide two estimation methods. One is to let $C_1 = 1$, since the four terms are of the same order. The other is to solve the equation about C_1 , $C_1^7 S_{K,1}^{10}/S_{K,2}^7 - [27 \alpha^2(W) \tilde{\nu}_{\mu_{(K)}}(C_1)/\{\beta^2(W) \tilde{\theta}_{\mu}\}]^3 = 0$, where $\tilde{\nu}_{\mu_{(K)}}(C_1)$ is the function of C_1 with the coefficients based on (c.1)-(c.4).
- (d) $\tilde{\gamma}_{(K)}^{(1,0)}(s, t)$ for the semi-dense functional data. According to Case (ii) in Corollary A.1, we have

$$\begin{aligned} \nu_{\gamma} = & \{1 + I(s = t)\} \frac{R(W)U(W)}{\alpha^2(W)} \iint E_1(s, t) ds dt \\ & + \{1 + I(s = t)\} \frac{C_1}{C_0^{1/2}} \frac{U(W)}{\alpha^2(W)} \iint E_2(s, t) f(t) ds dt \end{aligned}$$

$$\begin{aligned}
& + C_1^2 [E_\Phi \{\Phi^2(s)\Phi^2(t)\} - \gamma^2(s, t)] \left\{ f^{(1)}(s) \right\}^2 \\
& + C_1^2 \left[E_\Phi \left[\left\{ \Phi^{(1)}(s) \right\}^2 \Phi^2(t) \right] - \left\{ \gamma^{(1,0)}(s, t) \right\}^2 \right] \\
& + 2C_1^2 \frac{E_\Phi \{\Phi(s)\Phi^{(1)}(s)\Phi^2(t)\} - \gamma(s, t)\gamma^{(1,0)}(s, t)}{f(s)} f^{(1)}(s).
\end{aligned}$$

It is the sum of the following five terms. We use the online method to estimate each term.

- (d.1) Obtain the first term with the same techniques as (b).
- (d.2) Estimate the second term. Specifically, under model (1), $E_2(s, t) = E_\Phi \{\Phi^2(s)\Phi^2(t)\} + \sigma^2 \gamma(t, t)$ and $\tilde{E}_{2(K)}(T_{kij}, T_{kil}) = \{\tilde{C}_{ki,j,l} - \tilde{\gamma}_{(K)}(T_{kij}, T_{kil})\}^2 + \tilde{\gamma}_{(K)}^2(T_{kij}, T_{kil}) - \tilde{\sigma}_{(K)}^4 - 2\tilde{\sigma}_{(K)}^2 \tilde{\gamma}_{(K)}^2(T_{kij}, T_{kij})$. Obtain $\tilde{E}_{2(K)}(s, t)$ with the similar techniques as (b).
- (d.3) Estimate the third term. Specifically, obtain $E_\Phi \{\Phi^2(s)\Phi^2(t)\} - \gamma^2(s, t)$ by applying the online local linear regression to $\{\tilde{C}_{ki,j,l} - \tilde{\gamma}_{(K)}(T_{kij}, T_{kil})\}^2 - \tilde{\sigma}_{(K)}^4 - 2\tilde{\sigma}_{(K)}^2 \tilde{\gamma}_{(K)}^2(T_{kij}, T_{kij})$ with $h_{\tilde{\gamma}_{(K)}}$ and $\{\eta_{\gamma_{(K)}, \tilde{l}}\}_{\tilde{l}=1}^J$. Obtain $\tilde{f}_{(K)}^{(1)}(s)$ by applying the online local linear regression to $\tilde{f}_{(K)}(s)$ with $h_{\tilde{\mu}_{(K)}}$ and $\{\eta_{\mu_{(K)}, \tilde{l}}\}_{\tilde{l}=1}^J$. When estimating θ_γ online, we have $\tilde{f}_{(K)}(s)$.
- (d.4) Estimate the fourth term. Specifically, obtain $E_\Phi [\{\Phi^{(1)}(s)\}^2 \Phi^2(t)]$ by applying the online local linear regression to $\tilde{C}_{ki,j,l}$ with $h_{\tilde{\gamma}_{(K)}}$ and $\{\eta_{\gamma_{(K)}, \tilde{l}}\}_{\tilde{l}=1}^J$. Obtain $\tilde{\gamma}_{(K)}(s, t)$ by (b.1). For $\tilde{\gamma}_{(K)}^{(1,0)}(s, t)$, apply the online local linear regression to $\tilde{\gamma}_{(K)}(s, t)$ with $h_{\tilde{\gamma}_{(K)}}$ and $\{\eta_{\gamma_{(K)}, \tilde{l}}\}_{\tilde{l}=1}^J$.
- (d.5) Estimate the fifth term. Specifically, obtain $E_\Phi \{\Phi(s)\Phi^{(1)}(s)\Phi^2(t)\} - \gamma(s, t)\gamma^{(1,0)}(s, t)$ by applying the online local quadratic regression to $\{\tilde{C}_{ki,j,l} - \tilde{\gamma}_{(K)}(T_{kij}, T_{kil})\}^2 - \tilde{\sigma}_{(K)}^4 - 2\tilde{\sigma}_{(K)}^2 \tilde{\gamma}_{(K)}^2(T_{kij}, T_{kij})$ where the pilot bandwidth is $h_{\tilde{\gamma}_{\nu_1(K)}} = G_{\gamma_{\nu_1}} S_{K,2}^{-1/8}$ and the candidate bandwidth sequence is $\{\eta_{\gamma_{\nu_1(K)}, \tilde{l}}\}_{\tilde{l}=1}^J$ with $\{\eta_{\gamma_{\nu_1(K)}, \tilde{l}}\}_{\tilde{l}=1}^J = \{(J - \tilde{l} + 1)/J\}^{1/8} h_{\tilde{\gamma}_{\nu_1(K)}}$.
- (d.6) Based on the expressions of $\tilde{h}_{\gamma_{(K)}}^{opt}$, C_1, C_0 and ν_γ , we have $H^8 \alpha^2(W) \tilde{\theta}_\gamma / \beta^2(W) - 27 \tilde{\nu}_{\gamma_{(K)}}(H) = 0$, where $H = [27 \alpha^2(W) \nu_\gamma / \{\beta^2(W) \theta_\gamma\}]^{1/8}$ and $\tilde{\nu}_{\gamma_{(K)}}(H)$ is the function of H with the coefficients obtained by (d.1)-(d.5).
- (e) $\tilde{\mu}_{(K)}^{(1)}(t)$ and $\tilde{\gamma}_{(K)}^{(1,0)}(s, t)$ for the ultra-dense functional data. According to Case (iii) in Corollaries 1 and A.1, the unknown quantities ν_μ and ν_γ can be obtained with the same techniques as the semi-dense functional data.

The above processes refer to the methods and theories proposed by Wand and Jones (1994), Fan and Yao (1998), Zhang and Chen (2007) and Ruppert et al. (1995). There are many other methods to estimate the unknown quantities. For the variance σ^2 , Wang and Yu (2017) provided two bias-corrected versions, which perform better in the cases of immense oscillation of regression function and small size sample. The proposed bandwidth selection procedure involves further bandwidth selection. In order to avoid the computational time caused by iteration, we specify the bandwidths in advance. Refer to Sheather and Jones (1991), Wand and Jones (1994) and Yang and Yao (2023), when $G_\mu, G_{\mu_{\nu_1}}, G_{\mu_{\nu_2}}, G_{\theta_\mu}, G_\gamma, G_{\theta_\gamma}, G_{\gamma_{\nu_1}} = O_p(1)$,

the online learning bandwidths $\tilde{h}_{\mu(K)}^{opt}$ and $\tilde{h}_{\gamma(K)}^{opt}$ can achieve the convergence rates in Theorem 3. For practical use, we set $G = 0.8^{1/d}$ where $d = 1$ for $\mu^{(1)}(t)$ and $d = 2$ for $\gamma^{(1,0)}(s, t)$. We substitute $\tilde{\theta}_\mu, \tilde{\theta}_\gamma, \tilde{\nu}_\mu$ and $\tilde{\nu}_\gamma$ in (7) with the offline estimation $\hat{\theta}_\mu, \hat{\theta}_\gamma, \hat{\nu}_\mu$ and $\hat{\nu}_\gamma$ to get the offline learning bandwidths $\hat{h}_{\mu(K)}^{opt}$ and $\hat{h}_{\gamma(K)}^{opt}$. Similar to the online bandwidth selection described above, $\hat{\theta}_\mu, \hat{\theta}_\gamma, \hat{\nu}_\mu$ and $\hat{\nu}_\gamma$ are obtained by the offline local polynomial regression.

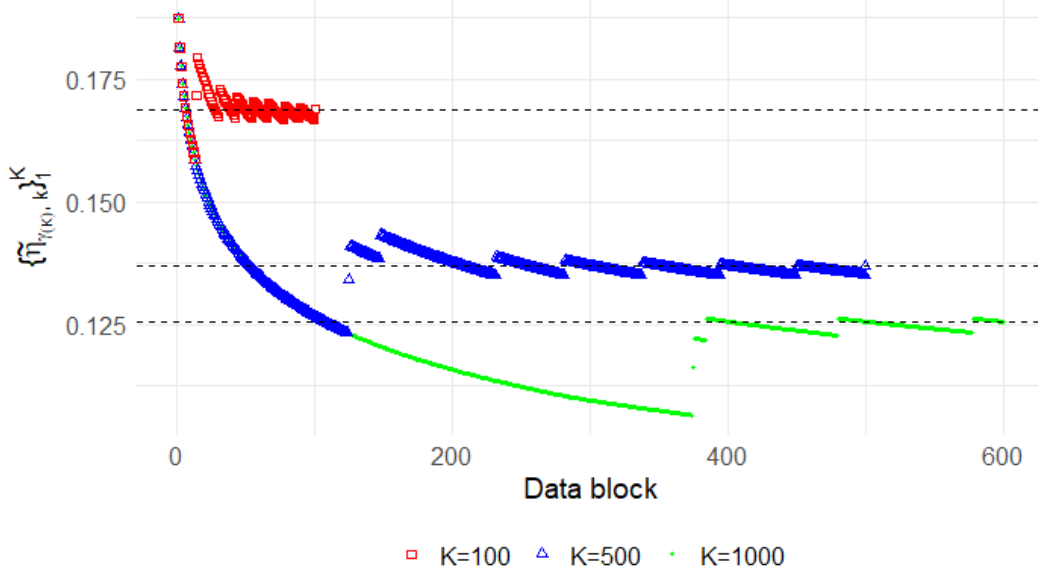


Figure S1: The pseudo-bandwidth sequence $\{\tilde{\eta}_{\gamma(K),k}\}_1^K$ with $L = 10$ under the dense sampling scheme based on the first 100, 500, 1000 data blocks.

S4. Additional Simulation Results

In this section, we provide extra simulation results, including a full sensitivity analysis of OCRL to the pilot parameter G , the convergence simulation on the pilot online estimates, as well as the parameter settings and comparison results of OCRL versus the spline, wavelet and stochastic gradient descent methods.

Figure S1 displays the pseudo-bandwidth sequence $\{\tilde{\eta}_{\gamma(K),k}\}_1^K$ with $L = 10$. We also illustrate the pseudo-bandwidth sequence $\{\tilde{\eta}_{(K),k}\}_1^K$ with $L = 10$ for $\tilde{\mu}_{(K)}^{(1)}(t)$ and $\tilde{\gamma}_{(K)}^{(1,0)}(s, t)$ of the non-dense sampling scheme based on the first 100, 500, 1000 data blocks in Figure S2. Furthermore, we assess the sensitivity of OCRL for G , considering $G \in [0.6, 1]$ for $\tilde{\mu}_{(K)}^{(1)}(t)$ and $G \in [0.7, 1.1]$ for $\tilde{\gamma}_{(K)}^{(1,0)}(s, t)$. Figure S3 presents the empirical mean integrated squared errors (eMISEs) under the dense and non-dense sampling schemes based on the first 300 data blocks. Over these ranges, OCRL is only weakly sensitive to G and the eMISE curves are relatively flat with a minimum near the recommended choice. The stopping rule for K can be determined by Theorem 3. Given a target convergence level $\tau > 0$ for the bandwidth error, the stopping index can be computed as $K_{stop} = \lceil K \geq 1 : S_{K,2} \geq \tau^{1/\alpha} \rceil$, where $\lceil \cdot \rceil$ denotes rounding up to the nearest integer. In our simulation setting, the cumulative information $S_{300,2}$ of the first 300 data blocks is

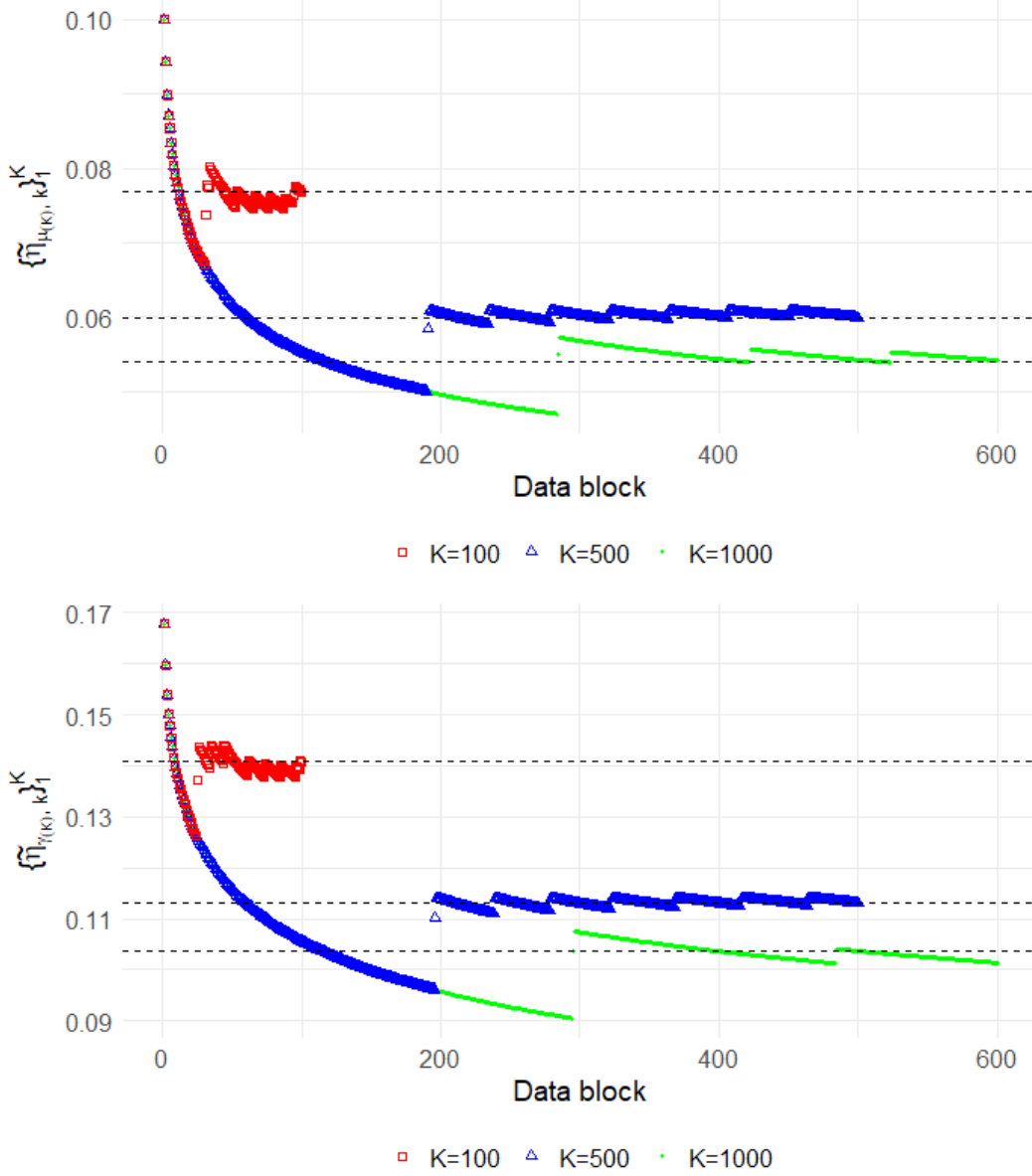


Figure S2: The pseudo-bandwidth sequence $\{\tilde{\eta}_{(K),k}^K\}$ with $L = 10$ for $\tilde{\mu}_{(K)}^{(1)}(t)$ and $\tilde{\gamma}_{(K)}^{(1,0)}(s, t)$ under the non-dense sampling scheme based on the first 100, 500, 1000 data blocks.

sufficiently large and the pilot-induced relative error in estimating $\gamma^{(1,0)}(s, t)$ falls below 5% by $K = 300$ regardless of dense or non-dense data. We also have conducted a convergence simulation on the pilot online estimates. Figure S4 reports that the convergence of $G_{\tilde{\gamma}^{(1,0)}}$, where $G_{\tilde{\gamma}^{(1,0)}} = \{27\alpha^2(W)\tilde{\nu}_\gamma/(\beta^2(W)\tilde{\theta}_\gamma)\}^{1/8}$. The L^1 -norm for the pilot differences of successive blocks falls below 10^{-4} by $K = 300$, and changes from $K = 300$ to $K = 500$ are negligible. This stabilization indicates that further updates would bring negligible statistical gains while adding disproportionate computational cost. Thus, we set $K = 300$ as the cutoff for the pilot updates, a choice that balances the statistical efficiency with computational efficiency.

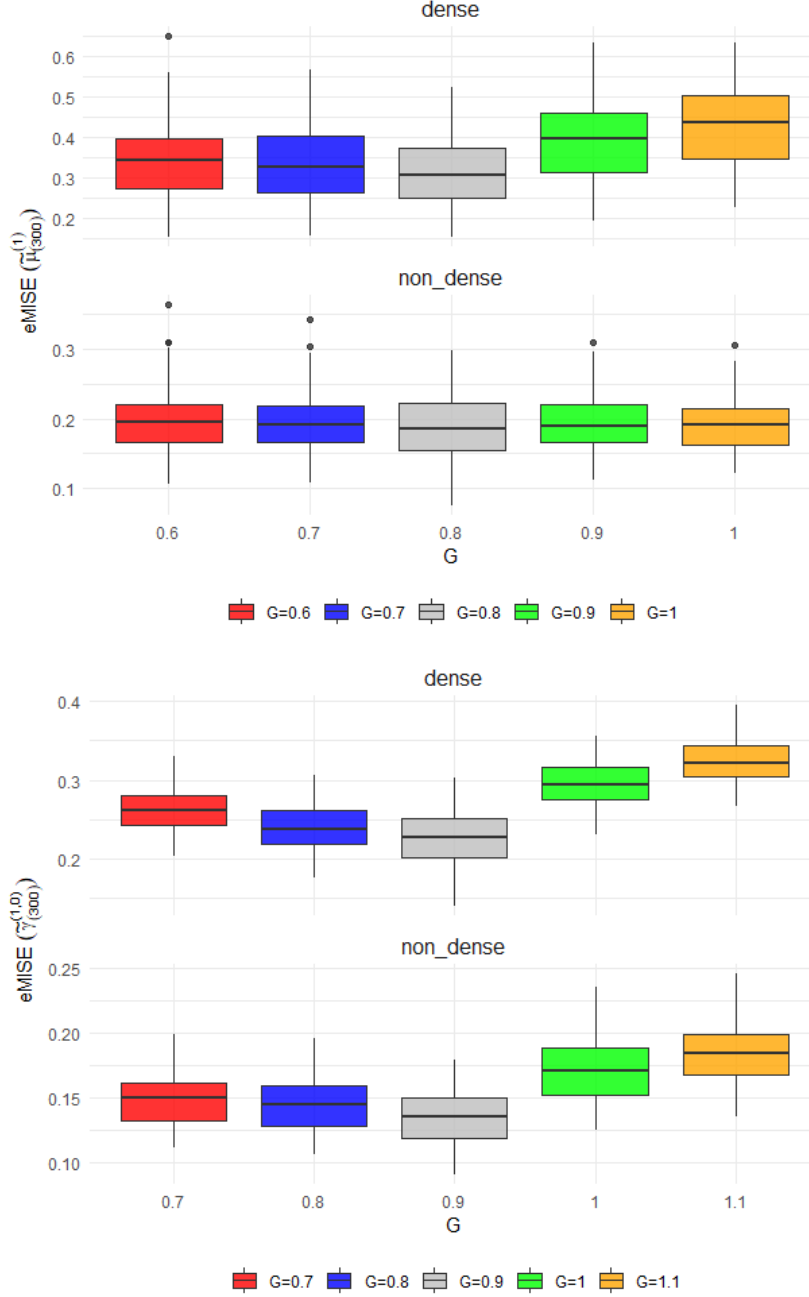


Figure S3: The eMISEs of $\tilde{\mu}^{(1)}(t)$ and $\tilde{\gamma}^{(1,0)}(s, t)$ for different G under the dense and non-dense sampling schemes based on the first 300 data blocks.

For the spline method, we employ the penalized B-spline smoother with the quartic basis and the third-order roughness penalty to estimate $\mu^{(1)}(t)$ and the tensor product of two cubic B-spline basis with penalties on the second-order partial derivatives in both directions to estimate $\gamma^{(1,0)}(s, t)$ (Sangalli et al., 2009; Wood et al., 2016). The smoothing parameters are chosen by the cross-validation. For the wavelet

method, we utilize the Daubechies extremal-phase wavelet to estimate $\mu^{(1)}(t)$ and the Daubechies least-asymmetric wavelet to estimate $\gamma^{(1,0)}(s, t)$, and moreover denoise the wavelet coefficients using universal hard-thresholding (Donoho and Johnstone, 1994; Pigoli and Sangalli, 2012). When estimating $\mu^{(1)}(t)$ and $\gamma^{(1,0)}(s, t)$ via SGD, we fit the penalized cubic B-spline with the second-order roughness penalty and adopt several stabilization techniques to improve the statistical accuracy (Bottou, 2010; Sutskever et al., 2013). Figures S5-S10 present the relative efficiencies and computational time of OCRL versus the spline, wavelet and SGD methods.

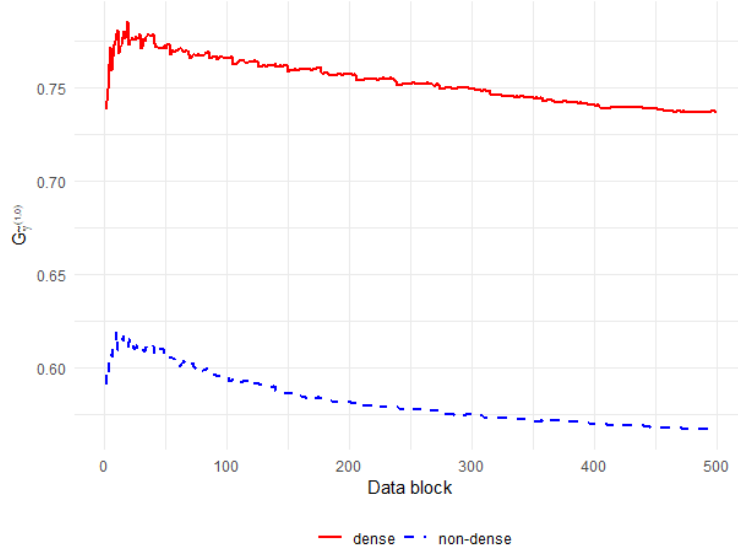


Figure S4: The convergence of $G_{\gamma^{(1,0)}}$ across the data block K under the dense and non-dense sampling schemes.

S5. Application to United States Traffic Accidents Data

We apply the proposed OCRL method to analyze the traffic accident data from the United States, which is known as a major public safety concern. Our findings highlight the significant role of the road conditions and the environmental factors in reducing accidents. Traffic accidents are a critical global public safety issue, with over a million lives lost annually and growing numbers each year. These accidents not only lead to a significant loss of life but also impose substantial economic costs and social burdens.

In this study, we analyze the traffic accident data from the United States between January 1, 2018 and March 31, 2023. The dataset, which includes approximately 6.6 million accident records, is publicly available at US Accidents Dataset https://smoosavi.org/datasets/us_accidents (Moosavi, Samavatian, Parthasarathy and Ramnath, 2019; Moosavi, Samavatian, Parthasarathy, Teodorescu and Ramnath, 2019). Our primary focus is on the frequency of traffic accidents per hour. For the analysis, we treat each day as a block and each state as a subject. We apply a logarithmic transformation to the number of traffic

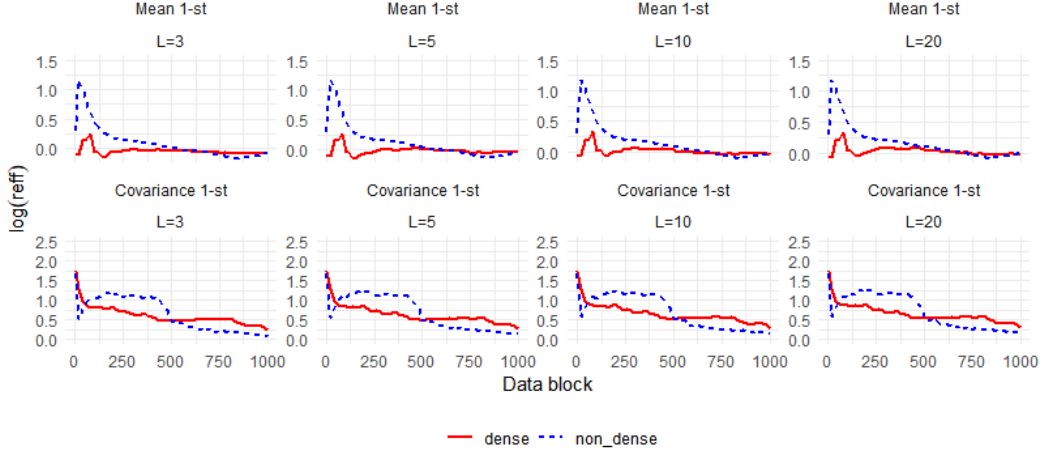


Figure S5: The empirical relative efficiencies of OCRL versus the spline method for $\tilde{\mu}_{(K)}^{(1)}(t)$ and $\tilde{\gamma}_{(K)}^{(1,0)}(s, t)$ under the dense (solid lines) and the non-dense data (dashed lines) with $L \in \{3, 5, 10, 20\}$. The y -axis is plotted on a logarithmic scale.

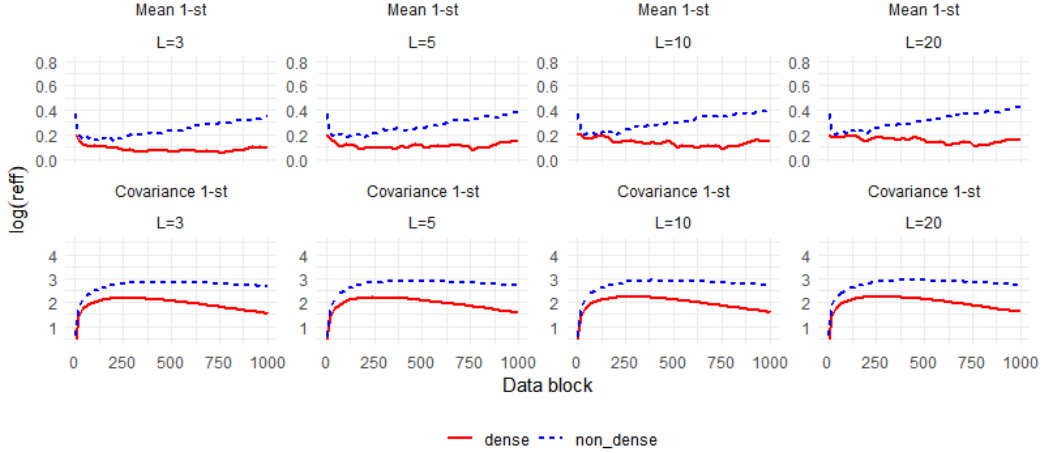


Figure S6: The empirical relative efficiencies of OCRL versus the wavelet method for $\tilde{\mu}_{(K)}^{(1)}(t)$ and $\tilde{\gamma}_{(K)}^{(1,0)}(s, t)$ under the dense (solid lines) and the non-dense data (dashed lines) with $L \in \{3, 5, 10, 20\}$. The y -axis is plotted on a logarithmic scale.

accidents for each hour, denoted as Y_{kij}^0 , with the transformed measurement given by $Y_{kij} = \log(Y_{kij}^0 + 1)$, as specified in model (1). We assume that the number of traffic accidents between days is independent. To ensure the accuracy of the estimation, we excluded 68 dates with fewer than 8 observations and 6 subjects, resulting in a total of $K = 1848$ data blocks used in the analysis. The number of subjects n_k follows the discrete uniform distribution on the set $\{6, \dots, 10\}$. For the dense sampling scheme, we set the number

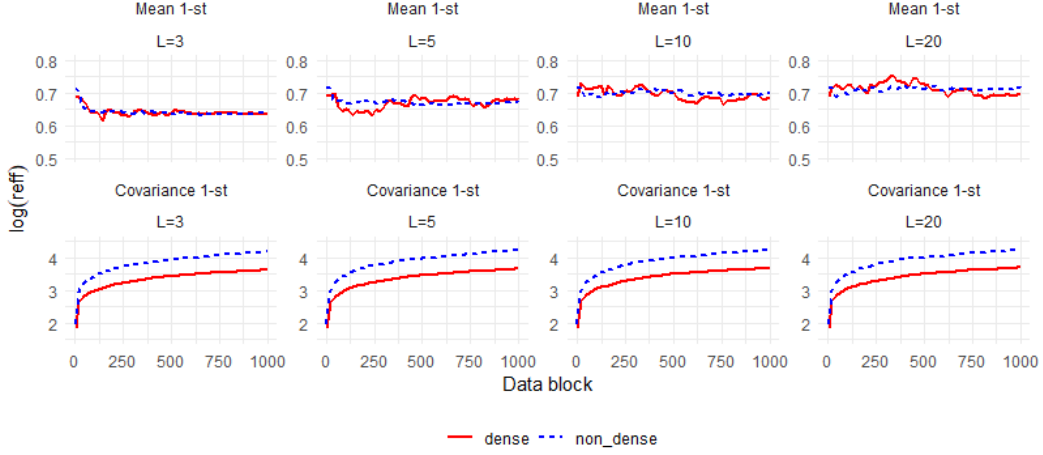


Figure S7: The empirical relative efficiencies of OCRL versus the SGD method for $\tilde{\mu}_{(K)}^{(1)}(t)$ and $\tilde{\gamma}_{(K)}^{(1,0)}(s, t)$ under the dense (solid lines) and the non-dense data (dashed lines) with $L \in \{3, 5, 10, 20\}$. The y -axis is plotted on a logarithmic scale.

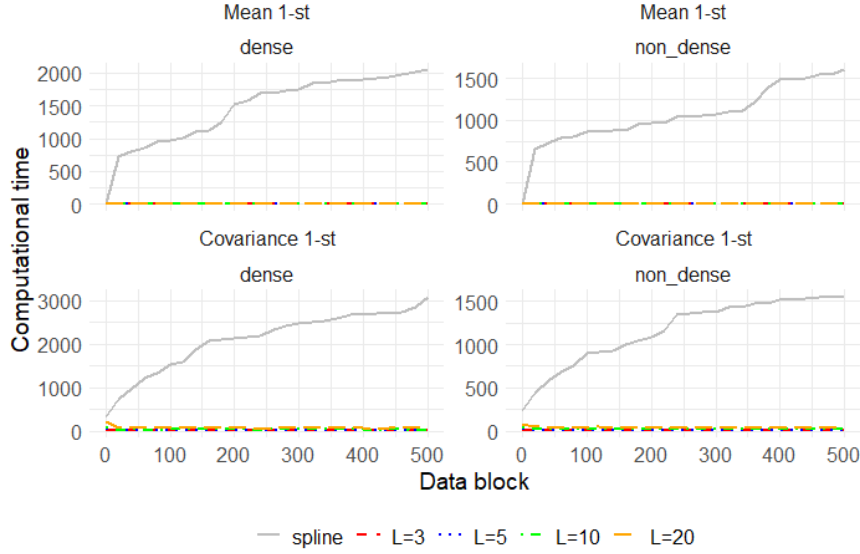


Figure S8: The comparison of the computational time in seconds between the OCRL and offline spline methods for $\mu^{(1)}(t)$ and $\gamma^{(1,0)}(s, t)$ of the dense and non-dense sampling schemes.

of measurements m_{ki} to 24, while for the non-dense sampling scheme, the number of measurements is randomly selected from $\{8, 9, 10\}$. The dense sampling scheme uses the same subjects in each block as non-dense sampling scheme, and the dense measurements include non-dense measurements for each sub-

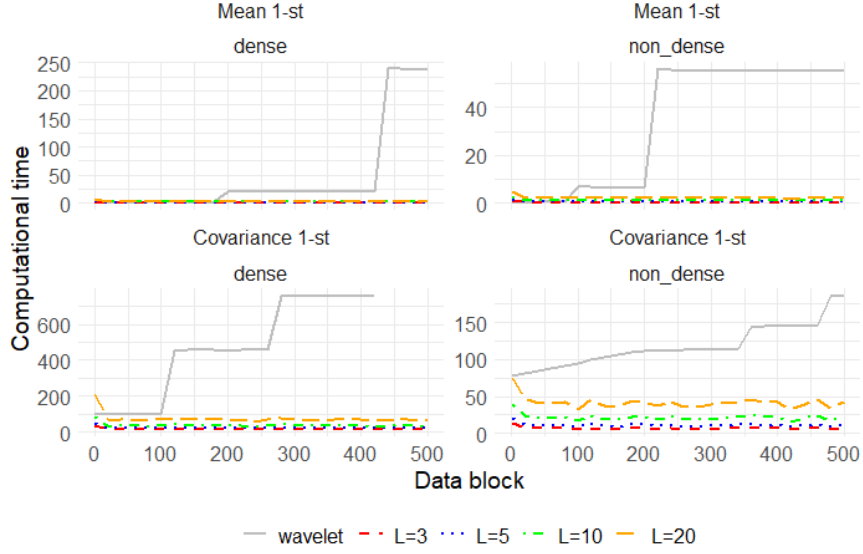


Figure S9: The comparison of the computational time in seconds between the OCRL and offline wavelet methods for $\mu^{(1)}(t)$ and $\gamma^{(1,0)}(s, t)$ of the dense and non-dense sampling schemes.

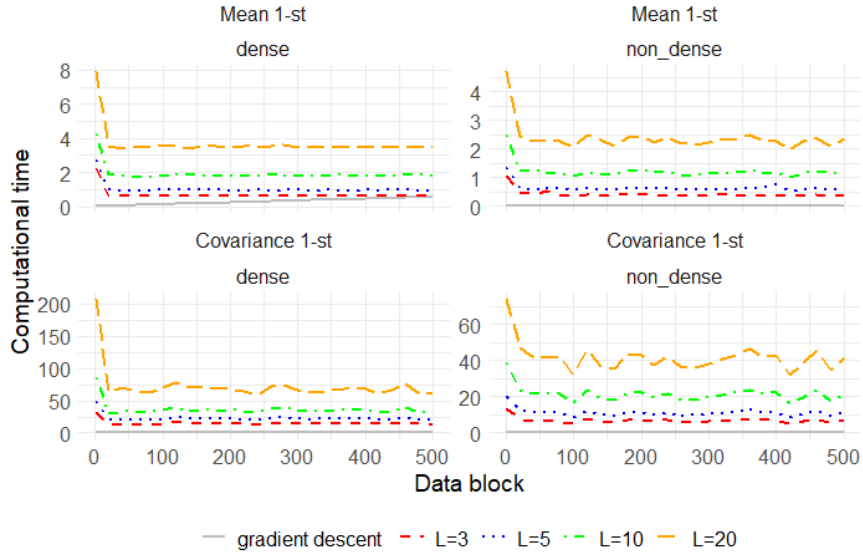


Figure S10: The comparison of the computational time in seconds between the OCRL and SGD methods for $\mu^{(1)}(t)$ and $\gamma^{(1,0)}(s, t)$ of the dense and non-dense sampling schemes.

ject. We use the offline change rate learning method for all 1848 data blocks as a baseline for comparison. The bandwidth selection follows the procedure outlined in Section S3.

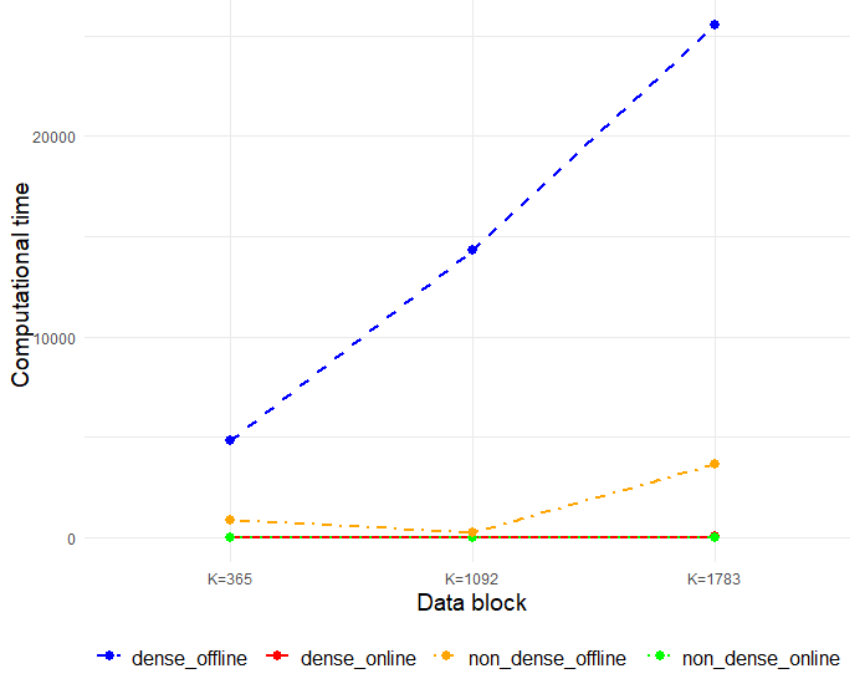


Figure S11: The comparison of the computational time in seconds between the OCRL and offline kernel methods for $\mu^{(1)}(t)$ and $\gamma^{(1,0)}(s, t)$ based on U.S. traffic accidents data, where $K = 365, 1092$ and 1783 correspond to the end of year 2018, 2020 and 2022, respectively.

We compare the computational time between the proposed OCRL method and offline kernel methods in Figure S11. The computational time required by the offline kernel method is at least several tens of times greater than that required by our new method. As illustrated by the example, the proposed online method greatly improves the computational speed and realizes real-time output of results. Figure S12 shows the first-order derivative estimation of the mean function at the end of year 2018, 2020 and 2022. OCRL demonstrates a significant trend towards the baseline for both the dense and non-dense data.

By examining the change rate of the mean function, we offer the following analysis. As more data are incorporated into the models, this trend becomes increasingly evident. Between approximately 2:00 and 5:00, the derivative is much greater than zero, indicating a rise in the number of traffic accidents, which tends to spike. During this period, the body's biological rhythms are at a low ebb, and drivers often experience fatigue and drowsiness. In addition, the reaction times and vision may be impaired when driving at night. The number of traffic accidents during the morning rush hour, which spans from 5:00 to 7:00, also increases due to the substantial rise in both vehicle and pedestrian traffic. However, the rate of increase during this period slows down. A small peak in traffic accidents is also observed around lunchtime. This may be attributed to people rushing to and from restaurants, thereby increasing the risk of accidents. The third peak occurs between 18:00 and 21:00. During this period, the frequency of traffic accidents shows a declining trend as people's off-work hours and evening entertainment activities become

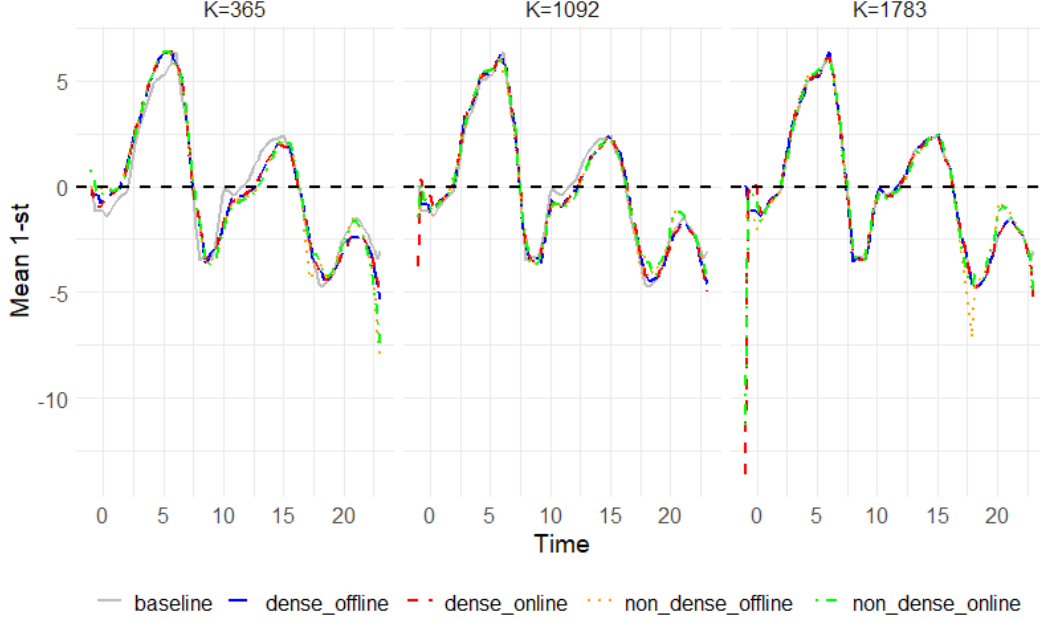


Figure S12: The first-order derivative estimation of the mean function using the OCRL and offline kernel methods based on the dense and non-dense cases of U.S. traffic accidents data. The baseline, represented by the solid line, shows $\hat{\mu}_{(1)}(t)$ using the offline method based on the full data from 2018 to 2023.

more spread out. Although this trend is decreasing, the rate of decline slows over time, likely due to the persistently high traffic flow.

S6. Lemma A.1 and Its Proof

Lemma A.1. *Following the same conditions as in Theorem 2, we let*

$$Q_{0R_{1,k}} = Q_{0,k'}(T_{kil}) = \frac{1}{s_{k',1}} \sum_{i'=1}^{n_{k'}} \sum_{j'=1}^{m_{k'i'}} \frac{1}{\tilde{\eta}_{\mu(K),k'}} W\left(\frac{T_{k'i'j'} - T_{kil}}{\tilde{\eta}_{\mu(K),k'}}\right),$$

$$\Delta\tilde{R}_{1,k} = -\frac{1}{s_{k',1}} \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \sum_{i'=1}^{n_{k'}} \sum_{j'=1}^{m_{k'i'}} \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W\left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) W\left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}}\right) \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) \frac{1}{\tilde{\eta}_{\mu(K),k'}}$$

$$W\left(\frac{T_{k'i'j'} - T_{kil}}{\tilde{\eta}_{\mu(K),k'}}\right) \{Y_{kij} - \mu(T_{kij})\} \left\{Y_{k'i'j'} - \mu(T_{k'i'j'}) + \frac{1}{2}\mu^{(2)}(T_{k'i'j'}) (T_{k'i'j'} - T_{kil})^2\right\}.$$

Then the mean and variance of $\Delta\tilde{R}_{1,k}/Q_{0R_{1,k}}$ are given as

$$E\left(\frac{\Delta\tilde{R}_{1,k}}{Q_{0R_{1,k}}}\right) = O_p\left(s_{k',1}^{-1} s_{k,2}^{-1} s_{k,3} \tilde{\eta}_{\gamma(K),k}\right),$$

$$\begin{aligned} \text{Var} \left(\frac{\Delta \tilde{R}_{1,k}}{Q_{0R_{1,k}}} \right) &= O_p \left(s_{k',1}^{-2} s_{k,2}^{-2} \left\{ s_{k,3}^2 \tilde{\eta}_{\gamma(K),k}^2 + s_{k,2} \left(s_{k,4} \tilde{\eta}_{\gamma(K),k}^2 + s_{k,3} \tilde{\eta}_{\gamma(K),k}^{-1} \right) + \tilde{\eta}_{\gamma(K),k}^{-2} \tilde{\eta}_{\mu(K),k}^{-1} \right. \right. \\ &\quad \left. \left. + s_{k',1} \left(s_{k,4} \tilde{\eta}_{\gamma(K),k}^{-1} + s_{k,3} \tilde{\eta}_{\gamma(K),k} \tilde{\eta}_{\mu(K),k}^{-1} \right) + s_{k,6} \tilde{\eta}_{\gamma(K),k}^2 + s_{k',1}^{-1} s_{k,3}^2 \tilde{\eta}_{\gamma(K),k}^2 \tilde{\eta}_{\mu(K),k}^{-1} \right\} \right). \end{aligned}$$

Proof. Firstly, we have $E(Q_{0R_{1,k}}) = f(T_{kil}) + \alpha(W) f^{(2)}(T_{kil}) \tilde{\eta}_{\mu(K),k}^2 / 2 + o_p(\tilde{\eta}_{\mu(K),k}^2)$ and

$$\begin{aligned} E \left(\Delta \tilde{R}_{1,k} \right) &= -\frac{1}{s_{k',1}} \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \sum_{i'=1}^{n_{k'}} \sum_{j'=1}^{m_{k'i'}} E \left[\frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}} \right) W \left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \right. \\ &\quad \left. \frac{1}{\tilde{\eta}_{\mu(K),k'}} W \left(\frac{T_{k'i'j'} - T_{kil}}{\tilde{\eta}_{\mu(K),k'}} \right) \{Y_{kij} - \mu(T_{kij})\} \left\{ Y_{k'i'j'} - \mu(T_{k'i'j'}) + \frac{1}{2} \mu^{(2)}(T_{k'i'j'}) (T_{k'i'j'} - T_{kil})^2 \right\} \right] \\ &= -\frac{1}{s_{k',1}} \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \sum_{i'=1}^K \sum_{i'=1}^{n_{k'}} \sum_{j'=1}^{m_{k'i'}} E \left[\frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}} \right) W \left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \right. \\ &\quad \left. \frac{1}{\tilde{\eta}_{\mu(K),k'}} W \left(\frac{T_{k'i'j'} - T_{kil}}{\tilde{\eta}_{\mu(K),k'}} \right) \{Y_{kij} - \mu(T_{kij})\} \left\{ Y_{k'i'j'} - \mu(T_{k'i'j'}) \right\} \right] \\ &= -\frac{1}{s_{k',1}} \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \sum_{j'=1}^{m_{ki}} E \left[\frac{1}{\tilde{\eta}_{\gamma(K),k}^2} \frac{1}{\tilde{\eta}_{\mu(K),k'}} W \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}} \right) W \left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \right. \\ &\quad \left. W \left(\frac{T_{kij'} - T_{kil}}{\tilde{\eta}_{\mu(K),k}} \right) \{Y_{kij} - \mu(T_{kij})\} \left\{ Y_{kij'} - \mu(T_{kij'}) \right\} \right] \\ &= O_p \left(s_{k',1}^{-1} s_{k,2}^{-1} s_{k,3} \tilde{\eta}_{\gamma(K),k} \right). \end{aligned}$$

Then by the delta method (Cox, 2005), it follows that $E(\Delta \tilde{R}_{1,k} / Q_{0R_{1,k}}) = \{E(\Delta \tilde{R}_{1,k}) / E(Q_{0R_{1,k}})\} \{1 + o_p(1)\} = O_p(s_{k',1}^{-1} s_{k,2}^{-1} s_{k,3} \tilde{\eta}_{\gamma(K),k})$.

Next, to derive the variance of $\Delta \tilde{R}_{1,k} / Q_{0R_{1,k}}$, we first compute the second moment of $\Delta \tilde{R}_{1,k}$ which can be expressed as

$$\begin{aligned} E \left(\Delta \tilde{R}_{1,k} \right)^2 &= \frac{1}{s_{k',1}^2} \frac{1}{s_{k,2}^2} E \left[\sum_{i_1=1}^{n_k} \sum_{1 \leq j_1 \neq l_1 \leq m_{ki_1}} \sum_{i'_1=1}^{n_{k'}} \sum_{j'_1=1}^{m_{k'i'_1}} \sum_{i_2=1}^{n_k} \sum_{1 \leq j_2 \neq l_2 \leq m_{ki_2}} \sum_{i'_2=1}^{n_{k'}} \sum_{j'_2=1}^{m_{k'i'_2}} \right. \\ &\quad \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) W \left(\frac{T_{ki_1l_1} - t}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \frac{1}{\tilde{\eta}_{\mu(K),k'}} W \left(\frac{T_{k'i'_1j'_1} - T_{ki_1l_1}}{\tilde{\eta}_{\mu(K),k'}} \right) \\ &\quad \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W \left(\frac{T_{ki_2j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) W \left(\frac{T_{ki_2l_2} - t}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_2j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \frac{1}{\tilde{\eta}_{\mu(K),k'}} W \left(\frac{T_{k'i'_2j'_2} - T_{ki_2l_2}}{\tilde{\eta}_{\mu(K),k'}} \right) \\ &\quad \left. \{Y_{ki_1j_1} - \mu(T_{ki_1j_1})\} \left\{ Y_{k'i'_1j'_1} - \mu(T_{k'i'_1j'_1}) + \frac{1}{2} \mu^{(2)}(T_{k'i'_1j'_1}) (T_{k'i'_1j'_1} - T_{ki_1l_1})^2 \right\} \right] \end{aligned}$$

$$\{Y_{ki_2j_2} - \mu(T_{ki_2j_2})\} \left\{ Y_{k'i'_2j'_2} - \mu(T_{k'i'_2j'_2}) + \frac{1}{2}\mu^{(2)}(T_{k'i'_2j'_2}) (T_{k'i'_2j'_2} - T_{ki_2l_2})^2 \right\} \Bigg].$$

We further decompose $E(\Delta\tilde{R}_{1,k})^2$ into four terms as follows:

$$\begin{aligned} (a) &= \frac{1}{s_{k',1}^2} \frac{1}{s_{k,2}^2} E \left\{ \sum \frac{1}{\tilde{\eta}_{\gamma(K),k}^4} \frac{1}{\tilde{\eta}_{\mu(K),k'}^2} W_{ki_1j_1}(s) W_{ki_1l_1}(t) W_{k'i'_1j'_1}(T_{ki_1l_1}) W_{ki_2j_2}(s) W_{ki_2l_2}(t) W_{k'i'_2j'_2}(T_{ki_2l_2}) \right. \\ &\quad \left. \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_2j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) (Y_{ki_1j_1} - \mu_{ki_1j_1}) (Y_{k'i'_1j'_1} - \mu_{k'i'_1j'_1}) (Y_{ki_2j_2} - \mu_{ki_2j_2}) (Y_{k'i'_2j'_2} - \mu_{k'i'_2j'_2}) \right\}, \\ (b) &= \frac{1}{2} \frac{1}{s_{k',1}^2} \frac{1}{s_{k,2}^2} E \left\{ \sum \frac{1}{\tilde{\eta}_{\gamma(K),k}^4} \frac{1}{\tilde{\eta}_{\mu(K),k'}^2} W_{ki_1j_1}(s) W_{ki_1l_1}(t) W_{k'i'_1j'_1}(T_{ki_1l_1}) W_{ki_2j_2}(s) W_{ki_2l_2}(t) W_{k'i'_2j'_2}(T_{ki_2l_2}) \right. \\ &\quad \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_2j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) (Y_{ki_1j_1} - \mu_{ki_1j_1}) (Y_{k'i'_1j'_1} - \mu_{k'i'_1j'_1}) (Y_{ki_2j_2} - \mu_{ki_2j_2}) \mu^{(2)}(T_{k'i'_2j'_2}) \\ &\quad \left. (T_{k'i'_2j'_2} - T_{ki_2l_2})^2 \right\}, \\ (c) &= \frac{1}{2} \frac{1}{s_{k',1}^2} \frac{1}{s_{k,2}^2} E \left\{ \sum \frac{1}{\tilde{\eta}_{\gamma(K),k}^4} \frac{1}{\tilde{\eta}_{\mu(K),k'}^2} W_{ki_1j_1}(s) W_{ki_1l_1}(t) W_{k'i'_1j'_1}(T_{ki_1l_1}) W_{ki_2j_2}(s) W_{ki_2l_2}(t) W_{k'i'_2j'_2}(T_{ki_2l_2}) \right. \\ &\quad \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_2j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) (Y_{ki_1j_1} - \mu_{ki_1j_1}) (Y_{k'i'_2j'_2} - \mu_{k'i'_2j'_2}) (Y_{ki_2j_2} - \mu_{ki_2j_2}) \mu^{(2)}(T_{k'i'_1j'_1}) \\ &\quad \left. (T_{k'i'_1j'_1} - T_{ki_1l_1})^2 \right\}, \\ (d) &= \frac{1}{4} \frac{1}{s_{k',1}^2} \frac{1}{s_{k,2}^2} E \left\{ \sum \frac{1}{\tilde{\eta}_{\gamma(K),k}^4} \frac{1}{\tilde{\eta}_{\mu(K),k'}^2} W_{ki_1j_1}(s) W_{ki_1l_1}(t) W_{k'i'_1j'_1}(T_{ki_1l_1}) W_{ki_2j_2}(s) W_{ki_2l_2}(t) W_{k'i'_2j'_2}(T_{ki_2l_2}) \right. \\ &\quad \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_2j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) (Y_{ki_1j_1} - \mu_{ki_1j_1}) (Y_{ki_2j_2} - \mu_{ki_2j_2}) \mu^{(2)}(T_{k'i'_2j'_2}) (T_{k'i'_1j'_1} - T_{ki_1l_1})^2 \\ &\quad \left. \mu^{(2)}(T_{k'i'_2j'_2}) (T_{k'i'_2j'_2} - T_{ki_2l_2})^2 \right\}, \end{aligned}$$

where $\sum = \sum_{i_1=1}^{n_k} \sum_{1 \leq j_1 \neq l_1 \leq m_{ki_1}} \sum_{i'_1=1}^{n_{k'}} \sum_{j'_1=1}^{m_{k'i'_1}} \sum_{i_2=1}^{n_k} \sum_{1 \leq j_2 \neq l_2 \leq m_{ki_2}} \sum_{i'_2=1}^{n_{k'}} \sum_{j'_2=1}^{m_{k'i'_2}}$, $W_{kij}(s) = W((T_{kij} - s)/\tilde{\eta}_{\gamma(K),k})$, $W_{k'i'_1j'_1}(T_{kil}) = W((T_{k'i'_1j'_1} - T_{kil})/\tilde{\eta}_{\mu(K),k'})$, and $\mu_{kij} = \mu_{(K)}(T_{kij})$.

For (a), we divide it into four nonzero main terms based on the k th block and the i th subject, yielding that

$$\begin{aligned}\Lambda_1^a &= \{(k, i_1) = (k', i'_1), (k, i_2) = (k', i'_2), (k, i_1) \neq (k, i_2)\}, \\ \Lambda_2^a &= \{(k, i_1) = (k, i_2), (k', i'_1) = (k', i'_2), (k, i_1) \neq (k', i'_1)\}, \\ \Lambda_3^a &= \{(k, i_1) = (k', i'_2), (k, i_2) = (k', i'_1), (k, i_1) \neq (k, i_2)\}, \\ \Lambda_4^a &= \{(k, i_1) = (k, i_2) = (k', i'_1) = (k', i'_2)\}.\end{aligned}$$

(i) Under Λ_1^a , we have

$$\begin{aligned}(a)|_{\Lambda_1^a} &= \frac{1}{s_{k,1}^2} \frac{1}{s_{k,2}^2} E \left\{ \sum \frac{1}{\tilde{\eta}_{\gamma(K),k}^4} \frac{1}{\tilde{\eta}_{\mu(K),k}^2} W_{ki_1j_1}(s) W_{ki_1l_1}(t) W_{ki_1j'_1}(T_{ki_1l_1}) W_{ki_2j_2}(s) W_{ki_2l_2}(t) W_{ki_2j'_2}(T_{ki_2l_2}) \right. \\ &\quad \left. \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_2j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) (Y_{ki_1j_1} - \mu_{ki_1j_1}) (Y_{ki_1j'_1} - \mu_{ki_1j'_1}) (Y_{ki_2j_2} - \mu_{ki_2j_2}) (Y_{ki_2j'_2} - \mu_{ki_2j'_2}) \right\} \\ &= \frac{1}{s_{k,1}^2} \frac{1}{s_{k,2}^2} \sum E \left\{ \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} \frac{1}{\tilde{\eta}_{\mu(K),k}} W_{ki_1j_1}(s) W_{ki_1l_1}(t) W_{ki_1j'_1}(T_{ki_1l_1}) \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) (Y_{ki_1j_1} - \mu_{ki_1j_1}) \right. \\ &\quad \left. (Y_{ki_1j'_1} - \mu_{ki_1j'_1}) \right\} E \left\{ \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} \frac{1}{\tilde{\eta}_{\mu(K),k}} W_{ki_2j_2}(s) W_{ki_2l_2}(t) W_{ki_2j'_2}(T_{ki_2l_2}) \left(\frac{T_{ki_2j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \right. \\ &\quad \left. (Y_{ki_2j_2} - \mu_{ki_2j_2}) (Y_{ki_2j'_2} - \mu_{ki_2j'_2}) \right\} \\ &= E^2 \left(\Delta \tilde{R}_{1,k} \right) \\ &= O_p \left(s_{k,1}^{-2} s_{k,2}^{-2} s_{k,3}^2 \tilde{\eta}_{\gamma(K),k}^2 \right).\end{aligned}$$

(ii) Under Λ_2^a , we have

$$\begin{aligned}(a)|_{\Lambda_2^a} &= \frac{1}{s_{k',1}^2} \frac{1}{s_{k,2}^2} E \left\{ \sum \frac{1}{\tilde{\eta}_{\gamma(K),k}^4} \frac{1}{\tilde{\eta}_{\mu(K),k'}^2} W_{ki_1j_1}(s) W_{ki_1l_1}(t) W_{k'i'_1j'_1}(T_{ki_1l_1}) W_{ki_1j_2}(s) W_{ki_1l_2}(t) W_{k'i'_1j'_2}(T_{ki_1l_2}) \right. \\ &\quad \left. \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_1j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) (Y_{ki_1j_1} - \mu_{ki_1j_1}) (Y_{k'i'_1j'_1} - \mu_{k'i'_1j'_1}) (Y_{ki_1j_2} - \mu_{ki_1j_2}) (Y_{k'i'_1j'_2} - \mu_{k'i'_1j'_2}) \right\} \\ &= \frac{1}{s_{k',1}^2} \frac{1}{s_{k,2}^2} \sum E \left\{ \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W_{ki_1j_1}(s) W_{ki_1j_2}(s) \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_1j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) (Y_{ki_1j_1} - \mu_{ki_1j_1}) \right. \\ &\quad \left. (Y_{ki_1j_2} - \mu_{ki_1j_2}) \right\} E \left\{ \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} \frac{1}{\tilde{\eta}_{\mu(K),k'}^2} W_{ki_1l_1}(t) W_{k'i'_1j'_1}(T_{ki_1l_1}) W_{ki_1l_2}(t) W_{k'i'_1j'_2}(T_{ki_1l_2}) \right. \\ &\quad \left. (Y_{k'i'_1j'_1} - \mu_{k'i'_1j'_1}) (Y_{k'i'_1j'_2} - \mu_{k'i'_1j'_2}) \right\}\end{aligned}$$

$$= O_p \left(s_{k,1}^{-2} s_{k,2}^{-1} \left(s_{k,4} \tilde{\eta}_{\gamma(K),k}^2 + s_{k,3} \tilde{\eta}_{\gamma(K),k}^{-1} \right) \right).$$

(iii) Under Λ_3^a , we have

$$\begin{aligned} (a)|_{\Lambda_3^a} &= \frac{1}{s_{k,1}^2} \frac{1}{s_{k,2}^2} E \left\{ \sum \frac{1}{\tilde{\eta}_{\gamma(K),k}^4} \frac{1}{\tilde{\eta}_{\mu(K),k}^2} W_{ki_1j_1}(s) W_{ki_1l_1}(t) W_{ki_2j_1'}(T_{ki_1l_1}) W_{ki_2j_2}(s) W_{ki_2l_2}(t) W_{ki_1j_2'}(T_{ki_2l_2}) \right. \\ &\quad \left. \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_2j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) (Y_{ki_1j_1} - \mu_{ki_1j_1}) (Y_{ki_2j_1'} - \mu_{ki_2j_1'}) (Y_{ki_2j_2} - \mu_{ki_2j_2}) (Y_{ki_1j_2'} - \mu_{ki_1j_2'}) \right\} \\ &= \frac{1}{s_{k,1}^2} \frac{1}{s_{k,2}^2} \sum E \left\{ \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} \frac{1}{\tilde{\eta}_{\mu(K),k}^2} W_{ki_1j_1}(s) W_{ki_2l_2}(t) \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) W_{ki_1j_2'}(T_{ki_2l_2}) (Y_{ki_1j_1} - \mu_{ki_1j_1}) \right. \\ &\quad \left. (Y_{ki_1j_2'} - \mu_{ki_1j_2'}) \right\} E \left\{ \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} \frac{1}{\tilde{\eta}_{\mu(K),k}^2} W_{ki_2j_2}(s) W_{ki_1l_1}(t) \left(\frac{T_{ki_2j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) W_{ki_2j_1'}(T_{ki_1l_1}) \right. \\ &\quad \left. (Y_{ki_2j_2} - \mu_{ki_2j_2}) (Y_{ki_2j_1'} - \mu_{ki_2j_1'}) \right\} \\ &= E^2 \left(\Delta \tilde{R}_{1,k} \right) \\ &= O_p \left(s_{k,1}^{-2} s_{k,2}^{-2} s_{k,3}^2 \tilde{\eta}_{\gamma(K),k}^2 \right). \end{aligned}$$

(iv) Under Λ_4^a , we have

$$\begin{aligned} (a)|_{\Lambda_4^a} &= \frac{1}{s_{k,1}^2} \frac{1}{s_{k,2}^2} E \left\{ \sum \frac{1}{\tilde{\eta}_{\gamma(K),k}^4} \frac{1}{\tilde{\eta}_{\mu(K),k}^2} W_{ki_1j_1}(s) W_{ki_1l_1}(t) W_{ki_1j_1'}(T_{ki_1l_1}) W_{ki_1j_2}(s) W_{ki_1l_2}(t) W_{ki_1j_2'}(T_{ki_1l_2}) \right. \\ &\quad \left. \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_1j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) (Y_{ki_1j_1} - \mu_{ki_1j_1}) (Y_{ki_1j_1'} - \mu_{ki_1j_1'}) (Y_{ki_1j_2} - \mu_{ki_1j_2}) (Y_{ki_1j_2'} - \mu_{ki_1j_2'}) \right\}. \end{aligned}$$

Moreover, it can be divided into the following three cases:

$$\begin{aligned} \Lambda_{4.1}^a &= \{j_1 = j_2, j_1' = j_2', l_1 = l_2, j_1 \neq l_1, j_2 \neq l_2\}, \\ \Lambda_{4.2}^a &= \{j_1 = j_2, j_1' \neq j_2', j_1 \neq l_1, j_2 \neq l_2\} \cup \{j_1 \neq j_2, j_1' = j_2', l_1 = l_2, j_1 \neq l_1, j_2 \neq l_2\}, \\ \Lambda_{4.3}^a &= \{j_1, j_2, j_1', j_2' \text{ are all different}, j_1 \neq l_1, j_2 \neq l_2\}. \end{aligned}$$

Under $\Lambda_{4.1}^a$, we have

$$\begin{aligned} (a)|_{\Lambda_{4.1}^a} &= \frac{1}{s_{k,1}^2} \frac{1}{s_{k,2}^2} E \left\{ \sum \frac{1}{\tilde{\eta}_{\gamma(K),k}^4} \frac{1}{\tilde{\eta}_{\mu(K),k}^2} W_{ki_1j_1}^2(s) W_{ki_1l_1}^2(t) W_{ki_1j_1'}^2(T_{ki_1l_1}) \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right)^2 (Y_{ki_1j_1} - \mu_{ki_1j_1})^2 \right. \\ &\quad \left. (Y_{ki_1j_1'} - \mu_{ki_1j_1'})^2 \right\} \end{aligned}$$

$$= O_p \left(s_{k,1}^{-2} s_{k,2}^{-2} s_{k,3} \tilde{\eta}_{\gamma(K),k}^{-2} \tilde{\eta}_{\mu(K),k}^{-1} \right).$$

Under $\Lambda_{4,2}^a$, we have

$$\begin{aligned} (a)|_{\Lambda_{4,2}^a} &= \frac{1}{s_{k,1}^2} \frac{1}{s_{k,2}^2} E \left\{ \sum \frac{1}{\tilde{\eta}_{\gamma(K),k}^4} \frac{1}{\tilde{\eta}_{\mu(K),k}^2} W_{ki_1j_1}^2(s) W_{ki_1l_1}(t) W_{ki_1j_1'}(T_{ki_1l_1}) W_{ki_1l_2}(t) W_{ki_1j_2'}(T_{ki_1l_2}) \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right)^2 \right. \\ &\quad \left. (Y_{ki_1j_1} - \mu_{ki_1j_1})^2 (Y_{ki_1j_1'} - \mu_{ki_1j_1'}) (Y_{ki_1j_2'} - \mu_{ki_1j_2'}) \right\} \\ &\quad + \frac{1}{s_{k,1}^2} \frac{1}{s_{k,2}^2} E \left\{ \sum \frac{1}{\tilde{\eta}_{\gamma(K),k}^4} \frac{1}{\tilde{\eta}_{\mu(K),k}^2} W_{ki_1j_1}(s) W_{ki_1l_1}^2(t) W_{ki_1j_1'}^2(T_{ki_1l_1}) W_{ki_1j_2}(s) \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_1j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \right. \\ &\quad \left. (Y_{ki_1j_1} - \mu_{ki_1j_1}) (Y_{ki_1j_1'} - \mu_{ki_1j_1'})^2 (Y_{ki_1j_2} - \mu_{ki_1j_2}) \right\} \\ &= O_p \left(s_{k,1}^{-1} s_{k,2}^{-2} \left(s_{k,4} \tilde{\eta}_{\gamma(K),k}^{-1} + s_{k,3} \tilde{\eta}_{\gamma(K),k} \tilde{\eta}_{\mu(K),k}^{-1} \right) \right). \end{aligned}$$

Under $\Lambda_{4,3}^a$, we have

$$\begin{aligned} (a)|_{\Lambda_{4,3}^a} &= \frac{1}{s_{k,1}^2} \frac{1}{s_{k,2}^2} E \left\{ \sum \frac{1}{\tilde{\eta}_{\gamma(K),k}^4} \frac{1}{\tilde{\eta}_{\mu(K),k}^2} W_{ki_1j_1}(s) W_{ki_1l_1}(t) W_{ki_1j_1'}(T_{ki_1l_1}) W_{ki_1j_2}(s) W_{ki_1l_2}(t) W_{ki_1j_2'}(T_{ki_1l_2}) \right. \\ &\quad \left. \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_1j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) (Y_{ki_1j_1} - \mu_{ki_1j_1}) (Y_{ki_1j_1'} - \mu_{ki_1j_1'}) (Y_{ki_1j_2} - \mu_{ki_1j_2}) (Y_{ki_1j_2'} - \mu_{ki_1j_2'}) \right\} \\ &= O_p \left(s_{k,1}^{-2} s_{k,2}^{-2} s_{k,6} \tilde{\eta}_{\gamma(K),k}^2 \right). \end{aligned}$$

Finally, we have

$$\begin{aligned} (a) &= (a)|_{\Lambda_1^a} + (a)|_{\Lambda_2^a} + (a)|_{\Lambda_3^a} + (a)|_{\Lambda_{4,1}^a} + (a)|_{\Lambda_{4,2}^a} + (a)|_{\Lambda_{4,3}^a} \\ &= O_p \left(s_{k,1}^{-2} s_{k,2}^{-2} s_{k,3} \tilde{\eta}_{\gamma(K),k}^{-2} \tilde{\eta}_{\mu(K),k}^{-1} \right) + O_p \left(s_{k,1}^{-1} s_{k,2}^{-2} \left(s_{k,4} \tilde{\eta}_{\gamma(K),k}^{-1} + s_{k,3} \tilde{\eta}_{\gamma(K),k} \tilde{\eta}_{\mu(K),k}^{-1} \right) \right) \\ &\quad + O_p \left(s_{k,1}^{-2} s_{k,2}^{-2} s_{k,6} \tilde{\eta}_{\gamma(K),k}^2 \right). \end{aligned}$$

For (b) and (c), it requires $\{(k, i_1) = (k, i_2) = (k', i_1')\}$ to ensure that these two terms are not equal to zero. Consequently, it leads to

$$\begin{aligned} (b) &= \frac{1}{2} \frac{1}{s_{k,1}^2} \frac{1}{s_{k,2}^2} E \left\{ \sum \frac{1}{\tilde{\eta}_{\gamma(K),k}^4} \frac{1}{\tilde{\eta}_{\mu(K),k}^2} W_{ki_1j_1}(s) W_{ki_1l_1}(t) W_{ki_1j_1'}(T_{ki_1l_1}) W_{ki_1j_2}(s) W_{ki_2l_2}(t) W_{ki_2j_2'}(T_{ki_1l_2}) \right. \\ &\quad \left. \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_1j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) (Y_{ki_1j_1} - \mu_{ki_1j_1}) (Y_{ki_1j_1'} - \mu_{ki_1j_1'}) (Y_{ki_1j_2} - \mu_{ki_1j_2}) \mu^{(2)}(T_{ki_1l_2}) (T_{ki_2j_2'} - T_{ki_1l_2})^2 \right\} \\ &= O_p \left(s_{k,1}^{-2} s_{k,2}^{-1} \left(s_{k,4} \tilde{\eta}_{\gamma(K),k}^2 + s_{k,3} \tilde{\eta}_{\gamma(K),k}^{-1} \right) \right) \end{aligned}$$

and

$$\begin{aligned}
(c) &= \frac{1}{2} \frac{1}{s_{k,1}^2} \frac{1}{s_{k,2}^2} E \left\{ \sum \frac{1}{\tilde{\eta}_{\gamma(K),k}^4} \frac{1}{\tilde{\eta}_{\mu(K),k'_1}} \frac{1}{\tilde{\eta}_{\mu(K),k}} \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_1j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) W_{ki_1j_1}(s) W_{ki_1l_1}(t) W_{k'i_1j'_1}(T_{ki_1l_1}) \right. \\
&\quad W_{ki_1j_2}(s) W_{ki_1l_2}(t) W_{k'i_1j'_2}(T_{ki_1l_2}) (Y_{ki_1j_1} - \mu_{ki_1j_1}) (Y_{k'i_1j'_2} - \mu_{k'i_1j'_2}) (Y_{ki_1j_2} - \mu_{ki_1j_2}) \mu^{(2)}(T_{ki_1l_1}) \\
&\quad \left. \left(T_{k'i_1j'_1} - T_{ki_1l_1} \right)^2 \right\}, \\
&= O_p \left(s_{k,1}^{-2} s_{k,2}^{-1} \left(s_{k,4} \tilde{\eta}_{\gamma(K),k}^2 + s_{k,3} \tilde{\eta}_{\gamma(K),k}^{-1} \right) \right).
\end{aligned}$$

Lastly for (d), it requires that $\{(k, i_1) = (k, i_2)\}$, yielding that

$$\begin{aligned}
(d) &= \frac{1}{4} \frac{1}{s_{K,1}^2} \frac{1}{s_{k,2}^2} E \left\{ \sum \frac{1}{\tilde{\eta}_{\gamma(K),k}^4} \frac{1}{\tilde{\eta}_{\mu(K),k'_1}} \frac{1}{\tilde{\eta}_{\mu(K),k'_2}} W_{ki_1j_1}(s) W_{ki_1l_1}(t) W_{k'i_1j'_1}(T_{ki_1l_1}) W_{ki_1j_2}(s) W_{ki_1l_2}(t) W_{k'i_2j'_2}(T_{ki_1l_2}) \right. \\
&\quad \left(\frac{T_{ki_1j_1} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{ki_1j_2} - s}{\tilde{\eta}_{\gamma(K),k}} \right) (Y_{ki_1j_1} - \mu_{ki_1j_1}) (Y_{ki_1j_2} - \mu_{ki_1j_2}) \mu^{(2)}(T_{ki_1l_1}) \left(T_{k'i_1j'_1} - T_{ki_1l_1} \right)^2 \mu^{(2)}(T_{ki_1l_2}) \\
&\quad \left. \left(T_{k'i_2j'_2} - T_{ki_1l_2} \right)^2 \right\} \\
&= O_p \left(s_{k',1}^{-2} s_{k,2}^{-1} \left(s_{k,4} \tilde{\eta}_{\gamma(K),k}^2 + s_{k,3} \tilde{\eta}_{\gamma(K),k}^{-1} \right) \tilde{\eta}_{\mu(K),k}^4 \right).
\end{aligned}$$

Taken together, we have

$$\begin{aligned}
\text{Var} \left(\Delta \tilde{R}_{1,k} \right) &= E \left(\Delta \tilde{R}_{1,k} \right)^2 - E^2 \left(\Delta \tilde{R}_{1,k} \right) \\
&= O_p \left(s_{k,1}^{-2} s_{k,2}^{-2} s_{k,3}^2 \tilde{\eta}_{\gamma(K),k}^2 \right) + O_p \left(s_{k',1}^{-2} s_{k,2}^{-1} \left(s_{k,4} \tilde{\eta}_{\gamma(K),k}^2 + s_{k,3} \tilde{\eta}_{\gamma(K),k}^{-1} \right) \right) \\
&\quad + O_p \left(s_{k,1}^{-2} s_{k,2}^{-2} s_{k,3} \tilde{\eta}_{\gamma(K),k}^{-2} \tilde{\eta}_{\mu(K),k}^{-1} \right) + O_p \left(s_{k,1}^{-1} s_{k,2}^{-2} \left(s_{k,4} \tilde{\eta}_{\gamma(K),k}^{-1} + s_{k,3} \tilde{\eta}_{\gamma(K),k} \tilde{\eta}_{\mu(K),k}^{-1} \right) \right) \\
&\quad + O_p \left(s_{k,1}^{-2} s_{k,2}^{-2} s_{k,6} \tilde{\eta}_{\gamma(K),k}^2 \right) + O_p \left(s_{k,1}^{-2} s_{k,2}^{-1} \left(s_{k,4} \tilde{\eta}_{\gamma(K),k}^2 + s_{k,3} \tilde{\eta}_{\gamma(K),k}^{-1} \right) \right) \\
&\quad + O_p \left(s_{k',1}^{-2} s_{k,2}^{-1} \left(s_{k,4} \tilde{\eta}_{\gamma(K),k}^2 + s_{k,3} \tilde{\eta}_{\gamma(K),k}^{-1} \right) \tilde{\eta}_{\mu(K),k}^4 \right).
\end{aligned}$$

While for $Q_{0R_{1,k}}$, by simple calculation it can be derived that

$$\begin{aligned}
E \left(Q_{0R_{1,k}} \right) &= f(T_{kil}) + \frac{1}{2} \alpha(W) f^{(2)}(T_{kil}) \tilde{\eta}_{\mu(K),k}^2 + o_p \left(\tilde{\eta}_{\mu(K),k}^2 \right), \\
\text{Var} \left(Q_{0R_{1,k}} \right) &= \frac{1}{s_{k',1}} R(W) f(T_{kil}) \frac{1}{\tilde{\eta}_{\mu(K),k}} + \frac{1}{2s_{k',1}} U(W) f^{(2)}(T_{kil}) \tilde{\eta}_{\mu(K),k}, \\
\text{Cov} \left(Q_{0R_{1,k}}, \Delta \tilde{R}_{1,k} \right) &= O_p \left(s_{k',1}^{-2} s_{k,2}^{-1} s_{k,3} \tilde{\eta}_{\gamma(K),k} \tilde{\eta}_{\mu(K),k}^{-1} \right).
\end{aligned}$$

Finally, we apply the delta method and it yields that

$$\begin{aligned}
\text{Var}\left(\frac{\Delta\tilde{R}_{1,k}}{Q_{0R_{1,k}}}\right) &= \left\{ \frac{\text{Var}(\Delta\tilde{R}_{1,k})}{E^2(Q_{0R_{1,k}})} - 2 \frac{E(\Delta\tilde{R}_{1,k})\text{Cov}(Q_{0R_{1,k}}, \Delta\tilde{R}_{1,k})}{E^3(Q_{0R_{1,k}})} + \frac{E^2(\Delta\tilde{R}_{1,k})\text{Var}(Q_{0R_{1,k}})}{E^4(Q_{0R_{1,k}})} \right\} \{1 + o_p(1)\} \\
&= O_p\left(s_{k,1}^{-2}s_{k,2}^{-2}s_{k,3}^2\tilde{\eta}_{\gamma(K),k}^2\right) + O_p\left(s_{k',1}^{-2}s_{k,2}^{-1}\left(s_{k,4}\tilde{\eta}_{\gamma(K),k}^2 + s_{k,3}\tilde{\eta}_{\gamma(K),k}^{-1}\right)\right) \\
&\quad + O_p\left(s_{k,1}^{-2}s_{k,2}^{-2}s_{k,3}\tilde{\eta}_{\gamma(K),k}^{-2}\tilde{\eta}_{\mu(K),k}^{-1}\right) + O_p\left(s_{k,1}^{-1}s_{k,2}^{-2}\left(s_{k,4}\tilde{\eta}_{\gamma(K),k}^{-1} + s_{k,3}\tilde{\eta}_{\gamma(K),k}\tilde{\eta}_{\mu(K),k}^{-1}\right)\right) \\
&\quad + O_p\left(s_{k,1}^{-2}s_{k,2}^{-2}s_{k,6}\tilde{\eta}_{\gamma(K),k}^2\right) + O_p\left(s_{k,1}^{-2}s_{k,2}^{-1}\left(s_{k,4}\tilde{\eta}_{\gamma(K),k}^2 + s_{k,3}\tilde{\eta}_{\gamma(K),k}^{-1}\right)\right) \\
&\quad + O_p\left(s_{k',1}^{-3}s_{k,2}^{-2}s_{k,3}^2\tilde{\eta}_{\gamma(K),k}^2\tilde{\eta}_{\mu(K),k}^{-1}\right) \\
&= O_p\left(s_{k',1}^{-2}s_{k,2}^{-2}\left\{s_{k,3}^2\tilde{\eta}_{\gamma(K),k}^2 + s_{k,2}\left(s_{k,4}\tilde{\eta}_{\gamma(K),k}^2 + s_{k,3}\tilde{\eta}_{\gamma(K),k}^{-1}\right) + \tilde{\eta}_{\gamma(K),k}^{-2}\tilde{\eta}_{\mu(K),k}^{-1}\right.\right. \\
&\quad \left.\left.+ s_{k',1}\left(s_{k,4}\tilde{\eta}_{\gamma(K),k}^{-1} + s_{k,3}\tilde{\eta}_{\gamma(K),k}\tilde{\eta}_{\mu(K),k}^{-1}\right) + s_{k,6}\tilde{\eta}_{\gamma(K),k}^2 + s_{k',1}^{-1}s_{k,3}^2\tilde{\eta}_{\gamma(K),k}^2\tilde{\eta}_{\mu(K),k}^{-1}\right\}\right).
\end{aligned}$$

S7. Proof of Theorem 1

To derive the asymptotic normality of $\tilde{\mu}_{(K)}^{(1)}(t)$, we first express it as

$$\begin{aligned}
&\tilde{\mu}_{(K)}^{(1)}(t) \\
&= \sum_{k=1}^K \omega_{\mu(K),k} \frac{1}{\tilde{\eta}_{\mu(K),k}} \frac{R_{0,k}(Q_{2,k}Q_{3,k} - Q_{1,k}Q_{4,k}) + R_{1,k}(Q_{0,k}Q_{4,k} - Q_{2,k}^2) + R_{2,k}(Q_{1,k}Q_{2,k} - Q_{0,k}Q_{3,k})}{Q_{0,k}(Q_{2,k}Q_{4,k} - Q_{3,k}^2) - Q_{1,k}(Q_{1,k}Q_{4,k} - Q_{2,k}Q_{3,k}) + Q_{2,k}(Q_{1,k}Q_{3,k} - Q_{2,k}^2)},
\end{aligned}$$

where $\omega_{\mu(K),k} = s_{k,1}/S_{k,1}$, $S_{K,j} = \sum_{k=1}^K s_{k,1}$, $s_{k,1} = \sum_{i=1}^{n_k} m_{ki}$, and

$$\begin{aligned}
Q_{r,k} &= \frac{1}{s_{k,1}} \sum_{i=1}^{n_k} \sum_{j=1}^{m_{ki}} \frac{1}{\tilde{\eta}_{\mu(K),k}} W\left(\frac{T_{kij} - t}{\tilde{\eta}_{\mu(K),k}}\right) \left(\frac{T_{kij} - t}{\tilde{\eta}_{\mu(K),k}}\right)^r, \quad r = 0, 1, 2, 3, 4, \\
R_{r,k} &= \frac{1}{s_{k,1}} \sum_{i=1}^{n_k} \sum_{j=1}^{m_{ki}} \frac{1}{\tilde{\eta}_{\mu(K),k}} W\left(\frac{T_{kij} - t}{\tilde{\eta}_{\mu(K),k}}\right) \left(\frac{T_{kij} - t}{\tilde{\eta}_{\mu(K),k}}\right)^r Y_{kij}, \quad r = 0, 1, 2.
\end{aligned}$$

Denoting

$$\begin{aligned}
\tilde{\mu}_k(t) &= \frac{R_{0,k}(Q_{2,k}Q_{4,k} - Q_{3,k}^2) + R_{1,k}(Q_{2,k}Q_{3,k} - Q_{1,k}Q_{4,k}) + R_{2,k}(Q_{1,k}Q_{3,k} - Q_{2,k}^2)}{Q_{0,k}(Q_{2,k}Q_{4,k} - Q_{3,k}^2) - Q_{1,k}(Q_{1,k}Q_{4,k} - Q_{2,k}Q_{3,k}) + Q_{2,k}(Q_{1,k}Q_{3,k} - Q_{2,k}^2)}, \\
\tilde{\mu}_k^{(2)}(t) &= \frac{2}{\tilde{\eta}_{\mu(K),k}^2} \frac{R_{0,k}(Q_{1,k}Q_{3,k} - Q_{2,k}^2) + R_{1,k}(Q_{1,k}Q_{2,k} - Q_{0,k}Q_{3,k}) + R_{2,k}(Q_{0,k}Q_{2,k} - Q_{1,k}^2)}{Q_{0,k}(Q_{2,k}Q_{4,k} - Q_{3,k}^2) - Q_{1,k}(Q_{1,k}Q_{4,k} - Q_{2,k}Q_{3,k}) + Q_{2,k}(Q_{1,k}Q_{3,k} - Q_{2,k}^2)},
\end{aligned}$$

we have

$$\tilde{\mu}_{(K)}^{(1)}(t) = \sum_{k=1}^K \omega_{\mu(K),k} \frac{R_{1,k} - Q_{1,k}\tilde{\mu}_k(t) - \frac{1}{2}Q_{3,k}\tilde{\mu}_k^{(2)}(t)\tilde{\eta}_{\mu(K),k}^2}{Q_{2,k}\tilde{\eta}_{\mu(K),k}}.$$

We further replace $\tilde{\mu}_k(t)$ and $\tilde{\mu}_k^{(2)}(t)$ with their true quantities and obtain the following definition,

$$\bar{\mu}_{(K)}^{(1)}(t) = \sum_{k=1}^K \omega_{\mu_{(K)},k} \frac{R_{1,k} - Q_{1,k}\mu(t) - \frac{1}{2}Q_{3,k}\mu^{(2)}(t)\tilde{\eta}_{\mu_{(K)},k}^2}{Q_{2,k}\tilde{\eta}_{\mu_{(K)},k}},$$

where $\mu(t)$ is the mean function and $\mu^{(2)}(t)$ is the second-order derivative of $\mu(t)$.

Moreover, by simple algebraic calculations,

$$\begin{aligned} \tilde{\mu}_{(K)}^{(1)}(t) - \bar{\mu}_{(K)}^{(1)}(t) = & - \sum_{k=1}^K \omega_{\mu_{(K)},k} \frac{1}{\tilde{\eta}_{\mu_{(K)},k}} \frac{1}{Q_{2,k}A_k} [Q_{2,k}(Q_{1,k}Q_{4,k} - Q_{2,k}Q_{3,k})G_{0,k} \\ & + \{Q_{1,k}(Q_{2,k}Q_{3,k} - Q_{1,k}Q_{4,k}) + Q_{3,k}(Q_{1,k}Q_{2,k} - Q_{0,k}Q_{3,k})\}G_{1,k} \\ & + Q_{2,k}(Q_{0,k}Q_{3,k} - Q_{1,k}Q_{2,k})G_{2,k}], \end{aligned}$$

where

$$\begin{aligned} A_k &= Q_{0,k}(Q_{2,k}Q_{4,k} - Q_{3,k}^2) - Q_{1,k}(Q_{1,k}Q_{4,k} - Q_{2,k}Q_{3,k}) + Q_{2,k}(Q_{1,k}Q_{3,k} - Q_{2,k}^2), \\ G_{0,k} &= R_{0,k} - Q_{0,k}\mu(t) - Q_{1,k}\mu^{(1)}(t)\tilde{\eta}_{\mu_{(K)},k} - \frac{1}{2}Q_{2,k}\mu^{(2)}(t)\tilde{\eta}_{\mu_{(K)},k}^2, \\ G_{1,k} &= R_{1,k} - Q_{1,k}\mu(t) - Q_{2,k}\mu^{(1)}(t)\tilde{\eta}_{\mu_{(K)},k} - \frac{1}{2}Q_{3,k}\mu^{(2)}(t)\tilde{\eta}_{\mu_{(K)},k}^2, \\ G_{2,k} &= R_{2,k} - Q_{2,k}\mu(t) - Q_{3,k}\mu^{(1)}(t)\tilde{\eta}_{\mu_{(K)},k} - \frac{1}{2}Q_{4,k}\mu^{(2)}(t)\tilde{\eta}_{\mu_{(K)},k}^2. \end{aligned}$$

It is easy to show that $A_k > 0$, $Q_{2,k} > 0$ and $P(A_k \neq 0) \rightarrow 1$, $P(Q_{2,k} \neq 0) \rightarrow 1$. In addition, we have $G_{0,k} = O_p(\tilde{\eta}_{\mu_{(K)},k}^4 + (1/(s_{k,1}\tilde{\eta}_{\mu_{(K)},k}))^{1/2} + (s_{k,2}/s_{k,1}^2)^{1/2})$, $G_{1,k} = O_p(\tilde{\eta}_{\mu_{(K)},k}^3 + (1/(s_{k,1}\tilde{\eta}_{\mu_{(K)},k}))^{1/2} + (s_{k,2}/s_{k,1}^2\tilde{\eta}_{\mu_{(K)},k}^2)^{1/2})$, and $G_{2,k} = O_p(\tilde{\eta}_{\mu_{(K)},k}^4 + \{1/(s_{k,1}\tilde{\eta}_{\mu_{(K)},k})\}^{\frac{1}{2}} + (s_{k,2}/s_{k,1}^2)^{1/2})$. Taken together, it leads to

$$\left\{ \min \left(S_{K,1}\rho_{\mu_{(K)},1}, \frac{S_{K,1}^2}{S_{K,2}\rho_{\mu_{(K)},2}} \right) \right\}^{1/2} \left(\tilde{\mu}_{(K)}^{(1)}(t) - \bar{\mu}_{(K)}^{(1)}(t) \right) = o_p(1).$$

This shows that $\tilde{\mu}_{(K)}^{(1)}(t)$ and $\bar{\mu}_{(K)}^{(1)}(t)$ have the same asymptotic distribution. Observing this, it suffices to derive the asymptotic joint normality for $\bar{\mu}_{(K)}^{(1)}(t)$.

Similar to the proof of Theorem 4.1 in Bhattacharya and Müller (1993) and by the Cramér-Wold device (Cramér and Wold, 1936) and the Lyapunov condition, we can achieve the asymptotic joint normality of $(R_{1,k} - E(R_{1,k}), Q_{1,k} - E(Q_{1,k}), Q_{3,k} - E(Q_{3,k}))$, where the rate of convergence is $[\min\{1/(S_{K,1}\rho_{\mu_{(K)},1}), S_{K,2}\rho_{\mu_{(K)},2}/S_{K,1}^2\}]^{1/2}$. Moreover, we can also deduce that

$$\begin{aligned} E(R_{1,k}) &= \alpha(W) \left\{ \mu(t)f^{(1)}(t) + \mu^{(1)}(t)f(t) \right\} \tilde{\eta}_{\mu_{(K)},k} \\ &+ \frac{1}{6}\beta(W) \left\{ \mu(t)f^{(3)}(t) + 3\mu^{(1)}(t)f^{(2)}(t) + 3\mu^{(2)}(t)f^{(1)}(t) + \mu^{(3)}(t)f(t) \right\} \tilde{\eta}_{\mu_{(K)},k}^3 + o_p(\tilde{\eta}_{\mu_{(K)},k}^3), \end{aligned}$$

$$\begin{aligned}
E(Q_{1,k}) &= \alpha(W)f^{(1)}(t)\tilde{\eta}_{\mu(K),k} + \frac{1}{6}\beta(W)f^{(3)}(t)\tilde{\eta}_{\mu(K),k}^3 + o_p\left(\tilde{\eta}_{\mu(K),k}^3\right), \\
E(Q_{2,k}) &= \alpha(W)f(t) + \frac{1}{2}\beta(W)f^{(2)}(t)\tilde{\eta}_{\mu(K),k}^2 + o_p\left(\tilde{\eta}_{\mu(K),k}^2\right), \\
E(Q_{3,k}) &= \beta(W)f^{(1)}(t)\tilde{\eta}_{\mu(K),k} + \frac{1}{6}c(W)f^{(3)}(t)\tilde{\eta}_{\mu(K),k}^3 + o_p\left(\tilde{\eta}_{\mu(K),k}^3\right), \\
\text{Var}(R_{1,k}) &= \frac{1}{s_{k,1}\tilde{\eta}_{\mu(K),k}}U(W)\left\{\mu^2(t) + \gamma(t,t) + \sigma^2\right\}f(t) \\
&\quad + \frac{s_{k,2}}{s_{k,1}^2}\tilde{\eta}_{\mu(K),k}^2\alpha^2(W)\left[\gamma(t,t)\left\{f^{(1)}(t)\right\}^2 + \gamma^{(1,1)}(t,t)f^2(t) + 2\gamma^{(1,0)}(t,t)f(t)f^{(1)}(t)\right] \\
&\quad + o_p\left(\frac{1}{s_{k,1}\tilde{\eta}_{\mu(K),k}} + \frac{s_{k,2}}{s_{k,1}^2}\tilde{\eta}_{\mu(K),k}^2\right), \\
\text{Var}(Q_{1,k}) &= \frac{1}{s_{k,1}\tilde{\eta}_{\mu(K),k}}U(W)f(t) + o_p\left(\frac{1}{s_{k,1}\tilde{\eta}_{\mu(K),k}}\right), \\
\text{Var}(Q_{2,k}) &= \frac{1}{s_{1,k}\tilde{\eta}_{\mu(K),k}}V(W)f(t) + o_p\left(\frac{1}{s_{k,1}\tilde{\eta}_{\mu(K),k}}\right), \\
\text{Var}(Q_{3,k}) &= \frac{1}{s_{k,1}\tilde{\eta}_{\mu(K),k}}Z(W)f(t) + o_p\left(\frac{1}{s_{k,1}\tilde{\eta}_{\mu(K),k}}\right), \\
\text{Cov}(R_{1,k}, Q_{1,k}) &= \frac{1}{s_{k,1}\tilde{\eta}_{\mu(K),k}}U(W)\mu(t)f(t) + o_p\left(\frac{1}{s_{1,k}\tilde{\eta}_{\mu(K),k}}\right), \\
\text{Cov}(R_{1,k}, Q_{2,k}) &= \frac{1}{s_{k,1}}V(W)\left\{\mu(t)f^{(1)}(t) + \mu^{(1)}(t)f(t)\right\} + o_p\left(\frac{1}{s_{1,k}\tilde{\eta}_{\mu(K),k}}\right), \\
\text{Cov}(R_{1,k}, Q_{3,k}) &= \frac{1}{s_{k,1}\tilde{\eta}_{\mu(K),k}}V(W)\mu(t)f(t) + o_p\left(\frac{1}{s_{k,1}\tilde{\eta}_{\mu(K),k}}\right), \\
\text{Cov}(Q_{1,k}, Q_{2,k}) &= \frac{1}{s_{k,1}}V(W)f^{(1)}(t) + o_p\left(\frac{1}{s_{k,1}}\right), \\
\text{Cov}(Q_{1,k}, Q_{3,k}) &= \frac{1}{s_{k,1}\tilde{\eta}_{\mu(K),k}}V(W)f(t) + o_p\left(\frac{1}{s_{k,1}\tilde{\eta}_{\mu(K),k}}\right), \\
\text{Cov}(Q_{2,k}, Q_{3,k}) &= \frac{1}{s_{k,1}}Z(W)f^{(1)}(t) + o_p\left(\frac{1}{s_{k,1}}\right),
\end{aligned}$$

where $\alpha(W) = \int x^2 W(x)dx$, $\beta(W) = \int x^4 W(x)dx$, $c(W) = \int x^6 W(x)dx$, $R(W) = \int W^2(x)dx$, $U(W) = \int x^2 W^2(x)dx$, $V(W) = \int x^4 W^2(x)dx$, $Z(W) = \int x^6 W^2(x)dx$. We also note that the kernel function is Lipschitz continuous by Assumption 2, which implies that the moments of the kernel function are finite.

Using the delta method, we can now derive the mean and variance of $\bar{\mu}_{(K)}^{(1)}(t)$ as

$$\begin{aligned}
E\left(\bar{\mu}_{(K)}^{(1)}(t)\right) &= \mu^{(1)}(t) + \frac{1}{6}\frac{\beta(W)}{\alpha(W)}\mu^{(3)}(t)\rho_{\mu(K),2} + o_p\left(\rho_{\mu(K),2}\right), \\
\text{Var}\left(\bar{\mu}_{(K)}^{(1)}(t)\right) &= \frac{\rho_{\mu(K),-3}}{S_{K,1}}\frac{U(W)}{\alpha^2(W)}\frac{\gamma(t,t) + \sigma^2}{f(t)} + \frac{S_{K,2}}{S_{K,1}^2}\frac{1}{f^2(t)}\left[\gamma(t,t)\left\{f^{(1)}(t)\right\}^2 + \gamma^{(1,1)}(t,t)f^2(t)\right]
\end{aligned}$$

$$\begin{aligned}
& + 2\gamma^{(1,0)}(t, t)f(t)f^{(1)}(t) \Big] + o_p \left(\frac{\rho_{\mu(K),-3}}{S_{K,1}} + \frac{S_{K,2}}{S_{K,1}^2} \right), \\
& = \frac{S_{K,1}}{\Gamma_\mu(t)\rho_{\mu(K),-3}} + o_p \left(\frac{\rho_{\mu(K),-3}}{S_{K,1}} + \frac{S_{K,2}}{S_{K,1}^2} \right),
\end{aligned}$$

where $\Gamma_\mu(t)$ is the same as defined in Theorem 1. Finally, the asymptotic normality of $\bar{\mu}_{(K)}^{(1)}(t)$ follows that as $K \rightarrow \infty$,

$$\sqrt{\frac{S_{K,1}}{\Gamma_\mu(t)\rho_{\mu(K),-3}}} \left\{ \bar{\mu}_{(K)}^{(1)}(t) - \mu^{(1)}(t) - \frac{1}{6} \frac{\beta(W)}{\alpha(W)} \mu^{(3)}(t) \rho_{\mu(K),2} \right\} \xrightarrow{d} N(0, 1),$$

and so is for $\tilde{\mu}_{(K)}^{(1)}(t)$. This proves Theorem 1.

S8. Proof of Theorem 2

To derive the asymptotic distribution of $\tilde{\gamma}_{(K)}^{(1,0)}(s, t)$, we first rewrite it as

$$\begin{aligned}
& \tilde{\gamma}_{(K)}^{(1,0)}(s, t) \\
& = \sum_{k=1}^K \omega_{\gamma(K),k} \frac{1}{\tilde{\eta}_{\gamma(K),k}} \frac{R_{00,k}B_{12,k} + R_{10,k}B_{22,k} + R_{01,k}B_{32,k} + R_{20,k}B_{42,k} + R_{02,k}B_{52,k} + R_{11,k}B_{62,k}}{B_k},
\end{aligned}$$

where $\omega_{\gamma(K),k} = s_{k,2}/S_{k,2}$ with $S_{K,2} = \sum_{k=1}^K s_{k,2}$ and $s_{k,2} = \sum_{i=1}^{n_k} m_{ki}(m_{ki} - 1)$, and

$$R_{pq,k} = \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \frac{1}{\tilde{\eta}_{\gamma(K),k}} W \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \frac{1}{\tilde{\eta}_{\gamma(K),k}} W \left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}} \right)^p \left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}} \right)^q \tilde{C}_{ki,j,l}^{(K)}$$

with $p, q = 0, 1, 2$. Let also

$$Q_{pq,k} = \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \frac{1}{\tilde{\eta}_{\gamma(K),k}} W \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}} \right) \frac{1}{\tilde{\eta}_{\gamma(K),k}} W \left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}} \right)^p \left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}} \right)^q$$

for $p, q = 0, 1, 2, 3, 4$ and that Q_k is a 6×6 matrix given by $Q_k = \{Q_{pq,k}\}_{0 \leq p, q \leq 4}$. Lastly, B_k is the determinant of Q_k and $B_{pq,k}$ is the algebraic cofactor of Q_k .

Denoting

$$\begin{aligned}
\tilde{\gamma}_k(s, t) &= \frac{R_{00,k}B_{11,k} + R_{10,k}B_{21,k} + R_{01,k}B_{31,k} + R_{20,k}B_{41,k} + R_{02,k}B_{51,k} + R_{11,k}B_{61,k}}{B_k}, \\
\tilde{\gamma}_k^{(0,1)}(s, t) &= \frac{1}{\tilde{\eta}_{\gamma(K),k}} \frac{R_{00,k}B_{13,k} + R_{10,k}B_{23,k} + R_{01,k}B_{33,k} + R_{20,k}B_{43,k} + R_{02,k}B_{53,k} + R_{11,k}B_{63,k}}{B_k}, \\
\tilde{\gamma}_k^{(2,0)}(s, t) &= \frac{2}{\tilde{\eta}_{\gamma(K),k}^2} \frac{R_{00,k}B_{14,k} + R_{10,k}B_{24,k} + R_{01,k}B_{34,k} + R_{20,k}B_{44,k} + R_{02,k}B_{54,k} + R_{11,k}B_{64,k}}{B_k},
\end{aligned}$$

$$\begin{aligned}\tilde{\gamma}_k^{(0,2)}(s, t) &= \frac{2}{\tilde{\eta}_{\gamma(K),k}^2} \frac{R_{00,k}B_{15,k} + R_{10,k}B_{25,k} + R_{01,k}B_{35,k} + R_{20,k}B_{45,k} + R_{02,k}B_{55,k} + R_{11,k}B_{65,k}}{B_k}, \\ \tilde{\gamma}_k^{(1,1)}(s, t) &= \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} \frac{R_{00,k}B_{16,k} + R_{10,k}B_{26,k} + R_{01,k}B_{36,k} + R_{20,k}B_{46,k} + R_{02,k}B_{56,k} + R_{11,k}B_{66,k}}{B_k},\end{aligned}$$

we have

$$\begin{aligned}\tilde{\gamma}_{(K)}^{(1,0)}(s, t) &= \sum_{k=1}^K \omega_{\gamma(K),k} \frac{1}{Q_{20,k} \tilde{\eta}_{\gamma(K),k}} \left\{ R_{10,k} - Q_{10,k} \tilde{\gamma}_k(s, t) - Q_{11,k} \tilde{\gamma}_k^{(0,1)}(s, t) \tilde{\eta}_{\gamma(K),k} - \frac{1}{2} Q_{30,k} \tilde{\gamma}_k^{(2,0)}(s, t) \tilde{\eta}_{\gamma(K),k}^2 \right. \\ &\quad \left. - \frac{1}{2} Q_{12,k} \tilde{\gamma}_k^{(0,2)}(s, t) \tilde{\eta}_{\gamma(K),k}^2 - Q_{21,k} \tilde{\gamma}_k^{(1,1)}(s, t) \tilde{\eta}_{\gamma(K),k}^2 \right\}.\end{aligned}$$

We further replace the quantities related to the online estimation with the true quantities, and as well as define,

$$\begin{aligned}\bar{\gamma}_{(K)}^{(1,0)}(s, t) &= \sum_{k=1}^K \omega_{\gamma(K),k} \frac{1}{Q_{20,k} h} \left\{ R_{10,k} - Q_{10,k} \gamma_k(s, t) - Q_{11,k} \gamma_k^{(0,1)}(s, t) h - \frac{1}{2} Q_{30,k} \gamma_k^{(2,0)}(s, t) h^2 \right. \\ &\quad \left. - \frac{1}{2} Q_{12,k} \gamma_k^{(0,2)}(s, t) h^2 - Q_{21,k} \gamma_k^{(1,1)}(s, t) h^2 \right\}, \\ \bar{\gamma}_{(K)}^{*(1,0)}(s, t) &= \sum_{k=1}^K \omega_{\gamma(K),k} \frac{1}{Q_{20,k} h} \left\{ R_{10,k}^* - Q_{10,k} \gamma_k(s, t) - Q_{11,k} \gamma_k^{(0,1)}(s, t) h - \frac{1}{2} Q_{30,k} \gamma_k^{(2,0)}(s, t) h^2 \right. \\ &\quad \left. - \frac{1}{2} Q_{12,k} \gamma_k^{(0,2)}(s, t) h^2 - Q_{21,k} \gamma_k^{(1,1)}(s, t) h^2 \right\},\end{aligned}$$

where h is the true $h_{\gamma(K)}$, $\gamma(s, t)$ is the covariance function, $\gamma^{(0,1)}(s, t)$, $\gamma^{(2,0)}(s, t)$, $\gamma^{(0,2)}(s, t)$ and $\gamma^{(1,1)}(s, t)$ are the true partial derivatives of $\gamma(s, t)$, and

$$R_{10,k}^* = \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \frac{1}{\tilde{\eta}_{\gamma(K),k}} W\left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) \frac{1}{\tilde{\eta}_{\gamma(K),k}} W\left(\frac{T_{kij} - t}{\tilde{\eta}_{\gamma(K),k}}\right) \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) C_{ki,j,l}^{(K)},$$

with $C_{ki,j,l}^{(K)} = \{Y_{kij} - \mu_{(K)}(T_{kij})\} \{Y_{kil} - \mu_{(K)}(T_{kil})\}$.

Next, by simple algebraic calculations,

$$\begin{aligned}&\tilde{\gamma}_K^{(1,0)}(s, t) - \bar{\gamma}_{(K)}^{(1,0)}(s, t) \\ &= - \sum_{k=1}^K \omega_{\gamma(K),k} \frac{1}{\tilde{\eta}_{\gamma(K),k}} \frac{1}{Q_{20,k} B_k} \{ G_{00,k} (Q_{10,k} B_{11,k} + Q_{11,k} B_{13,k} + Q_{30,k} B_{14,k} + Q_{12,k} B_{15,k} + Q_{11,k} B_{16,k}) \\ &\quad + G_{10,k} (Q_{10,k} B_{21,k} + Q_{11,k} B_{23,k} + Q_{30,k} B_{24,k} + Q_{12,k} B_{25,k} + Q_{11,k} B_{26,k}) \\ &\quad + G_{01,k} (Q_{10,k} B_{31,k} + Q_{11,k} B_{33,k} + Q_{30,k} B_{34,k} + Q_{12,k} B_{35,k} + Q_{11,k} B_{36,k}) \\ &\quad + G_{20,k} (Q_{10,k} B_{41,k} + Q_{11,k} B_{43,k} + Q_{30,k} B_{44,k} + Q_{12,k} B_{45,k} + Q_{11,k} B_{46,k}) \\ &\quad + G_{02,k} (Q_{10,k} B_{51,k} + Q_{11,k} B_{53,k} + Q_{30,k} B_{54,k} + Q_{12,k} B_{55,k} + Q_{11,k} B_{56,k}) \end{aligned}$$

$$+ G_{11,k} (Q_{10,k} B_{61,k} + Q_{11,k} B_{63,k} + Q_{30,k} B_{64,k} + Q_{12,k} B_{65,k} + Q_{11,k} B_{66,k}) \},$$

where

$$\begin{aligned} G_{00,k} &= R_{00,k} - Q_{00,k} \gamma(s, t) - Q_{10,k} \gamma^{(1,0)}(s, t) \tilde{\eta}_{\gamma(K),k} - Q_{01,k} \gamma^{(0,1)}(s, t) \tilde{\eta}_{\gamma(K),k} \\ &\quad - \frac{1}{2} Q_{20,k} \gamma^{(2,0)}(s, t) \tilde{\eta}_{\gamma(K),k}^2 - \frac{1}{2} Q_{02,k} \gamma^{(0,2)}(s, t) \tilde{\eta}_{\gamma(K),k}^2 - Q_{11,k} \gamma^{(1,1)}(s, t) \tilde{\eta}_{\gamma(K),k}^2, \\ G_{10,k} &= R_{10,k} - Q_{10,k} \gamma(s, t) - Q_{20,k} \gamma^{(1,0)}(s, t) \tilde{\eta}_{\gamma(K),k} - Q_{11,k} \gamma^{(0,1)}(s, t) \tilde{\eta}_{\gamma(K),k} \\ &\quad - \frac{1}{2} Q_{30,k} \gamma^{(2,0)}(s, t) \tilde{\eta}_{\gamma(K),k}^2 - \frac{1}{2} Q_{12,k} \gamma^{(0,2)}(s, t) \tilde{\eta}_{\gamma(K),k}^2 - Q_{21,k} \gamma^{(1,1)}(s, t) \tilde{\eta}_{\gamma(K),k}^2, \\ G_{01,k} &= R_{01,k} - Q_{01,k} \gamma(s, t) - Q_{11,k} \gamma^{(1,0)}(s, t) \tilde{\eta}_{\gamma(K),k} - Q_{02,k} \gamma^{(0,1)}(s, t) \tilde{\eta}_{\gamma(K),k} \\ &\quad - \frac{1}{2} Q_{21,k} \gamma^{(2,0)}(s, t) \tilde{\eta}_{\gamma(K),k}^2 - \frac{1}{2} Q_{03,k} \gamma^{(0,2)}(s, t) \tilde{\eta}_{\gamma(K),k}^2 - Q_{12,k} \gamma^{(1,1)}(s, t) \tilde{\eta}_{\gamma(K),k}^2, \\ G_{20,k} &= R_{20,k} - Q_{20,k} \gamma(s, t) - Q_{30,k} \gamma^{(1,0)}(s, t) \tilde{\eta}_{\gamma(K),k} - Q_{21,k} \gamma^{(0,1)}(s, t) \tilde{\eta}_{\gamma(K),k} \\ &\quad - \frac{1}{2} Q_{40,k} \gamma^{(2,0)}(s, t) \tilde{\eta}_{\gamma(K),k}^2 - \frac{1}{2} Q_{22,k} \gamma^{(0,2)}(s, t) \tilde{\eta}_{\gamma(K),k}^2 - Q_{31,k} \gamma^{(1,1)}(s, t) \tilde{\eta}_{\gamma(K),k}^2, \\ G_{02,k} &= R_{02,k} - Q_{02,k} \gamma(s, t) - Q_{12,k} \gamma^{(1,0)}(s, t) \tilde{\eta}_{\gamma(K),k} - Q_{03,k} \gamma^{(0,1)}(s, t) \tilde{\eta}_{\gamma(K),k} \\ &\quad - \frac{1}{2} Q_{22,k} \gamma^{(2,0)}(s, t) \tilde{\eta}_{\gamma(K),k}^2 - \frac{1}{2} Q_{04,k} \gamma^{(0,2)}(s, t) \tilde{\eta}_{\gamma(K),k}^2 - Q_{13,k} \gamma^{(1,1)}(s, t) \tilde{\eta}_{\gamma(K),k}^2, \\ G_{11,k} &= R_{11,k} - Q_{11,k} \gamma(s, t) - Q_{21,k} \gamma^{(1,0)}(s, t) \tilde{\eta}_{\gamma(K),k} - Q_{12,k} \gamma^{(0,1)}(s, t) \tilde{\eta}_{\gamma(K),k} \\ &\quad - \frac{1}{2} Q_{31,k} \gamma^{(2,0)}(s, t) \tilde{\eta}_{\gamma(K),k}^2 - \frac{1}{2} Q_{13,k} \gamma^{(0,2)}(s, t) \tilde{\eta}_{\gamma(K),k}^2 - Q_{22,k} \gamma^{(1,1)}(s, t) \tilde{\eta}_{\gamma(K),k}^2. \end{aligned}$$

It is easy to show that $B_k > 0$, $Q_{20,k} > 0$ and $P(B_k \neq 0) \rightarrow 1$, $P(Q_{20,k} \neq 0) \rightarrow 1$. In addition, we have $G_{00,k}$, $G_{20,k}$, $G_{02,k} = O_p(\tilde{\eta}_{\mu(K),k}^4 + \{1/(s_{k,2} \tilde{\eta}_{\mu(K),k}^2)\}^{1/2} + \{s_{k,3}/(s_{k,2}^2 \tilde{\eta}_{\mu(K),k})\}^{1/2} + (s_{k,4}/s_{k,2}^2)^{1/2})$, $G_{10,k}$, $G_{01,k} = O_p(\tilde{\eta}_{\mu(K),k}^3 + \{1/(s_{k,2} \tilde{\eta}_{\mu(K),k}^2)\}^{1/2} + \{s_{k,3}/(s_{k,2}^2 \tilde{\eta}_{\mu(K),k})\}^{1/2} + (s_{k,4} \tilde{\eta}_{\mu(K),k}^2/s_{k,2}^2)^{1/2})$, and $G_{11,k} = O_p(\tilde{\eta}_{\mu(K),k}^4 + \{1/(s_{k,2} \tilde{\eta}_{\mu(K),k}^2)\}^{1/2} + (s_{k,3} \tilde{\eta}_{\mu(K),k}/s_{k,2}^2)^{1/2} + (s_{k,4} \tilde{\eta}_{\mu(K),k}^4/s_{k,2}^2)^{1/2})$. Taken together, it leads to

$$\left\{ \min \left(S_{K,2} \rho_{\gamma(K),2}, \frac{S_{k,2}^2 \rho_{\gamma(K),1}}{S_{K,3}}, \frac{S_{K,2}^2}{S_{K,4} \rho_{\gamma(K),2}} \right) \right\}^{1/2} \left(\tilde{\gamma}_K^{(1,0)}(s, t) - \bar{\gamma}_{(K)}^{(1,0)}(s, t) \right) = o_p(1).$$

Moreover, by following the above notations, we have

$$\tilde{\gamma}_{(K)}^{(1,0)}(s, t) - \bar{\gamma}_{(K)}^{*(1,0)}(s, t) = \sum_{k=1}^K \omega_{\gamma(K),k} \frac{1}{\tilde{\eta}_{\gamma(K),k}} \frac{R_{10,k} - R_{10,k}^*}{Q_{20,k}}.$$

We further decompose $R_{10,k}$ into four terms as follows:

$$\begin{aligned} R_{10,k} &= R_{10,k}^* \\ &\quad + \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}} \right) W \left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}} \right) \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}} \right) (Y_{kij} - \mu_{kij}) (\mu_{kil} - \tilde{\mu}_{kil}) \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W\left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) W\left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}}\right) \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) (\mu_{kij} - \tilde{\mu}_{kij}) (Y_{kil} - \mu_{kil}) \\
& + \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W\left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) W\left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}}\right) \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) (\mu_{kij} - \tilde{\mu}_{kij}) (\mu_{kil} - \tilde{\mu}_{kil}),
\end{aligned}$$

where $\tilde{\mu}_{kij} = \tilde{\mu}_{(K)}(T_{kij})$. Using Theorem 1 in Yang and Yao (2023), we have $\tilde{\mu}_{(K)}(t) - \bar{\mu}_{(K)}(t) = o_p(\bar{\mu}_{(K)}(t) - \mu(t))$, where $\bar{\mu}_{(K)}(t) = \sum_{k=1}^K \omega_{\mu_{(K)},k} \left\{ R_{0,k}/Q_{0,k} - \mu^{(1)}(t) \tilde{\eta}_{\mu_{(K)},k} Q_{1,k}/Q_{0,k} \right\}$. Then we get

$$\begin{aligned}
& \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W\left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) W\left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}}\right) \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) (Y_{kij} - \mu_{kij}) (\mu_{kil} - \tilde{\mu}_{kil}) \\
& = \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W\left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) W\left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}}\right) (Y_{kij} - \mu_{kij}) (\mu_{kil} - \bar{\mu}_{kil}) \\
& = \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W\left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) W\left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}}\right) \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) (Y_{kij} - \mu_{kij}) \\
& \quad \left[\mu_{kil} - \sum_{k'=1}^K \omega_{\mu_{(K)},k'} \left\{ \frac{R_{0,k'}(T_{kil})}{Q_{0,k'}(T_{kil})} - \frac{Q_{1,k'}(T_{kil})}{Q_{0,k'}(T_{kil})} \mu^{(1)}(T_{kil}) \tilde{\eta}_{\mu_{(K)},k'} \right\} \right] \\
& = \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W\left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) W\left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}}\right) \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) (Y_{kij} - \mu_{kij}) \\
& \quad \left[\sum_{k'=1}^K \omega_{\mu_{(K)},k'} Q_{0,k'}^{-1}(T_{kil}) \frac{1}{s_{k',1}} \sum_{i'=1}^{n_{k'}} \sum_{j'=1}^{m_{k'i'}} \frac{1}{\tilde{\eta}_{\mu_{(K)},k'}} W\left(\frac{T_{k'i'j'} - T_{kil}}{\tilde{\eta}_{\mu_{(K)},k'}}\right) \right. \\
& \quad \left. \left\{ \mu_{kil} - Y_{k'i'j'} + (T_{k'i'j'} - T_{kil}) \mu^{(1)}(T_{kil}) \tilde{\eta}_{\mu_{(K)},k'} \right\} \right].
\end{aligned}$$

With the second-order Taylor expansion of $\mu(T_{kil})$ at $T_{k'i'j'}$ and the first-order Taylor expansion of $\mu^{(1)}(T_{kil})$ at $T_{k'i'j'}$, we have

$$\begin{aligned}
& \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W\left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) W\left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}}\right) \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) (Y_{kij} - \mu_{kij}) (\mu_{kil} - \tilde{\mu}_{kil}) \\
& = - \sum_{k'=1}^K \omega_{\mu_{(K)},k'} Q_{0,k'}^{-1}(T_{kil}) \frac{1}{s_{k',1}} \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1 \leq j \neq l \leq m_{ki}} \sum_{i'=1}^{n_{k'}} \sum_{j'=1}^{m_{k'i'}} \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W\left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) W\left(\frac{T_{kil} - t}{\tilde{\eta}_{\gamma(K),k}}\right) \left(\frac{T_{kij} - s}{\tilde{\eta}_{\gamma(K),k}}\right) \\
& \quad \frac{1}{\tilde{\eta}_{\mu_{(K)},k'}} W\left(\frac{T_{k'i'j'} - T_{kil}}{\tilde{\eta}_{\mu_{(K)},k'}}\right) \left\{ Y_{kij} - \mu(T_{kij}) \right\} \left\{ Y_{k'i'j'} - \mu(T_{k'i'j'}) + \frac{1}{2} \mu^{(2)}(T_{k'i'j'}) (T_{k'i'j'} - T_{kil})^2 \right\}.
\end{aligned}$$

Denoting

$$\begin{aligned}\Delta\tilde{R}_{1,k} &= -\frac{1}{s_{k',1}}\frac{1}{s_{k,2}}\sum_{i=1}^{n_k}\sum_{1\leq j\neq l\leq m_{ki}}\sum_{i'=1}^{n_{k'}}\sum_{j'=1}^{m_{k',i'}}\frac{1}{\tilde{\eta}_{\gamma(K),k}^2}W\left(\frac{T_{kij}-s}{\tilde{\eta}_{\gamma(K),k}}\right)W\left(\frac{T_{kil}-t}{\tilde{\eta}_{\gamma(K),k}}\right)\left(\frac{T_{kij}-s}{\tilde{\eta}_{\gamma(K),k}}\right) \\ &\quad \frac{1}{\tilde{\eta}_{\mu(K),k'}}W\left(\frac{T_{k'i'j'}-T_{kil}}{\tilde{\eta}_{\mu(K),k'}}\right)\{Y_{kij}-\mu(T_{kij})\}\left\{Y_{k'i'j'}-\mu(T_{k'i'j'})+\frac{1}{2}\mu^{(2)}(T_{k'i'j'})\left(T_{k'i'j'}-T_{kil}\right)^2\right\}, \\ \Delta\tilde{R}_{2,k} &= -\frac{1}{s_{k',1}}\frac{1}{s_{k,2}}\sum_{i=1}^{n_k}\sum_{1\leq j\neq l\leq m_{ki}}\sum_{i'=1}^{n_{k'}}\sum_{j'=1}^{m_{k',i'}}\frac{1}{\tilde{\eta}_{\gamma(K),k}^2}W\left(\frac{T_{kij}-s}{\tilde{\eta}_{\gamma(K),k}}\right)W\left(\frac{T_{kil}-t}{\tilde{\eta}_{\gamma(K),k}}\right)\left(\frac{T_{kij}-s}{\tilde{\eta}_{\gamma(K),k}}\right) \\ &\quad \frac{1}{\tilde{\eta}_{\mu(K),k'}}W\left(\frac{T_{k'i'j'}-T_{kij}}{\tilde{\eta}_{\mu(K),k'}}\right)\{Y_{kil}-\mu(T_{kil})\}\left\{Y_{k'i'j'}-\mu(T_{k'i'j'})+\frac{1}{2}\mu^{(2)}(T_{k'i'j'})\left(T_{k'i'j'}-T_{kij}\right)^2\right\},\end{aligned}$$

we have

$$\begin{aligned}R_{10,k} &= R_{10,k}^* + \sum_{k'=1}^K \omega_{\mu(K),k'} \frac{\Delta\tilde{R}_{1,k}}{Q_{0R_{1,k}}} + \sum_{k'=1}^K \omega_{\mu(K),k'} \frac{\Delta\tilde{R}_{2,k}}{Q_{0R_{2,k}}} \\ &\quad + \frac{1}{s_{k,2}} \sum_{i=1}^{n_k} \sum_{1\leq j\neq l\leq m_{ki}} \frac{1}{\tilde{\eta}_{\gamma(K),k}^2} W\left(\frac{T_{kij}-s}{\tilde{\eta}_{\gamma(K),k}}\right) W\left(\frac{T_{kil}-t}{\tilde{\eta}_{\gamma(K),k}}\right) \left(\frac{T_{kij}-s}{\tilde{\eta}_{\gamma(K),k}}\right) (\mu_{kij} - \tilde{\mu}_{kij}) (\mu_{kil} - \tilde{\mu}_{kil}).\end{aligned}$$

By Theorem 1 in Yang and Yao (2023), we get $(\mu_{kij} - \tilde{\mu}_{kij})(\mu_{kil} - \tilde{\mu}_{kil}) = O_p(\tilde{\eta}_{\mu(K),k}^4)$, which can be ignored compared to $\Delta\tilde{R}_{1,k}/Q_{0R_{1,k}}$. Moreover by the delta method and Lemma A.1, we can obtain that

$$\begin{aligned}\bar{\gamma}_{(K)}^{(1,0)}(s,t) - \bar{\gamma}_{(K)}^{*(1,0)}(s,t) &= \sum_{k=1}^K \omega_{\gamma(K),k} \frac{1}{\tilde{\eta}_{\gamma(K),k}} \frac{1}{Q_{20,k}} \left(\sum_{k'=1}^K \omega_{\mu(K),k'} \frac{\Delta\tilde{R}_{1,k}}{Q_{0R_{1,k}}} + \sum_{k'=1}^K \omega_{\mu(K),k'} \frac{\Delta\tilde{R}_{2,k}}{Q_{0R_{2,k}}} \right) \\ &= O_p \left(S_{k,1}^{-2} S_{K,2}^{-2} \left\{ S_{K,3}^2 + S_{K,2} \left(S_{K,4} + S_{K,3} \rho_{\gamma(K),-3} \right) + \rho_{\gamma(K),-4} \rho_{\mu(K),-1} \right. \right. \\ &\quad \left. \left. + S_{K,1} \left(S_{K,4} \rho_{\gamma(K),-3} + S_{k,3} \rho_{\gamma(K),-1} \rho_{\mu(K),1}^{-1} \right) + S_{K,6} + S_{K,1}^{-1} S_{K,3}^2 \rho_{\mu(K),-1} \right\} \right).\end{aligned}$$

This shows that $\bar{\gamma}_{(K)}^{(1,0)}(s,t)$ and $\bar{\gamma}_{(K)}^{*(1,0)}(s,t)$ have the same asymptotic distribution.

In what follows, we derive the asymptotic distribution of $\bar{\gamma}_{(K)}^{*(1,0)}(s,t)$. We use the Cramér-Wold device and the Lyapunov condition to establish the asymptotic joint normality of $(R_{10,k} - E(R_{10,k}), Q_{10,k} - E(Q_{10,k}), Q_{11,k} - E(Q_{11,k}), Q_{30,k} - E(Q_{30,k}), Q_{12,k} - E(Q_{12,k}), Q_{21,k} - E(Q_{21,k}))$. The rate of convergence is $[\min\{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2, s_{k,2}^2\tilde{\eta}_{\gamma(K),k}/s_{k,3}, s_{k,2}^2/(s_{k,4}\tilde{\eta}_{\gamma(K),k}^2)\}]^{1/2}$. In addition, we can deduce that

$$\begin{aligned}E(R_{10,k}^*) &= \alpha(W) \left\{ \gamma^{(1,0)}(s,t) f(s) f(t) + \gamma(s,t) f^{(1)}(s) f(t) \right\} \tilde{\eta}_{\gamma(K),k} + \beta(W) \left\{ \frac{1}{6} \gamma^{(3,0)}(s,t) f(s) f(t) \right. \\ &\quad \left. + \frac{1}{6} \gamma(s,t) f^{(3)}(s) f(t) + \frac{1}{2} \gamma^{(1,0)}(s,t) f^{(2)}(s) f(t) + \frac{1}{2} \gamma^{(2,0)}(s,t) f^{(1)}(s) f(t) \right\} \tilde{\eta}_{\gamma(K),k}^3 \\ &\quad + \alpha^2(W) \left\{ \gamma^{(1,1)}(s,t) f(s) f^{(1)}(t) + \gamma^{(0,1)}(s,t) f^{(1)}(s) f^{(1)}(t) + \frac{1}{2} \gamma^{(1,2)}(s,t) f(s) f(t) \right\}\end{aligned}$$

$$\begin{aligned}
& + \frac{1}{2} \gamma(s, t) f^{(1)}(s) f^{(2)}(t) + \frac{1}{2} \gamma^{(0,2)}(s, t) f^{(1)}(s) f(t) + \frac{1}{2} \gamma^{(1,0)}(s, t) f(s) f^{(2)}(t) \Big\} \tilde{\eta}_{\gamma(K),k}^3 \\
& + o_p \left(\tilde{\eta}_{\gamma(K),k}^3 \right), \\
E(Q_{10,k}) &= \alpha(W) f^{(1)}(s) f(t) \tilde{\eta}_{\gamma(K),k} + \frac{1}{2} \alpha^2(W) f^{(1)}(s) f^{(2)}(t) \tilde{\eta}_{\gamma(K),k}^3 + \frac{1}{6} \beta(W) f^{(3)}(s) f(t) \tilde{\eta}_{\gamma(K),k}^3 + o_p \left(\tilde{\eta}_{\gamma(K),k}^3 \right), \\
E(Q_{11,k}) &= \alpha^2(W) f^{(1)}(s) f^{(1)}(t) \tilde{\eta}_{\gamma(K),k}^2 + \frac{1}{6} \alpha(W) \beta(W) \left\{ f^{(1)}(s) f^{(3)}(t) + f^{(3)}(s) f^{(1)}(t) \right\} \tilde{\eta}_{\gamma(K),k}^4 + o_p \left(\tilde{\eta}_{\gamma(K),k}^4 \right), \\
E(Q_{12,k}) &= \alpha^2(W) f^{(1)}(s) f(t) \tilde{\eta}_{\gamma(K),k} + \alpha(W) \beta(W) \left\{ \frac{1}{6} f^{(3)}(s) f(t) + \frac{1}{2} f^{(1)}(s) f^{(2)}(t) \right\} \tilde{\eta}_{\gamma(K),k}^3 + o_p \left(\tilde{\eta}_{\gamma(K),k}^3 \right), \\
E(Q_{20,k}) &= \alpha(W) f(s) f(t) + \frac{1}{2} \beta(W) f^{(2)}(s) f(t) \tilde{\eta}_{\gamma(K),k}^2 + \frac{1}{2} \alpha^2(W) f(s) f^{(2)}(t) \tilde{\eta}_{\gamma(K),k}^2 \\
& + \frac{1}{4} \alpha(W) \beta(W) f^{(2)}(s) f^{(2)}(t) \tilde{\eta}_{\gamma(K),k}^4 + o_p \left(\tilde{\eta}_{\gamma(K),k}^4 \right), \\
E(Q_{21,k}) &= \alpha^2(W) f(s) f^{(1)}(t) \tilde{\eta}_{\gamma(K),k} + \frac{1}{6} \alpha(W) \beta(W) f(s) f^{(3)}(t) \tilde{\eta}_{\gamma(K),k}^3 + \frac{1}{2} \alpha(W) \beta(W) f^{(2)}(s) f^{(1)}(t) \tilde{\eta}_{\gamma(K),k}^3 \\
& + o_p \left(\tilde{\eta}_{\gamma(K),k}^3 \right), \\
\text{Var}(R_{10,k}^*) &= \{1 + I(s = t)\} \left\{ \frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} R(W) U(W) E_1(s, t) f(s) f(t) + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} U(W) E_2(s, t) f(s) f^2(t) \right\} \\
& + \frac{s_{k,4}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k}^2 \alpha^2(W) \left[E_\Phi \left[\left\{ \Phi(s) \Phi(t) f^{(1)}(s) + \Phi^{(1)}(s) \Phi(t) f(s) \right\}^2 \right] \right. \\
& \left. - \left\{ \gamma(s, t) f^{(1)}(s) + \gamma^{(1,0)}(s, t) f(s) \right\}^2 \right] f^2(t) + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} + \frac{s_{k,4}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k}^2 \right),
\end{aligned}$$

where

$$\begin{aligned}
E_1(s, t) &= E \left[\{Y_1 - \mu(T_1)\}^2 \{Y_2 - \mu(T_2)\}^2 | T_1 = s, T_2 = t \right], \\
E_2(s, t) &= E \left[\{Y_1 - \mu(T_1)\}^2 \{Y_2 - \mu(T_2)\} \{Y_3 - \mu(T_3)\} | T_1 = s, T_2 = t, T_3 = t \right],
\end{aligned}$$

and $\{(T_i, Y_i) : i = 1, 2, 3\}$ are the observations for the same subject,

$$\begin{aligned}
\text{Cov}(R_{10,k}^*, Q_{10,k}) &= \{1 + I(s = t)\} \left\{ \frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} R(W) U(W) \gamma(s, t) f(s) f(t) + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} U(W) \gamma(s, t) f(s) f^2(t) \right\} \\
& + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} \right), \\
\text{Cov}(R_{10,k}^*, Q_{11,k}) &= \{1 + I(s = t)\} \left[\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} U^2(W) \left\{ \gamma(s, t) f(s) f^{(1)}(t) + \gamma^{(0,1)}(s, t) f(s) f(t) \right\} \right. \\
& \left. + \frac{s_{k,3}}{s_{k,2}^2} \alpha(W) U(W) \gamma(s, t) f(s) f(t) f^{(1)}(t) \right] + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} + \frac{s_{k,3}}{s_{k,2}^2} \right), \\
\text{Cov}(R_{10,k}^*, Q_{12,k}) &= \{1 + I(s = t)\} \left\{ \frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} U^2(W) \gamma(s, t) f(s) f(t) + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} \alpha(W) U(W) \gamma(s, t) f(s) f^2(t) \right\}
\end{aligned}$$

$$\begin{aligned}
& + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} \right), \\
\text{Cov} (R_{10,k}^*, Q_{20,k}) &= \{1 + I(s=t)\} \left[\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} R(W) V(W) \left\{ \gamma(s,t) f^{(1)}(s) f(t) + \gamma^{(1,0)}(s,t) f(s) f(t) \right\} \right. \\
& + \frac{s_{k,3}}{s_{k,2}^2} V(W) \left\{ \gamma^{(1,0)}(s,t) f(s) f^2(t) + \gamma(s,t) f^{(1)}(s) f^2(t) \right\} \\
& + \left. \frac{s_{k,3}}{s_{k,2}^2} \alpha^2(W) R(W) \left\{ \gamma(s,t) f(s) f^{(1)}(s) f(t) + \gamma^{(1,0)}(s,t) f^2(s) f(t) \right\} \right] \\
& + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} + \frac{s_{k,3}}{s_{k,2}^2} \right), \\
\text{Cov} (R_{10,k}^*, Q_{21,k}) &= \{1 + I(s=t)\} \left[\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} U(W) V(W) \left\{ \gamma(s,t) f^{(1)}(s) f^{(1)}(t) + \gamma^{(1,1)}(s,t) f(s) f(t) \right. \right. \\
& + \left. \gamma^{(1,0)}(s,t) f(s) f^{(1)}(t) + \gamma^{(0,1)}(s,t) f^{(1)}(s) f(t) \right\} \\
& + \frac{s_{k,3}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k} \alpha(W) V(W) \left\{ \gamma^{(1,0)}(s,t) f(s) f(t) f^{(1)}(t) + \gamma(s,t) f^{(1)}(s) f(t) f^{(1)}(t) \right\} \\
& + \frac{s_{k,3}}{s_{k,2}^2} \alpha^2(W) U(W) \left\{ \gamma(s,t) f(s) f^{(1)}(s) f^{(1)}(t) + \gamma^{(1,0)}(s,t) f^2(s) f^{(1)}(t) \right. \\
& + \left. \gamma^{(0,1)}(s,t) f(s) f^{(1)}(s) f(t) + \gamma^{(1,1)}(s,t) f^2(s) f(t) \right\} \left. \right] + o_p \left(\frac{1}{s_{k,2}} + \frac{s_{k,3}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k} \right), \\
\text{Cov} (R_{10,k}^*, Q_{30,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} R(W) U(W) \gamma(s,t) f(s) f(t) + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} V(W) \gamma(s,t) f(s) f^2(t) \right\} \\
& + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} \right), \\
\text{Var} (Q_{10,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} R(W) U(W) f(s) f(t) + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} U(W) f(s) f^2(t) \right\} \\
& + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} \right), \\
\text{Var} (Q_{11,k}) &= \{1 + I(s=t)\} \left[\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} R(W) U(W) f(s) f(t) + \frac{s_{k,3}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k} \alpha^2(W) U(W) f(s) \left\{ f^{(1)}(t) \right\}^2 \right. \\
& + \left. \frac{s_{k,3}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k} \alpha^2(W) U(W) \left\{ f^{(1)}(s) \right\}^2 f(t) \right] + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} + \frac{s_{k,3}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k} \right), \\
\text{Var} (Q_{12,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} U(W) V(W) f(s) f(t) + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} \alpha^2(W) U(W) f(s) f^2(t) \right\} \\
& + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} \right),
\end{aligned}$$

$$\begin{aligned}
\text{Var}(Q_{20,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} R(W)V(W)f(s)f(t) + \frac{s_{k,3}}{s_{k,2}^2\tilde{\eta}_{\gamma(K),k}^2} \alpha^2(W)R(W)f^2(s)f(t) \right. \\
&\quad \left. + \frac{s_{k,3}}{s_{k,2}^2\tilde{\eta}_{\gamma(K),k}^2} V(W)f(s)f^2(t) \right\} + o_p \left(\frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2\tilde{\eta}_{\gamma(K),k}^2} \right), \\
\text{Var}(Q_{21,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} U(W)V(W)f(s)f(t) + \frac{s_{k,3}}{s_{k,2}^2\tilde{\eta}_{\gamma(K),k}^2} \alpha^2(W)U(W)f^2(s)f(t) \right\} \\
&\quad + o_p \left(\frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2\tilde{\eta}_{\gamma(K),k}^2} \right), \\
\text{Var}(Q_{30,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} R(W)V(W)f(s)f(t) + \frac{s_{k,3}}{s_{k,2}^2\tilde{\eta}_{\gamma(K),k}^2} W(W)f(s)f^2(t) \right\} \\
&\quad + o_p \left(\frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2\tilde{\eta}_{\gamma(K),k}^2} \right), \\
\text{Cov}(Q_{10,k}, Q_{11,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} U^2(W)f(s)f^{(1)}(t) + \frac{s_{k,3}}{s_{k,2}^2} \alpha(W)U(W)f(s)f(t)f^{(1)}(t) \right\} \\
&\quad + o_p \left(\frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2} \right), \\
\text{Cov}(Q_{10,k}, Q_{12,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} U^2(W)f(s)f(t) + \frac{s_{k,3}}{s_{k,2}^2\tilde{\eta}_{\gamma(K),k}^2} \alpha(W)U(W)f(s)f^2(t) \right\} \\
&\quad + o_p \left(\frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2\tilde{\eta}_{\gamma(K),k}^2} \right), \\
\text{Cov}(Q_{10,k}, Q_{20,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} R(W)V(W)f^{(1)}(s)f(t) + \frac{s_{k,3}}{s_{k,2}^2} \alpha^2(W)R(W)f(s)f^{(1)}(s)f(t) \right. \\
&\quad \left. + \frac{s_{k,3}}{s_{k,2}^2} V(W)f(s)f^{(1)}(s)f^2(t) \right\} + o_p \left(\frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2} \right), \\
\text{Cov}(Q_{10,k}, Q_{21,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2}} U(W)V(W)f^{(1)}(s)f^{(1)}(t) + \frac{s_{k,3}}{s_{k,2}^2}\tilde{\eta}_{\gamma(K),k}\alpha^2(W)U(W)f(s)f^{(1)}(s)f^{(1)}(t) \right. \\
&\quad \left. + \frac{s_{k,3}}{s_{k,2}^2}\tilde{\eta}_{\gamma(K),k}\alpha(W)V(W)f^{(1)}(s)f(t)f^{(1)}(t) \right\} + o_p \left(\frac{1}{s_{k,2}} + \frac{s_{k,3}}{s_{k,2}^2}\tilde{\eta}_{\gamma(K),k} \right), \\
\text{Cov}(Q_{10,k}, Q_{30,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} U(W)V(W)f(s)f^{(1)}(t) + \frac{s_{k,3}}{s_{k,2}^2} \alpha(W)V(W)f(s)f(t)f^{(1)}(t) \right\} \\
&\quad + o_p \left(\frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2} \right), \\
\text{Cov}(Q_{11,k}, Q_{12,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} U(W)V(W)f(s)f^{(1)}(t) + \frac{s_{k,3}}{s_{k,2}^2} \alpha^2(W)U(W)f(s)f(t)f^{(1)}(t) \right\} \\
&\quad + o_p \left(\frac{1}{s_{k,2}\tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2} \right)
\end{aligned}$$

$$\begin{aligned}
& + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} + \frac{s_{k,3}}{s_{k,2}^2} \right), \\
\text{Cov}(Q_{11,k}, Q_{20,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2}} U(W) V(W) f^{(1)}(s) f^{(1)}(t) + \frac{s_{k,3}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k} \alpha(W) V(W) f^{(1)}(s) f(t) f^{(1)}(t) \right. \\
& \quad \left. + \frac{s_{k,3}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k} \alpha^2(W) U(W) f(s) f^{(1)}(s) f(t) \right\} + o_p \left(\frac{1}{s_{k,2}} + \frac{s_{k,3}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k} \right), \\
\text{Cov}(Q_{11,k}, Q_{21,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} U(W) V(W) f^{(1)}(s) f(t) + \frac{s_{k,3}}{s_{k,2}^2} \alpha^2(W) U(W) f(s) f^{(1)}(s) f(t) \right\} \\
& \quad + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} + \frac{s_{k,3}}{s_{k,2}^2} \right), \\
\text{Cov}(Q_{11,k}, Q_{30,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} U(W) V(W) f(s) f^{(1)}(t) + \frac{s_{k,3}}{s_{k,2}^2} \alpha(W) V(W) f(s) f(t) f^{(1)}(t) \right\} \\
& \quad + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} + \frac{s_{k,3}}{s_{k,2}^2} \right), \\
\text{Cov}(Q_{12,k}, Q_{20,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} U(W) V(W) f^{(1)}(s) f(t) + \frac{s_{k,3}}{s_{k,2}^2} \alpha^2(W) U(W) f(s) f^{(1)}(s) f(t) \right. \\
& \quad \left. + \frac{s_{k,3}}{s_{k,2}^2} \alpha(W) V(W) f^{(1)}(s) f^2(t) \right\} + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} + \frac{s_{k,3}}{s_{k,2}^2} \right), \\
\text{Cov}(Q_{12,k}, Q_{21,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2}} V^2(W) f^{(1)}(s) f^{(1)}(t) + \frac{s_{k,3}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k} \alpha^2(W) V(W) f(s) f^{(1)}(s) f^{(1)}(t) \right. \\
& \quad \left. + \frac{s_{k,3}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k} \alpha^2(W) V(W) f^{(1)}(s) f(t) f^{(1)}(t) \right\} + o_p \left(\frac{1}{s_{k,2}} + \frac{s_{k,3}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k} \right), \\
\text{Cov}(Q_{12,k}, Q_{30,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} U(W) V(W) f(s) f(t) + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} \alpha(W) V(W) f(s) f^2(t) \right\} \\
& \quad + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} \right), \\
\text{Cov}(Q_{20,k}, Q_{21,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} U(W) V(W) f(s) f^{(1)}(t) + \frac{s_{k,3}}{s_{k,2}^2} \alpha^2(W) U(W) f^2(s) f^{(1)}(t) \right. \\
& \quad \left. + \frac{s_{k,3}}{s_{k,2}^2} \alpha(W) V(W) f(s) f(t) f^{(1)}(t) \right\} + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} + \frac{s_{k,3}}{s_{k,2}^2} \right), \\
\text{Cov}(Q_{20,k}, Q_{30,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} R(W) Z(W) f^{(1)}(s) f(t) + \frac{s_{k,3}}{s_{k,2}^2} \alpha(W) \beta(W) R(W) f(s) f^{(1)}(s) \right. \\
& \quad \left. f^{(1)}(t) + \frac{s_{k,3}}{s_{k,2}^2} Z(W) f^{(1)}(s) f^2(t) \right\} + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}} + \frac{s_{k,3}}{s_{k,2}^2} \right),
\end{aligned}$$

$$\begin{aligned} \text{Cov}(Q_{21,k}, Q_{30,k}) &= \{1 + I(s=t)\} \left\{ \frac{1}{s_{k,2}} U(W) Z(W) f^{(1)}(s) f^{(1)}(t) + \frac{s_{k,3}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k} \alpha(W) Z(W) f(s) f(t) f^{(1)}(t) \right. \\ &\quad \left. + \frac{s_{k,3}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k} \alpha(W) \beta(W) U(W) f(s) f^{(1)}(s) f^{(1)}(t) \right\} + o_p \left(\frac{1}{s_{k,2}} + \frac{s_{k,3}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k} \right). \end{aligned}$$

By the delta method, we can then derive the mean and variance of $\bar{\gamma}_{(K)}^{*(1,0)}(s, t)$ as

$$\begin{aligned} E \left(\bar{\gamma}_{(K)}^{*(1,0)}(s, t) \right) &= \gamma^{(1,0)}(s, t) + \frac{1}{6} \frac{\beta(W)}{\alpha(W)} \gamma^{(3,0)}(s, t) \rho_{\gamma(K),2} + \frac{1}{2} \alpha(W) \gamma^{(1,2)}(s, t) \rho_{\gamma(K),2} + o_p \left(\rho_{\gamma(K),2} \right), \\ \text{Var} \left(\bar{\gamma}_{(K)}^{*(1,0)}(s, t) \right) &= \{1 + I(s=t)\} \left\{ \frac{\rho_{\gamma(K),-4}}{S_{K,2}} \frac{R(W) U(W)}{\alpha^2(W)} \frac{E_1(s, t)}{f(s) f(t)} + \frac{S_{K,3} \rho_{\gamma(K),-3}}{S_{K,2}^2} \frac{U(W)}{\alpha^2(W)} \frac{E_2(s, t)}{f(s)} \right\} \\ &\quad + \frac{S_{K,4}}{S_{K,2}^2} \frac{E_{\Phi} \left[\left\{ \Phi(s) \Phi(t) f^{(1)}(s) + \Phi^{(1)}(s) \Phi(t) f(s) \right\}^2 \right]}{f^2(s)} - \left\{ \gamma^{(1,0)}(s, t) f(s) + \gamma(s, t) f^{(1)}(s) \right\}^2 \\ &\quad + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} + \frac{s_{k,4}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k}^2 \right) \\ &= \frac{S_{K,2}}{\Gamma_{\gamma}(s, t) \rho_{\gamma(K),-4}} + o_p \left(\frac{1}{s_{k,2} \tilde{\eta}_{\gamma(K),k}^2} + \frac{s_{k,3}}{s_{k,2}^2 \tilde{\eta}_{\gamma(K),k}} + \frac{s_{k,4}}{s_{k,2}^2} \tilde{\eta}_{\gamma(K),k}^2 \right), \end{aligned}$$

where $\Gamma_{\gamma}(s, t)$ is the same as defined in Theorem 2. Finally, the asymptotic normality of $\bar{\gamma}_{(K)}^{*(1,0)}(s, t)$ follows that as $K \rightarrow \infty$,

$$\begin{aligned} &\sqrt{\frac{S_{K,2}}{\Gamma_{\gamma}(s, t) \rho_{\gamma(K),-4}}} \left\{ \bar{\gamma}_{(K)}^{*(1,0)}(s, t) - \gamma^{(1,0)}(s, t) - \frac{1}{6} \frac{\beta(W)}{\alpha(W)} \gamma^{(3,0)}(s, t) \rho_{\gamma(K),2} - \frac{1}{2} \alpha(W) \gamma^{(1,2)}(s, t) \rho_{\gamma(K),2} \right\} \\ &\xrightarrow{d} N(0, 1), \end{aligned}$$

and so is for $\bar{\gamma}_{(K)}^{(1,0)}(s, t)$. This proves Theorem 2.

S9. Proof of Theorem 3

To deduce the convergence rate of the online learning bandwidth, we can generalize Theorem 1 and Theorem 2 both to the d -dimensional setting. Suppose that Assumptions 1, 2 and the assumptions for the asymptotic normality in the d -dimensional setting hold. If $\tilde{h}_{(K)} - \tilde{h}_{(K)}^{\text{opt}} = o_p(S_{K,d}^{-1/(d+6)})$, for any point $x \in [\tilde{h}_{(K)}, 1 - \tilde{h}_{(K)}]$, as $K \rightarrow \infty$, we have

$$\left\{ \frac{S_{K,d}}{\Gamma(x) \rho_{(K),-(2+d)}} \right\}^{1/2} \left\{ \tilde{m}_{(K)}^{(1)}(x) - m^{(1)}(x) - \frac{1}{6} \frac{\beta(W)}{\alpha(W)} \left(\sum_{j=1}^d m_j^{(3)}(x) \right) \rho_{(K),2} \right\} \xrightarrow{d} N(0, 1).$$

Denote $\theta = \int \{\sum_{j=1}^d m_j^{(3)}(x)\}^2 f(x) dx$ and $\nu = \int \Gamma(x) f(x) dx$. We use the online local linear regression to estimate $f(x)$. For the online estimation $\tilde{\theta}$ and $\tilde{\nu}$, we apply the online local polynomial regression of order 4 and p' , respectively. Let $\{\eta_{\theta_{(K)}, \tilde{l}}\}_{\tilde{l}=1}^J$ and $\{\eta_{\nu_{(K)}, \tilde{l}}\}_{\tilde{l}=1}^J$ be the candidate bandwidths of $\tilde{\theta}$ and $\tilde{\nu}$, respectively. Based on the MISE criterion, the optimal bandwidths and the online learning bandwidths are

$$\tilde{h}_{(K)}^{*opt} = \left\{ 9(2+d) \frac{\alpha^2(W)}{\beta^2(W)} \frac{\nu}{\theta} \right\}^{1/(d+6)} S_{K,d}^{-1/(d+6)} \text{ and } \tilde{h}_{(K)}^{opt} = \left\{ 9(2+d) \frac{\alpha^2(W)}{\beta^2(W)} \frac{\tilde{\nu}}{\tilde{\theta}} \right\}^{1/(d+6)} S_{K,d}^{-1/(d+6)}.$$

If the online learning bandwidths of $\tilde{\theta}$ and $\tilde{\nu}$ are $h_{\tilde{\theta}_{(K)}} = G_{\theta} S_{K,d}^{-1/(d+8)}$ and $h_{\tilde{\nu}_{(K)}} = G_{\nu} S_{K,d}^{-1/(d+2p'+2)}$ with $G_{\theta}, G_{\nu} = O_p(1)$, and the candidate bandwidths are $\eta_{\theta_{(K)}, \tilde{l}} = \{(J - \tilde{l} + 1)/J\}^{1/(d+8)} h_{\tilde{\theta}_{(K)}}$ and $\eta_{\nu_{(K)}, \tilde{l}} = \{(J - \tilde{l} + 1)/J\}^{1/(d+2p'+2)} h_{\tilde{\nu}_{(K)}}$, we can prove that as $K \rightarrow \infty$, $\tilde{h}_{(K)}$ satisfies

$$\frac{\tilde{h}_{(K)}^{opt} - \tilde{h}_{(K)}^{*opt}}{\tilde{h}_{(K)}^{*opt}} = O_p\left(\frac{\tilde{\nu} - \nu}{\nu}\right) + O_p\left(\frac{\tilde{\theta} - \theta}{\theta}\right).$$

The details of the proof are set out as follows. Based on the above notations,

$$\begin{aligned} \frac{\tilde{h}_{(K)}^{opt} - \tilde{h}_{(K)}^{*opt}}{\tilde{h}_{(K)}^{*opt}} &= \frac{\left(\frac{\tilde{\nu}}{\tilde{\theta}}\right)^{1/(d+6)} - \left(\frac{\nu}{\theta}\right)^{1/(d+6)}}{\left(\frac{\nu}{\theta}\right)^{1/(d+6)}} \\ &= \left\{ \frac{\theta}{\tilde{\theta}} \left(1 + \frac{\tilde{\nu} - \nu}{\nu}\right) \right\}^{1/(d+6)} - 1 \\ &= \left(\frac{\theta}{\tilde{\theta}}\right)^{1/(d+6)} \left(1 + \frac{1}{d+6} \frac{\tilde{\nu} - \nu}{\nu}\right) - 1 \\ &= \frac{1}{d+6} \left(\frac{\theta}{\tilde{\theta}}\right)^{1/(d+6)} \frac{\tilde{\nu} - \nu}{\nu} + \left(\frac{\theta}{\tilde{\theta}}\right)^{1/(d+6)} - 1 \\ &= \frac{1}{d+6} \left(\frac{\theta}{\tilde{\theta}}\right)^{1/(d+6)} \frac{\tilde{\nu} - \nu}{\nu} + \left(1 + \frac{\theta - \tilde{\theta}}{\tilde{\theta}}\right)^{1/(d+6)} - 1 \\ &= \frac{1}{d+6} \left(\frac{\theta}{\tilde{\theta}}\right)^{1/(d+6)} \frac{\tilde{\nu} - \nu}{\nu} + \frac{1}{d+6} \frac{\theta - \tilde{\theta}}{\tilde{\theta}}. \end{aligned}$$

The first-order derivatives of the mean and covariance functions correspond to $d = 1$ and $d = 2$, respectively. Denoting $\mathbb{X} = \{X_{ki} \in R^d, k = 1, \dots, K, i = 1, \dots, n_k\}$, we have

$$\begin{aligned} \theta - \tilde{\theta} &= O_p\left(\int \left\{m_j^{(3)}(x)\right\}^2 - \left\{\tilde{m}_j^{(3)}(x)\right\}^2 f(x) dx\right) \\ &= O_p\left(\int \left\{m_j^{(3)}(x) - \tilde{m}_j^{(3)}(x)\right\} \left\{m_j^{(3)}(x) + \tilde{m}_j^{(3)}(x)\right\} f(x) dx\right) \\ &= O_p\left(\left[\int \left\{m_j^{(3)}(x) - \tilde{m}_j^{(3)}(x)\right\}^2 f(x) dx\right]^{1/2} \left[\int \left\{m_j^{(3)}(x) + \tilde{m}_j^{(3)}(x)\right\}^2 f(x) dx\right]^{1/2}\right) \end{aligned}$$

$$= O_p \left(\left[\int \left\{ m_j^{(3)}(x) - \tilde{m}_j^{(3)}(x) \right\}^2 dx \right]^{1/2} f(x) \right).$$

Then, $E\{(\theta - \tilde{\theta})^2 | \mathbb{X}\} = O_p(E\{\{\tilde{m}_j^{(3)}(x) - m_j^{(3)}(x)\}^2 | \mathbb{X}\})$. We can choose the online local cubic regression with the bandwidth $h_{\tilde{\theta}(K)} = O_p(S_{K,d}^{-1/(d+8)})$ to estimate $m_j^{(3)}(x)$. This leads to $\tilde{\theta} - \theta = O_p(S_{K,d}^{-(d+4)/(d+8)})$. Under the different degree of the data density, ν in the asymptotic normal distribution is different. The order of $\tilde{\nu} - \nu$ is determined by the online bandwidth selection. For the dense functional data, when $d = 1$, we have $\tilde{\nu} - \nu = O_p(S_{K,1}^{-2/9})$. When $d = 2$, we have $\tilde{\nu} - \nu = O_p(S_{K,2}^{-1/4})$. For the non-dense functional data, when $d = 1$, we have $\tilde{\nu} - \nu = O_p(S_{K,1}^{-2/5})$. When $d = 2$, we have $\tilde{\nu} - \nu = O_p(S_{K,2}^{-1/3})$. This shows that $\tilde{\nu} - \nu$ dominates the convergence rate of $(\tilde{h}_{(K)}^{opt} - \tilde{h}_{(K)}^{*opt})/\tilde{h}_{(K)}^{opt}$, which thus proves Theorem 3.

S10. Proof of Theorem 4

Since $\tilde{h}_{(K)} = \eta_{(K),1} > \dots > \eta_{(K),L}$, without loss of generality, we set the expression of $\eta_{(K),\tilde{l}}$, $\eta_{(K),\tilde{l}} = g(\tilde{l})\tilde{h}_{(K)}$, where $g(\tilde{l})$ decreases with respect to $\tilde{l}$, $\tilde{l} = 1, \dots, L$ and $g(1) = 1$. Using $\tilde{h}_{(K)} - \hat{h}_{(K)} = o_p(S_{K,d}^{-1/(d+6)})$, we have

$$\frac{\eta_{(K),\tilde{l}}}{\hat{h}_{(K)}} = g(\tilde{l}) \left(\frac{S_{k,d}}{S_{K,d}} \right)^{-1/(d+6)} + o_p(1) = \left\{ g(\tilde{l})^{-(d+6)} \frac{S_{k,d}}{S_{K,d}} \right\}^{-1/(d+6)} + o_p(1).$$

Under Assumption 5, $S_{k,d}/S_{K,d}$ increases linearly with respect to k . This shows that the form of $g(\tilde{l})^{-(d+6)}$ should be linear and $g(\tilde{l})^{-(d+6)} \in (0, 1]$, yielding the candidate bandwidth as

$$g(\tilde{l}) = \left(\frac{L - \tilde{l} + 1}{L} \right)^{1/(d+6)}.$$

Based on the optimal form of the candidate bandwidths, we derive the lower bound for the relative efficiency ratio to assess the performance of the proposed online derivative estimation. When the data stream in continuously, there exists a breakpoint K_0 in the candidate bandwidth sequence selected for K data block $\{\tilde{\eta}_{(K),k}\}_1^K$. To be specific, when $k \leq K_0$, because of the restricted length of the candidates, $\tilde{\eta}_{(K),k}$ can only select the L th candidate bandwidth so as to be closer to $\tilde{h}_{(K)}$, then $\tilde{\eta}_{(K),k} \geq \tilde{h}_{(K)}$. When $k > K_0$, there are enough candidates and we can choose one of them to almost replace $\tilde{h}_{(K)}$. Hence, we have

$$\frac{\tilde{\eta}_{(K),k}}{\tilde{h}_{(K)}} = \begin{cases} L^{-1/(d+6)} \left(\frac{S_{k,d}}{S_{K,d}} \right)^{-1/(d+6)}, & k \leq K_0 \\ \left(1 + \left| \frac{S_{k,d}}{S_{K,d}} - \frac{L - \tilde{l}_k + 1}{L} \right| \right)^{1/(d+6)}, & k > K_0 \end{cases}$$

where $\tilde{l}_k = \arg \min_{1 \leq \tilde{l} \leq L} |S_{k,d}/S_{K,d} - (L - \tilde{l}_k + 1)/L|$. We further deduce that

$$\begin{aligned} \text{reff} \left(\tilde{m}_{(K)}^{(1)} \right)^{-1} &= \frac{\text{MISE} \left(\tilde{m}_{(K)}^{(1)} \right)}{\text{MISE} \left(\hat{m}_{(K)}^{(1)} \right)} \\ &= \frac{d+2}{d+6} \left\{ \sum_{k=1}^K \frac{s_{k,d}}{S_{K,d}} \left(\frac{\tilde{\eta}_{(K),k}}{\tilde{h}_{(K)}} \right)^2 \right\}^2 + \frac{4}{d+6} \left\{ \sum_{k=1}^K \frac{s_{k,d}}{S_{K,d}} \left(\frac{\tilde{\eta}_{(K),k}}{\tilde{h}_{(K)}} \right)^{-(d+2)} \right\} + O_p \left(\frac{\tilde{\nu} - \nu}{\nu} \right). \end{aligned}$$

When $k \leq K_0$, we have

$$\begin{aligned} \sum_{k=1}^{K_0} \frac{s_{k,d}}{S_{K,d}} \left(\frac{\tilde{\eta}_{(K),k}}{\tilde{h}_{(K)}} \right)^2 &= \frac{d+6}{d+4} L^{-2/(d+6)} \left(\frac{S_{K_0,d}}{S_{K,d}} \right)^{1-2/(d+6)}, \\ \sum_{k=1}^{K_0} \frac{s_{k,d}}{S_{K,d}} \left(\frac{\tilde{\eta}_{(K),k}}{\tilde{h}_{(K)}} \right)^{-(d+2)} &= \frac{d+6}{2d+8} L^{(d+2)/(d+6)} \left(\frac{S_{K_0,d}}{S_{K,d}} \right)^{1+(d+2)/(d+6)}. \end{aligned}$$

When $k > K_0$, we have

$$\begin{aligned} \sum_{k=K_0+1}^K \frac{s_{k,d}}{S_{K,d}} \left(\frac{\tilde{\eta}_{(K),k}}{\tilde{h}_{(K)}} \right)^2 &= 1 - \frac{S_{K_0,d}}{S_{K,d}}, \\ \sum_{k=K_0+1}^K \frac{s_{k,d}}{S_{K,d}} \left(\frac{\tilde{\eta}_{(K),k}}{\tilde{h}_{(K)}} \right)^{-(d+2)} &= 1 - \frac{S_{K_0,d}}{S_{K,d}}. \end{aligned}$$

The breakpoint K_0 requires that

$$L^{-\frac{1}{d+6}} \sum_{k=1}^{K_0} \frac{s_{k,d}}{S_{K_0,d}} \tilde{h}_{(k)} \geq \tilde{h}_{(K)}.$$

Hence, under Assumption 5, we have

$$\frac{S_{K_0,d}}{S_{K,d}} \leq \left(\frac{d+6}{d+5} \right)^{d+6} \frac{1}{L}.$$

Letting $\rho = S_{K_0,d}/S_{K,d} \in [0, \{(d+6)/(d+5)\}^{d+6}/L]$, we can thus obtain that

$$\begin{aligned} \text{reff} \left(\tilde{m}_{(K)}^{(1)} \right)^{-1} &= \frac{d+2}{d+6} \left(\frac{d+6}{d+4} L^{-2/(d+6)} \rho^{1-2/(d+6)} + 1 - \rho \right)^2 \\ &\quad + \frac{4}{d+6} \left(\frac{d+6}{2d+8} L^{(d+2)/(d+6)} \rho^{1+(d+2)/(d+6)} + 1 - \rho \right) + O_p \left(\frac{\tilde{\nu} - \nu}{\nu} \right), \end{aligned}$$

which has the maximum value at $\rho = \{(d+6)/(d+5)\}^{d+6}/L$. Finally, we have

$$\text{reff} \left(\tilde{m}_{(K)}^{(1)} \right) \geq (1 + c_1 L^{-1} + c_2 L^{-2})^{-1} + O_p \left(\frac{\tilde{\nu} - \nu}{\nu} \right),$$

where

$$c_1 = \frac{2}{d+4} \left\{ \left(\frac{d+6}{d+5} \right)^{2d+8} - \frac{(2d^2 + 19d + 46)(d+6)^{d+4}}{(d+5)^{d+6}} \right\},$$

$$c_2 = \frac{(d+2)(d+6)^{2d+9}}{(d+4)^2(d+5)^{2d+12}}.$$

Especially, $\tilde{\mu}_{(K)}^{(1)}(t)$ and $\tilde{\gamma}_{(K)}^{(1,0)}(s, t)$ respectively correspond to $d = 1$ and $d = 2$,

$$\text{reff} \left(\tilde{\mu}_{(K)}^{(1)} \right) \geq (1 + 0.2596L^{-1} + 0.0030L^{-2})^{-1} + O_p \left(\frac{\tilde{\nu}_\mu - \nu_\mu}{\nu_\mu} \right),$$

$$\text{reff} \left(\tilde{\gamma}_{(K)}^{(1,0)} \right) \geq (1 + 0.2604L^{-1} + 0.0018L^{-2})^{-1} + O_p \left(\frac{\tilde{\nu}_\gamma - \nu_\gamma}{\nu_\gamma} \right).$$

By Theorem 3, regardless of whether the functional data are dense or not, we have $O_p((\tilde{\nu}_\mu - \nu_\mu)/\nu_\mu) = O_p(S_{K,1}^{-2/9})$ and $O_p((\tilde{\nu}_\gamma - \nu_\gamma)/\nu_\gamma) = O_p(S_{K,1}^{-1/4})$. This proves Theorem 4.

References

- Antoniadis, A. and Fan, J. (2001). Regularization of wavelet approximations, *Journal of the American Statistical Association* **96**(455): 939–967.
- Bhattacharya, P. and Müller, H. G. (1993). Asymptotics for nonparametric regression, *Sankhyā: The Indian Journal of Statistics, Series A* **55**(3): 420–441.
- Boker, S. M., Deboeck, P. R., Edler, C. and Keel, P. K. (2011). Generalized local linear approximation of derivatives from time series, *Statistical Methods for Modeling Human Dynamics*, Routledge, pp. 161–178.
- Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent, *Proceedings of COMP-STAT'2010: 19th International Conference on Computational Statistics, Paris France, August 22-27, 2010*, Springer, pp. 177–186.
- Brabanter, K. D., Brabanter, J. D., Suykens, J. A. and Moor, B. D. (2011). Kernel regression in the presence of correlated errors, *Journal of Machine Learning Research* **12**(6): 1955–1976.

- Cao, G., Wang, L., Li, Y. and Yang, L. (2016). Oracle-efficient confidence envelopes for covariance functions in dense functional data, *Statistica Sinica* **26**(1): 359–383.
- Chu, Y. and Chan, S. (2015). A new local polynomial modeling-based variable forgetting factor RLS algorithm and its acoustic applications, *IEEE/ACM Transactions on Audio, Speech, and Language Processing* **23**(11): 2059–2069.
- Claeskens, G., Krivobokova, T. and Opsomer, J. D. (2009). Asymptotic properties of penalized spline estimators, *Biometrika* **96**(3): 529–544.
- Cox, C. (2005). Delta method, *Encyclopedia of Biostatistics*, John Wiley & Sons, Ltd, Chichester, pp. 1125–1127.
- Cramér, H. and Wold, H. (1936). Some theorems on distribution functions, *Journal of the London Mathematical Society* **1**(4): 290–294.
- Donoho, D. L. and Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage, *Biometrika* **81**(3): 425–455.
- Fan, J. and Yao, Q. (1998). Efficient estimation of conditional variance functions in stochastic regression, *Biometrika* **85**(3): 645–660.
- Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications*, Vol. 200, Springer, New York.
- 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**(1): 3321–3351.
- Mateu, J. and Romano, E. (2017). Advances in spatial functional statistics, *Stochastic Environmental Research and Risk Assessment* **31**(1): 1–6.
- Moosavi, S., Samavatian, M. H., Parthasarathy, S. and Ramnath, R. (2019). A countrywide traffic accident dataset, *arXiv preprint arXiv:1906.05409*.
- Moosavi, S., Samavatian, M. H., Parthasarathy, S., Teodorescu, R. and Ramnath, R. (2019). Accident risk prediction based on heterogeneous sparse data: new dataset and insights, *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 33–42.

- Pigoli, D. and Sangalli, L. M. (2012). Wavelets in functional data analysis: estimation of multidimensional curves and their derivatives, *Computational Statistics & Data Analysis* **56**(6): 1482–1498.
- Ruppert, D., Sheather, S. J. and Wand, M. P. (1995). An effective bandwidth selector for local least squares regression, *Journal of the American Statistical Association* **90**(432): 1257–1270.
- Sangalli, L. M., Secchi, P., Vantini, S. and Veneziani, A. (2009). Efficient estimation of three-dimensional curves and their derivatives by free-knot regression splines, applied to the analysis of inner carotid artery centrelines, *Journal of the Royal Statistical Society: Series C* **58**(3): 285–306.
- Sheather, S. J. and Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation, *Journal of the Royal Statistical Society: Series B* **53**(3): 683–690.
- Sutskever, I., Martens, J., Dahl, G. and Hinton, G. (2013). On the importance of initialization and momentum in deep learning, *International Conference on Machine Learning*, PMLR, pp. 1139–1147.
- Wand, M. P. and Jones, M. C. (1994). *Kernel Smoothing*, Chapman & Hall/CRC, New York.
- Wang, W. and Yu, P. (2017). Asymptotically optimal differenced estimators of error variance in nonparametric regression, *Computational Statistics & Data Analysis* **105**: 125–143.
- Wood, S. N., Pya, N. and Säfken, B. (2016). Smoothing parameter and model selection for general smooth models, *Journal of the American Statistical Association* **111**(516): 1548–1563.
- Yang, Y. and Yao, F. (2023). Online estimation for functional data, *Journal of the American Statistical Association* **118**(543): 1630–1644.
- Yang, Y., Yao, F. and Zhao, P. (2024). Online smooth backfitting for generalized additive models, *Journal of the American Statistical Association* **119**(546): 1215–1228.
- Zhang, H. and Li, Y. (2022). Unified principal component analysis for sparse and dense functional data under spatial dependency, *Journal of Business & Economic Statistics* **40**(4): 1523–1537.
- Zhang, J.-T. and Chen, J. (2007). Statistical inferences for functional data, *The Annals of Statistics* **35**(3): 1052–1079.
- Zhang, X. and Wang, J.-L. (2016). From sparse to dense functional data and beyond, *The Annals of Statistics* **44**(5): 2281–2321.