Supplementary file of the paper titled "A Three-Step Local Smoothing Approach for Estimating the Mean and Covariance Functions of Spatio-Temporal Data"

Kai Yang and Peihua Qiu

Department of Biostatistics, University of Florida

Gainesville, FL 32610

To save some space in the paper of the above title, the proofs of all the three Lemmas and one Theorem in the manuscript and some additional simulation results are presented in this supplementary file. The equations in this supplementary file are labeled as (A.1), (A.2), and so on; while we use labels like (1), (2), and so on, to refer to the equations in the manuscript.

Appendix A: Proof of Lemma 1

From the definition of $g_n(\theta; t, \boldsymbol{s})$, we have

$$g_n(\theta; t, \boldsymbol{s}) = V(t, t; \boldsymbol{s}, \boldsymbol{s}) + 2\sum_{k=1}^{\tau_n} V(t, t + k/n; \boldsymbol{s}, \boldsymbol{s}) \cos(2\theta k\pi),$$

where $\tau_n = \lfloor n(1-t) \rfloor$ is the greatest integer less than or equal to $n(1-t)$. Next we will show that $g(\theta; t, \boldsymbol{s}) = \lim_{n\to\infty} g_n(\theta; t, \boldsymbol{s})$ is well-defined for all $(t, \boldsymbol{s}) \in [0, 1] \times \Omega$. When $t = 1$, it is clear that $g_n(\theta; 1, \boldsymbol{s}) = V(1, 1; \boldsymbol{s}, \boldsymbol{s})$ for all n. So, $g(\theta; 1, \boldsymbol{s}) = \lim_{n\to\infty} g_n(\theta; 1, \boldsymbol{s}) = V(1, 1; \boldsymbol{s}, \boldsymbol{s})$ is well-defined for all $\boldsymbol{s}$. Next we consider the case when $(t, \boldsymbol{s}) \in [0, 1) \times \Omega$. For any $\theta \in [-0.5, 0.5]$, and two integers $n_1 \leq n_2$, when n_1 is large enough, we have $\log(n_1) \leq n_1(1-t)$. Then, it can be checked that

$$\begin{aligned}|g_{n_2}(\theta; t, \boldsymbol{s}) - g_{n_1}(\theta; t, \boldsymbol{s})| \leq\ & 2\sum_{k=1}^{\widetilde{\tau}_{n_1}} |V(t, t + k/n_2; \boldsymbol{s}, \boldsymbol{s}) - V(t, t + k/n_1; \boldsymbol{s}, \boldsymbol{s})| \\ & + 2\sum_{k=\widetilde{\tau}_{n_1}+1}^{\tau_{n_1}} |V(t, t + k/n_1; \boldsymbol{s}, \boldsymbol{s})| + 2\sum_{k=\widetilde{\tau}_{n_1}+1}^{\tau_{n_2}} |V(t, t + k/n_2; \boldsymbol{s}, \boldsymbol{s})|,\end{aligned}$$

where $\widetilde{\tau}_{n_1} = \lfloor \log(n_1) \rfloor$. Since $\{\varepsilon_0(t_i, \boldsymbol{s})\}$ is a strong mixing sequence, by the Davydov's inequality, we have

$$|\mathrm{Cov}(\varepsilon_0(t_i, \boldsymbol{s}), \varepsilon_0(t_k, \boldsymbol{s}))| \leq 12 C_\varepsilon^{2/\delta} C_0^{(\delta-2)/\delta} \exp(-C_1|k - i|(\delta - 2)/\delta),$$

for all i, k and $\boldsymbol{s}$, where C_ε, C_0, C_1 and δ are defined in Lemma 1. Thus, $V(t, t+\tau/n; \boldsymbol{s}, \boldsymbol{s}) = O(\exp(-C_1\tau(\delta-2)/\delta))$, for all $(t, \boldsymbol{s})$ and τ. Then, we have

$$\begin{aligned}
|g_{n_2}(\theta; t, \boldsymbol{s}) - g_{n_1}(\theta; t, \boldsymbol{s})| \leq & 2\sum_{k=1}^{\widetilde{\tau}_{n_1}} |V(t, t+k/n_2; \boldsymbol{s}, \boldsymbol{s}) - V(t, t+k/n_1; \boldsymbol{s}, \boldsymbol{s})| \\
& + O(\exp(-C_1\widetilde{\tau}_{n_1}(\delta-2)/\delta)) + O(\exp(-C_1\widetilde{\tau}_{n_1}(\delta-2)/\delta)) \\
= & O((n_2 - n_1)\widetilde{\tau}_{n_1}^2/(n_1 n_2)) + O(\exp(-C_1\widetilde{\tau}_{n_1}(\delta-2)/\delta)), \\
= & O(\log(n_1)^2/n_1) + O(\exp(-C_1\log(n_1)(\delta-2)/\delta)) \\
= & O(\log(n_1)^2/n_1 + \exp(-C_1\log(n_1)(\delta-2)/\delta))
\end{aligned}$$

Note that $\log(n_1)^2/n_1 \to 0$ and $\exp(-C_1\log(n_1)(\delta-2)/\delta) \to 0$ as $n_1 \to \infty$. Thus, $\{g_n(\theta; t, \boldsymbol{s}), n \geq 1\}$ is a Cauchy sequence. So, by the Cauchy convergence criterion, we know $g_n(\theta; t, \boldsymbol{s})$ is convergent. Thus, $g(\theta; t, \boldsymbol{s})$ is well-defined. For any $\theta, \theta' \in [-0.5, 0.5]$ and $\epsilon > 0$, there exists a postive integer $\widetilde{n}$, when $n \geq \widetilde{n}$, we have $|g(\theta; t, \boldsymbol{s}) - g_n(\theta; t, \boldsymbol{s})| < \epsilon$ and $|g(\theta'; t, \boldsymbol{s}) - g_n(\theta'; t, \boldsymbol{s})| < \epsilon$. It follows that, when $n \geq \widetilde{n}$,

$$\begin{aligned}
|g(\theta; t, \boldsymbol{s}) - g(\theta'; t, \boldsymbol{s})| \leq & |g_n(\theta; t, \boldsymbol{s}) - g_n(\theta'; t, \boldsymbol{s})| + 2\epsilon \\
\leq & 4\pi|\theta - \theta'| \sum_{k=1}^{\tau_n} k|V(t, t+k/n; \boldsymbol{s}, \boldsymbol{s})| + 2\epsilon \\
\leq & 48\pi C_\varepsilon^{2/\delta} C_0^{(\delta-2)/\delta} |\theta - \theta'| \sum_{k=1}^{\tau_n} k\exp(-C_1 k(2-\delta)/\delta) + 2\epsilon \\
\leq & C_\theta|\theta - \theta'| + 2\epsilon,
\end{aligned}$$

where $C_\theta = 48\pi C_\varepsilon^{2/\delta} C_0^{(\delta-2)/\delta} \sum_{k=1}^{\infty} k\exp(-C_1 k(2-\delta)/\delta)$. Since $\epsilon > 0$ is arbitrary, we have $g(\theta; t, \boldsymbol{s})$ is Lipchitz-1 continuous with respect to θ.

Next, we show that $g(\theta; t, \boldsymbol{s}) \geq 0$, for all $\theta \in [-0.5, 0.5]$ and $(t, \boldsymbol{s})$, under the conditions in Lemma 1. If this is not true, then there exist $\theta_0 \in [-0.5, 0.5]$ and $(t_0, \boldsymbol{s}_0) \in [0, 1] \times \Omega$ such that $g(\theta_0; t_0, \boldsymbol{s}_0) < 0$. Consider the $M \times M$ Toeplitz matrix $\mathbf{M}_T$, where the (i, j)-th element of $\mathbf{M}_T$ is $V(t_0, t_0 + \tau/n; \boldsymbol{s}_0, \boldsymbol{s}_0)$ and $\tau = |i - j|$. For the matrix $\mathbf{M}_T$, Grenander and Szegö (1958) proved that its smallest eigenvalue would decrease with M and converge to $\inf\{g(\theta; t_0, \boldsymbol{s}_0), \theta \in [-0.5, 0.5]\}$. Thus, when M is large, we have $\lambda_0 < g(\theta_0; t_0, \boldsymbol{s}_0)/2$, where λ_0 is the smallest eigenvalue of $\mathbf{M}_T$. The corresponding standardized eigenvector is denoted as $\boldsymbol{x}_0$ which has the property that $\boldsymbol{x}_0^T\boldsymbol{x}_0 = 1$. Let $1 \leq i_0 \leq n$ be the closest integer to $n \times t_0$. Then, $|t_{i_0} - t_0| \leq 1/n$, where $t_{i_0} = i_0/n$. Consider the covariance matrix of $\{\varepsilon_0(t_{i_0}, \boldsymbol{s}_0), \ldots, \varepsilon_0(t_{i_0+M-1}, \boldsymbol{s}_0)\}$, denoted as $\mathbf{M}_T^*$. Because the covariance function is assumed to be twice differentiable, it can be shown that all elements of $\mathbf{M}_T^* - \mathbf{M}_T$ are of the order $O(M/n)$ uniformly. So, the largest eigenvalue of $\mathbf{M}_T^* - \mathbf{M}_T$ is of the order $O(M^2/n)$, and we can find some constant $C_{11} > 0$ such that $C_{11}M^2/n\mathbf{I}_M - (\mathbf{M}_T^* - \mathbf{M}_T)$ is positive-semidefinite, where $\mathbf{I}_M$ is a $M \times M$ identity matrix. It follows that $\boldsymbol{x}_0^T[C_{11}M^2/n\mathbf{I}_M - (\mathbf{M}_T^* - \mathbf{M}_T)]\boldsymbol{x}_0 = C_{11}M^2/n - \boldsymbol{x}_0^T\mathbf{M}_T^*\boldsymbol{x}_0 + \lambda_0 \geq 0$. Hence, we have $C_{11}M^2/n \geq -\lambda_0$ for all n. On the other hand, when n is large enough, $C_{11}M^2/n < -g(\theta_0; t_0, \boldsymbol{s}_0)/2 < -\lambda_0$. Thus, we have a contradiction. Therefore, we must have $g(\theta; t, \boldsymbol{s}) \geq 0$, for all $\theta \in [-0.5, 0.5]$ and $(t, \boldsymbol{s})$. The result in Lemma 1 has then been proved.

Appendix B: Proof of Lemma 2

For simplicity of expression, we use y_{ij} and ε_{ij} to denote $y(t_i, \boldsymbol{s}_{ij})$ and $\varepsilon(t_i, \boldsymbol{s}_{ij})$, respectively. At a given point $(t, \boldsymbol{s}) \in [0, 1] \times \Omega$, from (4), we have

$$\begin{aligned}\widetilde{\lambda}(t, \boldsymbol{s}) &= \boldsymbol{e}_1^T(\mathbf{X}^T\boldsymbol{D}_K\mathbf{X})^{-1}\mathbf{X}^T\boldsymbol{D}_K\mathbf{Y} \\ &= \boldsymbol{e}_1^T(\mathbf{X}^T\boldsymbol{D}_K\mathbf{X})^{-1}\mathbf{X}^T\boldsymbol{D}_K\boldsymbol{\lambda} + \boldsymbol{e}_1^T(\mathbf{X}^T\boldsymbol{D}_K\mathbf{X})^{-1}\mathbf{X}^T\boldsymbol{D}_K\boldsymbol{\varepsilon} \\ &= \Pi_1 + \Pi_2,\end{aligned} \tag{A.1}$$

where $\boldsymbol{\lambda} = E(\mathbf{Y})$ and $\boldsymbol{\varepsilon} = (\varepsilon_{11}, \ldots, \varepsilon_{1m_1}, \ldots, \varepsilon_{nm_n})^T$. For Π_1, by the Taylor's expansion, it can be shown that

$$\Pi_1 = \boldsymbol{e}_1^T(\mathbf{X}^T\boldsymbol{D}_K\mathbf{X})^{-1}\mathbf{X}^T\boldsymbol{D}_K(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\mathcal{R}}) = \lambda(t, \boldsymbol{s}) + \Pi_3, \tag{A.2}$$

where $\boldsymbol{\beta} = (\lambda(t, \boldsymbol{s}), \partial\lambda(t, \boldsymbol{s})/\partial t, \partial\lambda(t, \boldsymbol{s})/\partial\boldsymbol{s})^T$, $\boldsymbol{\mathcal{R}} = (r_{11}, \ldots, r_{1m_1}, \ldots, r_{nm_n})^T$, $r_{ij} = ((t_i - t), (\boldsymbol{s}_{ij} - \boldsymbol{s})^T)\ \boldsymbol{\mathcal{H}}(t'_{ij}, \boldsymbol{s}'_{ij})((t_i - t), (\boldsymbol{s}_{ij} - \boldsymbol{s})^T)^T$, $\boldsymbol{\mathcal{H}}$ is the Hessian matrix of $\lambda(t, \boldsymbol{s})$, and $t'_{ij} \in [0, 1], \boldsymbol{s}'_{ij} \in \Omega$, for $j = 1, \ldots, m_i, i = 1, \ldots, n$. From (A.1) and (A.2), we have $\widetilde{\lambda}(t, \boldsymbol{s}) = \lambda(t, \boldsymbol{s}) + \Pi_2 + \Pi_3$. For Π_2, it can be checked that

$$\begin{aligned}\Pi_2 &= \boldsymbol{e}_1^T(\mathbf{X}^T\boldsymbol{D}_K\mathbf{X})^{-1}\mathbf{X}^T\boldsymbol{D}_K\boldsymbol{\varepsilon} \\ &= \boldsymbol{e}_1^T(\mathcal{D}(m, n)\mathbf{X}^T\boldsymbol{D}_K\mathbf{X})^{-1}\mathcal{D}(m, n)\mathbf{X}^T\boldsymbol{D}_K\boldsymbol{\varepsilon} \\ &= \boldsymbol{e}_1^T\mathbf{A}^{-1}(t, \boldsymbol{s})\mathbf{B}(t, \boldsymbol{s}),\end{aligned} \tag{A.3}$$

where $\mathcal{D}(m, n) = (nh_{t,0}mh_{s,0}^2f(\boldsymbol{s}))^{-1}$, $\mathbf{A}(t, \boldsymbol{s}) = \mathcal{D}(n, m)\mathbf{X}^T\boldsymbol{D}_K\mathbf{X}$ is a 4×4 matrix and $\mathbf{B}(t, \boldsymbol{s}) = \mathcal{D}(n, m)\mathbf{X}^T\boldsymbol{D}_K\boldsymbol{\varepsilon}$ is a vector of size 4. Next, we focus on the first element of the vector $\mathbf{B}(t, \boldsymbol{s})$, i.e.,

$$\mathbf{B}_1(t, \boldsymbol{s}) := (nh_{t,0})^{-1}\sum_{i=1}^n K_1((t_i - t)/h_{t,0})\boldsymbol{\varepsilon}_i(\boldsymbol{s}), \tag{A.4}$$

where $\boldsymbol{\varepsilon}_i(s) = \{mh_{s,0}^2 f(\boldsymbol{s})\}^{-1} \sum_{j=1}^{m_i} K_2(d_E(\boldsymbol{s}_{ij}, \boldsymbol{s})/h_{s,0})\varepsilon_{ij}$. We will show below that $\mathbf{B}_1(t, \boldsymbol{s}) = O_p(a(n, m))$ uniformly for $(t, \boldsymbol{s}) \in [0, 1] \times \Omega$, where $a(n, m) = \{\log(n)^2/(nh_{t,0}^2)\}^{1/2}$.

First, note that the spatial location of interest Ω is bounded, then it is clear that $[0, 1] \times \Omega$ can be covered by $\widetilde{N} = O(\{a(n, m)h_{t,0}\}^{-3})$ regions $\{R_l, l = 1, ..., \widetilde{N}\}$, where $R_l = \{(t, \boldsymbol{s}) : |t - t_l^*| \leq a(n, m)h_{t,0},\ d_E(\boldsymbol{s}, \boldsymbol{s}_l^*) \leq a(n, m)h_{t,0}\}$ and $\{(t_l^*, \boldsymbol{s}_l^*), l = 1, \ldots, \widetilde{N}\}$ are the centroids of the $\widetilde{N}$ regions. Since both kernel functions $K_1(x)$ and $K_2(x)$ are Lipschitz-1 continuous, let $0 < L_K < \infty$ be their Lipschitz constant. Because it is assumed that $h_{t,0}/h_{s,0} = O(1)$, we can find some constant $C_4 > 0$ such that $h_{t,0} \leq C_4 h_{s,0}$. Define $C_K = \sup_{x\in\mathbb{R}}\{K_1(x), K_2(x)\}$. Then, for any $(t, \boldsymbol{s}) \in R_l$ and a sufficiently large n, we have

$$\begin{aligned}
&\left| K_1\left(\frac{t_i - t}{h_{t,0}}\right) K_2\left(\frac{d_E(\boldsymbol{s}_{ij}, \boldsymbol{s})}{h_{s,0}}\right) - K_1\left(\frac{t_i - t_l^*}{h_{t,0}}\right) K_2\left(\frac{d_E(\boldsymbol{s}_{ij}, \boldsymbol{s}_l^*)}{h_{s,0}}\right)\right| \\
\leq & C_K L_K h_{t,0}^{-1}\{|t - t_l^*| + C_4 d_E(\boldsymbol{s}, \boldsymbol{s}_l^*)\} I\left(\frac{|t_i - t_l^*|}{h_{t,0}} \leq 2L_1\right) I\left(\frac{d_E(\boldsymbol{s}_{ij}, \boldsymbol{s}_l^*)}{h_{s,0}} \leq 2L_2\right),
\end{aligned} \tag{A.5}$$

where $[-L_1, L_1]$ and $[-L_2, L_2]$ are the compact supports for $K_1(x)$ and $K_2(x)$, respectively. Define $\widetilde{K}_1(x) = 1/(2L_1)I(|x| \leq 2L_1)$ and $\widetilde{K}_2(x) = 1/(4\pi L_2^2)I(|x| \leq 2L_2)$. Then, by (A.5), there exists a constant $C_5 > 0$ such that

$$\begin{aligned}
&\left| K_1\left(\frac{t_i - t}{h_{t,0}}\right) K_2\left(\frac{d_E(\boldsymbol{s}_{ij}, \boldsymbol{s})}{h_{s,0}}\right) - K_1\left(\frac{t_i - t_l^*}{h_{t,0}}\right) K_2\left(\frac{d_E(\boldsymbol{s}_{ij}, \boldsymbol{s}_l^*)}{h_{s,0}}\right)\right| \\
&\qquad \leq C_5 a(m, n)\widetilde{K}_1\left(\frac{t_i - t_l^*}{h_{t,0}}\right) \widetilde{K}_2\left(\frac{d_E(\boldsymbol{s}_{ij}, \boldsymbol{s}_l^*)}{h_{s,0}}\right).
\end{aligned} \tag{A.6}$$

Define

$$\begin{aligned}
\widetilde{\mathbf{B}}_1(t, \boldsymbol{s}) =& \{nmh_{t,0}h_{s,0}^2 f(\boldsymbol{s})\}^{-1} \sum_{i=1}^{n} \sum_{j=1}^{m_i} \widetilde{K}_1((t_i - t)/h_{t,0}) \\
& \times \widetilde{K}_2(d_E(\boldsymbol{s}_{ij}, \boldsymbol{s})/h_{s,0})|\varepsilon_{ij}|.
\end{aligned}$$

Since $\widetilde{K}_1(\cdot)$ and $\widetilde{K}_2(\cdot)$ satisfy the assumptions about the kernel function in Lemma 2, it can be checked that

$$E(\widetilde{\mathbf{B}}_1(t, \boldsymbol{s})) \leq C_\varepsilon^{1/\delta}(1 + O(h_{s,0}^2 + 1/(nh_{t,0}))) < \infty,$$

where δ and C_ε are defined in Lemma 2. Based on the result in (A.6), it can be shown that

$$\begin{aligned}
&\sup_{(t,\boldsymbol{s})\in R_l} \big|\mathbf{B}_1(t,\boldsymbol{s}) - E\{\mathbf{B}_1(t,\boldsymbol{s})\}\big| \\
\le&\big|\mathbf{B}_1(t_l^*,\boldsymbol{s}_l^*) - E\{\mathbf{B}_1(t_l^*,\boldsymbol{s}_l^*)\}\big| + C_5 a(n,m)\big[\widetilde{\mathbf{B}}_1(t_l^*,\boldsymbol{s}_l^*) + E\{\widetilde{\mathbf{B}}_1(t_l^*,\boldsymbol{s}_l^*)\}\big] \\
\le&\big|\mathbf{B}_1(t_l^*,\boldsymbol{s}_l^*) - E\{\mathbf{B}_1(t_l^*,\boldsymbol{s}_l^*)\}\big| + C_5 a(n,m)\big[|\widetilde{\mathbf{B}}_1(t_l^*,\boldsymbol{s}_l^*) - E\{\widetilde{\mathbf{B}}_1(t_l^*,\boldsymbol{s}_l^*)\}|\big] \\
&+ 2C_5 a(n,m) E\{\widetilde{\mathbf{B}}_1(t_l^*,\boldsymbol{s}_l^*)\} \\
\le&\big|\mathbf{B}_1(t_l^*,\boldsymbol{s}_l^*) - E\{\mathbf{B}_1(t_l^*,\boldsymbol{s}_l^*)\}\big| + C_5\big[|\widetilde{\mathbf{B}}_1(t_l^*,\boldsymbol{s}_l^*) - E\{\widetilde{\mathbf{B}}_1(t_l^*,\boldsymbol{s}_l^*)\}|\big] \\
&+ 2C_5 a(n,m) T,
\end{aligned} \tag{A.7}$$

where the final inequality is obtained because $a(n,m) < 1$ and $T > E(\widetilde{\mathbf{B}}_1(t,\boldsymbol{s}))$ when n, m and T are large enough. By (A.7), it can be shown that

$$\begin{aligned}
&\Pr\left(\sup_{(t,\boldsymbol{s})\in[0,1]\times\Omega} |\mathbf{B}_1(t,\boldsymbol{s}) - E\{\mathbf{B}_1(t,\boldsymbol{s})\}| > (2+4C_5)Ta(n,m)\right) \\
\le&\widetilde{N} \max_{1\le l\le \widetilde{N}} \Pr\left(|\mathbf{B}_1(t_l^*,\boldsymbol{s}_l^*) - E\{\mathbf{B}_1(t_l^*,\boldsymbol{s}_l^*)\}| > 2Ta(n,m)\right) \\
&+ \widetilde{N} \max_{1\le l\le \widetilde{N}} \Pr\left(|\widetilde{\mathbf{B}}_1(t_l^*,\boldsymbol{s}_l^*) - E\{\widetilde{\mathbf{B}}_1(t_l^*,\boldsymbol{s}_l^*)\}| > 2Ta(n,m)\right).
\end{aligned} \tag{A.8}$$

For the two parts on the right-hand side of (A.8), we can use similar arguments to find their upper bounds, because both $(K_1(x), K_2(x))$ and $(\widetilde{K}_1(x), \widetilde{K}_2(x))$ satisfy the assumptions about the kernel functions given in Lemma 2.

Second, for any $(t,\boldsymbol{s}) \in [0,1]\times\Omega$, by the fact that $E\{\mathbf{B}_1(t,\boldsymbol{s})\} = 0$, we have

$$\Pr\big(|\mathbf{B}_1(t,\boldsymbol{s}) - E\{\mathbf{B}_1(t,\boldsymbol{s})\}| > 2Ta(n,m)\big) = \Pr\big(|\mathbf{B}_1(t,\boldsymbol{s})| > 2Ta(n,m)\big). \tag{A.9}$$

Next, we replace $\boldsymbol{\varepsilon}_i(\boldsymbol{s})$ in $\mathbf{B}_1(t,\boldsymbol{s})$ by its truncated version $\boldsymbol{\varepsilon}_i(\boldsymbol{s})I(|\boldsymbol{\varepsilon}_i(\boldsymbol{s})| \le \varphi_n)$, and evaluate the error caused by this truncation, where $\varphi_n = \{n/\log(n)^2\}^{1/2}$ and $\boldsymbol{\varepsilon}_i(\boldsymbol{s})$ is defined in (A.4). Based on the assumption that $E|\varepsilon(t,\boldsymbol{s})|^\delta \le C_\varepsilon$ for some $\delta > 5$, it can be checked that $E\big(|\boldsymbol{\varepsilon}_i(\boldsymbol{s})|^5\big) \le \Theta_0$, for some constant $\Theta_0 > 0$. Define

$$\mathrm{TR}(t,\boldsymbol{s}) = \frac{1}{nh_{t,0}} \sum_{i=1}^n K_1((t_i - t)/h_{t,0})\boldsymbol{\varepsilon}_i(\boldsymbol{s}) I(|\boldsymbol{\varepsilon}_i(\boldsymbol{s})| > \varphi_n),$$

then we have

$$\begin{aligned}
E(|\mathrm{TR}(t,\boldsymbol{s})|) &\le \frac{1}{nh_{t,0}} \sum_{i=1}^n K_1((t_i - t)/h_{t,0}) E\{|\boldsymbol{\varepsilon}_i(\boldsymbol{s})| I(|\boldsymbol{\varepsilon}_i(\boldsymbol{s})| > \varphi_n)\} \\
&\le \frac{\Theta_0}{nh_{t,0}} \sum_{i=1}^n K_1((t_i - t)/h_{t,0})\varphi_n^{-4} = O(\varphi_n^{-4}).
\end{aligned} \tag{A.10}$$

By the Markov's inequality, it is clear that $|\mathrm{TR}(t, \boldsymbol{s}) - E(\mathrm{TR}(t, \boldsymbol{s}))| = O_p(\varphi_n^{-4})$, for any $(t, \boldsymbol{s}) \in [0, 1] \times \Omega$.

Let $\widetilde{\varepsilon}_i(t, \boldsymbol{s}) = \varepsilon_i(\boldsymbol{s}) I(|\varepsilon_i(\boldsymbol{s})| \leq \varphi_n) K_1((t_i - t)/h_{t,0})$, $Z_i(t, \boldsymbol{s}) = \widetilde{\varepsilon}_i(t, \boldsymbol{s}) - E\{\widetilde{\varepsilon}_i(t, \boldsymbol{s})\}$ and $\mathbf{B}_1^*(t, \boldsymbol{s}) = \mathbf{B}_1(t, \boldsymbol{s}) - \mathrm{TR}(t, \boldsymbol{s})$. It can be checked that $\mathbf{B}_1^*(t, \boldsymbol{s}) - E\{\mathbf{B}_1^*(t, \boldsymbol{s})\} = (nh_{t,0})^{-1} \sum_{i=1}^n Z_i(t, \boldsymbol{s})$, for $(t, \boldsymbol{s}) \in [0, 1] \times \Omega$. Given any possible values of $\{\boldsymbol{s}_{ij}, i = 1, \ldots, n, j = 1, \ldots, m_i\}$, we have $|Z_i(t, \boldsymbol{s})| \leq 2C_K \varphi_n$, and for any positive integer $L \leq n$, it can be checked that

$$E\Big[\sum_{i=1}^{L} Z_i(t, \boldsymbol{s})^2 \mid \mathcal{S}_\sigma\Big] \leq \Theta_1 L D(\boldsymbol{s})^2,$$

for some constant $\Theta_1 > 0$, where $\mathcal{S}_\sigma$ is the σ-algebra generated by the spatial locations $\mathcal{S} = \{\boldsymbol{s}_{11}, \ldots, \boldsymbol{s}_{nm_n}\}$, $D(\boldsymbol{s}) = \max_{1 \leq i \leq n} D_i(\boldsymbol{s})$, and $D_i(\boldsymbol{s}) = (mh_{s,0}^2 f(\boldsymbol{s}))^{-1} \sum_{j=1}^{m_i} K_2(d_E(\boldsymbol{s}_{ij}, \boldsymbol{s})/h_{s,0})$. Let the strong mixing coefficient of $\{\varepsilon_{ij}\}$ be define as

$$\alpha^*(k) = \sup_{n \geq 1, 1 \leq i \leq n-k} \sup_{A,B} \{|P(AB) - P(A)P(B)| : A \in \mathcal{F}_1^i, B \in \mathcal{F}_{i+k}^n\},$$

where $\mathcal{F}_{k_0}^{k_1}$ is the σ-algebra generated by $\{\varepsilon_{kl}, k_0 \leq k \leq k_1, l = 1, \ldots, m_k\}$. Since $\{\varepsilon_0(t_i, \boldsymbol{s}_{ij})\}$ are independent of $\{\varepsilon_1(t_i, \boldsymbol{s}_{ij})\}$, $\{\varepsilon_1(t_i, \boldsymbol{s}_{ij})\}$ are independent at different times and/or locations, and $\varepsilon_{ij} = \varepsilon_0(t_i, \boldsymbol{s}_{ij}) + \varepsilon_1(t_i, \boldsymbol{s}_{ij})$, from Theorem 5.2 in Bradley (2005), it can be checked that $\alpha^*(k) \leq \alpha(k)$. Note that $\{\boldsymbol{s}_{ij}, j = 1, \ldots, m_i, i = 1, \ldots, n\}$ are independent of the random errors $\{\varepsilon_{11}, \ldots, \varepsilon_{nm_n}\}$. So, for given $\{\boldsymbol{s}_{ij}, i = 1, \ldots, n, j = 1, ..., m_i\}$, $\{Z_i(t, \boldsymbol{s}) : i = 1, \ldots, n\}$ is a strong mixing sequence with the strong mixing coefficient $\{\widetilde{\alpha}(k), k = 0, 1, \ldots\}$, and $\widetilde{\alpha}(k) \leq \alpha^*(k) \leq \alpha(k)$, for $k \geq 1$. Let L_n be an integer closest to $\{T^{1/2} \log(n)\}/(10C_K)$. Then, we have $nh_{t,0} a(n, m) T > 8C_K L_n \varphi_n$ when n is large. By Theorem 2.1 in Liebscher (1996), for $1 \leq l \leq \widetilde{N}$, it can be shown that

$$\begin{aligned} &\Pr\left(|\mathbf{B}_1^*(t_l^*, \boldsymbol{s}_l^*) - E(\mathbf{B}_1^*(t_l^*, \boldsymbol{s}_l^*))| > a(n, m)T \mid \mathcal{S}_\sigma\right) \\ \leq &4 \exp\left(-\frac{T^2 \log(n)}{64\Theta_1 D(\boldsymbol{s})^2 + T^{3/2}}\right) + \frac{40nC_0C_K}{T^{1/2}\log(n)} \exp\left(-\frac{C_1 T^{1/2} \log(n)}{10C_K}\right), \end{aligned} \tag{A.11}$$

when $\log(n) > 1$. Note that the second term on the right-hand side of (A.11) is independent of the choice of $\{\boldsymbol{s}_{ij}, i = 1, \ldots, n, j = 1, ..., m_i\}$. Let $C_{\max} = \max_{1 \leq i \leq n} m_i/m$ and $C_{\min} = \min_{1 \leq i \leq n} m_i/m$, then, by the Bernstein's inequality, we have

$$\begin{aligned}
&\Pr\left(|\mathbf{B}_1^*(t_l^*, \boldsymbol{s}_l^*) - E(\mathbf{B}_1^*(t_l^*, \boldsymbol{s}_l^*))| > a(n,m)T\right) \\
\leq & E\left\{4\exp\left(-\frac{T^2\log(n)}{64\Theta_1 D(\boldsymbol{s})^2 + T^{3/2}}\right) I(D(\boldsymbol{s}) \leq C_{\max}T^{1/2})\right\} \\
&+ 4\Pr\left(D(\boldsymbol{s}) > C_{\max}T^{1/2}\right) + \frac{40nC_K}{T^{1/2}\log(n)}\exp\left(-\frac{C_1 T^{1/2}\log(n)}{10C_K}\right) \\
\leq & 4\exp\left(-\frac{T^2\log(n)}{64\Theta_1 C_{\max}^2 T + T^{3/2}}\right) + O\left(n\exp(-mh_{s,0}^2 C_{\min}C_{\max}(T^{1/2}-1))\right) \\
&+ \frac{40nC_K}{T^{1/2}\log(n)}\exp\left(-\frac{C_1 T^{1/2}\log(n)}{10C_K}\right).
\end{aligned} \tag{A.12}$$

In addition, from (A.10), by the Markov's inequality, we have

$$\Pr\left(|\mathrm{TR}(t_l^*, \boldsymbol{s}_l^*) - E\{\mathrm{TR}(t_l^*, \boldsymbol{s}_l^*)\}| > a(n,m)T\right) = O\left(\{a(n,m)T\varphi_n^4\}^{-1}\right). \tag{A.13}$$

Therefore, by combining (A.12) with (A.13), when T is large enough, we have

$$\begin{aligned}
&\Pr\left(|\mathbf{B}_1(t_l^*, \boldsymbol{s}_l^*) - E(\mathbf{B}_1(t_l^*, \boldsymbol{s}_l^*))| > 2a(n,m)T\right) = O\left(\{a(n,m)T\varphi_n^4\}^{-1}\right) \\
&+ O\left(n^{-T^{1/2}/65}\right) + O\left(n\exp(-mh_{s,0}^2 C_{\min}C_{\max}(T^{1/2}-1))\right) \\
&+ O\left(n\exp\left\{-\log(n)\frac{C_1 T^{1/2}}{10C_K}\right\}\right).
\end{aligned} \tag{A.14}$$

By (A.8) and (A.14), it can be shown that, when T is large enough,

$$\begin{aligned}
&\Pr\left(\sup_{(t,\boldsymbol{s})\in[0,1]\times\Omega} |\mathbf{B}_1(t,\boldsymbol{s}) - E(\mathbf{B}_1(t,\boldsymbol{s}))| > (2+4C_5)Ta(n,m)\right) \\
\leq & O\left(\{a(n,m)^4 h_{t,0}^3 T\varphi_n^4\}^{-1}\right) + O\left(a(n,m)^{-3}h_{t,0}^{-3}n^{-T^{1/2}/65}\right) \\
&+ O\left(a(n,m)^{-3}h_{t,0}^{-3}n\exp\{-mh_{s,0}^2 C_{\min}C_{\max}(T^{1/2}-1)\}\right) \\
&+ O\left(a(n,m)^{-3}h_{t,0}^{-3}n\exp\left\{-\log(n)\frac{C_1 T^{1/2}}{20C_K}\right\}\right) = o(1).
\end{aligned} \tag{A.15}$$

Note that $E(\mathbf{B}_1(t,\boldsymbol{s})) = 0$. So, by (A.15), we have $\mathbf{B}_1(t,\boldsymbol{s}) = O_p(a(n,m))$, which is uniformly true for all $(t,\boldsymbol{s}) \in [0,1]\times\Omega$. The vector of the remaining elements of $\mathbf{B}(t,\boldsymbol{s})$ can be proved in a similarly way to be of the order $\boldsymbol{H}\mathbf{1}O_p(a(n,m))$, where $\boldsymbol{H} = \mathrm{diag}\{h_{t,0}, h_{s,0}, h_{s,0}\}$ and $\mathbf{1} = (1,1,1)^T$. Thus, we have

$$\mathbf{B}(t,\boldsymbol{s}) = \begin{pmatrix} O_p(a(n,m)) \\ \boldsymbol{H}\mathbf{1}O_p(a(n,m)) \end{pmatrix},$$

which are uniformly true for all $(t,\boldsymbol{s}) \in [0,1]\times\Omega$.

Next, we will study the properties of $\mathbf{A}(t, \boldsymbol{s})$. To this end, let $b(n,m) = h_{t,0}^2 + h_{s,0}^2 + \{\log(n)^2/(nh_{t,0}^2)\}^{1/2}$, $b^*(n,m) = b(n,m)+\{\log(m)/(mh_{s,0}^2)\}^{1/2}$, and $\boldsymbol{\mu}_2(K) = \mathrm{diag}\{\mu_{21}(K), \mu_{22}(K), \mu_{22}(K)\}$, where $\mu_{21}(K) = \int x^2 K_1(x)dx$, $\mu_{22}(K) = \int u_1^2 K_2(d_E(\boldsymbol{u}, \mathbf{0}))d\boldsymbol{u}$ and $\boldsymbol{u} = (u_1, u_2)^T$. Then, it can be shown by similar arguments to those for deriving (A.5)-(A.15) that

$$\mathbf{A}(t, \boldsymbol{s}) = \begin{pmatrix} a + O_P(b^*(n,m)) & \mathbf{1}^T \boldsymbol{H} O_p(b^*(n,m)) \\ \boldsymbol{H}\mathbf{1} O_p(b^*(n,m)) & \boldsymbol{C}(1 + O_p(b^*(n,m))) \end{pmatrix},$$

where $a \in [C_{\min}, C_{\max}]$, and all elements of the 3×3 matrix $\boldsymbol{C}$ are in the same order of the corresponding elements of $\boldsymbol{H}^2 \boldsymbol{\mu}_2(K)$. By combining the above results, we have

$$\Pi_2 = \boldsymbol{e}_1^T \mathbf{A}(t, \boldsymbol{s})^{-1} \mathbf{B}(t, \boldsymbol{s}) = O_p(a(n,m)), \tag{A.16}$$

which is uniformly true for all $(t, \boldsymbol{s}) \in [0,1] \times \Omega$. For Π_3 defined in (A.2), in a similar way that we study the property of $\mathbf{B}(t, \boldsymbol{s})$, it can be checked that

$$\Pi_3 = O_p\left(h_{t,0}^2 + h_{s,0}^2\right), \tag{A.17}$$

which is uniformly true for all $(t, \boldsymbol{s}) \in [0,1] \times \Omega$. By combining the results in (A.1), (A.2), (A.16) and (A.17), the result (12) in Lemma 2 has been proved.

Appendix C: Proof of Lemma 3

First, we derive the convergence property of $\widehat{\sigma}^2(t, \boldsymbol{s})$ in (13). For simplicity, denote $\widetilde{\varepsilon}_{ij} = \widetilde{\varepsilon}(t_i, \boldsymbol{s}_j) = y(t_i, \boldsymbol{s}_{ij}) - \widetilde{\lambda}(t_i, \boldsymbol{s}_{ij})$. From Lemma 2, we know that $|\varepsilon_{ij} - \widetilde{\varepsilon}_{ij}|$ is bounded by a term of the order $O_p(b(n,m))$ uniformly for all i and j, where $b(n,m) = h_{t,0}^2 + h_{s,0}^2 + \{\log(n)^2/(nh_{t,0}^2)\}^{1/2}$. Let $G_0(t, \boldsymbol{s}) = \mathcal{D}(n,m) \sum_{i=1}^n \sum_{j=1}^{m_i} w_1(i,j)$, $G_1(t, \boldsymbol{s}) = \mathcal{D}(n,m) \sum_{i=1}^n \sum_{j=1}^{m_i} w_1(i,j) \varepsilon_{ij}^2$, and $G_2(t, \boldsymbol{s}) = \mathcal{D}(n,m) \sum_{i=1}^n \sum_{j=1}^{m_i} w_1(i,j) \widetilde{\varepsilon}_{ij}^2$, where $\mathcal{D}(n,m)$ is defined in the proof of Lemma 2. Then, it is clear that $\widehat{\sigma}^2(t, \boldsymbol{s}) = G_2(t, \boldsymbol{s})/G_0(t, \boldsymbol{s})$. For the difference between $G_2(t, \boldsymbol{s})$ and $G_1(t, \boldsymbol{s})$, we have

$$\begin{aligned} |G_2(t, \boldsymbol{s}) - G_1(t, \boldsymbol{s})| \leq & 2\mathcal{D}(n,m) \sum_{i=1}^n \sum_{j=1}^{m_i} w_1(i,j) |\varepsilon_{ij}| |\widetilde{\varepsilon}_{ij} - \varepsilon_{ij}| \\ & + \mathcal{D}(n,m) \sum_{i=1}^n \sum_{j=1}^{m_i} w_1(i,j) (\widetilde{\varepsilon}_{ij} - \varepsilon_{ij})^2. \end{aligned}$$

Since $|\varepsilon_{ij} - \widetilde{\varepsilon}_{ij}|$ is uniformly bounded by a term of the order $O_p(b(n,m))$ and it can be easily checked that $\mathcal{D}(n,m) \sum_{i=1}^n \sum_{j=1}^{m_i} w_1(i,j) |\varepsilon_{ij}| / G_0(t, \boldsymbol{s})$ is uniformly bounded by $O_p(1)$ for all $(t, \boldsymbol{s})$, we have

$$\sup_{t \in [0,1], \boldsymbol{s} \in \Omega} |(G_2(t, \boldsymbol{s}) - G_1(t, \boldsymbol{s}))/G_0(t, \boldsymbol{s})| = O_p(b(n,m)). \tag{A.18}$$

Let

$$G_1(t,\boldsymbol{s})/G_0(t,\boldsymbol{s}) = G_3(t,\boldsymbol{s})/G_0(t,\boldsymbol{s}) + G_4(t,\boldsymbol{s})/G_0(t,\boldsymbol{s}) = \Lambda_1 + \Lambda_2, \tag{A.19}$$

where $G_3(t,\boldsymbol{s}) = \mathcal{D}(n,m)\sum_{i=1}^{n}\sum_{j=1}^{m_i} w_1(i,j)\sigma^2(t_i,\boldsymbol{s}_j)$ and $G_4(t,\boldsymbol{s}) = G_1(t,\boldsymbol{s}) - G_3(t,\boldsymbol{s})$. For the random errors $\{\varepsilon(t,\boldsymbol{s})\}$, from the condition that $\Pr(|\varepsilon(t,\boldsymbol{s})| \geq k) \leq C_2 k^{\vartheta}\exp(-C_3 k)$, we have

$$E(|\varepsilon(t,\boldsymbol{s})|^{\delta}) \leq \sum_{k=1}^{\infty} k^{\delta}\Pr(|\varepsilon(t,\boldsymbol{s})| \geq k-1) \leq C_2 \sum_{k=1}^{\infty} k^{\delta+\vartheta}\exp\{-C_3(k-1)\} < \infty,$$

where $\delta > 0$ is any constant. So, we can find a constant $C_{\varepsilon,\delta} > 0$ such that $E(|\varepsilon_{ij}|^{\delta}) \leq C_{\varepsilon,\delta}$, for all i and j, when δ is pre-specified. It follows that $E(|\varepsilon_{ij}|^{10}) \leq C_{\varepsilon,10}$. By some similar arguments to those in the proof of Lemma 2, it can be checked that $|\Lambda_1 - \sigma^2(t,\boldsymbol{s})| = O_p(h_{t,1}^2 + h_{s,1}^2)$, and $\Lambda_2 = O_p(h_{t,1}^2 + h_{s,1}^2 + \{\log(n)^2/(nh_{t,1}^2)\}^{1/2})$, which are uniformly true for all $(t,\boldsymbol{s})$. Therefore, from (A.18) and (A.19), the result in (13) is true.

Next, we derive the property of $\widehat{V}(t,t';\boldsymbol{s},\boldsymbol{s}')$ given in (14) of Lemma 3. To this end, let $\mathcal{D}^*(n,m) = \{n^2h_{t,1}^2 m^2 h_{s,1}^4 f(\boldsymbol{s})f(\boldsymbol{s}')\}^{-1}$,

$$Q_0(t,t';\boldsymbol{s},\boldsymbol{s}') = \mathcal{D}^*(n,m)\sum_{i,j}\sum_{(k,l)\neq(i,j)} w_2(i,j,k,l),$$

$$Q_1(t,t';\boldsymbol{s},\boldsymbol{s}') = \mathcal{D}^*(n,m)\sum_{i,j}\sum_{(k,l)\neq(i,j)} w_2(i,j,k,l)\varepsilon_{ij}\varepsilon_{kl}, \quad \text{and}$$

$$Q_2(t,t';\boldsymbol{s},\boldsymbol{s}') = \mathcal{D}^*(n,m)\sum_{i,j}\sum_{(k,l)\neq(i,j)} w_2(i,j,k,l)\widetilde{\varepsilon}_{ij}\widetilde{\varepsilon}_{kl}.$$

Then it is clear that $\widehat{V}(t,t';\boldsymbol{s},\boldsymbol{s}') = Q_2(t,t';\boldsymbol{s},\boldsymbol{s}')/Q_0(t,t';\boldsymbol{s},\boldsymbol{s}')$. Similar to the arguments for deriving (A.18) above, it can be checked that

$$\left|(Q_2(t,t';\boldsymbol{s},\boldsymbol{s}') - Q_1(t,t';\boldsymbol{s},\boldsymbol{s}'))/Q_0(t,t';\boldsymbol{s},\boldsymbol{s}')\right| = O_p(b(n,m)), \tag{A.20}$$

which is uniformly true for all t, t', $\boldsymbol{s}$ and $\boldsymbol{s}'$. Let

$$\begin{aligned} Q_1(t,t';\boldsymbol{s},\boldsymbol{s}')/Q_0(t,t';\boldsymbol{s},\boldsymbol{s}') &= \{Q_3(t,t';\boldsymbol{s},\boldsymbol{s}') + Q_4(t,t';\boldsymbol{s},\boldsymbol{s}')\}/Q_0(t,t';\boldsymbol{s},\boldsymbol{s}') \\ &= \Lambda_1^* + \Lambda_2^*, \end{aligned} \tag{A.21}$$

where $Q_3(t,t';\boldsymbol{s},\boldsymbol{s}') = \mathcal{D}^*(n,m)\sum_{i,j}\sum_{(k,l)\neq(i,j)} w_2(i,j,k,l)V(t_i,t_k;\boldsymbol{s}_{ij},\boldsymbol{s}_{kl})$, and $Q_4(t,t';\boldsymbol{s},\boldsymbol{s}') = Q_1(t,t';\boldsymbol{s},\boldsymbol{s}') - Q_3(t,t';\boldsymbol{s},\boldsymbol{s}')$. For the first part Λ_1^*, since the covariance function V is twice continuously differentiable, it can be checked that $|\Lambda_1^* - V(t,t';\boldsymbol{s},\boldsymbol{s}')| = O_p(h_{t,1}^2 + h_{s,1}^2)$ uniformly. Next, we will show that $Q_4(t,t';\boldsymbol{s},\boldsymbol{s}') = O_p(\{\log(n)^2/(nh_{t,1}^2)\}^{1/2})$ uniformly. To this end, denote

$$Y_{i,k}^*(\boldsymbol{s},\boldsymbol{s}') := \{m^2 h_{s,1}^4 f(\boldsymbol{s}) f(\boldsymbol{s}')\}^{-1} \sum_{j=1}^{m_i} \sum_{l=1}^{m_k} K_2(d_E(\boldsymbol{s}_{ij},\boldsymbol{s})/h_{s,1}) \times K_2(d_E(\boldsymbol{s}_{kl},\boldsymbol{s})/h_{s,1})(\varepsilon_{ij}\varepsilon_{kl} - V(t_i,t_k;\boldsymbol{s}_{ij},\boldsymbol{s}_{kl}))$$

and

$$Q_5(t,t';\boldsymbol{s},\boldsymbol{s}') := \{nh_{t,1}\}^{-1} \sum_{i=1}^{n} K_1((t_i - t)/h_{t,1}) X_i(t';\boldsymbol{s},\boldsymbol{s}'),$$

where $X_i(t';\boldsymbol{s},\boldsymbol{s}') = \{nh_{t,1}\}^{-1} \sum_{k=1}^{n} K_1((t_k - t')/h_{t,1}) Y_{i,k}^*(\boldsymbol{s},\boldsymbol{s}')$. It can shown that

$$\begin{aligned} \left|Q_5(t,t';\boldsymbol{s},\boldsymbol{s}') - Q_4(t,t';\boldsymbol{s},\boldsymbol{s}')\right| &= O_p\left(\{n^2 h_{t,1}^2 m^2 h_{s,1}^4\}^{-1}\right) \Big| \sum_{i=1}^{n}\sum_{j=1}^{m_i} K_1((t_i - t)/h_{t,1}) \\ &\quad \times K_1((t_i - t')/h_{t,1}) K_2(d_E(\boldsymbol{s}_{ij},\boldsymbol{s})/h_{s,1}) K_2(d_E(\boldsymbol{s}_{ij},\boldsymbol{s}')/h_{s,1})(\varepsilon_{ij}^2 - \sigma^2(t_i,\boldsymbol{s}_{ij}))\Big| \\ &= O_p\left(nh_{t,1} m h_{s,1}^2\right), \end{aligned} \tag{A.22}$$

which is uniformly true for all $t, t', \boldsymbol{s}$, and $\boldsymbol{s}'$. Thus, to prove the result that $Q_4(t,t';\boldsymbol{s},\boldsymbol{s}') = O_p(\{\log(n)^2/(nh_{t,1}^2)\}^{1/2})$, it suffices to show that $\mathcal{O}_5(t,t';\boldsymbol{s},\boldsymbol{s}') = O_p(\{\log(n)^2/(nh_{t,1}^2)\}^{1/2})$. Since $Q_5(t,t';\boldsymbol{s},\boldsymbol{s}')$ is a weighted average of $X_i(t';\boldsymbol{s},\boldsymbol{s}')$, it is enough to show that

$$X_i(t';\boldsymbol{s},\boldsymbol{s}') = O_p(\{\log(n)^2/(nh_{t,1}^2)\}^{1/2})$$

uniformly for all i and t', $\boldsymbol{s}$ and $\boldsymbol{s}'$, which is shown below.

For $X_i(t';\boldsymbol{s},\boldsymbol{s}')$ defined before, we will use some similar arguments to those for deriving (A.5)-(A.15) to find an uniform bound for it. To this end, we first divide the space $[0,1]\times\Omega^2$ into $N^* = O(\{a^*(n,m)h_{t,1}\}^{-5})$ regions $\{R_l^*, l = 1,\dots N^*\}$, where $a^*(n,m) = \{\log(n)^2/(nh_{t,1}^2)\}^{1/2}$. Let $(t_l^*, \boldsymbol{u}_l^*, \boldsymbol{v}_l^*)$ be the centroid of the region R_l^*, for $l = 1,\dots N^*$. Then, by the similar arguments those in (A.5)-(A.8), we can find some constants $C_6 > 0$ and $C_7 > 0$ such that

$$\begin{aligned} &\Pr\left(\sup_{t'\in[0,1]}\sup_{\boldsymbol{s},\boldsymbol{s}'\in\Omega} |X_i(t';\boldsymbol{s},\boldsymbol{s}') - E\{X_i(t';\boldsymbol{s},\boldsymbol{s}')\}| \geq C_6 T a^*(n,m)\right) \\ \leq & C_7 N^* \max_{1\leq l\leq N^*} \Pr\left(|X_i(t_l^*;\boldsymbol{u}_l^*,\boldsymbol{v}_l^*) - E\{X_i(t_l^*;\boldsymbol{u}_l^*,\boldsymbol{v}_l^*)\}| \geq 2Ta^*(n,m)\right). \end{aligned} \tag{A.23}$$

For any t', $\boldsymbol{s}$ and $\boldsymbol{s}'$, because $E(X_i(t';\boldsymbol{s},\boldsymbol{s}')) = 0$, we have

$$\Pr\Big(|X_i(t';\boldsymbol{s},\boldsymbol{s}') - E(X_i(t';\boldsymbol{s},\boldsymbol{s}'))| \geq 2Ta^*(n,m)\Big) = \Pr\Big(|X_i(t';\boldsymbol{s},\boldsymbol{s}')| \geq 2Ta^*(n,m)\Big). \tag{A.24}$$

Let $X_i^*(t'; \boldsymbol{s}, \boldsymbol{s}') = \{1/(nh_{t,1}) \sum_{k=1}^{n} K_1((t_k - t')/h_{t,1}) Y_{i,k}^*(\boldsymbol{s}, \boldsymbol{s}') I(|Y_{i,k}^*(\boldsymbol{s}, \boldsymbol{s}')| \leq \varphi_n)$, where φ_n is defined in the proof of Lemma 2. Since $E(|\varepsilon(t, \boldsymbol{s})|^{14}) \leq C_{\varepsilon,14}$, it can be checked that $E|Y_{i,k}^*(\boldsymbol{s}, \boldsymbol{s}')|^7 \leq C_8$, for some constant $C_8 > 0$. By the Markov's inequality, it follows that

$$\Pr\left(\left|\mathrm{TR}_i^*(t'; \boldsymbol{s}, \boldsymbol{s}') - E(\mathrm{TR}_i^*(t'; \boldsymbol{s}, \boldsymbol{s}'))\right| \leq a^*(n,m)T\right) = O(\{a^*(n,m)T\varphi_n^6\}^{-1}), \tag{A.25}$$

where $\mathrm{TR}_i^*(t'; \boldsymbol{s}, \boldsymbol{s}') = X_i(t'; \boldsymbol{s}, \boldsymbol{s}') - X_i^*(t'; \boldsymbol{s}, \boldsymbol{s}')$. For $X_i^*(t'; \boldsymbol{s}, \boldsymbol{s}')$, similar to the arguments in deriving (A.11)-(A.12), it can be checked that

$$\begin{aligned}
&\Pr\left(|X_i^*(t' ; \boldsymbol{s}, \boldsymbol{s}') - E(X_i^*(t'; \boldsymbol{s}, \boldsymbol{s}'))| \geq a^*(n,m)T\right) = O\left(\exp(-T^{1/2}\log(n)/C_9)\right) \\
&\qquad + O\left(n\exp(-T^{1/2}\log(n)/C_{10})\right) + O\left(n\exp(-mh_{s,1}^2(T^{1/2}-1))\right),
\end{aligned} \tag{A.26}$$

where $C_9, C_{10} > 0$ are some constants. By combining the results in (A.25) and (A.26) and the fact that $E(X_i(t'; \boldsymbol{s}, \boldsymbol{s}')) = 0$, we have

$$\begin{aligned}
&\Pr\left(|X_i(t'; \boldsymbol{s}, \boldsymbol{s}')| \geq 2a^*(n,m)T\right) = O\left(\{a^*(n,m)T\varphi_n^6\}^{-1}\right) + O\left(\exp(-T^{1/2}\log(n)/C_9)\right) \\
&\qquad + O\left(n\exp(-T^{1/2}\log(n)/C_{10})\right) + O\left(n\exp(-mh_{s,1}^2(T^{1/2}-1)\right).
\end{aligned} \tag{A.27}$$

By (A.23), (A.24) and (A.27), we have

$$\begin{aligned}
&\Pr\left(\sup_{t' \in [0,1]} \sup_{\boldsymbol{s}, \boldsymbol{s}' \in \Omega} |X_i(t'; \boldsymbol{s}, \boldsymbol{s}') - E\{X_i(t'; \boldsymbol{s}, \boldsymbol{s}')\}| \geq C_6 T a^*(n,m)\right) \\
\leq& O\left(\{h_{t,1}a^*(n,m)\}^{-5}[\{a^*(n,m)\varphi_n^6 T\}^{-1} + \exp(-T^{1/2}\log(n)/C_9]\right) \\
&+ O\left(\{h_{t,1}a^*(n,m)\}^{-5}\left[n\exp(-T^{1/2}\log(n)/C_{10}) + \exp\{-(T^{1/2}-1)mh_{s,1}^2\}\right]\right),
\end{aligned} \tag{A.28}$$

which is uniformly true for $1 \leq i \leq n$. It follows that

$$\begin{aligned}
&\Pr\left(\max_{1 \leq i \leq n} \sup_{t \in [0,1]} \sup_{\boldsymbol{s}, \boldsymbol{s}' \in \Omega} |X_i(t'; \boldsymbol{s}, \boldsymbol{s}') - E\{X_i(t'; \boldsymbol{s}, \boldsymbol{s}')\}| \geq C_6 T a^*(n,m)\right) \\
\leq& n \max_{1 \leq i \leq n} \Pr\left(\sup_{t \in [0,1]} \sup_{\boldsymbol{s}, \boldsymbol{s}' \in \Omega} |X_i(t'; \boldsymbol{s}, \boldsymbol{s}') - E\{X_i(t'; \boldsymbol{s}, \boldsymbol{s}')\}| \geq C_6 T a^*(n,m)\right) = o(1).
\end{aligned} \tag{A.29}$$

Since $E\{X_i(t'; \boldsymbol{s}, \boldsymbol{s}')\} = 0$, from (A.29), we have $X_i(t'; \boldsymbol{s}, \boldsymbol{s}') = O_p(a^*(n,m))$ uniformly for all $i, t, \boldsymbol{s}$ and $\boldsymbol{s}'$. So, $Q_5(t, t'; \boldsymbol{s}, \boldsymbol{s}') = O_p(\{\log(n)^2/(nh_{t,1}^2)\}^{1/2})$ uniformly for all $t, t', \boldsymbol{s}$ and $\boldsymbol{s}'$. Similarly, it can be checked that $|Q_4(t, t'; \boldsymbol{s}, \boldsymbol{s}') - Q_5(t, t'; \boldsymbol{s}, \boldsymbol{s}')| = O_p(a^*(n,m))$ uniformly. Thus, we have $Q_4(t, t'; \boldsymbol{s}, \boldsymbol{s}') = O_p(a^*(n,m))$ uniformly. In addition, it can be shown that $C_{\min} + O_p(a^*(n,m) + h_{t,1}^2 + h_{s,1}^2) \leq Q_0(t, t'; \boldsymbol{s}, \boldsymbol{s}') \leq C_{\max} + O_p(a^*(n,m) + h_{t,1}^2 + h_{s,1}^2)$ uniformly. Therefore, we have $\Lambda_2^* = Q_4(t, t'; \boldsymbol{s}, \boldsymbol{s}')/Q_0(t, t'; \boldsymbol{s}, \boldsymbol{s}') = O_p(a^*(n,m))$, where $C_{\min} = \min_{1 \leq i \leq n} m_i/m$ and $C_{\max} = \max_{1 \leq i \leq n} m_i/m$. By combining this result with (A.20)-(A.21) and the result that $\Lambda_1^* = V(t, t'; \boldsymbol{s}, \boldsymbol{s}') + O_p(h_{t,1}^2 + h_{s,1}^2)$, the result in (14) can then be proved.

Appendix D: Proof of Theorem 1

For any $(t, \boldsymbol{s}) \in [0,1] \times \Omega$, let Υ be the number of elements in the set $\{i, |t_i - t| \leq h_{t,2}\}$. We can change the order of $\{i, i = 1, \ldots, n\}$ to obtain a new sequence $\{l_i, i = 1, \ldots, n\}$ such that the first Υ elements in $\{l_i\}$ are $\{i, |t_i - t| \leq h_{t,2}\}$ and $l_1 < \cdots < l_\Upsilon$. Let $\mathbf{X}_\nu = (\mathbf{X}_{l_1 1}, \ldots, \mathbf{X}_{l_\Upsilon, m_{l_\Upsilon}})^T$ and $\mathbf{Y}_\nu = (y(t_{l_1}, \boldsymbol{s}_{l_1,1}), \ldots, y(t_{l_\Upsilon}, \boldsymbol{s}_{l_\Upsilon, m_{l_\Upsilon}}))^T$. Then, it can be checked that

$$\widehat{\lambda}(t, \boldsymbol{s}) = \boldsymbol{e}_1^T \left(\mathbf{X}_\nu^T \widetilde{\boldsymbol{\Sigma}}_{K,\nu}^{-1} \mathbf{X}_\nu \right)^{-1} \mathbf{X}_\nu^T \widetilde{\boldsymbol{\Sigma}}_{K,\nu}^{-1} \mathbf{Y}_\nu,$$

where $\boldsymbol{e}_1 = (1,0,0,0)^T$, $\widetilde{\boldsymbol{\Sigma}}_{K,\nu}^{-1} = \boldsymbol{D}_{K,\nu}^{1/2} (\boldsymbol{I}_\nu \widetilde{\boldsymbol{\Sigma}}_{\mathbf{Y},\nu} \boldsymbol{I}_\nu)^{-1} \boldsymbol{D}_{K,\nu}^{1/2}$, $\boldsymbol{I}_\nu = \mathrm{diag}\big\{ I\big(|t_{l_1} - t| \leq h_{t,2}\big) I\big(d_E(\boldsymbol{s}_{l_1,1}, \boldsymbol{s}) \leq h_{s,2}\big), \ldots, I\big(|t_{l_\Upsilon} - t| \leq h_{t,2}\big) I\big(d_E(\boldsymbol{s}_{l_\Upsilon, m_{l_\Upsilon}}, \boldsymbol{s}) \leq h_{s,2}\big) \big\}$, $\boldsymbol{D}_{K,\nu} = \mathrm{diag}\{w_0(l_1, 1), \ldots, w_0(l_\Upsilon, m_{l_\Upsilon})\}$, $w_0(i,j)$ is defined in (3), and $\widetilde{\boldsymbol{\Sigma}}_{\mathbf{Y},\nu}$ is the estimated covariance matrix of $\mathbf{Y}_\nu$. Let $\boldsymbol{\lambda}_\nu = E(\mathbf{Y}_\nu)$ and $\boldsymbol{\varepsilon}_\nu = \mathbf{Y}_\nu - \boldsymbol{\lambda}_\nu$. By the Taylor's expansion, it can be shown that

$$\begin{aligned}
\widehat{\lambda}(t, \boldsymbol{s}) =& \boldsymbol{e}_1^T (\mathbf{X}_\nu^T \widetilde{\boldsymbol{\Sigma}}_{K,\nu}^{-1} \mathbf{X}_\nu)^{-1} \mathbf{X}_\nu^T \widetilde{\boldsymbol{\Sigma}}_{K,\nu}^{-1} \mathbf{Y}_\nu \\
=& \boldsymbol{e}_1^T (\mathbf{X}_\nu^T \widetilde{\boldsymbol{\Sigma}}_{K,\nu}^{-1} \mathbf{X}_\nu)^{-1} \mathbf{X}_\nu^T \widetilde{\boldsymbol{\Sigma}}_{K,\nu}^{-1} \boldsymbol{\lambda}_\nu + \boldsymbol{e}_1^T (\mathbf{X}_\nu^T \widetilde{\boldsymbol{\Sigma}}_{K,\nu}^{-1} \mathbf{X}_\nu)^{-1} \mathbf{X}_\nu^T \widetilde{\boldsymbol{\Sigma}}_{K,\nu}^{-1} \boldsymbol{\varepsilon}_\nu \\
=& \boldsymbol{e}_1^T (\mathbf{X}_\nu^T \widetilde{\boldsymbol{\Sigma}}_{K,\nu}^{-1} \mathbf{X}_\nu)^{-1} \mathbf{X}_\nu^T \widetilde{\boldsymbol{\Sigma}}_{K,\nu}^{-1} \mathbf{X}_\nu \boldsymbol{\beta}_\nu + \boldsymbol{e}_1^T (\mathbf{X}_\nu^T \widetilde{\boldsymbol{\Sigma}}_{K,\nu}^{-1} \mathbf{X}_\nu)^{-1} \mathbf{X}_\nu^T \widetilde{\boldsymbol{\Sigma}}_{K,\nu}^{-1} \widetilde{\mathcal{R}} \\
&+ \boldsymbol{e}_1^T (\mathbf{X}_\nu^T \widetilde{\boldsymbol{\Sigma}}_{K,\nu}^{-1} \mathbf{X}_\nu)^{-1} \mathbf{X}_\nu^T \widetilde{\boldsymbol{\Sigma}}_{K,\nu}^{-1} \boldsymbol{\varepsilon}_\nu = \Lambda_1^{**} + \Lambda_2^{**} + \Lambda_3^{**},
\end{aligned} \tag{A.30}$$

where $\widetilde{\boldsymbol{\mathcal{R}}} = (\widetilde{r}_{11}, \ldots, \widetilde{r}_{\Upsilon, m_{l_\Upsilon}})$, $\widetilde{r}_{ij} = ((t_{l_i} - t), (\boldsymbol{s}_{l_i,j} - \boldsymbol{s})^T) \boldsymbol{\mathcal{H}}(t''_{ij}, \boldsymbol{s}''_{ij}) ((t_{l_i} - t), (\boldsymbol{s}_{l_i,j} - \boldsymbol{s})^T)^T$, $\boldsymbol{\beta}_\nu = (\lambda(t, \boldsymbol{s}), \partial\lambda(t, \boldsymbol{s})/\partial t, \partial\lambda(t, \boldsymbol{s})/\partial \boldsymbol{s})^T$, $\boldsymbol{\mathcal{H}}$ is the Hessian matrix of $\lambda(t, \boldsymbol{s})$, $t''_{ij} \in [0,1]$ and $\boldsymbol{s}''_{ij} \in \Omega$, for $i = 1, \ldots, \Upsilon$ and $j = 1, \ldots, m_{l_i}$. For Λ_1^{**}, it is clear that $\Lambda_1^{**} = \lambda(t, \boldsymbol{s})$. Next, we will show that $\Lambda_3^{**} = O_p(\{1/(nh_{t,2})\}^{1/2})$, and it can be shown similarly that $\Lambda_2^{**} = O_p(h_{t,2}^2 + h_{s,2}^2)$. From Lemma 3, we have

$$\begin{aligned}
|\widehat{V}(t, t'; \boldsymbol{s}, \boldsymbol{s}') - V(t, t'; \boldsymbol{s}, \boldsymbol{s}')| = O_p\big(& h_{t,0}^2 + h_{s,0}^2 + \{\log(n)^2/(nh_{t,0}^2)\}^{1/2} \\
&+ h_{t,1}^2 + h_{s,1}^2 + \{\log(n)^2/(nh_{t,1}^2)\}^{1/2}\big)
\end{aligned}$$

uniformly for all $(t, \boldsymbol{s})$ and $(t', \boldsymbol{s}')$. Define $||\mathbf{A}||_{\max} = \max_{1 \leq i,j \leq N} |a_{ij}|$, where $\mathbf{A} = (a_{ij})$ is a $N \times N$ matrix. Then it can be shown that

$$\begin{aligned}
||\widehat{\boldsymbol{\Sigma}}_{\mathbf{Y}} - \boldsymbol{\Sigma}_{\mathbf{Y}}||_{\max} = O_p\big((& h_{t,0}^2 + h_{s,0}^2 + \{\log(n)^2/(nh_{t,0}^2)\}^{1/2} \\
&+ h_{t,1}^2 + h_{s,1}^2 + \{\log(n)^2/(nh_{t,1}^2)\}^{1/2})\big),
\end{aligned}$$

where $\boldsymbol{\Sigma}_{\mathbf{Y}}$ is the covariance matrix of $\mathbf{Y}$, $\mathbf{Y} = (y(t_1, \boldsymbol{s}_{11}), \ldots, y(t_n, \boldsymbol{s}_{nm_n}))^T$, and $\widehat{\boldsymbol{\Sigma}}_{\mathbf{Y}}$ is the estimated covariance matrix computed from $\widehat{V}(t, t'; \boldsymbol{s}, \boldsymbol{s}')$. In the paragraph immediately before Expression

(7) in Section 2.2, we defined the projection of $\widehat{\mathbf{\Sigma}}_{\mathbf{Y}}$ to the set of all symmetric positive definite matrices to be $\widetilde{\mathbf{\Sigma}}_{\mathbf{Y}}$. For this matrix, because $\mathbf{\Sigma}_{\mathbf{Y}}$ is positive definite, we have

$$\begin{aligned}||\widetilde{\mathbf{\Sigma}}_{\mathbf{Y}} - \mathbf{\Sigma}_{\mathbf{Y}}||_{\max} =& O_p((h_{t,0}^2 + h_{s,0}^2 + \{\log(n)^2/(nh_{t,0}^2)\}^{1/2} \\ &+ h_{t,1}^2 + h_{s,1}^2 + \{\log(n)^2/(nh_{t,1}^2)\}^{1/2})).\end{aligned}$$

Note that $\widetilde{\mathbf{\Sigma}}_{\mathbf{Y},\nu}$ is a submatrix of $\widetilde{\mathbf{\Sigma}}_{\mathbf{Y}}$, we have $\widetilde{\mathbf{\Sigma}}_{\mathbf{Y},\nu} = \mathbf{\Sigma}_{\mathbf{Y},\nu}(1+o_p(1))$, where $\mathbf{\Sigma}_{\mathbf{Y},\nu}$ is the covariance matrix of $\mathbf{Y}_\nu$. Then it can be checked that

$$\Lambda_3^{**} = (1 + o_p(1))\Pi_3^*, \tag{A.31}$$

where $\Pi_3^* = \boldsymbol{e}_1^T(\mathbf{X}_\nu^T\mathbf{\Sigma}_{K,\nu}^{-1}\mathbf{X}_\nu)^{-1}\mathbf{X}_\nu^T\mathbf{\Sigma}_{K,\nu}^{-1}\boldsymbol{\varepsilon}_\nu$, $\mathbf{\Sigma}_{K,\nu}^{-1} = \boldsymbol{D}_{K,\nu}^{1/2}(\boldsymbol{I}_\nu\mathbf{\Sigma}_{\mathbf{Y},\nu}\boldsymbol{I}_\nu)^{-1}\boldsymbol{D}_{K,\nu}^{1/2}$. For the matrix $\mathbf{\Sigma}_{\mathbf{Y},\nu}$, the $(\sum_{i'=1}^{i-1} m_{l_{i'}} + j, \sum_{k'=1}^{k-1} m_{l_{k'}} + l)$-th element is $V(t_{l_i}, \boldsymbol{s}_{l_i,j}; t_{l_k}, \boldsymbol{s}_{l_k,l})$, for $1 \le i,k \le \Upsilon$, $1 \le j \le m_{l_i}$ and $1 \le l \le m_{l_k}$. Thus, we have $||\boldsymbol{I}_\nu(\mathbf{\Sigma}_{\mathbf{Y},\nu} - \mathbf{\Sigma}_{\mathbf{Y},\nu}^*)\boldsymbol{I}_\nu||_{\max} = O((h_{t,2} + h_{s,2}))$, where the $(\sum_{i'=1}^{i-1} m_{l_{i'}} + j, \sum_{k'=1}^{k-1} m_{l_{k'}} + l)$th element of $\mathbf{\Sigma}_{\mathbf{Y},\nu}^*$ is $V(t, t+|i-k|/n; \boldsymbol{s}, \boldsymbol{s})$ if $(i,j) \ne (k,l)$, and the $(\sum_{i'=1}^{i-1} m_{l_{i'}} + j, \sum_{i'=1}^{i-1} m_{l_{i'}} + j)$-th element is $\sigma^2(t,\boldsymbol{s})$. Hence, we have

$$\Pi_3^* = (1 + o_p(1))\Pi_3^{**}, \tag{A.32}$$

where $\Pi_3^{**} = \boldsymbol{e}_1^T(\mathbf{X}_\nu^T\mathbf{W}_\nu\mathbf{X}_\nu)^{-1}\mathbf{X}_\nu^T\mathbf{W}_\nu\boldsymbol{\varepsilon}_\nu$, $\mathbf{W}_\nu = \boldsymbol{D}_{K,\nu}^{1/2}(\boldsymbol{I}_\nu\mathbf{\Sigma}_{\mathbf{Y},\nu}^*\boldsymbol{I}_\nu)^{-1}\boldsymbol{D}_{K,\nu}^{1/2}$.

Let $m_0 = C_{\max}m$, denote $\mathbf{\Sigma}_{\mathbf{Y},0}^* = \mathbf{\Sigma}_t \otimes \mathbf{\Sigma}_s$, where $C_{\max} = \max_{1\le i\le n} m_i/m$, $\mathbf{\Sigma}_t = (\sigma_t(i_1,i_2))$ is a $\Upsilon \times \Upsilon$ matrix with $\sigma_t(i_1,i_2) = V(t, t+|i_1-i_2|/n; \boldsymbol{s}, \boldsymbol{s})$, $\mathbf{\Sigma}_s = (\sigma_s(j_1,j_2))$ is a $m_0 \times m_0$ matrix with $\sigma_s(j_1,j_1) = \sigma^2(t,\boldsymbol{s})/V(t,t;\boldsymbol{s},\boldsymbol{s})$ and $\sigma_s(j_1,j_2) = 1$ when $j_1 \ne j_2$, and $\otimes$ is the Kronecker product. It is clear that $\mathbf{\Sigma}_{\mathbf{Y},\nu}^*$ is a principal submatrix of $\mathbf{\Sigma}_{\mathbf{Y},0}^*$. By the assumption that $g(\theta;t,\boldsymbol{s}) > 0$, for all $\theta \in [-0.5, 0.5]$, we can find two constants $\omega_{t,0} > 0$ and $\omega_{t,1} > 0$ such that $\omega_{t,0} \le g(\theta;t,\boldsymbol{s}) \le \omega_{t,1}$, for $\theta \in [-0.5, 0.5]$. By Lemma 1 in Xiao and Wu (2012), we have $\omega_{t,0} \le \lambda_{\min}(\mathbf{\Sigma}_t) \le \lambda_{\max}(\mathbf{\Sigma}_t) \le \omega_{t,1}$, where $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ denote the smallest and largest eigenvalues of a matrix. On the other hand, it is clear that $\mathbf{\Sigma}_s = \mathbf{1}\mathbf{1}^T + (\sigma^2(t,\boldsymbol{s})/V(t,t;\boldsymbol{s},\boldsymbol{s}) - 1)\mathbf{I}$, where $\mathbf{I}$ is a $m_0 \times m_0$ identity matrix and $\mathbf{1} = (1,\ldots,1)^T$. Here, we focus on the cases when $\mathrm{Var}(\varepsilon_1(t,\boldsymbol{s})) > 0$. If $\mathrm{Var}(\varepsilon_1(t,\boldsymbol{s})) = 0$, we can consider the generalized inverse matrix and prove the theorem similarly. When $\mathrm{Var}(\varepsilon_1(t,\boldsymbol{s})) > 0$, we have $\lambda_{\min}(\mathbf{\Sigma}_s) > \omega_{s,0}$, for some constant $\omega_{s,0} > 0$. Since $\mathbf{\Sigma}_{\mathbf{Y},0}^* = \mathbf{\Sigma}_t \otimes \mathbf{\Sigma}_s$, it is clear that $\lambda_{\min}(\mathbf{\Sigma}_{\mathbf{Y},0}^*) \ge \omega_{t,0} \times \omega_{s,0} > 0$. Notice that $\mathbf{\Sigma}_{\mathbf{Y},\nu}^*$ is a principal submatrix of the positive definite matrix $\mathbf{\Sigma}_{\mathbf{Y},0}^*$, it can be checked that $\lambda_{\min}(\mathbf{\Sigma}_{\mathbf{Y},\nu}^*) \ge \lambda_{\min}(\mathbf{\Sigma}_{\mathbf{Y},0}^*) \ge \omega_{t,0} \times \omega_{s,0}$.

Let $\widetilde{\mathbf{A}}(t,\boldsymbol{s}) = \mathbf{X}_\nu^T\mathbf{W}_\nu\mathbf{X}_\nu$ and $\widetilde{\mathbf{B}}(t,\boldsymbol{s}) = \mathbf{X}_\nu^T\mathbf{W}_\nu\boldsymbol{\varepsilon}_\nu$. Then, we have

$$\Pi_3^{**} = \boldsymbol{e}_1^T\widetilde{\mathbf{A}}(t,\boldsymbol{s})^{-1}\widetilde{\mathbf{B}}(t,\boldsymbol{s}). \tag{A.33}$$

For $\widetilde{\mathbf{B}}(t,\boldsymbol{s})$, we first consider its first element $\widetilde{\mathbf{B}}_1(t,\boldsymbol{s})$. Obviously, we have $E(\widetilde{\mathbf{B}}_1(t,\boldsymbol{s}))=0$. For the variance of $\widetilde{\mathbf{B}}_1(t,\boldsymbol{s})$, note that $E(\widetilde{\mathbf{B}}_1(t,\boldsymbol{s})|\mathcal{S}_\sigma)=0$, where $\mathcal{S}_\sigma$ is the σ-field generated by $\mathcal{S}=\{\boldsymbol{s}_{11},\ldots,\boldsymbol{s}_{nm_n}\}$, and $\mathrm{Var}(\widetilde{\mathbf{B}}_1(t,\boldsymbol{s}))=\mathrm{Var}(E(\widetilde{\mathbf{B}}_1(t,\boldsymbol{s})|\mathcal{S}_\sigma))+E(\mathrm{Var}(\widetilde{\mathbf{B}}_1(t,\boldsymbol{s})|\mathcal{S}_\sigma))$. So, to find $\mathrm{Var}(\widetilde{\mathbf{B}}_1(t,\boldsymbol{s}))$, we only need to find $\mathrm{Var}(\widetilde{\mathbf{B}}_1(t,\boldsymbol{s})|\mathcal{S}_\sigma)$. Let $\mathbf{1}_\nu=(1,\ldots,1)^T$ and $m(\mathcal{S})=\max_{1\le i\le n} m_i(\mathcal{S})$, where $m_i(\mathcal{S})$ is the number of elements in $\{\boldsymbol{s}_{ij}, d_E(\boldsymbol{s}_{ij},\boldsymbol{s})\le h_{s,2}, j=1,\ldots,m_i\}$. Then, we have

$$\begin{aligned}\mathrm{Var}(\widetilde{\mathbf{B}}_1(t,\boldsymbol{s})|\mathcal{S}_\sigma)&=\mathbf{1}_\nu^T\mathbf{W}_\nu\mathbf{\Sigma}_{\mathbf{Y},\nu}\mathbf{W}_\nu\mathbf{1}_\nu\le C_{12}m(\mathcal{S})\mathbf{1}_\nu^T\mathbf{W}_\nu^2\mathbf{1}_\nu\\&\le C_{12}C_K^2m(\mathcal{S})\mathbf{1}_\nu^T\boldsymbol{D}_K^{1/2}(\boldsymbol{I}_\nu\mathbf{\Sigma}^*_{\mathbf{Y},\nu}\boldsymbol{I}_\nu)^{-2}\boldsymbol{D}_K^{1/2}\mathbf{1}_\nu\\&\le\{\omega_{t,0}\omega_{s,0}\}^{-2}C_{12}C_K^2m(\mathcal{S})\mathbf{1}_\nu^T\boldsymbol{D}_K\mathbf{1}_\nu=m(\mathcal{S})^2O(nh_{t,2}),\end{aligned}\tag{A.34}$$

for some constant $C_{12}>0$. For $E(m(\mathcal{S})^2)$, note that $n\exp(-mh_{s,2}^2)=O(1)$. So, by the Bernstein's inequality, we have

$$\begin{aligned}&E(m(\mathcal{S})^2/(mh_{s,2}^2)^2)=E[(\max_{1\le i\le n}m_i(\mathcal{S}))^2/(mh_{s,2}^2)^2]\\=&E[\max_{1\le i\le n}m_i(\mathcal{S})^2/(mh_{s,2}^2)^2]\le\sum_{k=0}^{\infty}(k+1)\Pr(\max_{1\le i\le n}m_i(\mathcal{S})^2/(mh_{s,2}^2)^2\ge k)\\\le&O(1)+n\sum_{k=1}^{\infty}(k+1)\max_{1\le i\le n}\Pr(m_i(\mathcal{S})/(mh_{s,2}^2)\ge\sqrt{k})\\\le&O(1)+2n\sum_{k=1}^{\infty}(k+1)\exp(-mh_{s,2}^2C_{\min}(\sqrt{k}-C_{\max}))\\\le&O(1)+O(1)\times\sum_{k=4}^{\infty}(k+1)\exp\left(-mh_{s,2}^2\left(C_{\min}(\sqrt{k}-C_{\max})-1\right)\right)<\infty.\end{aligned}\tag{A.35}$$

From (A.34) and (A.35), we have $\mathrm{Var}(\widetilde{\mathbf{B}}_1(t,\boldsymbol{s}))=O(nh_{t,2}m^2h_{s,2}^4)$. Since $E(\widetilde{\mathbf{B}}_1(t,\boldsymbol{s}))=0$, we have $\widetilde{\mathbf{B}}_1(t,\boldsymbol{s})=O_p(\{nh_{t,2}\}^{1/2}mh_{s,2}^2)$. Similarly, it can be shown that

$$\widetilde{\mathbf{B}}(t,\boldsymbol{s})/(nh_{t,2}mh_{s,2}^2)=\begin{pmatrix}\{nh_{t,2}\}^{-1/2}O_p(1)\\\widetilde{\boldsymbol{H}}\mathbf{1}\{nh_{t,2}\}^{-1/2}O_p(1)\end{pmatrix},\tag{A.36}$$

and

$$\widetilde{\mathbf{A}}(t,\boldsymbol{s})/(nh_{t,2}mh_{s,2}^2)=\begin{pmatrix}\mathcal{C}_1+o_p(1)&\mathbf{1}^T\widetilde{\boldsymbol{H}}o_p(1)\\\widetilde{\boldsymbol{H}}\mathbf{1}o_p(1)&\widetilde{\boldsymbol{H}}^2\boldsymbol{\mathcal{C}}_2(1+o_p(1))\end{pmatrix},\tag{A.37}$$

where $\widetilde{\boldsymbol{H}}=\mathrm{diag}\{h_{t,2},h_{s,2},h_{s,2}\}$, $\mathbf{1}=(1,1,1)^T$, $0<\mathcal{C}_1<\infty$ is a constant, and $\boldsymbol{\mathcal{C}}_2$ is a matrix with every element being a positive number. From (A.31)-(A.33), (A.36) and (A.37), we have $\Lambda_3^{**}=\Pi_3^{**}=O_p(\{nh_{t,2}\}^{-1/2})$. By some similar arguments, we have $\Lambda_2^{**}=O_p(h_{t,2}^2+h_{s,2}^2)$. By combining these results with (A.30) and the fact that $\Lambda_1^{**}=\lambda(t,\boldsymbol{s})$, the result in (15) of Theorem 1 has been proved.

Appendix D: Additional simulation results about computation time

In Figure S.1, we investigate the effect of (m, n) on the proposed method Step2 and the setup is described in Section 4 of the main paper.

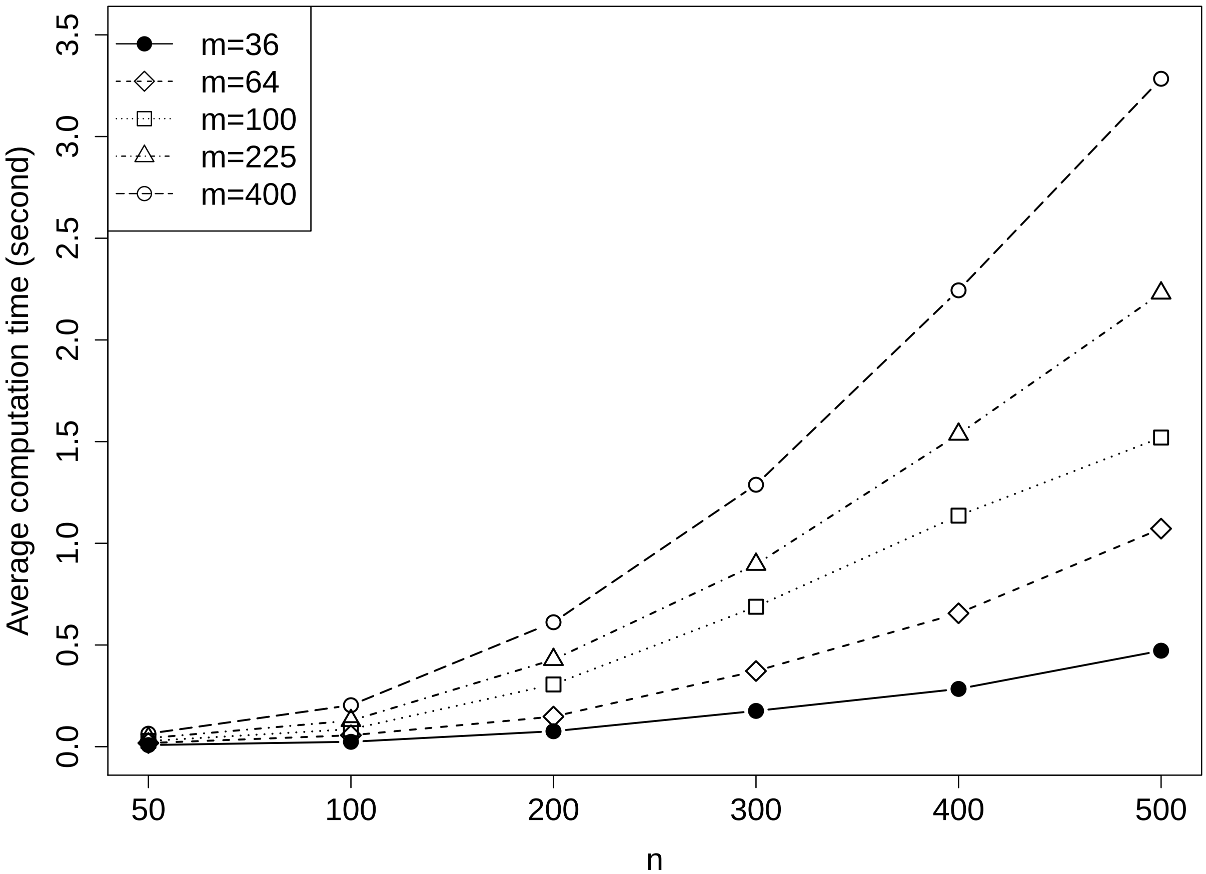

Figure S.1: Average computation times (in seconds) of the proposed method Step2 in cases when $(\phi_t, \phi_s) = (0.6, 3)$.

References

Bradley, R.C. (2005), "Basic properties of strong mixing conditions. A survey and some open questions," *Probability Surveys*, **2**, 107–144.

Grenander, U., and Szegö, G. (1958), *Toeplitz forms and their applications*, Oakland: University of California Press.

Liebscher, E. (1996), "Strong convergence of sums of α-mixing random variables with applications to density estimation," *Stochastic Processes and Their Applications*, **65**, 69–80.

Xiao, H., and Wu, W.B. (2012), "Covariance matrix estimation for stationary time series," *The Annals of Statistics*, **40**, 466–493.