SUPPLEMENTARY MATERIALS

Supplementary Materials for Nonparametric Estimation of the Spatio-Temporal Covariance Structure

Kai Yang  | Peihua Qiu*

Department of Biostatistics, University of Florida, Gainesville, Florida, USA

Correspondence
Peihua Qiu, Department of Biostatistics, University of Florida, Gainesville, Florida, USA.
Email: pqiu@ufl.edu

Abstract
To save some space in the paper with the above title, the proofs of the theorems are presented in this supplementary file.

PROOF OF THEOREM 1

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

$$
\hat{\lambda}(t, s) = e_i^T(X^TW_0X)^{-1}X^TW_0Y
= e_i^T(X^TW_0X)^{-1}X^TW_0\lambda + e_i^T(X^TW_0X)^{-1}X^TW_0\varepsilon
$$

(A.1)

where $\lambda = E(Y)$ and $\varepsilon = (\varepsilon_{11}, \ldots, \varepsilon_{1m_1}, \ldots, \varepsilon_{nm_n})^T$. For $\Pi_1$, by the Taylor’s expansion, it can be shown that

$$
\Pi_1 = e_i^T(X^TW_0X)^{-1}X^TW_0(X\beta + R) = \lambda(t, s) + \Pi_3,
$$

(A.2)

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

$$
\Pi_2 = e_i^T(X^TW_0X)^{-1}X^TW_0\varepsilon
= e_i^T(D(m, n)X^TW_0X)^{-1}D(m, n)X^TW_0\varepsilon
= e_i^TA_{-1}(t, s)B(t, s),
$$

(A.3)

where $D(m, n) = (nh_1m_1^2f(s))^{-1}$, $A(t, s) = D(n, m)X^TW_0X$ is a $4 \times 4$ matrix and $B(t, s) = D(n, m)X^TW_0\varepsilon$ is a vector of length 4. Next, we consider the first element of the vector $B(t, s)$, i.e.,

$$
B_1(t, s) = (nh_1)^{-1}\sum_{i=1}^{n} K_1((t_i - t)/h_1) \varepsilon_i(s),
$$

(A.4)

where $\varepsilon_i(s) = \{mh_2^2f(s)\}^{-1}\sum_{j=1}^{m_i} K_2(d_E(s_{ij}, s)/h_2) \varepsilon_{ij}$. We will show below that $B_1(t, s) = O_p(h_1^2 + h_2^2 + (1/(nh_1))^{1/2})$.

For the variance of $B_1(t, s)$ defined in (A.4), it is clear that

$$
\text{Var}(B_1(t, s)) = \{nh_1\}^{-2}\sum_{i=1}^{n} \sum_{k=1}^{n} K_1((t_i - t)/h_1) K_1((t_k - t)/h_1) \text{Cov}(\varepsilon_i(s), \varepsilon_k(s)).
$$

(A.5)
To calculate the covariance $\operatorname{Cov}(\varepsilon(s), \varepsilon_k(s))$, let $S_n$ be the $\sigma$-algebra generated by $\{s_{ij}, j = 1, \ldots, m_i, i = 1, \ldots, n\}$. Given any possible values of $S = \{s_{ij}, j = 1, \ldots, m_i, i = 1, \ldots, n\}$, by the Davydov’s inequality, it can be checked that

$$E(\varepsilon_i(s)\varepsilon_i(k)|S) = (mh^2_{ij}f(s))^{-2} \sum_{j=1}^{m_i} \sum_{k=1}^{m_j} K_2 \left( \frac{d_E(s_{ij}, s)}{h^2_i} \right) K_2 \left( \frac{d_E(s_{kj}, s)}{h^2_k} \right) \operatorname{Cov}(\varepsilon_i(s), \varepsilon_i(k))$$

$$\leq 12 (mh^2_{ij}f(s))^{-2} C_2^{ij} C_0^{(\delta-2)/\delta} \sum_{j=1}^{m_i} \sum_{k=1}^{m_j} K_2 \left( \frac{d_E(s_{ij}, s)}{h^2_i} \right) K_2 \left( \frac{d_E(s_{kj}, s)}{h^2_k} \right)$$

\[(A.6)\]

$$= D^2(s) \exp \left\{ -\frac{C_1(\delta-2)}{\delta} |k-i| \right\} \times O(1),$$

where $D(s) = \max_{1 \leq i \leq n} D_i(s)$ and $D_i(s) = (mh^2_{ij}f(s))^{-1} \sum_{j=1}^{m_i} K_2 \left( \frac{d_E(s_{ij}, s)}{h^2} \right)$. Denote $C_{\min} = \min_{1 \leq i \leq n} m_i/m$ and $C_{\max} = \max_{1 \leq i \leq n} m_i/m$. Note that the random variable $D_i(s)$ is nonnegative, by the Bernstein’s inequality, it can be shown that

$$E(D(s)^2) = E\left[ \left( \max_{1 \leq i \leq n} D_i(s) \right)^2 \right]$$

$$= E\left[ \max_{1 \leq i \leq n} D_i^2(s) \right] \leq \sum_{k=0}^{\infty} (k+1) \Pr\left( \max_{1 \leq i \leq n} D_i^2(s) \geq k \right)$$

$$\leq 1 + n \sum_{k=1}^{\infty} (k+1) \max_{1 \leq i \leq n} \Pr\left( D_i(s) \geq \sqrt{k} \right)$$

\[(A.7)\]

$$\leq 1 + n \sum_{k=1}^{\infty} (k+1) \exp \left\{ -mh^2_2 \left( C_{\min} \left( \sqrt{k} - C_{\max} \right) \right) \right\}$$

$$= O(1) + O(1) \times \sum_{k=1}^{\infty} (k+1) \exp \left\{ -mh^2_2 \left( C_{\min} \left( \sqrt{k} - C_{\max} \right) - 1 \right) \right\} = O(1).$$

It follows from (A.6) and (A.7) that

$$\operatorname{Cov}(\varepsilon(s), \varepsilon_k(s)) = E \left[ E(\varepsilon(s)\varepsilon_k(s)|S) \right]$$

$$= E(D^2(s)) \exp \left\{ -\frac{C_1(\delta-2)}{\delta} |k-i| \right\} \times O(1)$$

\[(A.8)\]

Then by (A.5), we have

$$\operatorname{Var}(B_1(t,s)) = O(1) \times [nh_1]^{-2} \sum_{i=1}^{n} \sum_{k=1}^{n} K_1 \left( (t_i-t)/h_1 \right) K_1 \left( (t_k-t)/h_1 \right) \exp \left\{ -\frac{C_1(\delta-2)}{\delta} |k-i| \right\}$$

$$= O(1) \times [nh_1]^{-2} \sum_{i=1}^{n} \sum_{k=1}^{n} K_1 \left( (t_i-t)/h_1 \right) \exp \left\{ -\frac{C_1(\delta-2)}{\delta} \tau \right\}$$

\[(A.9)\]

$$= O(1) \times [nh_1]^{-1} \sum_{\tau=0}^{\infty} \exp \left\{ -\frac{C_1(\delta-2)}{\delta} \tau \right\} = O(1/(nh_1)).$$

Note that $E(B_1(t,s)) = 0$, we have

$$B_1(t,s) = O_p\left( \left\{ 1/(nh_1) \right\}^{1/2} \right) = O_p(\nu(n,m)),$$

\[(A.10)\]

where $\nu(n,m) = h^2_1 + h^2_2 + \{1/(nh_1)\}^{1/2}$. After defining $\nu^2(n,m) = \nu(n,m) + \{1/(mh^2_2)\}^{1/2}$ and $\mu_2(K) = \text{diag}\{\mu_{21}(K), \mu_{22}(K), \mu_{22}(K)\}$, where $\mu_{21}(K) = \int x^2 K_1(x) dx$, $\mu_{22}(K) = \int u^2 K_2(d_E(u,0)) du$, and $u = (u_1, u_2)^T$. It can be shown similarly that

$$A(t,s) = \begin{pmatrix} a + O_p(\nu^2(n,m)) & \mathbf{1}^T \mathbf{H} \mathbf{O}_p(\nu(n,m)) \\ \mathbf{H} \mathbf{1} \mathbf{O}_p(\nu(n,m)) & C(1 + O_p(\nu^2(n,m))) \end{pmatrix},$$

$$\mathbf{B}(t,s) = \begin{pmatrix} O_p(\nu(n,m)) \\ \mathbf{H} \mathbf{1} \mathbf{O}_p(\nu(n,m)) \end{pmatrix},$$
where \( a \in [C_{\min}, C_{\max}] \), and all elements of the \( 3 \times 3 \) matrix \( C \) are in the same order of the corresponding elements of \( H^2 \mu_v(K), I = (1,1,1)^T \) and \( H = \text{diag}(h_1, h_2, h_3) \). It follows that

\[
\Pi_2 = e_i^T A^{-1}(t,s) B(t,s) = O_p(v(n,m)).
\] (A.11)

By using similar arguments, we have \( \Pi_3 = O_p(v(n,m)) \). By combining this result with (A.1), (A.2) and (A.11), the result in (11) of the paper is true.

**PROOF OF THEOREM 2**

From (A.1)-(A.3), we have \( \hat{\lambda}(t,s) = \lambda(t,s) + \Pi_2 + \Pi_3 \), where \( \Pi_2 = e_i^T A^{-1}(t,s) B(t,s) \), and \( A^{-1}(t,s) \) and \( B(t,s) \) are defined in the proof of Theorem 1. Next we will show that \( B_i(t,s) = O_p(a(n,m)) \) uniformly for \( (t,s) \in [0,1] \times \Omega \), where \( a(n,m) = \{\log^2(n)/(nh_1^2)\}^{1/2} \).

First, note that the spatial location of interest \( \Omega \) is bounded, then it is clear that \([0,1] \times \Omega \) can be covered by \( N^* = O(|a(n,m)h_1|^{-3}) \) regions \( \{R_l; l = 1, ..., N^*\} \), where \( R_l = \{(t,s) : |t - t_l^*| \leq a(n,m)h_1, \; d_E(s,s_l^*) \leq a(n,m)h_1 \} \) and \( \{(t_l^*, s_l^*), l = 1, ..., N^*\} \) are the centroids of the \( N^* \) 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 \( h_1/h_2 = O(1) \), we can find some constant \( C_2 > 0 \) such that \( h_1 \leq C_2h_2 \). Define \( C_K = \sup_{x \in \Omega} \{K_1(x), K_2(x)\} \). Then, for any \( (t,s) \in R_l \) and a sufficiently large \( n \), we have

\[
\left| K_1 \left( \frac{t - t_l^*}{h_1} \right) K_2 \left( \frac{d_E(s_{ij}, s)}{h_2} \right) - K_1 \left( \frac{t - t_l^*}{h_1} \right) K_2 \left( \frac{d_E(s_{ij}, s_l^*)}{h_2} \right) \right| \\
\leq C_L K_l h_1^{-1} \left( |t - t_l^*| + C_2 d_E(s_{ij}, s_l^*) \right) \left( \frac{|t - t_l^*|}{h_1} \leq 2L_1 \right) \left( \frac{d_E(s_{ij}, s_l^*)}{h_2} \leq 2L_2 \right),
\] (A.12)

where \([L_1, L_1]\) and \([-L_2, L_2]\) are the finite supports for \( K_1(x) \) and \( K_2(x) \), respectively. Define \( \tilde{K}_1(x) = 1/(2L_1)|x| \leq 2L_1 \) and \( \tilde{K}_2(x) = 1/(4\pi L_2^2)|x| \leq 2L_2 \). Then, by (A.12), there exists a constant \( C_3 > 0 \) such that

\[
\left| K_1 \left( \frac{t - t_l^*}{h_1} \right) K_2 \left( \frac{d_E(s_{ij}, s)}{h_2} \right) - K_1 \left( \frac{t - t_l^*}{h_1} \right) K_2 \left( \frac{d_E(s_{ij}, s_l^*)}{h_2} \right) \right| \\
\leq C_3 a(n,m) \tilde{K}_1 \left( \frac{t - t_l^*}{h_1} \right) \tilde{K}_2 \left( \frac{d_E(s_{ij}, s_l^*)}{h_2} \right).
\] (A.13)

Define

\[
\tilde{\mathcal{B}}_i(t,s) = \{nh_1^2 f(s)\}^{-1} \sum_{i=1}^{n} \sum_{j=1}^{m} \tilde{K}_1 \left( \frac{(t - t_l^*)}{h_1} \right) \tilde{K}_2 \left( \frac{d_E(s_{ij}, s_l^*)}{h_2} \right) |\epsilon_{ij}|
\]

Since \( \tilde{K}_1(\cdot) \) and \( \tilde{K}_2(\cdot) \) satisfy the assumptions about the kernel function in Theorem 1, it can be checked that

\[
E \left( \tilde{\mathcal{B}}_i(t,s) \right) \leq C^{1/6} \left( 1 + O(h_2^2 + 1/(nh_1)) \right) < \infty,
\]

where \( \delta \) and \( C_\epsilon \) are defined in Theorem 1. Based on the result in (A.13), it can be checked that

\[
\sup_{(t,s) \in R_l} \left| \mathcal{B}_i(t,s) - E[\mathcal{B}_i(t,s)] \right| \leq C_3 a(n,m) \left[ \tilde{\mathcal{B}}_i(t_l^*, s_l^*) + E \left( \tilde{\mathcal{B}}_i(t_l^*, s_l^*) \right) \right]
\]

\[
\leq C_3 a(n,m) \left[ \tilde{\mathcal{B}}_i(t_l^*, s_l^*) + E \left( \tilde{\mathcal{B}}_i(t_l^*, s_l^*) \right) \right] + 2C_3 a(n,m) E \left( \mathcal{B}_i(t_l^*, s_l^*) \right)
\]

\[
\leq 2C_3 a(n,m) M,
\] (A.14)

where the final inequality is obtained because \( a(n,m) < 1 \) and \( T > E(\mathcal{B}_i(t,s)) \) when \( n, m \) and \( T \) are large enough. By (A.14), it can be checked that

\[
\Pr \left( \sup_{(t,s) \in [0,1] \times \Omega} |B_i(t,s) - E[B_i(t,s)]| > (2 + 4C_3)T a(n,m) \right) \\
\leq N^* \max_{1 \leq i \leq N^*} \Pr \left( |B_i(t_l^*, s_l^*) - E[B_i(t_l^*, s_l^*)]| > 2T a(n,m) \right)
\]

\[
+ N^* \max_{1 \leq i \leq N^*} \Pr \left( |\tilde{B}_i(t_l^*, s_l^*) - E[\tilde{B}_i(t_l^*, s_l^*)]| > 2T a(n,m) \right).
\] (A.15)
For the two parts on the right-hand side of (A.15), we can use similar arguments to find their upper bounds, because both $(K_i(x), K_j(x))$ and $(\bar{K}_i(x), \bar{K}_j(x))$ satisfy the assumptions on the kernel functions given in Theorem 1.

Second, for any $(t, s) \in [0, 1] \times \Omega$, by the fact that $E(\epsilon_i(t, s)) = 0$, we have

$$\Pr \left( |\epsilon_i(t, s)| > 2T a(n, m) \right) = 4 \exp \left( \frac{-T^2 \log(n)}{64 \Theta_1 D(s)^2 + T^{3/2}} \right) + 4 \exp \left( \frac{-m C K T^{1/2}}{64 \Theta_1 D(s)^2 + T^{3/2}} \log(n) \right),$$

(A.18)

when $\log(n) > 1$. Note that the second term on the right-hand side of (A.18) is independent of the choice of $\{s_{ij}, j = 1, \ldots, m_i, i = 1, \ldots, n\}$. Then, by the Bernstein’s inequality, we have

$$\Pr \left( |\epsilon_i(t, s)| > a(n, m)T \right) \leq 4 \exp \left( \frac{-T^2 \log(n)}{64 \Theta_1 D(s)^2 + T^{3/2}} \right),$$

(A.19)

where $C_{\max}$ and $C_{\min}$ are defined in the proof of Theorem 1. In addition, from (A.17), by the Markov’s inequality, we have

$$\Pr \left( |\epsilon_i(t, s)| > a(n, m)T \right) = O \left( \left( \frac{a(n, m)T}{\varphi_n^4} \right)^{-1} \right).$$

(A.20)
Therefore, by combining (A.19) with (A.20), when $T$ is large enough, we have
\[
\Pr \left( |B_i(t^*, s^*_i) - E(B_i(t^*, s^*_i))| > 2a(n, m)T \right) = O \left( \left\{ a(n, m)T \phi_n^2 \right\}^{-1} \right) \\
+ O \left( n^{-T^{1/2}/65} \right) + O \left( n \exp \left( -mh_2^2 \left( C_{min} T^{1/2} - 1 \right) \right) \right) \\
+ O \left( n \exp \left\{ -\log(n) \frac{C_1 T^{1/2}}{10C_K} \right\} \right).
\] (A.21)

By (A.15) and (A.21), it can be shown that, when $T$ is large enough,
\[
\Pr \left( \sup_{(t, s) \in [0, 1] \times \Omega} |B_i(t, s) - E(B_i(t, s))| > (2 + 4C_T)Ta(n, m) \right) \\
= O \left( \{a(n, m)^4h_1T\phi_n^2\}^{-1} \right) + O \left( a(n, m)^{-3}h_1^{-3}n^{-T^{1/2}/65} \right) \\
+ O \left( a(n, m)^{-3}h_1^{-3}n \exp \left\{ -mh_2^2 \left( C_{min} T^{1/2} - 1 \right) \right\} \right) \\
+ O \left( a(n, m)^{-3}h_1^{-3}n \exp \left\{ -\log(n) \frac{C_1 T^{1/2}}{20C_K} \right\} \right) = o(1).
\] (A.22)

Note that $E(B_i(t, s)) = 0$. So, by (A.22), we have $B_i(t, s) = O_p(a(n, m))$, which is uniformly true for all $(t, s) \in [0, 1] \times \Omega$. The vector of the remaining elements of $B(t, s)$ can be proved in a similarly way to be of the order $H1O_p(a(n, m))$, where $H = \text{diag}(h_1, h_2, h_3)$ and $I = (1, 1, 1)^T$. Thus, we have
\[
B(t, s) = \begin{pmatrix} O_p(a(n, m)) \\ H1O_p(a(n, m)) \end{pmatrix},
\]
which are uniformly true for all $(t, s) \in [0, 1] \times \Omega$.

Next, we will study the properties of $A(t, s)$. To this end, let $b(n, m) = h_1^2 + h_2^2 + (\log(n)^2/(nh_2^2))^{1/2}$, and $b^*(n, m) = b(n, m) + (\log(m)/(mh_2^2))^{1/2}$. Then, it can be shown by similar arguments to those for deriving (A.12)-(A.22) that
\[
A(t, s) = \begin{pmatrix} a + O_p(b^*(n, m)) & 1^T H O_p(b^*(n, m)) \\ H1O_p(b^*(n, m)) & C(1 + O_p(b^*(n, m))) \end{pmatrix},
\]
where $a \in [C_{min}, C_{max}]$, all elements of the $3 \times 3$ matrix $C$ are in the same order of the corresponding elements of $H^2 \mu_2(K)$, and $\mu_2(K)$ is defined in the proof of Theorem 1. By combining the above results, we have
\[
\Pi_2 = e_1^T A(t, s)^{-1} B(t, s) = O_p(a(n, m)),
\] (A.23)
which is uniformly true for all $(t, s) \in [0, 1] \times \Omega$. For $\Pi_3$ defined in (A.2), in a similar way that we study the property of $B(t, s)$, it can be checked that
\[
\Pi_3 = O_p\left( h_1^2 + h_2^2 \right),
\] (A.24)
which is uniformly true for all $(t, s) \in [0, 1] \times \Omega$. By combining the results in (A.1), (A.2), (A.23) and (A.24), the result (12) in Theorem 2 has been proved.

**PROOF OF THEOREM 3**

First, we consider the convergence property of $\hat{\sigma}^2(t, s)$. Since
\[
\hat{e}_{ij} - e_{ij} = \left( y_{ij} - \hat{\lambda}(t, s_{ij}) \right) - \left( y_{ij} - \lambda(t, s_{ij}) \right) = \lambda(t, s_{ij}) - \hat{\lambda}(t, s_{ij}),
\]
we know from Theorem 2 that $\hat{e}_{ij} - e_{ij}$ is bounded by a term of the order $O_p(b(n, m))$ uniformly for all $i$ and $j$. Then we have
\[
\hat{\sigma}^2(t, s) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} w_{i, j} \hat{e}_{ij}^2}{\sum_{i=1}^{n} \sum_{j=1}^{m} w_{i, j}} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} w_{i, j} e_{ij}^2}{\sum_{i=1}^{n} \sum_{j=1}^{m} w_{i, j} (i, j)} + O_p(b(n, m)) =: \Pi_5/\Pi_4 + O_p(b(n, m)),
\] (A.25)
Similarly, it can be shown that

$$\Pi_5 = \frac{1}{nh_3^3} \sum_{i=1}^n \frac{m_i}{m} K_i (\frac{t_i - t}{h_3}) [\sigma^2(t, s)f(s) + O(h_3^2)]$$

(A.26)

By similar arguments to those in (A.6)-(A.9), it can be checked that $\text{Var}(\Pi_4) = O(1/(nh_3^3))$. By combining this result with that in (A.26), we have

$$\Pi_5 = \frac{1}{nh_3^3} \sum_{i=1}^n \frac{m_i}{m} K_i (\frac{t_i - t}{h_3}) + O_p(\hat{v}(n, m)).$$

(A.27)

Similarly, it can be shown that

$$\Pi_4 = \frac{1}{nh_3^3} \sum_{i=1}^n \frac{m_i}{m} K_i (\frac{t_i - t}{h_3}) + O_p(\hat{v}(n, m)) + \{mh_3^2\}^{-1/2},$$

(A.28)

It follows from (A.25), (A.27) and (A.28) that the result in (13) of the paper is true.

Second, for any $\sigma \geq 0$, $s, s' \in \Omega$, and $t', t \in [0, 1]$ such that $n(t' - t) = \sigma + o(1)$, we consider the convergence property of $\tilde{V}(t, t'; s, s')$. To this end, we first decompose $\tilde{V}(t, t'; s, s') - V(t, t'; s, s')$ into three parts. Define

$$\tilde{V}(t, t'; s, s') = \sum_{i,j} \sum_{(k,l) \neq (i,j)} \frac{w_2(i, j, k, l) \epsilon_{ij} \epsilon_{kl}}{\sum_{i,j} \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l)},$$

and

$$\beta^*(t, t'; s, s') = \sum_{i,j} \sum_{(k,l) \neq (i,j)} \frac{w_2(i, j, k, l) V(t, t'; s, s')}{\sum_{i,j} \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l)}.$$

Then, it is straightforward that

$$\tilde{V}(t, t'; s, s') - V(t, t'; s, s') = \{\tilde{V}(t, t'; s, s') - V^*(t, t'; s, s')\} + \{V^*(t, t'; s, s') - V(t, t'; s, s')\}.$$

(A.29)

where

$$\Lambda_1 = \sum_{i,j} \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l) \left( \epsilon_{ij} \epsilon_{kl} - V(t, t'; s, s') \right), \quad \Lambda_2 = \sum_{i,j} \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l),$$

$$\Lambda_3 = \sum_{i,j} \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l) \left( V(t, t'; s, s') \right), \quad \Lambda_4 = \sum_{i,j} \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l) \left( \epsilon_{ij} \epsilon_{kl} - \epsilon_{ij} \epsilon_{kl} \right).$$

Next, we will show that each of $\Lambda_1/\Lambda_2$, $\Lambda_3/\Lambda_2$ and $\Lambda_4/\Lambda_2$ can be bounded by a term of the order $O_p(\hat{v}(n, m) + b(n, m))$, where $\hat{v}(n, m) = h_3^2 + h_3^2 + \{nh_3\}^{-1/2}$ and $b(n, m) = h_3^2 + h_3^2 + \{\log(n^2/(nh_3^3))\}^{1/2}$. To this end, let us first consider $\Lambda_1$. From its definition, it can be checked that $E (\Lambda_1 | S_\sigma) = 0$, where $S_\sigma$ is the $\sigma$-algebra generated by $S = \{s_1, \ldots, s_{nm}\}$. It follows from this that

$$E (\Lambda_1) = E \left[ E (\Lambda_1 | S_\sigma) \right] = 0.$$  

(A.30)

Denote $R(i, j, k, l; i', j', k', l') = E \{ w_2(i, j, k, l) w_2(i', j', k', l') \} | \text{Cov}(\epsilon_{ij} \epsilon_{kl}, \epsilon_{i'j'} \epsilon_{k'l'})$. To calculate the variance of $\Lambda_1$, since $E (\Lambda_1 | S_\sigma) = 0$ and $\text{Var}(\Lambda_1) = \text{Var} (E (\Lambda_1 | S_\sigma)) + E \left[ \text{Var} (\Lambda_1 | S_\sigma) \right]$, we have

$$\text{Var}(\Lambda_1) = \sum_{i,j} \sum_{(k,l) \neq (i,j)} \sum_{i',j',k',l'} \sum_{i',j',k',l'} \text{Cov}(\epsilon_{ij} \epsilon_{kl}, \epsilon_{i'j'} \epsilon_{k'l'}) \text{Cov}(\epsilon_{ij} \epsilon_{kl}, \epsilon_{i'j'} \epsilon_{k'l'}) \leq \sum_{i,j} \sum_{k,l} \sum_{i',j',k',l'} R(i, j, k, l; i', j', k', l') \leq 2 \sum_{i,j} \sum_{k,l} \sum_{i',j',k',l'} R(i, j, k, l; i', j', k', l')$$

(A.31)

$$= O(1) \times \sum_{i,j} \sum_{k,l} \sum_{i',j',k',l'} R(i, j, k, l; i', j', k', l') = O(1) \times (\Lambda_{1,1} + \Lambda_{1,2} + \Lambda_{1,3}).$$

where
where
\[ \Lambda_{1,1} = \sum_{i,j} \sum_{i' \geq i} \sum_{k' \geq i'} \sum_{j \leq j'} R(i, j, k; i', j', k', l'), \]
\[ \Lambda_{1,2} = \sum_{i,j} \sum_{i' \geq i} \sum_{k' \geq i} \sum_{j' \leq j} R(i, j, k; i', j', k', l'), \]
\[ \Lambda_{1,3} = \sum_{i,j} \sum_{i' \geq i} \sum_{k' \geq i} \sum_{j' \leq j} R(i, j, k; i', j', k', l'). \] (A.32)

Next, we will find the upper bounds for \( \Lambda_{1,1}, \Lambda_{1,2}, \) and \( \Lambda_{1,3}, \) respectively. To this end, we first consider \( \text{Cov}(\epsilon_{ij,k}, \epsilon_{i'j',k'}) \), for any \( i, j, k, l, i', j', k', \) and \( l' \) such that \( i \leq k, \) \( i' \leq k' \) and \( i \leq l' \). Note that, if \( i \leq k \leq l' \), by the Davydov’s inequality, we have
\[ |\text{Cov}(\epsilon_{ij,k}, \epsilon_{i'j',k'})| \leq C_4 \exp(-C_3 |l'| - k), \] (A.33)
where \( C_4 = 12 C_\epsilon^3 C_0^{(\delta-4)/6} \) and \( C_3 = C_1(\delta - 4)/\delta. \) If \( i' < k' \), then it can be shown that
\[ |\text{Cov}(\epsilon_{ij,k}, \epsilon_{i'j',k'})| \leq |\text{Cov}(\epsilon_{ij,k}, \epsilon_{i'j',k'})| + |E(\epsilon_{ij,k})E(\epsilon_{i'j',k'})| + |E(\epsilon_{ij,k})E(\epsilon_{i'j',k'})| \]
\[ \leq C_4 \{ \exp(-C_3 |l' - k|) + \exp(-C_3 |l' - k')| + \exp(-C_3 |l'| - l|) \}. \] (A.34)

If \( k' < k \), it is clear that
\[ |\text{Cov}(\epsilon_{ij,k}, \epsilon_{i'j',k'})| \leq 2C_4. \] (A.35)

Denote \( \Delta_n(t, t') = \{|t - t'| - 1/n, |t - t'| + 1/n\}. \) For the quantity \( \Lambda_{1,1} \), it can be shown from (A.33) that
\[ \Lambda_{1,1} = O(1) \times m^3 h_3^4 \sum_{i,j} \sum_{i' \geq i} \sum_{k' \geq i'} \sum_{j \leq j'} K_i(t_i - t_i') K_k(t_k - t_k') K_j(t_j - t_j') \times \text{I}(\{|t_i - t_k| \in \Delta_n(t, t')\}) \text{I}(\{|t_{i'} - t_{k'}| \in \Delta_n(t, t')\}) \exp(-C_3 |l' - k|). \] (A.36)

Note that, for every integer number \( k \), the number of different \( i \)’s such that \( I(\{|t_i - t_k| \in \Delta_n(t, t')\}) = 1 \) cannot exceed 3. Meanwhile, given the value of \( i' \), we have at most 3 different \( k' \)’s such that \( I(\{|t_{i'} - t_{k'}| \in \Delta_n(t, t')\}) = 1 \). Thus, it follows from (A.36) that
\[ \Lambda_{1,1} = O(1) \times C_K^3 m^3 h_3^4 \sum_{i,j} \sum_{i' \geq i} \sum_{k' \geq i'} \sum_{j \leq j'} K_i(t_i - t_i') \exp(-C_3 |l' - k|) \]
\[ = O(m^3 h_3^4) \times \sum_{i,j} \sum_{i' \geq i} \sum_{k' \geq i'} \sum_{j \leq j'} \exp(-C_3 \tau) \]
\[ = O(m^3 h_3^4) \times \sum_{i,j} \sum_{i' \geq i} \sum_{k' \geq i'} \sum_{j \leq j'} \exp(-C_3 \tau), \] (A.37)

where \( C_K = \sup_{x \in \mathbb{R}} \{K_1(x), K_2(x)\} \). When \( t' \) is fixed, it has been well studied that \( (nh_3)^{-1} \sum_{k=1}^n K_1((t_k - t')/h_3) = 1 + O(1/(nh_3)). \) However, in the setup of Theorem 3 of the paper, \( t' \) may change with the number of observation times \( n. \) As a result, in order to obtain the upper bound of \( \Lambda_{1,1}, \) we have to calculate the upper bound of \( \sum_{u \in [0,1]} (t_k - u)/h_3 \). Define \( \Delta_k = (t_{k-1}, t_k], \) for \( k = 1, \ldots, n, \) where \( t_0 = 0 \) and \( \{t_k = k/n, k = 1, \ldots, n\}. \) Then, for any \( u \in [0,1], \) it can be shown that
\[ \frac{1}{n h_3} \sum_{k=1}^n K_1\left(\frac{t_k - u}{h_3}\right) = \frac{1}{h_3} \sum_{k=1}^n \int K_1(\frac{x - u}{h_3})I(x \in \Delta_k)dx = \frac{1}{h_3} \sum_{k=1}^n \int K_1(\frac{x - u}{h_3})I(x \in \Delta_k)dx 
+ \frac{1}{h_3} \sum_{k=1}^n \int \left\{K_1(\frac{t_k - u}{h_3}) - K_1(\frac{x - u}{h_3})\right\}I(x \in \Delta_k)dx 
\leq \int K_1(t)dz + \frac{1}{h_3} \sum_{k=1}^n \int \left\{K_1(\frac{t_k - u}{h_3}) - K_1(\frac{x - u}{h_3})\right\}I(x \in \Delta_k)dx 
= 1 + \frac{1}{h_3} \sum_{k=1}^n \int \left\{K_1(\frac{t_k - u}{h_3}) - K_1(\frac{x - u}{h_3})\right\}I(x \in \Delta_k)dx. \] (A.38)
Note that when \( x \in \Delta_k \), \( K_1 \left( (t_k - u)/h_3 \right) - K_1 \left( (x - u)/h_3 \right) \leq L_K |x - t_k| h_3^{-1} |l(t_k - u) \leq 2L_1 h_3 \), where \( L_K \) is the Lipschitz constant of the kernel functions and \( [−L_1, L_1] \) is the finite support of \( K_1(x) \). It follows that

\[
\left| \frac{1}{h_3} \sum_{k=1}^{n} \int \left\{ K_1 \left( \frac{t_k - u}{h_3} \right) - K_1 \left( \frac{x - u}{h_3} \right) \right\} I(x \in \Delta_k) dx \right|
\leq \frac{1}{h_3} \sum_{k=1}^{n} \int \left| K_1 \left( \frac{t_k - u}{h_3} \right) - K_1 \left( \frac{x - u}{h_3} \right) \right| I(x \in \Delta_k) dx
\leq \frac{1}{h_3} \sum_{k=1}^{n} \int L_K |x - t_k| h_3^{-1} |l(t_k - u) \leq 2L_1 h_3| I(x \in \Delta_k) dx
\leq L_K (nh_3^2)^{-1} \sum_{k=1}^{n} \int |l(t_k - u) \leq 2L_1 h_3| I(x \in \Delta_k) dx
\leq L_K (nh_3^2)^{-1} (4L_1 h_3 + 2/n).
\]

Combining the results in (A.39) and (A.38), it immediately follows that \( (nh_3)^{-1} \sum_{k=1}^{n} K_1 \left( (t_k - u)/h_3 \right) \leq 1 + L_K (nh_3^2)^{-1} (4L_1 h_3 + 2/n) \), for any \( u \in [0, 1] \). Thus,

\[
\sup_{u \in [0, 1]} (nh_3)^{-1} \sum_{k=1}^{n} K_1 \left( (t_k - u)/h_3 \right) \leq 1 + L_K (nh_3^2)^{-1} (4L_1 h_3 + 2/n) = 1 + O(nh_3)^{-1}.
\]

(A.40)

Based on the results in (A.37) and (A.40), we have

\[
\Lambda_1,1 = O(nm^4 h_3^4 h_4^8) \times (nh_3)^{-1} \sum_{k=1}^{n} K_1 \left( (t_k - t)/h_3 \right)
= O(nm^4 h_3^4 h_4^8) \times \sup_{u \in [0, 1]} (nh_3)^{-1} \sum_{k=1}^{n} K_1 \left( (t_k - u)/h_3 \right) = O(nm^4 h_3^4 h_4^8).
\]

(A.41)

To calculate the second part \( \Lambda_{1,2} \) in (A.32), note that the number of different \( k \)'s such that \( i' < k \leq k' \) is no more than \( |k' - i'| \) and \( |k' - i'| \leq r_n + 1 \), where \( r_n \) is the closest positive integer less than or equal to \( n|t - t'| \). By using the result in (A.34), we have

\[
\Lambda_{1,2} = O(1) \times m^4 h_3^8 \sum_{i} \sum_{i' < k' < k} \sum_{i < k < k'} K_1 \left( \frac{t_i - t}{h_3} \right) K_1 \left( \frac{t_k - t'}{h_3} \right) K_1 \left( \frac{t_{i'} - t'}{h_3} \right) K_1 \left( \frac{t_{i'} - t'}{h_3} \right) \times 1 \left( |t_{i'} - t_{k'}| \geq \Delta_n(t, t') \right)
\times \left( \exp \left( -C_5 |i' - i| \right) + \exp \left( -C_5 |i' - k'| \right) + \exp \left( -C_5 |i' - i| \right) \right)
= O(nh_3^4 m^4 h_4^8) + O(m^4 h_4^8) \sum_{i} (r_n + 1) K_1 \left( \frac{t_i - t}{h_3} \right) \exp \left( -C_5 |r_n - 1| \right)
+ O(nh_3^4 m^4 h_4^8) \sum_{i} \sum_{i'} K_1 \left( \frac{t_i - t}{h_3} \right) K_1 \left( \frac{t_{i'} - t}{h_3} \right) \exp \left( -C_5 |i' - i| \right)
= O(nh_3^4 m^4 h_4^8) + O(nh_3^4 m^4 h_4^8) \times \sup_{u \in [0, 1]} (nh_3)^{-1} \sum_{i}^{n} K_1 \left( \frac{t_i - u}{h_3} \right) \sum_{r=0}^{\infty} \exp \left( -C_5 r \right) = O(nh_3^4 m^4 h_4^8).
\]

(A.42)

Note that \( k - i \leq k' - i' + 2 \) if both \( |t_k - t_i| \) and \( |t_k - t_{i'}| \) are in \( \Delta_n(t, t') \), \( i \leq k \), and \( i' \leq k' \). It follows that \( k \leq k' + 2 - (i' - i) \). When \( i' \geq i \), the number of different \( k \)'s satisfying \( k > k' \) and \( k < k' + 2 - (i' - i) \) is less than or equal to 2. From the definition of \( \Lambda_{1,3} \) and the result in (A.35), it can be easily checked that

\[
\Lambda_{1,3} = O(nh_3 m^4 h_4^8).
\]

(A.43)

From the results in (A.41)-(A.43), we have \( \text{Var}(\Lambda_1) = O(nm^4 h_3^4 h_4^8) \). It follows from this result and (A.30) that

\[
\Lambda_1 = O_p(nm^2 h_3^4 h_4^8 \nu(n, m)).
\]

(A.44)

For \( \Lambda_2 \), by using the similar arguments in (A.30)-(A.44), we can obtain the result that

\[
\Lambda_2 = \sum_{i=1}^{n} \sum_{k=1}^{n} K_1 \left( \frac{t_i - t}{h_3} \right) K_1 \left( \frac{t_k - t'}{h_3} \right) \times m_m m_k f(s) f(s') h_4^8 + O_p(nm^2 h_3^4 h_4^8 \nu(n, m)),
\]

(A.45)
where \( \tilde{v}(n, m) = v(n, m) + \{ mh_4^2 \}^{-1/2} \). Since \( n(t - t') = o + o(1) \), when \( n \) is large enough, for arbitrary \( 1 \leq i \leq n \), we can find at least one integer \( 1 \leq k \leq n \) such that \( I(\{ |t_i - t_k| \} \in \Delta_n(t, t')) = 1 \). Moreover, it can be easily checked that \( |K_1((t, t')/h_3) - K_1((t_k - t')/h_3)| \leq L_K(nh_3)^{-1} \) when \( I(\{ |t_i - t_k| \} \in \Delta_n(t, t')) = 1 \). So, we have

\[
(nh_3)^{-1} \sum_{i=1}^{n} \sum_{k=1}^{n} K_1 \left( \frac{t_i - t_k}{h_3} \right) K_1 \left( \frac{t_k - t'}{h_3} \right) \mathbb{1}(\{ |t_i - t_k| \in \Delta_n(t, t') \}) \geq (nh_3)^{-1} \sum_{i=1}^{n} K_1 \left( \frac{t_i - t}{h_3} \right) - L_K(nh_3)^{-2} \sum_{i=1}^{n} K_1 \left( \frac{t_i - t}{h_3} \right)
\]

\[
\geq (nh_3)^{-1} \inf_{u \in [0, 1]} \sum_{i=1}^{n} K_1 \left( \frac{t_i - u}{h_3} \right) - L_K(nh_3)^{-2} \sup_{u \in [0, 1]} \sum_{i=1}^{n} K_1 \left( \frac{t_i - u}{h_3} \right) = \frac{1}{2} \mu(K_1^2) + O(\frac{1}{nh_3}),
\]

where \( \mu(K_1^2) = \int K_1^2(x) dx \). It follows from (A.44)-(A.46) that

\[
\Lambda_1/\Lambda_2 = O_p(\tilde{v}(n, m)). \tag{A.47}
\]

For \( \Lambda_3 \), note that \( V(t, t'; s, s') \) is twice continuously differentiable. By the Taylor’s expansion, the following result is true:

\[
\Lambda_3 = O_p(nm^2 h_3 h_4^2 \tilde{v}(n, m)). \tag{A.48}
\]

From (A.45), (A.46) and (A.48), we have

\[
\Lambda_3/\Lambda_2 = O_p(\tilde{v}(n, m)). \tag{A.49}
\]

So far, we have shown that \( \Lambda_1/\Lambda_2 + \Lambda_2/\Lambda_2 = O_p(\tilde{v}(n, m)) \). To prove that the result in (14) of the paper is true, it suffices to show that \( \Lambda_2/\Lambda_2 = O_p(\tilde{v}(n, m)) \). From the definition of \( \Lambda_2 \) and \( \Lambda_4 \) and the result that \( \hat{\epsilon}_{ij} - \epsilon_{ij} \) is bounded by a term of the order \( O_p(b(n, m)) \) uniformly for all \( i \) and \( j \) (see the arguments at the beginning of the proof), we have

\[
|\Lambda_2| \leq \sum_{i,j} \left| \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l) [\hat{\epsilon}_{ij} - \epsilon_{ij}] [\hat{\epsilon}_{kl} - \epsilon_{kl}] \right| \leq \sum_{i,j} \left| \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l) [\hat{\epsilon}_{ij} - \epsilon_{ij}] [\hat{\epsilon}_{kl} - \epsilon_{kl}] \right|
\]

\[
\leq \sum_{i,j} \left| \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l) \right| + \sum_{i,j} \left| \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l) [\hat{\epsilon}_{kl} - \epsilon_{kl}] \right|
\]

\[
= O_p(b(n, m)^2) + O_p(b(n, m)) \times \sum_{i,j} \left| \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l) \right|
\]

\[
+ O_p(b(n, m)) \times \sum_{i,j} \left| \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l) \hat{\epsilon}_{kl} \right|
\]

\[
\leq O_p(b(n, m)^2) + O_p(b(n, m)) \times \sum_{i,j} \left| \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l) \right| + O_p(b(n, m)) \times \sum_{i,j} \left| \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l) \right|
\]

\[
= O_p(b(n, m)^2) + O_p(b(n, m)) \times \sum_{i,j} \left| \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l) \right|.
\]

Similar to the arguments in (A.31)-(A.47), it can be shown that

\[
\sum_{i,j} \left| \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l) \epsilon_{kl} \right| = O_p(1), \quad \text{and}
\]

\[
\sum_{i,j} \left| \sum_{(k,l) \neq (i,j)} w_2(i, j, k, l) \epsilon_{ij} \right| = O_p(1).
\]

The results in (A.50) and (A.51) imply that

\[
\Lambda_4/\Lambda_2 = O_p(b(n, m)). \tag{A.52}
\]

By combining the results in (A.47), (A.49) and (A.52), the result in (14) of the paper has been proved.