MULTIVARIATE SPLINE ESTIMATION AND INference
FOR IMAGE-ON-SCALAR REGRESSION

Shan Yu\textsuperscript{1}, Guannan Wang\textsuperscript{2}, Li Wang\textsuperscript{1} and Lijian Yang\textsuperscript{3}

\textsuperscript{1}Iowa State University, Ames, IA 50011, USA.

\textsuperscript{2}College of William \& Mary, Williamsburg, VA 23187, USA.

\textsuperscript{3}Tsinghua University, Beijing 100084, China.

Supplementary Material

This supplementary document includes detailed proofs of the theoretical results in the main article and additional results from simulation and application studies. In Appendix 1, we first describe the basic properties of the bivariate splines, then we present the asymptotic properties of the penalized/unpenalized bivariate spline estimators and the piecewise constant bivariate spline estimators of the coefficient functions, and investigate the convergence of the proposed covariance estimator of the coefficient functions. In Appendix 2, we provide more results from the simulation studies and ADNI data analysis.

S1. Appendix 1

In the following, we use $c$, $C$, $c_1$, $c_2$, $C_1$, $C_2$, etc. as generic constants, which may be different even in the same line. For any sequence $a_n$ and $b_n$, we write $a_n \asymp b_n$ if there exist two positive constants $c_1, c_2$ such that $c_1|a_n| \leq |b_n| \leq c_2|a_n|$, for all $n \geq 1$. For a real valued vector $\mathbf{a}$, denote $\|\mathbf{a}\|$ its Euclidean norm. For a matrix $\mathbf{A} = (a_{ij})$, denote $\|\mathbf{A}\|_\infty = \max_{i,j} |a_{ij}|$. For any positive definite matrix $\mathbf{A}$,
let $\lambda_{\text{min}}(A)$ and $\lambda_{\text{max}}(A)$ be the smallest and largest eigenvalues of $A$. For a vector valued function $g = (g_0, \ldots, g_p)^\top$, denote $\|g\|_{L^2(\Omega)} = \left\{ \sum_{\ell=0}^p \|g_\ell\|_{L^2(\Omega)}^2 \right\}^{1/2}$ and $\|g\|_{\infty,\Omega} = \max_{0 \leq \ell \leq p} \|g_\ell\|_{\infty,\Omega}$, where $\|g_\ell\|_{L^2,\Omega}$ and $\|g_\ell\|_{\infty,\Omega}$ are the $L^2$ norm and supremum norm of $g_\ell$ defined at the beginning of Section 2.3. Further denote $\|g\|_{u,\infty,\Omega} = \max_{0 \leq \ell \leq p} \parallel g_\ell \parallel_{u,\infty,\Omega}$, where $\parallel g_\ell \parallel_{u,\infty,\Omega}$ is the $L^2$ norm of $g_\ell$ defined at the beginning of Section 2.3. Furthermore, let $\|\cdot\|_{n,N}$ be the norm introduced by the inner product $\langle \cdot, \cdot \rangle_{n,N}$, where, for $g^{(1)}(z)$ and $g^{(2)}(z)$, $\langle g^{(1)}, g^{(2)} \rangle_{n,N} = \frac{1}{nN} \sum_{\ell,\ell'=0}^p \sum_{i=1}^n \sum_{j=1}^N X_{i\ell}X_{i\ell'} g^{(1)}_\ell(z_j) g^{(2)}_{\ell'}(z_j)$.

and the theoretical inner product as

$$\langle g^{(1)}, g^{(2)} \rangle = \sum_{\ell,\ell'=0}^p E(X_{\ell}X_{\ell'}) \int_\Omega g^{(1)}_\ell(z) g^{(2)}_{\ell'}(z) dz,$$

and denote the corresponding empirical and theoretical norms $\|\cdot\|_{n,N}$ and $\|\cdot\|$.

Furthermore, let $\|\cdot\|_\varepsilon$ be the norm introduced by the inner product $\langle \cdot, \cdot \rangle_{\varepsilon}$, where, for $g^{(1)}(z)$ and $g^{(2)}(z)$,

$$\langle g^{(1)}, g^{(2)} \rangle_{\varepsilon} = \sum_{\ell,\ell'=0}^p \int_\Omega \left\{ \sum_{i+j=2}^2 \binom{2}{i} (\nabla_{z_1}^i \nabla_{z_2}^j g^{(1)}_\ell)^2 \right\}^{1/2} \left\{ \sum_{i+j=2}^2 \binom{2}{i} (\nabla_{z_1}^i \nabla_{z_2}^j g^{(2)}_{\ell'})^2 \right\}^{1/2} dz_1 dz_2.$$

Let $A(\Omega)$ be the area of the domain $\Omega$, and without loss of generality, we assume $A(\Omega) = 1$ in the rest of the article. Note that the triangulation for different coefficient function can be different from each other. For notational convenience in the proof below, we consider a common triangulation for all the explanatory variables: $B_0(z) = B_1(z) = \cdots = B_p(z) = B(z)$, and $\beta(\ell, z_j) = B^\top(z_j) \gamma_\ell$. 

2
S1. APPENDIX 1

S1.1 Properties of Bivariate Splines

We cite two important results from Lai and Schumaker (2007).

**Lemma S1** (Theorem 2.7, Lai and Schumaker (2007)). Let \( \{B_m\}_{m \in M} \) be the Bernstein polynomial basis for spline space \( S^r_d(\Delta) \) defined over a \( \pi \)-quasi-uniform triangulation \( \Delta \). Then there exist positive constants \( c, C \) depending on the smoothness \( r, d \), and the shape parameter \( \pi \) such that

\[
c|\Delta|^2 \sum_{m \in M} \gamma_m^2 \leq \left\| \sum_{m \in M} \gamma_m B_m \right\|_{L^2}^2 \leq C|\Delta|^2 \sum_{m \in M} \gamma_m^2.
\]

**Lemma S2** (Theorems 10.2 and 10.10, Lai and Schumaker (2007)). Suppose that \( |\Delta| \) is a \( \pi \)-quasi-uniform triangulation of a polygonal domain \( \Omega \), and \( g(\cdot) \in W^{d+1,\infty}(\Omega) \).

(i) For bi-integer \( (a_1, a_2) \) with \( 0 \leq a_1 + a_2 \leq d \), there exists a spline \( g^*(\cdot) \in S^0_d(\Delta) \) such that

\[
\| \nabla_{z_1}^{a_1} \nabla_{z_2}^{a_2} (g - g^*) \|_{\infty} \leq C|\Delta|^{d+1-a_1-a_2} |g|_{d+1,\infty},
\]

where \( C \) is a constant depending on \( d, r, \) and the shape parameter \( \pi \).

(ii) For bi-integer \( (a_1, a_2) \) with \( 0 \leq a_1 + a_2 \leq d \), there exists a spline \( g^{**}(\cdot) \in S^r_d(\Delta) \) (\( d \geq 3r + 2 \)) such that

\[
\| \nabla_{z_1}^{a_1} \nabla_{z_2}^{a_2} (g - g^{**}) \|_{\infty} \leq C|\Delta|^{d+1-a_1-a_2} |g|_{d+1,\infty},
\]

where \( C \) is a constant depending on \( d, r, \) and the shape parameter \( \pi \).

Lemma S2 shows that \( S^0_d(\Delta) \) has full approximation power, and \( S^r_d(\Delta) \) also has full approximation power if \( d \geq 3r + 2 \). For any \( g(\cdot) \) in Sobolev space \( C^{(0)}(\Omega) \), there exists a spline \( g^*(\cdot) \in PC(\Delta) \) such that

\[
\| g - g^* \|_{\infty} \leq C|\Delta|\| g \|_{\infty}.
\]

**Lemma S3.** Let \( g(z) = (g_0(z), \ldots, g_p(z))^\top \), where \( g_\ell(z) = \sum_{m \in M} \gamma_{\ell m} B_m(z) \). Then, under Assumptions (A3) and (A5), \( \| g \| \propto \sum_{\ell=0}^p \| g_\ell \|_{L^2} \).

**Proof.** By (S1),

\[
\| g \|^2 = \sum_{\ell, \ell' = 0}^p E(X_{\ell}X_{\ell'}) \int_\Omega g_{\ell}(z)g_{\ell'}(z)dz = \int_\Omega g^\top(z)\Sigma_X g(z)dz.
\]
According to Assumptions (A3) and (A5), \( \|g\|^2 \gc 1_{0}^{\infty} g^T(z)g(z)dz \leq \sum_{\ell=0}^{p} \|g_{\ell}\|_{L^2} \). \( \square \)

**Lemma S4.** Under Assumptions (A4) and (A5), for any Bernstein basis polynomials \( B_m(z) \), \( m \in \mathcal{M} \), of degree \( d \geq 0 \), we have

\[
\max_{m \in \mathcal{M}} \left| \frac{1}{N} \sum_{j=1}^{N} B^k_m(z_j) - \int_{\Omega} B^k_m(z)dz \right| = O \left( \|\Delta\|N^{-1/2} \right), \quad 1 \leq k < \infty, \quad (S3)
\]

\[
\max_{m,m' \in \mathcal{M}} \left| \frac{1}{N} \sum_{j=1}^{N} B_m(z_j)B_{m'}(z_j) - \int_{\Omega} B_m(z)B_{m'}(z)dz \right| = O \left( \|\Delta\|N^{-1/2} \right), \quad (S4)
\]

\[
\max_{m,m' \in \mathcal{M}} \left| \sum_{j,j'=1}^{N} G(z_j,z_{j'}) B_m(z_j)B_{m'}(z_{j'}) - \int_{\Omega^2} G(z,z') B_m(z)B_{m'}(z')dzdz' \right|
\]

\[
= O \left( N^{-1/2} \|\Delta\|^3 \right), \quad (S5)
\]

\[
\max_{m \in \mathcal{M}} \|\sigma B_m\|_{N,L^2}^2 - \|\sigma B_m\|_{L^2}^2 \leq \max_{m \in \mathcal{M}} \left| \frac{1}{N} \sum_{j=1}^{N} B^2_m(z_j)\sigma^2(z_j) - \int_{\Omega} \sigma^2(z)B^2_m(z)dz \right|
\]

\[
= O \left( N^{-1/2} \|\Delta\| \right). \quad (S6)
\]

**Proof.** Note that there are \( d^* = (d + 1)(d + 2)/2 \) Bernstein basis polynomials on each triangle and \( \int_{\Omega} B^k_m(z)dz = \int_{[m/d^*]} B^k_m(z)dz \), for any \( k \geq 1 \).

For piecewise constant basis functions, we have \( B_m(z) = I(z \in T_m) \), then

\[
\left| \frac{1}{N} \sum_{j=1}^{N} B^k_m(z_j) - \int_{\Omega} B^k_m(z)dz \right| = \left| \frac{1}{N} \sum_{j=1}^{N} I(z_j \in T_m) - A(T_m) \right|
\]

According to Assumption (A5),

\[
\max_{m \in \mathcal{M}} \left| \frac{1}{N} \sum_{j=1}^{N} B^k_m(z_j) - \int_{\Omega} B^k_m(z)dz \right| \leq CN^{-1/2}\|\Delta\|.
\]

For any \( j = 1, \ldots, N \), let \( V_j \) be the \( j \)th pixel, and it is clear that

\[
\left| \frac{1}{N} \sum_{j=1}^{N} B^k_m(z_j) - \int_{\Omega} B^k_m(z)dz \right| \leq \sum_{j=1}^{N} \int_{V_j} \{B^k_m(z_j) - B^k_m(z)\}dz + \int_{\Omega \cup \cup V_j} B^k_m(z)dz.
\]
If $d \geq 1$, by the properties of bivariate spline basis functions in [Lai and Schumaker 2007], \(\int_{\Omega \setminus \cup V_j} B^k_m(z) \, dz = O(N^{-1/2}|\Delta|)\), and

\[
\left| \sum_{j=1}^N \int_{V_j} \{ B^k_m(z_j) - B^k_m(z) \} \, dz \right| \leq \sum_{\{j: z_j \in T_{[m/d+1]} \}} \int_{V_j} |B^k_m(z_j) - B^k_m(z)| \, dz
\]

\[
\leq C(N|\Delta|^2) \times N^{-1} \times (N^{-1/2}|\Delta|^{-1}) \leq CN^{-1/2}|\Delta|.
\]

Thus, (S3) holds. The proof of (S4) is similar to the proof (S3), thus omitted.

Next, for any \(m, m' \in M\),

\[
\frac{1}{N^2} \sum_{j=1}^N \sum_{j'=1}^N G_\eta(z_j, z_{j'}) B_m(z_j) B_{m'}(z_{j'}) - \int_{\Omega^2} G_\eta(z, z') B_m(z) B_{m'}(z') \, dz \, dz' = \sum_{j=1}^N \sum_{j'=1}^N \int_{V_j \times V_{j'}} \{ G_\eta(z_j, z_{j'}) B_m(z_j) B_{m'}(z_{j'}) - G_\eta(z, z') B_m(z) B_{m'}(z') \} \, dz \, dz' + \int_{\Omega^2 \setminus \cup_{j,j'} V_j \times V_{j'}} \{ G_\eta(z_j, z_{j'}) B_m(z_j) B_{m'}(z_{j'}) - G_\eta(z, z') B_m(z) B_{m'}(z') \} \, dz \, dz'.
\]

As \(N \to \infty\),

\[
\int_{\Omega^2 \setminus \cup_{j,j'} V_j \times V_{j'}} \{ G_\eta(z_j, z_{j'}) B_m(z_j) B_{m'}(z_{j'}) - G_\eta(z, z') B_m(z) B_{m'}(z') \} \, dz \, dz' = O\left(\frac{|\Delta|^3}{\sqrt{N}}\right).
\]

Notice that

\[
\left| \sum_{j=1}^N \sum_{j'=1}^N \int_{V_j \times V_{j'}} \{ G_\eta(z_j, z_{j'}) B_m(z_j) B_{m'}(z_{j'}) - G_\eta(z, z') B_m(z) B_{m'}(z') \} \, dz \, dz' \right|
\]

\[
\leq \sum_{\{(j,j'): z_j \in T_{[m/d+1]} \}, z_{j'} \in T_{[m'/d+1]} \}} \int_{V_j \times V_{j'}} \omega_{j,j'}(G_\eta K_m, 2N^{-1/2}) \, dz \, dz',
\]

where \(K_m(z, z') = B_m(z) B_{m'}(z')\) and

\[
\omega_{j,j'}(g, \varrho) = \sup_{(z_1, z_1'), (z_2, z_2') \in V_j \times V_{j'}, \|z_1 - z_2\|^2 + \|z_1' - z_2'\|^2 = \varrho} |g(z_1, z_1') - g(z_2, z_2')|.
\]
is the modulus of continuity of \( g \) on \( V_j \times V_{j'} \). Therefore, by Assumption (A4), we have

\[
\left| \sum_{j=1}^{N} \sum_{j' = 1}^{N} \int_{V_j \times V_{j'}} \{ G_\eta(z_j, z_{j'}) B_m(z_j) B_{m'}(z_{j'}) - G_\eta(z, z') B_m(z) B_{m'}(z') \} \, dz \, dz' \right|
\leq (N|\Delta|^2)^2 \times N^{-2} \times (N^{-1/2}|\Delta|^{-1}) = O(N^{-1/2}|\Delta|^3).
\]

Thus, (S5) follows.

Finally, note that

\[
\left| \frac{1}{N} \sum_{j=1}^{N} B_m^2(z_j) \sigma^2(z_j) - \int_{\Omega} \sigma^2(z) \, dz \right| \leq \sum_{j=1}^{N} \int_{V_j} \{ B_m^2(z_j) \sigma^2(z_j) - B_m^2(z) \sigma^2(z) \} \, dz \leq \sum_{j=1}^{N} \int_{V_j} \{ B_m^2(z_j) \sigma^2(z_j) - B_m^2(z) \sigma^2(z) \} \, dz.
\]

It is easy to see that \( \int_{\Omega \setminus V_j} |B_m^2(z_j) \sigma^2(z_j) - B_m^2(z) \sigma^2(z)| \, dz = O(N^{-1/2}|\Delta|) \). Denote \( \omega_j(g, \varrho) = \sup_{z, z' \in V_j, \|z - z'\| = \varrho} |g(z) - g(z')| \) is the modulus of continuity of \( g \) on the \( j \)th pixel \( V_j \), then by Assumption (A4), we have

\[
\left| \sum_{j=1}^{N} \int_{V_j} \{ B_m^2(z_j) \sigma^2(z_j) - B_m^2(z) \sigma^2(z) \} \, dz \right| \leq \sum_{\{j: z_j \in T_{[m/\ell']}\}} \int_{V_j} \omega_j(B_m^2 \sigma^2, 2N^{-1/2}) \, dz
\leq C(N|\Delta|^2) \times N^{-1} \times (N^{-1/2}|\Delta|^{-1}) \leq CN^{-1/2}|\Delta|.
\]

We obtain (S6).

\[\square\]

**Lemma S5.** For any \( m \in \mathcal{M} \), \( 0 \leq \ell, \ell' \leq p \), let \( \Phi_{m, \ell, \ell'} = E(X_{\ell} X_{\ell'}) \int_{\Omega} B_m^2(z) \, dz \).

Suppose Assumptions (A3) and (A5) hold, and \( N^{1/2}|\Delta| \to \infty \) as \( N \to \infty \), then with probability 1, one has

\[
\max_{m \in \mathcal{M}} \max_{0 \leq \ell, \ell' \leq p} \left| \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} B_m^2(z_j) X_{i\ell} X_{i\ell'} - \Phi_{m, \ell, \ell'} \right| = O \left( n^{-1/2}|\Delta|^2 (\log n)^{1/2} + N^{-1/2}|\Delta| \right).
\]
Proof. Let $\varsigma_{i,m} \equiv \varsigma_{i,m,\ell,\ell'} = \frac{1}{N} \sum_{j=1}^{N} B_m^2(z_j) X_{i\ell} X_{i\ell'}$. If $N^{1/2} |\Delta| \to \infty$ as $N \to \infty$, then by [S3], we can show that $E(\varsigma_{i,m}) = \frac{1}{N} \sum_{j=1}^{N} B_m^2(z_j) E(X_\ell X_{\ell'}) \asymp |\Delta|^2$, and $E(\varsigma_{i,m})^2 = \left\{ \frac{1}{N} \sum_{j=1}^{N} B_m^2(z_j) \right\}^2 E(X_\ell X_{\ell'})^2 \asymp |\Delta|^4$.

Next define a sequence $D_n = n^\alpha$ with $\alpha \in (1/3, 1/2)$. We make use of the following truncated and tail decomposition $X_{i\ell'} = X_{i\ell} X_{i\ell'} = X_{i\ell}^{D_n} + X_{i\ell'}^{D_n}$, where $X_{i\ell}^{D_n} = X_{i\ell} X_{i\ell'} I \{ |E(X_{i\ell} X_{i\ell'}) > D_n \}, X_{i\ell'}^{D_n} = X_{i\ell} X_{i\ell'} I \{ |X_{i\ell} X_{i\ell'}| \leq D_n \}$. Correspondingly the truncated and tail parts of $\varsigma_{i,m}$ are $\varsigma_{i,m,v} \equiv \varsigma_{i,m,v,\ell,\ell'} = \frac{1}{N} \sum_{j=1}^{N} B_m^2(z_j) X_{i\ell'}^{D_n}, v = 1, 2$.

According to Assumption (A3), for any $\ell, \ell' = 0, \ldots, p$,

$$\sum_{n=1}^{\infty} P \{ |X_{n\ell} X_{n\ell'}| > D_n \} \leq \sum_{n=1}^{\infty} \frac{E |X_{n\ell} X_{n\ell'}|^3}{D_n^3} \leq C_b \sum_{n=1}^{\infty} D_n^{-3} < \infty.$$

By Borel-Cantelli Lemma, $\frac{1}{N} \sum_{j=1}^{N} B_m^2(z_j) X_{i\ell'}^{D_n,1} = 0$, almost surely. So for any $k \geq 1$, $\sup_{m,\ell,\ell'} |n^{-1} \sum_{i=1}^{n} \varsigma_{i,m,1}| = O_{a.s.}(n^{-k})$. Since $N^{1/2} |\Delta| \to \infty$ as $N \to \infty$,

$$|E(\varsigma_{i,m,1})| = |E(X_{i\ell}^{D_n})| \left\{ \frac{1}{N} \sum_{j=1}^{N} B_m^2(z_j) \right\} \leq D_n^{-2} E |X_{i\ell} X_{i\ell'}|^3 \left\{ \int_{\Omega} B_m(z) dz + O(N^{-1/2} |\Delta|) \right\} \leq C D_n^{-2} |\Delta|^2.$$

Next, we consider the truncated part $\varsigma_{i,m,2}$. Define $\varsigma_{i,m,2}^* = \varsigma_{i,m,2} - E(\varsigma_{i,m,2})$, then $E\varsigma_{i,m,2}^* = 0$, and

$$E(\varsigma_{i,m,2}^*)^2 = E(\varsigma_{i,m,2})^2 - (E\varsigma_{i,m,2})^2 = \left\{ \frac{1}{N} \sum_{j=1}^{N} B_m^2(z_j) \right\}^2 \left\{ E(X_{i\ell}^{D_n})^2 - (E X_{i\ell}^{D_n,2})^2 \right\}.$$

Note that $E(X_{i\ell}^{D_n})^2 \leq D_n^{-1} E |X_{i\ell} X_{i\ell'}|^3 \leq c D_n^{-1}$, thus, $E(X_{i\ell}^{D_n})^2 = E(X_{i\ell}^2) - E(X_{i\ell}^{D_n,1})^2 = E(X_{i\ell}^2) - o(1)$. Therefore, there exists $c_\varsigma$ such that for large $n$, we
have $E(\varsigma_{i,m,2}^*)^2 \geq c_\varsigma E(X_{i\ell'}')^2 \times \left\{ \frac{1}{N} \sum_{j=1}^{N} B_m^2(z_j) \right\}^2$. Next for any $k > 2$,

$$E|\varsigma_{i,m,2}^*|^k = E|\varsigma_{i,m,2} - E(\varsigma_{i,m,2})|^k \leq 2^{k-1} \left( E|\varsigma_{i,m,2}|^k + |E(\varsigma_{i,m,2})|^k \right)$$

$$= 2^{k-1} \left\{ E|X^s_{i\ell',2}|^k + O(1) \right\} \left\{ \frac{1}{N} \sum_{j=1}^{N} B_m^2(z_j) \right\}^k,$$

then there exists $C_\varsigma > 0$ such that for any $k > 2$ and large $n$,

$$E|\varsigma_{i,m,2}^*|^k \leq 2^{k-1} \left\{ D_n^{k-2} E(X_{i\ell'}')^2 + O(1) \right\} \left\{ \frac{1}{N} \sum_{j=1}^{N} B_m^2(z_j) \right\}^k$$

$$\leq 2^k D_n^{k-2} E(\varsigma_{i,m,2}^*)^2 \left\{ \frac{1}{N} \sum_{j=1}^{N} B_m^2(z_j) \right\}^{k-2} \leq (C_\varsigma D_n |\Delta|^2)^{k-2} k! E(\varsigma_{i,m,2}^*)^2,$$

which implies that $\{\varsigma_{i,m,2}^*\}_{i=1}^{n}$ satisfies Cramér’s condition with constant $C_\varsigma D_n |\Delta|^2$.

Applying Bernstein’s inequality to $\sum_{i=1}^{n} \varsigma_{i,m,2}^*$, for $k > 2$ and any large enough $\delta > 0$,

$$P\left\{ \left| \frac{1}{n} \sum_{i=1}^{n} \varsigma_{i,m,2}^* \right| \geq \delta n^{-1/2} |\Delta|^2 (\log n)^{1/2} \right\} \leq 2 \exp\left\{ -\frac{\delta^2 \log(n)}{4 + 2C_\varsigma D_n \delta (\log n)^{1/2} n^{-1/2}} \right\}.$$

Assume that $|\Delta|^{-2} \asymp n^\tau$ for some $0 < \tau < \infty$, we have

$$\sum_{n=1}^{\infty} P\left\{ \max_{m \in M, 0 \leq \ell, \ell' \leq p} \left| \frac{1}{n} \sum_{i=1}^{n} \varsigma_{i,m,2}^* \right| \geq \delta n^{-1/2} |\Delta|^2 (\log n)^{1/2} \right\} \leq 2 \sum_{n=1}^{\infty} \sum_{m \in M} \sum_{0 \leq \ell, \ell' \leq p} n^{-2-\tau} < \infty.$$

Thus, $\sup_{m,\ell,\ell'} |\sum_{i=1}^{n} \varsigma_{i,m,2}^*| = O_{a.s.} \left\{ n^{-1/2} |\Delta|^2 (\log n)^{1/2} \right\}$ as $n \to \infty$, by Borel-Cantelli Lemma. Furthermore,

$$\max_{m,\ell,\ell'} \left| n^{-1} \sum_{i=1}^{n} \varsigma_{i,m} - E\varsigma_{i,m} \right| \leq \max_{m,\ell,\ell'} \left| n^{-1} \sum_{i=1}^{n} \varsigma_{i,m,1} \right| + \max_{m,\ell,\ell'} \left| n^{-1} \sum_{i=1}^{n} \varsigma_{i,m,2}^* \right| + \max_{m,\ell,\ell'} |E\varsigma_{i,m,1}|

= O_{a.s.} (n^{-k}) + O_{a.s.} \left\{ n^{-1/2} |\Delta|^2 (\log n)^{1/2} \right\} + O \left( D_n^{-2} |\Delta|^2 \right) = O_{a.s.} \left\{ n^{-1/2} |\Delta|^2 (\log n)^{1/2} \right\}$. 
Finally, we notice that

\[
\max_{m \in M, 0 \leq \ell, \ell' \leq p} \left| \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} B_m^2(z_j) X_{it} X_{i\ell'} - \Phi_{m, \ell, \ell'} \right|
\]

\[
= \max_{m \in M, 0 \leq \ell, \ell' \leq p} \left| n^{-1} \sum_{i=1}^{n} \varsigma_{i,m} - E \varsigma_{i,m} \right| + |E X_{it} X_{i\ell'}| \max_{m \in M} \left| \frac{1}{N} \sum_{j=1}^{N} B_m^2(z_j) - \int_{\Omega} B_m^2(z) dz \right|
\]

\[
= O_{a.s.} \left\{ n^{-1/2} |\Delta|^2 (\log n)^{1/2} \right\} + O(N^{-1/2} |\Delta|).
\]

We obtain the desired result. \(\square\)

The following lemma provide the uniform convergence rate at which the empirical inner product in (S1) approximates the theoretical inner product in (S2).

**Lemma S6.** Let \(g_\ell^{(1)}(z) = \sum_{m \in M} c_{\ell m}^{(1)} B_m(z), g_\ell^{(2)}(z) = \sum_{m \in M} c_{\ell m}^{(2)} B_m(z)\) be any spline functions in \(S^r_{d}(\Delta)\). Denote \(g(z) = (g_0(z), \ldots, g_p(z))^\top\) with \(g_\ell \in S^r_{d}(\Delta), \ell = 0, \ldots, p\).

**Suppose Assumptions (A3) and (A5) hold, and \(N^{1/2} |\Delta| \to \infty\) as \(N \to \infty\), then**

\[
R_{n,N} = \sup_{g^{(1)}, g^{(2)} \in S^r_{d}(\Delta)} \left| \langle g^{(1)}, g^{(2)} \rangle_{n,N} - \langle g^{(1)}, g^{(2)} \rangle \right| \|g^{(1)}\| \|g^{(2)}\| = O_P \left\{ n^{-1/2} (\log n)^{1/2} + N^{-1/2} |\Delta|^{-1} \right\}.
\]

**Proof.** It is easy to see

\[
\langle g^{(1)}, g^{(2)} \rangle_{n,N} = \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} \left\{ \sum_{\ell=0}^{p} \sum_{m \in M} c_{\ell m}^{(1)} X_{it} B_m(z_j) \right\} \left\{ \sum_{\ell'=0}^{p} \sum_{m' \in M} c_{\ell' m'}^{(2)} X_{i\ell'} B_{m'}(z_j) \right\}
\]

\[
= \sum_{\ell, m, \ell', m'} c_{\ell m}^{(1)} c_{\ell' m'}^{(2)} \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} X_{it} X_{i\ell'} B_{m}(z_j) B_{m'}(z_j).
\]

Note that \(\|g^{(r)}\|^2 = \sum_{\ell, m} \sum_{\ell', m'} c_{\ell m}^{(r)} c_{\ell' m'}^{(r)} E(X_{it} X_{i\ell'}) \int_{\Omega} B_m(z) B_{m'}(z) dz, r = 1, 2\). It fol-
lows from Assumptions (A1), (A2), Lemmas S1 and S3 that,
\[ c_v |\Delta|^2 \sum_{\ell, m} \{c_{\ell m}^{(v)}\}^2 \leq \|g^{(v)}\|^2 \leq C_v |\Delta|^2 \sum_{\ell, m} \{c_{\ell m}^{(v)}\}^2; \]
\[ C_1 |\Delta|^2 \left[ \sum_{\ell, m} \{c_{\ell m}^{(1)}\}^2 \sum_{\ell', m'} \{c_{\ell' m'}^{(2)}\}^2 \right]^{1/2} \leq \|g^{(1)}\| \|g^{(2)}\| \leq C_2 |\Delta|^2 \left[ \sum_{\ell, m} \{c_{\ell m}^{(1)}\}^2 \sum_{\ell', m'} \{c_{\ell' m'}^{(2)}\}^2 \right]^{1/2}. \]

With the above preparation, we have
\[ R_{n,N} \leq \frac{\sum_{\ell, \ell', m-m' \leq (d+2)(d+1)/2} |c_{\ell m}^{(1)} c_{\ell' m'}^{(2)}|}{C_1 |\Delta|^2 \left[ \sum_{\ell, m} \{c_{\ell m}^{(1)}\}^2 \sum_{\ell', m'} \{c_{\ell' m'}^{(2)}\}^2 \right]^{1/2}} \times \max_{m, m' \in M} \left| \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N B_m(z_j) B_{m'}(z_j) X_{it} X_{it'} - E(X_{it} X_{it'}) \int_{\Omega} B_m(z) B_{m'}(z) dz \right| \]
\[ \leq C \frac{\max_{m, m' \in M}}{\max_{0 \leq \ell, \ell' \leq p}} \left| \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N B_m(z_j) B_{m'}(z_j) X_{it} X_{it'} - E(X_{it} X_{it'}) \int_{\Omega} B_m(z) B_{m'}(z) dz \right|. \]

The desired result follows from (S7) and Lemma S5.

As a direct result of Lemma S6, we can see that
\[ \sup_{g \in S_0'(\Delta)} \left| \frac{\|g\|^2_{n,N}}{\|g\|^2} - 1 \right| = O_P\{n^{-1/2}(\log n)^{1/2} + N^{-1/2} |\Delta|^{-1}\}. \] (S8)

**S1.2 Uniform Convergence of the Unpenalized Spline Estimators**

In this section, we consider the unpenalized spline smoothing approach. The unpenalized bivariate spline estimator of \( \beta^o = (\beta_0^o, \ldots, \beta_p^o)^T \) is defined as
\[ \tilde{\beta} = (\tilde{\beta}_0, \ldots, \tilde{\beta}_p)^T = \arg \min_{\beta \in \mathbb{R}^{p+1}} \sum_{i=1}^n \sum_{j=1}^N \left\{ Y_i(z_j) - \sum_{\ell=0}^p X_{it} \beta_{\ell}(z_j) \right\}^2. \] (S9)
Denote

\[ \tilde{\theta}_\mu = (\tilde{\theta}_{\mu,0}, \ldots, \tilde{\theta}_{\mu,p})^\top = \Gamma_n^{-1} \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N \left\{ \tilde{X}_i \otimes \tilde{B}(z_j) \right\} \tilde{X}_i^\top \theta^\alpha(z_j) \]

\[ \tilde{\theta}_\eta = (\tilde{\theta}_{\eta,0}, \ldots, \tilde{\theta}_{\eta,p})^\top = \Gamma_n^{-1} \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N \left\{ \tilde{X}_i \otimes \tilde{B}(z_j) \right\} \eta_h(z_j), \]

\[ \tilde{\theta}_\varepsilon = (\tilde{\theta}_{\varepsilon,0}, \ldots, \tilde{\theta}_{\varepsilon,p})^\top = \Gamma_n^{-1} \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N \left\{ \tilde{X}_i \otimes \tilde{B}(z_j) \right\} \sigma(z_j) \varepsilon_{ij}, \]

where

\[ \Gamma_{n,0} = \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N (\tilde{X}_i \tilde{X}_i^\top) \otimes \{ \tilde{B}(z_j) \tilde{B}^\top(z_j) \}. \tag{S10} \]

**Lemma S7.** Under Assumptions (A3) and (A5), if \( N^{1/2} |\Delta| \to \infty \) as \( N \to \infty \), then there exist constants \( 0 < c_\Gamma < C_\Gamma < \infty \), such that with probability approaching 1, as \( N \to \infty, n \to \infty \),

\[ c_\Gamma |\Delta|^2 \leq \lambda_{\min} (\Gamma_{n,0}) \leq \lambda_{\max} (\Gamma_{n,0}) \leq C_\Gamma |\Delta|^2, \]

where \( \Gamma_{n,0} \) is in (S10).

**Proof.** Note that for any vector \( \theta = (\theta_0^\top, \ldots, \theta_p^\top) \) with \( \gamma_\ell = (\gamma_{\ell m}, m \in \mathcal{M}) \),

\[ \theta^\top \Gamma_{n,0} \theta = \frac{1}{nN} \gamma^\top \sum_{i=1}^n \sum_{j=1}^N (\tilde{X}_i \tilde{X}_i^\top) \otimes \{ \tilde{B}(z_j) \tilde{B}^\top(z_j) \} \gamma = \| g_\gamma \|_{n,N}^2, \tag{S11} \]

where \( \gamma = Q_2 \theta \), and \( g_\gamma = (g_{\gamma_0}, \ldots, g_{\gamma_p}) \) with \( g_{\gamma_\ell} = \sum_{m \in \mathcal{M}} \gamma_{\ell m} B_m \). By (S8), we have

\[ c(1-\rho_{n,N})|\Delta|^2 \| \gamma \|^2 \leq (1-\rho_{n,N})\| g_\gamma \|^2 \leq \| g_\gamma \|_{n,N}^2 = (1+\rho_{n,N})\| g_\gamma \|^2 \leq C(1+\rho_{n,N})|\Delta|^2 \| \gamma \|^2, \]

in which we have used the stability conditions in Lemma [S1]. \( \square \)
Next, we consider the following decomposition $\tilde{\beta}(z) = \tilde{\beta}_\mu(z) + \tilde{\eta}(z) + \tilde{\epsilon}(z)$, where

$$\tilde{\beta}_\mu(z) = (\tilde{\beta}_{\mu,0}(z), \ldots, \tilde{\beta}_{\mu,p}(z))^\top = (I \otimes \tilde{B}(z))^\top \tilde{\theta}_\mu,$$  \hspace{1cm} (S12)

$$\tilde{\eta}(z) = (\tilde{\eta}_0(z), \ldots, \tilde{\eta}_p(z))^\top = (I \otimes \tilde{B}(z))^\top \tilde{\theta}_\eta,$$ \hspace{1cm} (S13)

$$\tilde{\epsilon}(z) = (\tilde{\epsilon}_0(z), \ldots, \tilde{\epsilon}_p(z))^\top = (I \otimes \tilde{B}(z))^\top \tilde{\theta}_\epsilon.$$  \hspace{1cm} (S14)

**Lemma S8.** Under Assumptions (A2)–(A5) and (C1), if $N^{1/2}|\Delta| \to \infty$ as $N \to \infty$,

$$\|\sum_{k=1}^\infty \lambda_k^{1/2}\psi_k\|_\infty < \infty \text{ and } n^{1/(4+\delta_2)} \ll n^{1/2}N^{-1/2}|\Delta|^{-1} \text{ for some } \delta_2,$$

then for $\tilde{\eta}$ and $\tilde{\epsilon}$ in (S13) and (S14), $\|\tilde{\eta}\|_\infty = O_P\{n^{-1/2}(\log n)^{1/2}\}$ and $\|\tilde{\epsilon}\|_\infty = O_P\{(nN)^{-1/2}(\log n)^{1/2}|\Delta|^{-1}\}$.\[1\]

**Proof.** Note that for any $\ell = 0, 1, \ldots, p$, $\tilde{\eta}_\ell(z) = \sum_{m \in \mathcal{M}} \tilde{\theta}_{\eta,\ell,m} \tilde{B}_m(z)$ for some coefficients $\tilde{\theta}_{\eta,\ell,m}$, so the order of $\tilde{\eta}_\ell(z)$ is related to that of $\tilde{\theta}_{\eta,\ell,m}$. In fact

$$\|\tilde{\eta}\|_\infty = \max_{0 \leq \ell \leq p} \|\tilde{\eta}_\ell\|_\infty \leq C_\eta \|\tilde{\theta}_{\eta,\ell}\|_\infty = \left\|\left(\mathbf{e}_\ell \otimes 1\right)^\top \Gamma_{n,0}^{-1} \sum_{i=1}^n \sum_{j=1}^N \{\tilde{X}_i \otimes \tilde{B}(z_j)\} \eta_i(z_j)\right\|_\infty,$$

where $\tilde{\theta}_\eta = (\tilde{\theta}_{\eta,\ell,m})_{m \in \mathcal{M}}$ with $\mathcal{M}$ being an index set of the transformed Bernstein basis polynomials $\tilde{B}_m(z)$ and $\Gamma_{n,0}$ is the symmetric positive definite matrix defined in (S10).

Thus, by Lemma S7

$$\|\tilde{\eta}\|_\infty \leq C|\Delta|^{-2} \max_{0 \leq \ell \leq p} \max_{m \in \mathcal{M}} \left|\frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N X_i \eta_i(z_j) \tilde{B}_m(z_j)\right|,$$

almost surely. Next, we show that with probability 1,

$$\max_{0 \leq \ell \leq p} \max_{m \in \mathcal{M}} \left|\frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N X_i \eta_i(z_j) \tilde{B}_m(z_j)\right| = O\left\{n^{-1/2}|\Delta|^{2}(\log n)^{1/2}\right\}. \hspace{1cm} (S15)$$

To prove (S15), let $\omega_i = \omega_{i,m} = \sum_{k=1}^\infty \lambda_k^{1/2} X_i \xi_{ik} \frac{1}{N} \sum_{j=1}^N \tilde{B}_m(z_j) \psi_k(z_j)$, where $E(\omega_i) = $
We decompose the random variable $\varpi_i$ into a tail part and a truncated part, 

$$\varpi_i^{D_n} = \sum_{k=1}^{\infty} \lambda_k^{1/2} \left\{ \frac{1}{N} \sum_{j=1}^{N} \tilde{B}_m(z_j) \psi_k(z_j) \right\} X_{it}\xi_{ik} I \{ |X_{it}\xi_{ik}| > D_n \},$$

$$\varpi_i^{D_n} = \sum_{k=1}^{\infty} \lambda_k^{1/2} \left\{ \frac{1}{N} \sum_{j=1}^{N} \tilde{B}_m(z_j) \psi_k(z_j) \right\} X_{it}\xi_{ik} I \{ |X_{it}\xi_{ik}| \leq D_n \} - \mu_i^{D_n},$$

$$\mu_i^{D_n} = \sum_{k=1}^{\infty} \lambda_k^{1/2} \left\{ \frac{1}{N} \sum_{j=1}^{N} \tilde{B}_m(z_j) \psi_k(z_j) \right\} E[X_{it}\xi_{ik} I \{ |X_{it}\xi_{ik}| \leq D_n \}],$$

where $D_n = n^{\alpha} (1/(4 + \delta_1) < \alpha < 1/2)$. At first, we show that tail part vanishes almost surely. Note that, for any $k \geq 1$,

$$\sum_{n=1}^{\infty} P \{ |X_{it}\xi_{nk}| > D_n \} \leq \sum_{n=1}^{\infty} \frac{E|X_{it}\xi_{nk}|^{4+\delta_1}}{D_n^{4+\delta_1}} \leq \delta_1 \sum_{n=1}^{\infty} D_n^{-(4+\delta_1)} < \infty.$$ 

By the Borel-Cantelli’s lemma, we can show that $E \left| \frac{1}{n} \sum_{i=1}^{n} \varpi_i^{D_n} \right| = O(n^{-r})$, for any $r > 0$. As $E(\varpi_i) = 0$, then it is straightforward to verify that $\mu_i^{D_n} = -E(\varpi_i^{D_n}) = O(D_n^{-2}|\Delta|^2)$.

Next, notice that $E(\varpi_i^{D_n}) = 0$. Then, $\text{Var}(\varpi_i^{D_n}) = E(\varpi_i^2) - E(\varpi_i^{D_n})^2 - (\mu_i^{D_n})^2 \approx |\Delta|^4$. Also, we have, for any $r \geq 3$,

$$E|\varpi_i^{D_n}|^r = E \left| \sum_{k=1}^{\infty} \lambda_k^{1/2} \left\{ \frac{1}{N} \sum_{j=1}^{N} \tilde{B}_m(z_j) \psi_k(z_j) \right\} X_{it}\xi_{ik} I \{ |X_{it}\xi_{ik}| \leq D_n \} - \mu_i^{D_n} \right|^r$$

$$\leq 2^{r-1} \left[ E \left| \sum_{k=1}^{\infty} \lambda_k^{1/2} \left\{ \frac{1}{N} \sum_{j=1}^{N} \tilde{B}_m(z_j) \psi_k(z_j) \right\} X_{it}\xi_{ik} I \{ |X_{it}\xi_{ik}| \leq D_n \} \right|^r + (\mu_i^{D_n})^r \right]$$

$$\leq \left\{ 2D_n \frac{1}{N} \sum_{j=1}^{N} \tilde{B}_m(z_j) \sum_{k=1}^{\infty} \lambda_k^{1/2} \psi_k(z_j) \right\}^{r-2} E|\varpi_i^{D_n}|^2 \leq (CD_n|\Delta|^2)^{r-2} E|\varpi_i^{D_n}|^2.$
Thus, \( E |\varpi_{i,2}/n|^r \leq \{Cn^{-1}D_n |\Delta|^2\} r^{-2} r! E(\varpi_{i,2}^2/n^2) < \infty \) with the Cramer constant \( c^* = Cn^{-1}D_n |\Delta|^2 \). By the Bernstein inequality, for any large enough \( \delta > 0 \),

\[
P \left\{ \left| \frac{1}{n} \sum_{i=1}^{n} \varpi_{i} \right| \geq \delta n^{-1/2} |\Delta|^2 (\log n)^{1/2} \right\} \leq 2 \exp \left\{ \frac{-\delta^2 \log n}{4c + 2\delta CD_n (\log n)^{1/2} n^{-1/2}} \right\} \leq 2n^{-3}.
\]

Hence,

\[
\sum_{n=1}^{\infty} P \left\{ \max_{0 \leq \ell \leq p} \max_{m \in M} \left| \frac{1}{n} \sum_{i=1}^{n} \varpi_{i} \right| \geq \delta n^{-1/2} |\Delta|^2 (\log n)^{1/2} \right\} \leq C|\Delta|^{-2} \sum_{n=1}^{\infty} n^{-3} < \infty
\]

for such \( \delta > 0 \). Thus, Borel-Cantelli’s lemma implies that \( \|\tilde{\eta}\|_\infty = O_P\{n^{-1/2}(\log n)^{1/2}\} \).

The result of \( \|\tilde{\varepsilon\|}_\infty = O_P\{(nN)^{-1/2}(\log n)^{1/2}|\Delta|^{-1}\} \) can be established similarly, thus omitted.

For \( \tilde{\beta}(z) \) defined in (S9), Theorem S1 below provides its uniform convergence rate to \( \beta^o \).

**Theorem S1.** Under Assumptions (A1)–(A6), for \( \tilde{\beta}(z) \) defined in (S9), \( \|\tilde{\beta} - \beta^o\|_\infty = O_P\{|\Delta|^{d+1} \|\beta^o\|_{d+1,\infty} + n^{-1/2}(\log n)^{1/2}\} \).

**Proof.** Note that \( \|\tilde{\beta} - \beta^o\|_\infty \leq \|\tilde{\beta}_\mu - \beta^o\|_\infty + \|\tilde{\eta}\|_\infty + \|\tilde{\varepsilon}\|_\infty \), where

\[
\tilde{\beta}_\mu = \arg \min_{\beta \in \mathcal{G}^{(p+1)}} \sum_{i=1}^{n} \sum_{j=1}^{N} \left\{ \sum_{\ell=0}^{p} X_{i\ell}(\beta^o_\ell - g_\ell)(z_j) \right\}^2 .
\]

Let \( \beta^* = (\beta^*_0, \ldots, \beta^*_p) \in \mathcal{G}^{(p+1)} \), where \( \beta^*_\ell \)'s are the best approximation to \( \beta^o_\ell \)'s with the approximation rate \( \|\hat{\beta}^*_\ell - \beta^o_\ell\|_\infty \leq C|\Delta|^{d+1} \|\beta^o\|_{d+1,\infty} \) for any \( \ell = 0, \ldots, p \). By Lai and Wang (2013),

\[
\|\tilde{\beta}_\mu - \beta^o\|_\infty \leq \|\tilde{\beta}_\mu - \beta^*\|_\infty + \|\beta^* - \beta^o\|_\infty \leq C|\Delta|^{d+1} \|\beta^o\|_{d+1,\infty}.
\]

The desired result follows from Lemma SS.
S1.3 Asymptotic Properties of Penalized Spline Estimators

Let \( \tilde{\mathbf{B}}(z) = Q_2^T \mathbf{B}(z) \), then for \( \mathbf{U} = \mathbf{X} \otimes (\mathbf{BQ}_2) \) defined in Section 2.2, we have

\[
\mathbf{U}^\top = (\tilde{\mathbf{X}}_1 \otimes \tilde{\mathbf{B}}(z_1), \ldots, \tilde{\mathbf{X}}_1 \otimes \tilde{\mathbf{B}}(z_N), \ldots, \tilde{\mathbf{X}}_n \otimes \tilde{\mathbf{B}}(z_1), \ldots, \tilde{\mathbf{X}}_n \otimes \tilde{\mathbf{B}}(z_N)),
\]

and \( \mathbf{U}^\top \mathbf{U} = \sum_{i=1}^n \sum_{j=1}^N (\tilde{\mathbf{X}}_i \tilde{\mathbf{X}}_i^\top) \otimes \{\tilde{\mathbf{B}}(z_j) \tilde{\mathbf{B}}^\top(z_j)\}, \mathbf{U}^\top \mathbf{Y} = \sum_{i=1}^n \sum_{j=1}^N (\tilde{\mathbf{X}}_i \otimes \tilde{\mathbf{B}}(z_j)) \mathbf{Y}_{ij} \).

Let

\[
\Gamma_{n,\rho} = \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N (\tilde{\mathbf{X}}_i \tilde{\mathbf{X}}_i^\top) \otimes \{\tilde{\mathbf{B}}(z_j) \tilde{\mathbf{B}}^\top(z_j)\} + \frac{\rho n}{nN} \mathbf{I}_p \otimes \mathbf{Q}_2 \mathbf{E} \mathbf{M} \mathbf{Q}_2, \tag{S17}
\]

which is a symmetric positive definite matrix.

Next, we define

\[
\hat{\theta}_\mu = (\hat{\theta}_{\mu,0}^\top, \ldots, \hat{\theta}_{\mu,p}^\top)^\top = \Gamma_{n,\rho}^{-1} \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N \{\tilde{\mathbf{X}}_i \otimes \tilde{\mathbf{B}}(z_j)\} \tilde{\mathbf{X}}_i^\top \beta^o(z_j),
\]

\[
\hat{\theta}_\eta = (\hat{\theta}_{\eta,0}^\top, \ldots, \hat{\theta}_{\eta,p}^\top)^\top = \Gamma_{n,\rho}^{-1} \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N \{\tilde{\mathbf{X}}_i \otimes \tilde{\mathbf{B}}(z_j)\} \sum_{k=1}^\infty \lambda_k^{1/2} \xi_{ik} \psi_k(z_j),
\]

\[
\hat{\theta}_\epsilon = (\hat{\theta}_{\epsilon,0}^\top, \ldots, \hat{\theta}_{\epsilon,p}^\top)^\top = \Gamma_{n,\rho}^{-1} \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N \{\tilde{\mathbf{X}}_i \otimes \tilde{\mathbf{B}}(z_j)\} \sigma(z_j) e_{ij}.
\]

Note that, for any \( \ell = 0, \ldots, p \), the penalized bivariate spline estimator \( \hat{\beta}_\ell \) can be written as:

\[
\hat{\beta}_\ell(z) = \hat{\beta}_{\mu,\ell}(z) + \hat{\eta}_\ell(z) + \hat{\epsilon}_\ell(z), \tag{S18}
\]

where

\[
\hat{\beta}_{\mu,\ell}(z) = \tilde{\mathbf{B}}(z)^\top \hat{\theta}_{\mu,\ell}, \quad \hat{\eta}_\ell(z) = \tilde{\mathbf{B}}(z)^\top \hat{\theta}_{\eta,\ell}, \quad \hat{\epsilon}_\ell(z) = \tilde{\mathbf{B}}(z)^\top \hat{\theta}_{\epsilon,\ell}.
\]

Therefore, we have

\[
\hat{\beta}_\ell(z) - \beta^o_\ell(z) = \hat{\beta}_{\mu,\ell}(z) - \beta^o_\ell(z) + \hat{\eta}_\ell(z) + \hat{\epsilon}_\ell(z). \tag{S19}
\]
Lemma S9. Under Assumptions (A3)–(A5), if \( N^{1/2}|\Delta| \to \infty \) as \( N \to \infty \), then there exist constants \( 0 < c_\Gamma < C_\Gamma < \infty \), such that with probability approaching 1 as \( N \to \infty \) and \( n \to \infty \),
\[
c_{\Gamma}|\Delta|^2 \leq \lambda_{\min}(\Gamma_{n,\rho}) \leq \lambda_{\max}(\Gamma_{n,\rho}) \leq C_\Gamma \left(|\Delta|^2 + \frac{\rho_n}{nN|\Delta|^2}\right).
\]

Proof. By (S11), it is easy to see that, for any vector \( \theta = (\theta_0^\top, \ldots, \theta_p^\top)^\top \),
\[
\theta^\top \Gamma_{n,\rho} \theta = \|g_\gamma\|_{n,N}^2 + \frac{\rho_n}{nN} \sum_{\ell=0}^p \gamma_{\ell}^\top \left[\langle B_m, B_{m'}\rangle_{\ell}\right]_{m,m' \in \mathcal{M}} \gamma_{\ell},
\]
where \( \gamma = (\gamma_0, \ldots, \gamma_p)^\top = Q_2 \theta \) with \( \gamma_{\ell} = (\gamma_{\ell m}, m \in \mathcal{M})^\top \). Using the Markov’s inequality in the supplement of Lai and Wang (2013) and Lemma S1, we have
\[
\sum_{\ell=0}^p \left\| \sum_{m \in \mathcal{M}} \gamma_{\ell m} B_m \right\|_E^2 \leq \frac{C}{|\Delta|^2} \|g_\gamma\|^2.
\]
Thus, the largest eigenvalue of the matrix \( \Gamma_{n,\rho} \) in (S17) satisfies that
\[
\lambda_{\max}(\Gamma_{n,\rho}) \leq C \left\{ (1 + R_{n,N})|\Delta|^2 + (nN|\Delta|^2)^{-1}\rho_n \right\}.
\]
Thus, we have with probability approaching 1,
\[
\lambda_{\max}(\Gamma_{n,\rho}) \leq C_\Gamma \left\{ |\Delta|^2 + (nN|\Delta|^2)^{-1}\rho_n \right\} \text{ for some positive constant } C_\Gamma.
\]
On the other hand, we use Lemma S1 and equation (S8) to have \( \|g_\gamma\|_{n,N}^2 = (1 - R_{n,N})\|g_\gamma\|^2 \geq c(1 - R_{n,N})|\Delta|^2\|g_\gamma\|^2 \).

Therefore, \( \lambda_{\min}(\Gamma_{n,\rho}) \geq c(1 - R_{n,N})|\Delta|^2 = c_\Gamma|\Delta|^2 \).

Lemma S10. Under Assumptions (A1), (A3) and (A5), if \( N^{1/2}|\Delta| \to \infty \), one has
\[
\|\hat{\beta}_\mu - \beta^o\|_{\infty} = O_P \left\{ \frac{\rho_n}{nN|\Delta|^2}\|\beta^o\|_{2,\infty} + \left(1 + \frac{\rho_n}{nN|\Delta|^2}\right)|\Delta|^{d+1}\|\beta^o\|_{d+1,\infty} \right\}.
\]

Proof. Define
\[
A_n = \sup_{g \in \mathcal{G}^{(p+1)}} \left\{ \frac{\|g\|_{\infty}}{\|g\|_{n,N}}, \|g\|_{n,N} \neq 0 \right\}, \quad \overline{A}_n = \sup_{g \in \mathcal{G}^{(p+1)}} \left\{ \frac{\|g\|_{\infty}}{\|g\|_{n,N}}, \|g\|_{n,N} \neq 0 \right\}, \quad (S20)
\]
where random variables \( A_n \) and \( \overline{A}_n \) depend on the collection of \( X_{it} \)'s, \( i = 1, \ldots, n, \).
\( \ell = 0, \ldots, p \). It is clear that
\[
\| \beta_o - \hat{\beta}_\mu \|_\infty \leq \| \beta_o - \tilde{\beta}_\mu \|_\infty + \| \tilde{\beta}_\mu - \hat{\beta}_\mu \|_\infty,
\]
where \( \tilde{\beta}_\mu \) is given in (S12), and \( \| \tilde{\beta}_\mu - \beta^o \|_\infty \leq C|\Delta|^{d+1}\| \beta^o \|_{d+1, \infty} \) according to (S16).

By the definition of \( A_n \) in (S20), we have
\[
\| \tilde{\beta}_\mu - \hat{\beta}_\mu \|_\infty \leq A_n \| \tilde{\beta}_\mu - \hat{\beta}_\mu \|_{n,N}.
\] (S21)

Note that the penalized spline \( \hat{\beta}_\mu \) of \( \beta^o \) is characterized by the orthogonality relations
\[
n N \langle \beta^o - \hat{\beta}_\mu, g \rangle_{n,N} = \rho_n \langle \hat{\beta}_\mu, g \rangle \varepsilon, \quad \text{for all } g \in \mathcal{G}^{(p+1)}, \tag{S22}
\]
while \( \tilde{\beta}_\mu \) is characterized by
\[
\langle \beta^o - \tilde{\beta}_\mu, g \rangle_{n,N} = 0, \quad \text{for all } g \in \mathcal{G}^{(p+1)}. \tag{S23}
\]

By (S22) and (S23), we have
\[
n N \| \tilde{\beta}_\mu - \hat{\beta}_\mu \|_{n,N}^2 = \rho_n \langle \hat{\beta}_\mu, \tilde{\beta}_\mu - \hat{\beta}_\mu \rangle \varepsilon. \tag{S24}
\]

Thus, by Cauchy-Schwarz inequality and the definition of \( A_n \),
\[
n N \| \tilde{\beta}_\mu - \hat{\beta}_\mu \|_{n,N}^2 \leq \rho_n \| \tilde{\beta}_\mu \|_\varepsilon \| \tilde{\beta}_\mu - \hat{\beta}_\mu \|_\varepsilon \leq \rho_n A_n \| \tilde{\beta}_\mu \|_\varepsilon \| \hat{\beta}_\mu - \hat{\beta}_\mu \|_{n,N}.
\]

Similarly, using (S24), \( n N \| \tilde{\beta}_\mu - \hat{\beta}_\mu \|_{n,N}^2 = \rho_n \{ \langle \hat{\beta}_\mu, \tilde{\beta}_\mu \rangle \varepsilon - \langle \hat{\beta}_\mu, \hat{\beta}_\mu \rangle \varepsilon \} \geq 0 \). Thus, by

Cauchy-Schwarz inequality, \( \| \tilde{\beta}_\mu \|_\varepsilon \leq \langle \hat{\beta}_\mu, \tilde{\beta}_\mu \rangle \varepsilon \leq \| \hat{\beta}_\mu \|_\varepsilon \| \tilde{\beta}_\mu \|_\varepsilon \), which implies that \( \| \tilde{\beta}_\mu \|_\varepsilon \leq \| \hat{\beta}_\mu \|_\varepsilon \). Therefore,
\[
\| \tilde{\beta}_\mu - \hat{\beta}_\mu \|_{n,N} \leq \rho_n (nN)^{-1} A_n \| \tilde{\beta}_\mu \|_\varepsilon. \tag{S25}
\]

Combining (S21) and (S25) yields that
\[
\| \tilde{\beta}_\mu - \hat{\beta}_\mu \|_\infty \leq A_n \| \tilde{\beta}_\mu - \hat{\beta}_\mu \|_{n,N} \leq \rho_n (nN)^{-1} A_n A_n \| \tilde{\beta}_\mu \|_\varepsilon.
\]
By Lemma S2, we have
\[
\|\tilde{\beta}_\mu\|_E = C_1\{\|\beta^0\|_{2,\infty} + \sum_{a_1+a_2=2} \|\nabla_{z_1}^{a_1}\nabla_{z_2}^{a_2}(\beta^0 - \tilde{\beta}_\mu)\|_\infty\} \leq C_2(\|\beta^0\|_{2,\infty} + |\Delta|^{d-1}\|\beta^0\|_{d+1,\infty}).
\]

It follows
\[
\|\tilde{\beta}_\mu - \hat{\beta}_\mu\|_\infty = \rho_n(nN)^{-1}A_n\overline{A}_nC_2(\|\beta^0\|_{2,\infty} + |\Delta|^{d-1}\|\beta^0\|_{d+1,\infty}). \tag{S26}
\]

Next we derive the order of \(A_n\) and \(\overline{A}_n\). By Markov’s inequality, for any \(g \in G^{(p+1)}\),
\[
\|g\|_\infty \leq C|\Delta|^{-1}\|g\|, \|g\|_\epsilon \leq C|\Delta|^{-2}\|g\|.
\]
Equation (S8) implies that
\[
\sup_{g \in G(\triangle)} \{\|g\|_{n,N}/\|g\|\} \geq [1 - O_P \{(\log n)^{1/2}n^{-1/2} + N^{-1/2}|\Delta|^{-1}\}]^{1/2}.
\]

Thus, we have
\[
A_n \leq C|\Delta|^{-1}\left[1 - O_P \{(\log n)^{1/2}n^{-1/2} + N^{-1/2}|\Delta|^{-1}\}\right]^{-1/2} = O_P (|\Delta|^{-1}),
\]
\[
\overline{A}_n \leq C|\Delta|^{-2}\left[1 - O_P \{(\log n)^{1/2}n^{-1/2} + N^{-1/2}|\Delta|^{-1}\}\right]^{-1/2} = O_P (|\Delta|^{-2}).
\]

Plugging the order of \(A_n\) and \(\overline{A}_n\) into (S26) yields that
\[
\|\tilde{\beta}_\mu - \hat{\beta}_\mu\|_\infty = O_P \left\{ \frac{C_2\rho_n}{nN|\Delta|^3} (\|\beta^0\|_{2,\infty} + |\Delta|^{d-1}\|\beta^0\|_{d+1,\infty}) \right\}.
\]

Hence,
\[
\|\tilde{\beta}_\mu - \beta^0\|_\infty \leq C_1|\Delta|^{d+1}\|\beta^0\|_{d+1,\infty} + O_P \left\{ \frac{C_2\rho_n}{nN|\Delta|^3} (\|\beta^0\|_{2,\infty} + |\Delta|^{d-1}\|\beta^0\|_{d+1,\infty}) \right\}.
\]

Therefore, Lemma S10 is established. \(\square\)

**Lemma S11.** Suppose Assumptions (A2)–(A5) hold and \(N^{1/2}|\Delta| \to \infty\) as \(N \to \infty\),
then \(\|\hat{\theta}_\eta\|^2 = O_P(n^{-1}|\Delta|^{-2}).\)
Proof. Note that $\hat{\theta}_\eta = \Gamma_{n,\rho}^{-1} \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} \left\{ \tilde{X}_i \otimes \tilde{B}(z_j) \right\} \sum_{k=1}^{\infty} \lambda_k^{1/2} \xi_{ik} \psi_k(z_j)$. According to Lemma S9,

$$\|\hat{\theta}_\eta\|^2 \asymp \frac{1}{n^2 N^2 |\Delta|^4} \sum_{i,i'=1}^{n} \sum_{j,j'=1}^{N} \left\{ \tilde{X}_i \otimes \tilde{B}(z_j) \right\}^\top \times \sum_{k=1}^{\infty} \lambda_k^{1/2} \xi_{ik} \psi_k(z_j) \sum_{k'=1}^{\infty} \lambda_{k'}^{1/2} \xi_{i'k'} \psi_k'(z_{j'}) .$$

Note that

$$\tilde{X}_i \otimes \tilde{B}(z_j) \sum_{k=1}^{\infty} \lambda_k^{1/2} \xi_{ik} \psi_k(z_j)$$

$$= \left( X_{i0} \tilde{B}(z_j) \right)^\top \sum_{k=1}^{\infty} \lambda_k^{1/2} \xi_{ik} \psi_k(z_j), \ldots, X_{ip} \tilde{B}(z_j) \right)^\top \sum_{k=1}^{\infty} \lambda_k^{1/2} \xi_{ik} \psi_k(z_j) \right)^\top ,$$

so one has

$$\|\hat{\theta}_\eta\|^2 \asymp \frac{1}{n^2 N^2 |\Delta|^4} \sum_{\ell=0}^{p} \sum_{i,i'=1}^{n} \sum_{j,j'=1}^{N} X_{i\ell} X_{i'\ell} \tilde{B}(z_j) \tilde{B}(z_{j'}) \sum_{k,k'=1}^{\infty} (\lambda_k \lambda_{k'})^{1/2} \xi_{ik} \psi_k(z_j) \xi_{i'k'} \psi_{k'}(z_{j'}) .$$

Because the eigenvalues of $Q_* Q_2^\top$ are either 0 or 1, under Assumptions (A2) and (A3), for any $\ell, i$,

$$\frac{1}{N^2} \sum_{j=1}^{N} \sum_{j'=1}^{N} E \left\{ X_{j\ell} \tilde{B}(z_j) \tilde{B}(z_{j'}) \sum_{k,k'=1}^{\infty} (\lambda_k \lambda_{k'})^{1/2} \xi_{ik} \psi_k(z_j) \xi_{i'k'} \psi_{k'}(z_{j'}) \right\}$$

$$\leq C \sum_{m \in M} \frac{1}{N^2} \sum_{j=1}^{N} \sum_{j'=1}^{N} B_m(z_j) B_m(z_{j'}) G_\eta(z_j, z_{j'}) .$$

Assumption (A4) and (S5) imply that

$$\frac{1}{N^2} \sum_{j \neq j'} B_m(z_j) B_m(z_{j'}) G_\eta(z_j, z_{j'}) = \int_{T_m \times T_m} G_\eta(z, z') B_m(z) B_m(z') dz dz'$$

$$\times \{ 1 + O(N^{-1/2} |\Delta|^3) \} = O(|\Delta|^4) .$$
Thus,

$$\frac{1}{N^2} \sum_{j=1}^{N} \sum_{j'=1}^{N} EX_{i\ell}^2 B(z_j)^\top B(z_{j'}) \sum_{k,k'=1}^{\infty} (\lambda_k \lambda_{k'})^{1/2} \xi_{ik} \psi_k(z_j) \xi_{i'k'} \psi_{k'}(z_{j'}) \leq C|\Delta|^2.$$  

Next for any $\ell, i \neq i', j, j'$, we have

$$E \left\{ X_{i\ell} X_{i'\ell} B(z_j)^\top B(z_{j'}) \sum_{k,k'=1}^{\infty} (\lambda_k \lambda_{k'})^{1/2} \xi_{ik} \psi_k(z_j) \xi_{i'k'} \psi_{k'}(z_{j'}) \right\} = 0.$$  

Therefore, $E\|\tilde{\theta}_\varepsilon\|^2 \leq CP(n^{-1} \Delta^{-2})$. The conclusion of the lemma follows. \hfill \blacksquare 

**Lemma S12.** Suppose Assumptions (A2)--(A5) hold and $N^{1/2}|\Delta| \to \infty$ as $N \to \infty$, then $\|\tilde{\theta}_\varepsilon\|^2 = O_P(n^{-1} N^{-1} |\Delta|^{-4}).$

**Proof.** By the definition of $\tilde{\theta}_\varepsilon$ in (S37), we have

$$\|\tilde{\theta}_\varepsilon\|^2 = \frac{1}{n^2 N^2 |\Delta|^4} \left( |\Delta|^{-2} \Gamma_{n,\rho} \right)^{-1} \sum_{i=1}^{n} \sum_{j=1}^{N} \left\{ \tilde{X}_i \otimes \tilde{B}(z_j) \right\}^\top \sigma(z_j) \varepsilon_{ij} \times \left( |\Delta|^{-2} \Gamma_{n,\rho} \right)^{-1} \sum_{i=1}^{n} \sum_{j'=1}^{N} \left\{ \tilde{X}_i \otimes \tilde{B}(z_{j'}) \right\} \sigma(z_{j'}) \varepsilon_{ij'}.$$

By Lemma S9

$$\|\tilde{\theta}_\varepsilon\|^2 \asymp \frac{1}{n^2 N^2 |\Delta|^4} \sum_{i, i'=1}^{n} \sum_{j, j'=1}^{N} \left\{ \tilde{X}_i \otimes \tilde{B}(z_j) \right\}^\top \sigma(z_j) \varepsilon_{ij} \left\{ X_{i'} \otimes \tilde{B}(z_{j'}) \right\} \sigma(z_{j'}) \varepsilon_{i'j'}.$$

Note that

$$\tilde{X}_i \otimes \tilde{B}(z_j) \sigma(z_j) \varepsilon_{ij} = \left( X_{i0} \tilde{B}(z_j)^\top \sigma(z_j) \varepsilon_{ij}, X_{i1} \tilde{B}(z_j)^\top \sigma(z_j) \varepsilon_{ij}, \ldots, X_{i_p} \tilde{B}(z_j)^\top \sigma(z_j) \varepsilon_{ij} \right)^\top,$$

so one has

$$\|\tilde{\theta}_\varepsilon\|^2 \asymp \frac{1}{n^2 N^2 |\Delta|^4} \sum_{\ell=0}^{p} \sum_{i, i'=1}^{n} \sum_{j, j'=1}^{N} X_{i\ell} X_{i'\ell} \tilde{B}(z_j)^\top \tilde{B}(z_{j'}) \sigma(z_j) \sigma(z_{j'}) \varepsilon_{ij} \varepsilon_{i'j'}.$$

20
Because the eigenvalues of $Q_2Q_2^\top$ are either 0 or 1, under Assumption (A2), for any $\ell, i$, by (S6),

\[
\frac{1}{N} \sum_{j=1}^{N} \tilde{B}(z_j)^\top \tilde{B}(z_j)\sigma^2(z_j) = B(z_j)^\top Q_2Q_2^\top B(z_j)\sigma^2(z_j) \leq C \sum_{m \in \mathcal{M}} \frac{1}{N} \sum_{j=1}^{N} B_m^2(z_j)\sigma^2(z_j) \\
\leq C \sum_{m \in \mathcal{M}} \int_{T_{|\mu|/d_n}} \sigma^2(z)B_m^2(z)dz\{1 + O(N^{-1/2}|\Delta|^{-1})\} \leq C.
\]

Next note that for any $\ell, i, j \neq j'$, $E\{X_{i\ell}^2\tilde{B}(z_j)^\top \tilde{B}(z_{j'})\varepsilon_{ij}\varepsilon_{ij'}\} = 0$, and for any $\ell, i \neq i'$, $j, j'$, $E\{X_{i\ell}X_{i'\ell}\tilde{B}(z_j)^\top \tilde{B}(z_{j'})\sigma(z_j)\sigma(z_{j'})\varepsilon_{ij}\varepsilon_{i'j'}\} = 0$. Therefore,

\[
E\|\hat{\theta}_\varepsilon\|^2 \leq \frac{1}{nN|\Delta|^4} \sum_{\ell=0}^{p} E\{X_{i\ell}^2\} \frac{1}{N} \sum_{j=1}^{N} \tilde{B}(z_j)^\top \tilde{B}(z_j)\sigma^2(z_j) \leq Cp(nN)^{-1}|\Delta|^{-4}.
\]

The conclusion of the lemma follows. \qed

**Proof of Theorem 1** By Lemma S11 and the properties of the bivariate spline basis functions in Lemma S1, $\|\hat{\eta}_\ell\|^2_{L_2} \asymp |\Delta|^2\|\hat{\theta}_{\eta,\ell}\|^2 = O_P(n^{-1})$ and $\|\hat{\varepsilon}_\ell\|^2_{L_2} \asymp |\Delta|^2\|\hat{\theta}_{\varepsilon,\ell}\|^2 = O_P(n^{-1}|\Delta|^{-2})$, for any $\ell = 0, 1, \ldots, p$. It is clear that $\|\hat{\beta}_\ell - \beta_\ell^0\|^2_{L_2} \leq \|\hat{\beta}_{\mu,\ell} - \beta_{\mu,\ell}^0\|^2_{L_2} + \|\hat{\eta}_\ell\|^2_{L_2} + \|\hat{\varepsilon}_\ell\|^2_{L_2}$, where the asymptotic order of $\|\hat{\beta}_{\mu,\ell} - \beta_{\mu,\ell}^0\|_{L_2}$ is the same as $\|\hat{\beta}_{\mu,\ell} - \beta_{\mu,\ell}^0\|_{\infty}$. The desired result follows from Lemma S10. \qed

**Lemma S13.** Under Assumptions (A1)–(A6), if for any $\ell = 0, 1, \ldots, p$, $|X_{i\ell}| \leq C_\ell < \infty$, then as $N \to \infty$ and $n \to \infty$, one has for any vector $a = (a_0^\top, \ldots, a_p^\top)^\top$ with $a^\top a = 1$, $[\text{Var}\{a^\top(\hat{\theta}_\eta + \hat{\theta}_\varepsilon)\}]^{-1/2}\{a^\top(\hat{\theta}_\eta + \hat{\theta}_\varepsilon)\} \xrightarrow{L} N(0, 1)$, where $\hat{\theta}_\eta$ and $\hat{\theta}_\varepsilon$ are given in (S37).

**Proof.** For coefficient vectors $\hat{\theta}_\eta$, $\hat{\theta}_\varepsilon$ and the matrix $\Gamma_{n,\rho}$ defined in (S17), $\text{Var}\{a^\top(\hat{\theta}_\eta + \hat{\theta}_\varepsilon)\}$
\( \hat{\theta}_\varepsilon \}) = a^\top \{ E(\hat{\theta}_\eta \hat{\theta}_\eta^\top) + E(\hat{\theta}_\varepsilon \hat{\theta}_\varepsilon^\top) \} a. \) Denote \( \Psi_\eta = (\Psi_{\eta,\ell,\ell'})_{\ell,\ell'} \) and \( \Psi_\varepsilon = (\Psi_{\varepsilon,\ell,\ell'})_{\ell,\ell'}, \) with

\[
\Psi_{\eta,\ell,\ell'} = \frac{1}{n^2 N^2} \sum_{i=1}^{n} \sum_{j=1}^{N} \sum_{j'=1}^{N} X_{i\ell} X_{i\ell'} B(z_j) \tilde{B}(z_{j'}) G_\eta(z_j, z_{j'}),
\]

\[
\Psi_{\varepsilon,\ell,\ell'} = \frac{1}{n^2 N^2} \sum_{i=1}^{n} \sum_{j=1}^{N} X_{i\ell} X_{i\ell'} \tilde{B}(z_j) \tilde{B}(z_j) \sigma^2(z_j),
\]

then we have

\[
a^\top E(\hat{\theta}_\eta \hat{\theta}_\eta^\top) a = E a^\top \Gamma_{n,\rho}^{-1} \frac{1}{n^2 N^2} \sum_{i=1}^{n} \sum_{j=1}^{N} \{ \tilde{X}_i \otimes \tilde{B}(z_j) \} \{ \tilde{X}_i \otimes \tilde{B}(z_{j'}) \}^\top G_\eta(z_j, z_{j'}) \Gamma_{n,\rho}^{-1} a
\]

\[
= E a^\top \Gamma_{n,\rho}^{-1} \Psi_\eta \Gamma_{n,\rho}^{-1} a,
\]

\[
a^\top E(\hat{\theta}_\varepsilon \hat{\theta}_\varepsilon^\top) a = E a^\top \Gamma_{n,\rho}^{-1} \frac{1}{n^2 N^2} \sum_{i=1}^{n} \sum_{j=1}^{N} \{ \tilde{X}_i \otimes \tilde{B}(z_j) \} \{ \tilde{X}_i \otimes \tilde{B}(z_{j'}) \}^\top \sigma^2(z_j) \Gamma_{n,\rho}^{-1} a
\]

\[
= E a^\top \Gamma_{n,\rho}^{-1} \Psi_\varepsilon \Gamma_{n,\rho}^{-1} a.
\]

Note that for any vector \( a \) with \( a^\top a = 1 \), we can rewrite as \( a^\top (\hat{\theta}_\eta + \hat{\theta}_\varepsilon) = \sum_{i=1}^{n} a_i^{\eta+\varepsilon} \mathbf{j}_i \), where

\[
(a_i^{\eta+\varepsilon})^2 = a^\top \Gamma_{n,\rho}^{-1} \frac{1}{n^2 N^2} \sum_{j=1}^{N} \sum_{j'=1}^{N} \{ \tilde{X}_i \otimes \tilde{B}(z_j) \} \{ \tilde{X}_i \otimes \tilde{B}(z_{j'}) \}^\top G_\eta(z_j, z_{j'}) \Gamma_{n,\rho}^{-1} a
\]

\[
+ a^\top \Gamma_{n,\rho}^{-1} \frac{1}{n^2 N^2} \sum_{j=1}^{N} \{ \tilde{X}_i \otimes \tilde{B}(z_j) \} \{ \tilde{X}_i \otimes \tilde{B}(z_{j'}) \}^\top \sigma^2(z_j) \Gamma_{n,\rho}^{-1} a = (a_i^{\eta})^2 + (a_i^{\varepsilon})^2,
\]

and conditional on \( \{ \tilde{X}_i, i = 1, \ldots, n \} \), \( \mathbf{j}_i \) are independent with mean zero and variance one. Thus, \( \sum_{i=1}^{n} (a_i^{\eta})^2 = a^\top \Gamma_{n,\rho}^{-1} \Psi_\eta \Gamma_{n,\rho}^{-1} a \) and \( \sum_{i=1}^{n} (a_i^{\varepsilon})^2 = a^\top \Gamma_{n,\rho}^{-1} \Psi_\varepsilon \Gamma_{n,\rho}^{-1} a \).

According to Lemma 8, Assumptions (A2) and (A4),

\[
E a^\top \Gamma_{n,\rho}^{-1} \Psi_\eta \Gamma_{n,\rho}^{-1} a \geq c_{\Gamma}^{-2} \left( |\Delta|^2 + \frac{\rho_n}{nN|\Delta|^2} \right)^{-2} E a^\top \Psi_\eta a,
\]

22
where

\[ a^\top \Psi_n a = \frac{1}{n^2 N^2} \sum_{\ell, \ell' = 0}^{p} \sum_{i=1}^{n} \sum_{j=1}^{N} \sum_{j'=1}^{N} X_{i\ell} X_{i\ell'} a_{\ell} \tilde{B}(z_j) \tilde{B}(z_{j'})^\top a_{\ell'} G_n(z_j, z_{j'}) \]

\[ = \frac{1}{n^2} \sum_{k=1}^{\infty} \sum_{i=1}^{n} \left\{ \frac{1}{N} \sum_{j=1}^{N} \sum_{j'=1}^{N} \lambda_j \frac{1}{2} X_{i\ell} g_{\ell}(z_j) \psi_k(z_j) \right\}^2 \]

with \( g_\ell(z) = a_\ell \tilde{B}(z) \). Therefore, by Assumption (A3), we have

\[ Ea^\top \Psi_n a = \frac{1}{n} \sum_{k=1}^{\infty} \sum_{\ell=0}^{p} \left\{ \frac{1}{N} \sum_{j=1}^{N} \lambda_j \frac{1}{2} g_{\ell}(z_j) \psi_k(z_j) \right\}^2 \]

\[ \geq \frac{c}{nN^2} \sum_{\ell=0}^{p} \sum_{j=1}^{N} \sum_{j'=1}^{N} g_{\ell}(z_j) g_{\ell}(z_{j'}) G_n(z_j, z_{j'}) \]

\[ \simeq \frac{1}{n} \sum_{\ell=0}^{p} \int_{\Omega^2} g_{\ell}(z) g_{\ell}(z') G_n(z, z') dz dz'. \]

Noting that the eigenvalues of \( G_n \) are strictly positive, we have

\[ Ea^\top \Psi_n a \geq c_1 n^{-1} \sum_{\ell=0}^{p} \int_{\Omega} g_{\ell}^2(z) dz \geq c_2 n^{-1} |\Delta|^2 \|a\|^2. \]

Therefore, we have \( Ea^\top \Gamma^{-1}_{n, \rho} \Psi_n \Gamma^{-1}_{n, \rho} a \geq cn^{-1} \left( 1 + \frac{\rho_n}{n N |\Delta|^4} \right)^{-2} |\Delta|^{-2} \). Similarly, one can show that \( Ea^\top \Gamma^{-1}_{n, \rho} \Psi_n \Gamma^{-1}_{n, \rho} a \geq c(nN)^{-1} \left( 1 + \frac{\rho_n}{n N |\Delta|^4} \right)^{-2} |\Delta|^{-2} \). In addition,

\[ \max(a_i^n)^2 \leq \frac{C}{|\Delta|^4} \frac{1}{n^2 N^2} \sum_{j=1}^{N} \sum_{j'=1}^{N} \left\{ (\tilde{X}_j, \tilde{X}_{j'}) \otimes \tilde{B}(z_j) \tilde{B}(z_{j'})^\top \right\} G_n(z_j, z_{j'}) a \]

\[ \leq \frac{C}{|\Delta|^4} \sum_{\ell, \ell'}^{p} \frac{1}{n^2 N^2} \max_i |X_{i\ell} X_{i\ell'}| \sum_{j=1}^{N} \sum_{j'=1}^{N} g_{\ell}(z_j) g_{\ell'}(z_{j'}) G_n(z_j, z_{j'}) \leq C n^{-2} |\Delta|^{-2}, \]

\[ \max(a_i^\ell)^2 \leq \frac{C}{|\Delta|^4} \frac{1}{n^2 N^2} \sum_{j=1}^{N} \left\{ (\tilde{X}_j, \tilde{X}_j^\top) \otimes \tilde{B}(z_j) \tilde{B}(z_j)^\top \right\} \sigma^2(z_j) a \]

\[ \leq \frac{C}{|\Delta|^4} \sum_{\ell, \ell'}^{p} \frac{1}{n^2 N^2} \max_i |X_{i\ell} X_{i\ell'}| \sum_{j=1}^{N} g_{\ell}(z_j) g_{\ell'}(z_{j'}) \sigma^2(z_j) \leq C n^{-2} N^{-1} |\Delta|^{-2}. \]

Thus, if \( \rho_n n^{-1} N^{-1} |\Delta|^{-4} \to 0 \), we have

\[ \frac{\max_{1 \leq i \leq n}(a_i^n + a_i^\ell)^2}{\sum_{i=1}^{n}(a_i^n + a_i^\ell)^2} \leq C n^{-1} \left( 1 + \frac{\rho_n}{n N |\Delta|^4} \right)^2 \to 0, \]
which satisfies the Lindeberg condition.

**Theorem S2.** Under Assumptions (A1)–(A6), if for any \( \ell = 0, 1, \ldots, p, |X_\ell| \leq C_\ell < \infty \), \( \sup_{z \in \Omega} [\text{Var}(\hat{\beta}_\ell(z))]^{-1/2}(\hat{\beta}_{\mu,\ell}(z) - \beta_\ell^\circ(z)) = o_p(1) \), for \( \ell = 0, \ldots, p \).

**Proof.** Using similar arguments as in the proof of Lemma S13 and the result of Lemma S9, we have for any \( \|a\| = 1 \),

\[
Ea^\top \Gamma_{n,\rho}^{-1} \Psi_{n,\rho}^{-1} \Gamma_{n,\rho}^{-1} a \leq C^{-2}_T |\Delta|^{-4} Ea^\top \Psi_{n,\rho}^{-1} a \leq Cn^{-1}|\Delta|^{-2},
\]

and \( \sup G_{n,\rho}^{-1} |\Delta|^{-2} \). Therefore, based on the proof of Lemma S13 for any \( \|a\| = 1 \),

\[
cn^{-1}|\Delta|^{-4} \left(1 + \frac{\rho_n}{nN|\Delta|^4}\right)^{-2} \leq Ea^\top \Gamma_{n,\rho}^{-1} \Psi_{n,\rho}^{-1} \Gamma_{n,\rho}^{-1} a \leq Cn^{-1}|\Delta|^{-2},
\]

\[
c(nN)^{-1} \left(1 + \frac{\rho_n}{nN|\Delta|^4}\right)^{-2} |\Delta|^{-2} \leq Ea^\top \Gamma_{n,\rho}^{-1} \Psi_{n,\rho}^{-1} \Gamma_{n,\rho}^{-1} a \leq C(nN)^{-1}|\Delta|^{-2}.
\]

Thus,

\[
\text{Var}(\hat{\beta}_\ell) = \{e_\ell \otimes \tilde{B}(z)\}^\top \text{E}\{\Gamma_{n,\rho}^{-1} (\Psi_{n,\rho} + \Psi_{n,\rho}^{-1}) \Gamma_{n,\rho}^{-1}\} \{e_\ell \otimes \tilde{B}(z)\}
\]

\[
\times \{e_\ell \otimes \tilde{B}(z)\}^\top \text{E}(\Gamma_{n,\rho}^{-1} \Psi_{n,\rho}^{-1}) \{e_\ell \otimes \tilde{B}(z)\} \times \{e_\ell \otimes \tilde{B}(z)\}^\top
\]

\[
\times E \left[ \Gamma_{n,\rho}^{-1} \frac{1}{n^2 N^2} \sum_{i=1}^n \sum_{j=1}^N \left\{ \tilde{X}_i \otimes \tilde{B}(z_j) \right\} \left\{ \tilde{X}_i \otimes \tilde{B}(z_{j'}) \right\}^\top G_{n,\rho}^{-1} \{e_\ell \otimes \tilde{B}(z)\}. \right.
\]

By Lemma S9, we have

\[
\text{Var}(\hat{\beta}_\ell) \leq \frac{1}{nN^2|\Delta|^2} \sum_{j=1}^N \sum_{j'=1}^N \{e_\ell \otimes \tilde{B}(z)\}^\top \tilde{B}(z_j) \tilde{B}(z_{j'})^\top \{e_\ell \otimes \tilde{B}(z)\} G_{n,\rho}^{-1}(z_j, z_{j'}),
\]

\[
\text{Var}(\hat{\beta}_\ell) \geq \frac{1}{nN^2|\Delta|^2} \sum_{j=1}^N \sum_{j'=1}^N \{e_\ell \otimes \tilde{B}(z)\}^\top \tilde{B}(z_j) \tilde{B}(z_{j'})^\top \{e_\ell \otimes \tilde{B}(z)\} G_{n,\rho}^{-1}(z_j, z_{j'})
\]

\[
\times \left(1 + \frac{\rho_n}{nN|\Delta|^4}\right)^{-2},
\]

and according to Lemmas S1 and S4, we have \( cn^{-1} \left(1 + \frac{\rho_n}{nN|\Delta|^4}\right)^{-2} \leq \text{Var}(\hat{\beta}_\ell) \leq Cn^{-1}. \)
According to Lemma S10 if \( \rho_n n^{-1/2} N^{-1} |\triangle|^{-3} \to 0 \) and \( n^{1/2} |\triangle|^{d+1} \to 0 \), the bias term in (S19) is negligible compared to the order of \([\text{Var}\{\hat{\beta}_\ell(z)\}]^{1/2}\).

Proof of Theorem 2. Theorem 2 follows from (S19), Lemma S13 and Theorem S2.

\[ \square \]

### S1.4 Asymptotic Properties of Piecewise Constant Spline Estimators

In this section, we study the asymptotic properties of the piecewise constant spline estimators defined in the spline space \( \mathcal{PC}(\triangle) \). Define piecewise constant bivariate spline functions

\[
\hat{\beta}_c^c(z) = (\hat{\beta}_{c,0}^c(z), \ldots, \hat{\beta}_{c,p}^c(z))^\top = \hat{V}_{m(z)}^{-1} \left\{ \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N B_m(z_j) X_{it} \sum_{\ell=0}^p \beta_\ell^o(z_j) X_{i\ell'} \right\}_{\ell=0}^p,
\]

\[
\hat{\eta}(z) = (\hat{\eta}_0(z), \ldots, \hat{\eta}_p(z))^\top = \hat{V}_{m(z)}^{-1} \left\{ \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N B_m(z_j) X_{it} \sum_{k=1}^{\infty} \xi_{ik} \psi_k(z_j) \right\}_{\ell=0}^p,
\]

\[
\hat{\varepsilon}(z) = (\hat{\varepsilon}_0(z), \ldots, \hat{\varepsilon}_p(z))^\top = \hat{V}_{m(z)}^{-1} \left\{ \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N B_m(z_j) X_{it} \hat{\varepsilon}_{ij} \right\}_{\ell=0}^p,
\]

where \( \hat{V}_{m(z)} \) is defined in (2.7).

The next two theorems concern the functions \( \hat{\beta}_{c,\ell}^c(z) = (\hat{\beta}_{c,\ell,0}^c(z), \ldots, \hat{\beta}_{c,\ell,p}^c(z))^\top \), \( \hat{\eta}_\ell(z) \), \( \hat{\varepsilon}_\ell(z) \), \( \ell = 0, \ldots, p \), given in (S27), (S28) and (S29). Theorem S3 gives the uniform convergence rate of \( \hat{\beta}_{c,\ell}^c(z) \) to \( \beta_\ell^o(z) \).

**Theorem S3.** Under Assumptions (A1'), (A2)–(A6), the constant spline functions \( \hat{\beta}_{c,\ell}^c(z), \ell = 0, \ldots, p \), satisfy \( \sup_{z \in \Omega} \sup_{0 \leq \ell \leq p} \left| \hat{\beta}_{c,\ell}^c(z) - \beta_\ell^o(z) \right| = O_P(|\triangle|) \).
In the following, we provide detailed proofs of Theorems S3. For the random matrix \( \tilde{V}_m \) defined in (2.7), the lemma below shows that its inverse can be approximated by the inverse of a deterministic matrix \( A_m^{-1}\Sigma_X^{-1} \), where \( A_m = \int_\Omega B_m(z)dz \).

**Lemma S14.** Under Assumptions (A3) and (A5), for any \( m \in \mathcal{M} \), we have

\[
\tilde{V}_m^{-1} = A_m^{-1}\Sigma_X^{-1} + O_P \{ n^{-1/2}|\Delta|^2(\log n)^{1/2} + N^{-1/2}|\Delta| \} .
\] (S30)

**Proof.** By Lemma S5 \( \| \tilde{V}_m - A_m\Sigma_X \|_\infty = O_P \{ n^{-1/2}|\Delta|^2(\log n)^{1/2} + N^{-1/2}|\Delta| \} \). Using the fact that for any matrices \( A \) and \( B \), \( (A + \delta B)^{-1} = A^{-1} - \delta A^{-1}BA^{-1} + O(\delta^2) \), we obtain (S30).

**Proof of Theorem S3.** According to Lemma S2, there exist functions \( \beta^*_\ell \in \mathcal{PC}(\Delta) \) that satisfies \( \| \beta^*_\ell - \beta_\ell \|_\infty = O(|\Delta|) \) for \( \ell = 0, 1, \ldots, p \). By the definition of \( \tilde{\beta}_\mu(z) \) in (S27), \( \tilde{\beta}_\mu(z) = (\tilde{\beta}_{\mu,0}(z), \tilde{\beta}_{\mu,1}(z), \ldots, \tilde{\beta}_{\mu,p}(z))^\top = (\tilde{\gamma}_m(z), \ldots, \tilde{\gamma}_m(z), p)^\top = \tilde{\gamma}_m(z) \), where \( \tilde{\gamma}_m = \tilde{V}_m^{-1} \left\{ (nN)^{-1} \sum_{i=1}^n \sum_{j=1}^N B_m(z_j)X_{i\ell} \sum_{\ell' = 0}^p \beta^*_\ell(z_j)X_{i\ell'} \right\}_{\ell=0}^p \) for \( \tilde{V}_m \) defined in (2.7).

Let

\[
\tilde{\beta}(z) = (\tilde{\beta}_0(z), \tilde{\beta}_1(z), \ldots, \tilde{\beta}_p(z))^\top = \tilde{V}_m^{-1}(z) \left[ \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N B_m(z_j)X_{i\ell} \sum_{\ell' = 0}^p \beta^*_\ell(z_j)X_{i\ell'} \right]_{\ell=0}^p ,
\]

then

\[
\tilde{\beta}^\varepsilon_\mu(z) - \tilde{\beta}(z) = \tilde{V}_m^{-1}(z) \left[ \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N B_m(z_j)X_{i\ell} \sum_{\ell' = 0}^p \{ \beta^\varepsilon_\ell(z_j) - \beta^*_\ell(z_j) \} X_{i\ell'} \right]_{\ell=0}^p .
\]

Observing that \( \tilde{\beta}_\ell \equiv \beta^*_\ell \in \mathcal{PC}(\Delta), \tilde{\beta}^\varepsilon_\mu(z) = \tilde{\beta}^\varepsilon_\mu(z) - \beta^*_\ell(z) + \beta^*_\ell(z) \), \( \ell = 0, 1, \ldots, p \).

It is easy to see \( \| \tilde{\beta}^\varepsilon_\mu - \tilde{\beta}_\ell \|_\infty = O_P(|\Delta|) \). Hence, for \( \ell = 0, 1, \ldots, p, \| \tilde{\beta}^\varepsilon_\mu - \beta^*_\ell \|_\infty \leq \| \tilde{\beta}^\varepsilon_\mu - \tilde{\beta}_\ell \|_\infty + \| \beta^*_\ell - \beta^*_\ell \|_\infty = O_P(|\Delta|) \), which completes the proof. \( \square \)
By Lemma S14, the inverse of the random matrix \( \hat{V}_m \) can be approximated by that of a deterministic matrix \( A_m \Sigma_X \). Substituting \( \hat{V}_m \) with \( A_m \Sigma_X \) in (S28) and (S29), we define the random vectors

\[
\hat{\eta}^*(z) = (\hat{\eta}_0^*(z), \ldots, \hat{\eta}_p^*(z))^\top = A_m^{-1} \Sigma_X^{-1} \left\{ \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} B_m(z_j) X_{i\ell} \sum_{k=1}^{\infty} \xi_{ik}\psi_{k}(z_j) \right\}_{\ell=0}^p.
\]

(S31)

\[
\hat{\epsilon}^*(z) = (\hat{\epsilon}_0^*(z), \ldots, \hat{\epsilon}_p^*(z))^\top = A_m^{-1} \Sigma_X^{-1} \left\{ \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} B_m(z_j) X_{i\ell}\epsilon_{ij} \right\}_{\ell=0}^p.
\]

(S32)

The next lemma implies that the difference between \( \hat{\eta}^*(z) \) and \( \hat{\eta}(z) \) and the difference between \( \hat{\epsilon}^*(z) \) and \( \hat{\epsilon}(z) \) are both negligible uniformly over \( z \in \Omega \).

**Lemma S15.** Under Assumptions (A2)–(A5) and (C1), if \( N^{1/2}|\Delta| \to \infty \) as \( N \to \infty \), \( \|\sum_{k=1}^{\infty} \lambda_k^{1/2} \psi_k \|_\infty < \infty \) and \( n^{1/(4+\delta_2)} \ll n^{1/2} N^{-1/2}|\Delta|^{-1} \) for some \( \delta_2 \), then for \( \hat{\eta}(z) \), \( \hat{\epsilon}(z) \) given in (S28), (S29) and \( \hat{\eta}^*(z) \), \( \hat{\epsilon}^*(z) \) given in (S31), (S32), as \( N \to \infty \) and \( n \to \infty \), we have

\[
\sup_{z \in \Omega} \left\| \hat{\eta}(z) - \hat{\eta}^*(z) \right\|_\infty = O_P \left\{ n^{-1}|\Delta| \frac{4 \log(n)}{n} + n^{-1/2} N^{-1/2}|\Delta|^{3}(\log n)^{1/2} \right\}, \quad \text{(S33)}
\]

\[
\sup_{z \in \Omega} \left\| \hat{\epsilon}(z) - \hat{\epsilon}^*(z) \right\|_\infty = O_P \left\{ n^{-1} N^{-1/2}|\Delta|^{3} \frac{\log(n)}{n} + n^{-1/2} N^{-1}|\Delta|^{2}(\log n)^{1/2} \right\}. \quad \text{(S34)}
\]

**Proof.** Comparing \( \hat{\eta}(z) \) and \( \hat{\eta}^*(z) \) given in (S28) and (S31), we have

\[
\hat{\eta}(z) - \hat{\eta}^*(z) = \left\{ \hat{V}_m^{-1} - A_m^{-1} \Sigma_X^{-1} \right\} \left\{ \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} B_m(z_j) X_{i\ell} \sum_{k=1}^{\infty} \xi_{ik}\psi_{k}(z_j) \right\}_{\ell=0}^p.
\]

Now let \( \zeta_{i,m,\ell} \equiv \zeta_{i} = n^{-1} \left[ X_{i\ell} \sum_{k=1}^{\infty} \left\{ \frac{1}{N} \sum_{j=1}^{N} B_m(z_j)\psi_{k}(z_j) \right\} \xi_{ik} \right] \), then it is easy to see that \( \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} B_m(z_j) X_{i\ell} \sum_{k=1}^{\infty} \xi_{ik}\psi_{k}(z_j) = \frac{1}{N} \sum_{i=1}^{n} \zeta_{i,m,\ell} \). It is easy to see that
\[ E(\zeta_i) = 0, \text{ and} \]
\[ \sigma_{\zeta_i}^2 = E(\zeta_i^2) = n^{-2}E(X_i^2) \int_{T_m \times T_m} G_\eta(u, v) \, du dv \{1 + O(N^{-1/2}|\Delta|^{-1})\}. \]

Note that \( \{\sigma_{\zeta_i}^{-1}\zeta_i\}_{i=1}^n \) are uncorrelated random variables with mean 0. Assume that \(|\Delta|^{-2} \asymp n^\tau\) for some \(0 < \tau < \infty\), we can show that for any large enough \(\delta > 0\),
\[ P \left[ \sum_{i=1}^n \zeta_i \geq \delta \{ C \log(n)n^{-1}|\Delta|^{4}E(X_i^2) \}^{1/2} \right] \leq 2n^{-2-\tau}. \]
Therefore,
\[ \sum_{n=1}^{\infty} P \left\{ \sup_{m \in M, 0 \leq \ell \leq p} \left| \sum_{i=1}^n \zeta_{i,m,\ell} \right| \geq \delta n^{-1/2} |\Delta|^{2}(\log n)^{1/2} \right\} < \infty. \]
Thus, \( \sup_{m,\ell} \left| \sum_{i=1}^n \zeta_{i,m,\ell} \right| = O_P \{ n^{-1/2}\Delta^{2}(\log n)^{1/2} \} \) as \(n \to \infty\) by Borel-Cantelli Lemma. It follows that \( \sup_{m,\ell} n^{-1} \sum_{i=1}^n \zeta_{i,m,\ell} = O_P \{ n^{-1/2}\Delta^{2}(\log n)^{1/2} \} \). Finally, according to (S30), we obtain (S33). The result in (S34) can be proved similarly. \qed

**Lemma S16.** For any \(z \in \Omega\), the covariance matrices of \(\hat{\eta}^*(z)\) and \(\hat{\varepsilon}^*(z)\) are
\[ \Sigma_\eta(z) = E \left\{ \hat{\eta}^*(z)\hat{\eta}^{\ast T}(z) \right\} = A_{m(z)}^{-2} \Sigma_X^{-1} \frac{1}{nN^2} \sum_{k=1}^{\infty} \lambda_k \left\{ \sum_{j=1}^{N} B_{m(z)}(z_j) \psi_k(z_j) \right\}^2, \]
\[ \Sigma_\varepsilon(z) = E \left\{ \hat{\varepsilon}^*(z)\hat{\varepsilon}^{\ast T}(z) \right\} = A_{m(z)}^{-2} \Sigma_X^{-1} \frac{1}{nN^2} \sum_{j=1}^{N} B_{m(z)}^2(z_j) \sigma^2(z_j), \]
in addition,
\[ \sup_{z \in \Omega} \| \Sigma_\eta(z) + \Sigma_\varepsilon(z) - \Sigma_n(z) \|_\infty = O(n^{-1}N^{-1/2}|\Delta|^{-1}), \quad (S35) \]
where \( \Sigma_n(z) \) is given in (2.9).

**Proof.** Note that \(A_{m(z)}^2 \hat{\eta}^*(z)\hat{\eta}^{\ast T}(z)\) is equal to
\[ \Sigma_X^{-1} \left\{ \frac{1}{n^2N^2} \sum_{i=1}^{n} \sum_{j=1}^{N} B_{m(z)}(z_j) X_{it} \sum_{k=1}^{\infty} \xi_{ik} \psi_k(z_j) \sum_{i'=1}^{n} \sum_{j'=1}^{N} B_{m(z)}(z_{j'}) X_{i't'} \sum_{k'=1}^{\infty} \xi_{i'k'} \psi_{k'}(z_{j'}) \right\} \sum_{\ell,\ell'=0}^{\infty} \Sigma_X^{-1}. \]

28
Thus,
\[ \Sigma_\eta(z) = E \left\{ \tilde{\eta}^*(z) \tilde{\eta}^T(z) \right\} = A_{m(z)}^{-2} \Sigma_X^{-1} \sum_{k=1}^{\infty} \frac{1}{nN^2} \lambda_k \left\{ \sum_{j=1}^{N} B_{m(z)}(z_j) \psi_k(z_j) \right\}^2. \]

Similarly, we can derive the covariance of \( \tilde{\varepsilon}^*(z) \): \( \Sigma_\varepsilon(z) = A_{m(z)}^{-2} \Sigma_X^{-1} \sum_{j=1}^{N} B_{m(z)}(z_j) \sigma^2(z_j). \)

Observe that
\[ \sum_{k=1}^{\infty} \lambda_k \left\{ \sum_{j=1}^{N} B_{m(z)}(z_j) \psi_k(z_j) \right\}^2 = \frac{1}{n^2} \sum_{j=1}^{N} \sum_{j'=1}^{N} G_\eta(z_j, z_{j'}) B_{m(z)}(z_j) B_{m(z)}(z_{j'}). \]

Hence, by (S5) and (S6) in Lemma S4, (S35) holds. Therefore,
\[ \Sigma_\eta(z) + \Sigma_\varepsilon(z) = \left( nA_{m(z)}^{-2} - 1 \right) \Sigma_X^{-1} \int_{T_{mn(z)}} G_\eta(u, v) \, dudv \{ 1 + O(N^{-1/2} |\Delta|^{-1}) \} \]
\[ + \left( nN A_{m(z)}^{-2} - 1 \right) \Sigma_X^{-1} \int_{T_{mn(z)}} \sigma^2(u) \, du \{ 1 + O(N^{-1/2} |\Delta|^{-1}) \} \]
\[ = n^{-1} \Sigma_X^{-1} G_\eta(z, z) \{ 1 + O(N^{-1/2} |\Delta|^{-1}) \}. \]

Therefore, \( \sup_{z \in \Omega} \left\| \Sigma_\eta(z) + \Sigma_\varepsilon(z) - n^{-1} \Sigma_X^{-1} G_\eta(z, z) \right\|_\infty = O(n^{-1} N^{-1/2} |\Delta|^{-1}). \) The desired result in (S35) follows.

**Proof of Theorem 3.** Note that, for any vector \( a = (a_0, \ldots, a_p)^T \in \mathcal{R}^{(p+1)}, \) we have
\[ E \left[ \sum_{\ell=0}^{p} a_\ell \{ \tilde{\eta}^*_\ell(z) + \tilde{\varepsilon}^*_\ell(z) \} \right] = 0, \]
and
\[ \sum_{\ell=0}^{p} a_\ell \tilde{\eta}^*_\ell(z) = a^T A_{m(z)}^{-1} \Sigma_X^{-1} \sum_{i=1}^{n} \sum_{j=1}^{N} B_{m(z)}(z_j) \sum_{k=1}^{\infty} \xi_{ijk} \psi_k(z_j) X_i = \sum_{i=1}^{n} a^T A_{m(z)}^{-1} \Sigma_X^{-1} \delta^\eta_i, \]
\[ \sum_{\ell=0}^{p} a_\ell \tilde{\varepsilon}^*_\ell(z) = a^T A_{m(z)}^{-1} \Sigma_X^{-1} \sum_{i=1}^{n} \sum_{j=1}^{N} B_{m(z)}(z_j) \varepsilon_{ij} X_i = \sum_{i=1}^{n} a^T A_{m(z)}^{-1} \Sigma_X^{-1} \delta^\varepsilon_i, \]
where \( \delta^\eta_i = \frac{1}{nN} \sum_{j=1}^{N} B_{m(z)}(z_j) \sum_{k=1}^{\infty} \xi_{ijk} \psi_k(z_j) X_i \) and \( \delta^\varepsilon_i = \frac{1}{nN} \sum_{j=1}^{N} B_{m(z)}(z_j) \varepsilon_{ij} X_i \) are independent sequences with variances \( \text{Var}(\delta^\eta_i) = \frac{1}{n^2N^2} \sum_{j,j'} B_{m(z)}(z_j) B_{m(z)}(z_{j'}) G_\eta(z_j, z_{j'}) \Sigma_X \)
and \( \text{Var}(\delta^\varepsilon_i) = \frac{1}{n^2N^2} \sum_{j=1}^{N} B_{m(z)}(z_j) B_{m(z)}(z_j) \sigma^2(z_j) \Sigma_X, \) respectively. Therefore, we
have

\[
\text{Var}\left( a^\top A_{m(z)}^{-1} \Sigma_X^{-1} \mathbf{1}_n^\top \right) = \frac{1}{n} a^\top \Sigma_n(z) a = \frac{A_{m(z)}^{-2}}{n^2 N^2} \sum_{j=1}^N B_m(z_j) B_m(z_{j'}) G_\eta(z_j, z_{j'}) a^\top \Sigma_X^{-1} a,
\]

\[
\text{Var}\left( a^\top A_{m(z)}^{-1} \Sigma_X^{-1} \mathbf{1}_n^\top \right) = \frac{1}{n} a^\top \Sigma_\epsilon(z) a = \frac{A_{m(z)}^{-2}}{n^2 N^2} \sum_{j=1}^N B_m(z_j) B_m(z_j) \sigma^2(z_j) a^\top \Sigma_X^{-1} a.
\]

Using central limit theorem, we have

\[
[a^\top \{ \Sigma_n(z) + \Sigma_\epsilon(z) \} a]^{-1/2} \sum_{\ell=0}^p a_\ell \{ \hat{\eta}_\ell(z) + \hat{\epsilon}_\ell(z) \} \xrightarrow{\mathcal{L}} N(0, 1).
\]

By (S35), as \( N \to \infty \) and \( n \to \infty \), \( \{ a^\top \Sigma_n(z) a\}^{-1/2} \sum_{\ell=0}^p a_\ell \{ \hat{\eta}_\ell(z) + \hat{\epsilon}_\ell(z) \} \xrightarrow{\mathcal{L}} N(0, 1) \). Therefore, \( \{ a^\top \Sigma_n(z) a\}^{-1/2} \sum_{\ell=0}^p a_\ell \{ \hat{\beta}_\ell(z) - \beta_\ell^0(z) \} \xrightarrow{\mathcal{L}} N(0, 1) \) follows from (S18), Theorem S3, Lemma S15 and Slutsky’s Theorem. Applying Cramér-Wold’s device, we obtain \( \Sigma_n^{-1/2}(z) \{ \hat{\beta}_\ell(z) - \beta_\ell^0(z) \}^p_{\ell=0} \xrightarrow{\mathcal{L}} N \left( 0, I_{(p+1)\times(p+1)} \right) \), as \( N \to \infty \) and \( n \to \infty \), and consequently, \( \sigma_{n,\ell}^{-1}(z) \{ \hat{\beta}_\ell(z) - \beta_\ell^0(z) \} \xrightarrow{\mathcal{L}} N(0,1) \), for any \( z \in \Omega \) and \( \ell = 0, \ldots, p \).

\[ \square \]

S1.5 Convergence of the Covariance Estimator

For any \( i = 1, \ldots, n \), and estimated residuals \( \hat{R}_{ij} = Y_{ij} - \sum_{\ell=0}^p X_{i\ell} \hat{\beta}_\ell(z_j) \), denote \( \hat{\mathbf{\theta}}_i = \arg\min_{\mathbf{\theta}} \sum_{j=1}^N \left\{ \hat{R}_{ij} - B_\eta(z_j) Q_{\eta,2} \mathbf{\theta} \right\}^2 \), where \( B_\eta(z) \) is the set of bivariate spline basis functions used to estimate \( \eta_i(z) \), and \( Q_{\eta,2} \) is given in the following QR decomposition of the transpose of the smoothness matrix \( H_\eta^\top \): \( H_\eta^\top = Q_\eta R_\eta = (Q_{\eta,1} Q_{\eta,2}) (R_{\eta,1}^{(R_{\eta,2})}) \). Then, the bivariate spline estimator of \( \eta_i(z) \) can be written as \( \hat{\eta}_i(z) = B_\eta(z)^\top Q_{\eta,2} \hat{\mathbf{\theta}}_i = \widetilde{\eta}(z)^\top \hat{\mathbf{\theta}}_i \). Let

\[
\mathbf{Y}_n = \frac{1}{N} \sum_{j=1}^N \widetilde{\mathbf{B}}_\eta(z_j) \widetilde{\mathbf{B}}_\eta^\top(z_j),
\]

\[ 30 \]
then we have

\[
\hat{\vartheta}_i = \Upsilon_n^{-1} \frac{1}{N} \sum_{j=1}^{N} \tilde{B}_\eta(z_j) \hat{R}_{ij} \\
= \Upsilon_n^{-1} \frac{1}{N} \sum_{j=1}^{N} \tilde{B}_\eta(z_j) \left[ \sum_{\ell=0}^{p} X_{i\ell} \{ \beta_\ell^o(z_j) - \hat{\beta}_\ell(z_j) \} + \eta(z_j) + \sigma(z_j) \varepsilon_{ij} \right].
\]  

(S36)

Lemma S17. Under Assumptions (A3)–(A5), if 

\( (N^{1/2} |\Delta_\eta|) / \log(|\Delta_\eta|^{-1}) \to \infty \) as 

\( N \to \infty \), then there exist constants 

\( 0 < c_T < C_T < \infty \), such that with probability 

approaching 1 as 

\( N \to \infty, n \to \infty \), 

\( c_T |\Delta_\eta|^2 \leq \lambda_{\min}(Y_n) \leq \lambda_{\max}(Y_n) \leq C_T |\Delta_\eta|^2 \).

The proof is similar to the proof of S7, thus omitted.

Next we define

\[
\tilde{b}_i(z) = \tilde{B}_\eta(z)^\top \Upsilon_n^{-1} \frac{1}{N} \sum_{j=1}^{N} \tilde{B}_\eta(z_j) \sum_{\ell=0}^{p} X_{i\ell} \{ \beta_\ell^o(z_j) - \hat{\beta}_\ell(z_j) \};
\]  

(S37)

\[
\tilde{\eta}_i(z) = \tilde{B}_\eta(z)^\top \Upsilon_n^{-1} \frac{1}{N} \sum_{j=1}^{N} \tilde{B}_\eta(z_j) \eta(z_j), \tilde{\varepsilon}_i(z) = \tilde{B}_\eta(z)^\top \Upsilon_n^{-1} \frac{1}{N} \sum_{j=1}^{N} \tilde{B}_\eta(z_j) \sigma(z_j) \varepsilon_{ij}.
\]

Then, the estimation error 

\( D_i(z) = \tilde{\eta}_i(z) - \eta_i(z) \) in (3.1) can be decomposed as the following:

\[
D_i(z) = \tilde{b}_i(z) + \nabla \eta_i(z) + \tilde{\varepsilon}_i(z).
\]

For any \( z, z' \in \Omega \), denote

\[
\tilde{G}_\eta(z, z') = n^{-1} \sum_{i=1}^{n} \eta_i(z) \eta_i(z').
\]

The following lemma shows the uniform convergence of \( \tilde{G}_\eta(z, z') \) to \( G_\eta(z, z') \) in probability over all \( (z, z') \in \Omega^2 \).

Lemma S18. Under Assumptions (A1)–(A5) and (C1)–(C3), \( \sup_{(z,z') \in \Omega^2} |\tilde{G}_\eta(z, z') - G_\eta(z, z')| = O_P\{n^{-1/2}(\log n)^{1/2}\} \).
Proof. Let \( \bar{\xi}_{kk'} = n^{-1} \sum_{i=1}^{n} \xi_{ik} \xi_{ik'} \), then

\[
\bar{G}_{\eta}(z, z') - G_{\eta}(z, z') = \sum_{k=1}^{\infty} \lambda_k \psi_k(z) \psi_k(z') (\bar{\xi}_{kk} - 1) + \sum_{k \neq k'} \bar{\xi}_{kk'} (\lambda_k \lambda_{k'})^{1/2} \psi_k(z) \psi_{k'}(z').
\]

As \( E \left[ \sum_{k=1}^{\infty} \lambda_k \psi_k(z) \psi_k(z') (\bar{\xi}_{kk} - 1) \right] = 0 \), then \( E \{ \bar{G}_{\eta}(z, z') - G_{\eta}(z, z') \} = 0. \) Note that \( E \{ \eta^2(z) \eta^2(z') \} = G_{\eta}(z, z) G_{\eta}(z', z') + 2 G_{\eta}^2(z, z') + \sum_{k=1}^{\infty} \lambda_k^2 E (\xi_{1k} - 3) \psi_k^2(z) \psi_k^2(z'). \)

Next,

\[
E \left\{ \bar{G}_{\eta}(z, z') - G_{\eta}(z, z') \right\}^2 = E \left\{ \frac{1}{n} \sum_{i=1}^{n} \eta_i(z) \eta_i(z') - G_{\eta}(z, z') \right\}^2
\]

\[
= \frac{1}{n} \left\{ G_{\eta}(z, z) G_{\eta}(z', z') + G_{\eta}^2(z, z') + \sum_{k=1}^{\infty} \lambda_k^2 E (\xi_{1k} - 3) \psi_k^2(z) \psi_k^2(z') \right\}.
\]

Therefore, \( E \left\{ \bar{G}_{\eta}(z, z') - G_{\eta}(z, z') \right\}^2 \propto n^{-1}. \) Hence, following from Bernstein inequality, \( \sup_{(z, z') \in \Omega^2} \left| \bar{G}_{\eta}(z, z') - G_{\eta}(z, z') \right| = O_P \{ n^{-1/2} (\log n)^{1/2} \}, \) and the desired result follows. \( \square \)

Proof of Theorem 4. Note that

\[
\sup_{(z, z') \in \Omega^2} \left| \bar{G}_{\eta}(z, z') - G_{\eta}(z, z') \right| \leq \sup_{(z, z') \in \Omega^2} \{ \left| \bar{G}_{\eta}(z, z') - \bar{G}_{\eta}(z, z') \right| + \left| \bar{G}_{\eta}(z, z') - G_{\eta}(z, z') \right| \},
\]

where \( \sup_{(z, z') \in \Omega^2} \left| \bar{G}_{\eta}(z, z') - G_{\eta}(z, z') \right| = o_P(1) \) according to Lemma S18 and

\[
\sup_{(z, z') \in \Omega^2} \left| \bar{G}_{\eta}(z, z') - \bar{G}_{\eta}(z, z') \right| \leq \sup_{(z, z') \in \Omega^2} \left| n^{-1} \sum_{i=1}^{n} \eta_i(z) D_i(z') \right|
\]

\[
+ \sup_{(z, z') \in \Omega^2} \left| n^{-1} \sum_{i=1}^{n} \eta_i(z') D_i(z) \right| + \sup_{(z, z') \in \Omega^2} \left| n^{-1} \sum_{i=1}^{n} D_i(z) D_i(z') \right|.
\]

With some simple calculations, we have

\[
\sum_{i=1}^{n} \eta_i(z) D_i(z') = \sum_{i=1}^{n} \eta_i(z) \tilde{b}_i(z') + \sum_{i=1}^{n} \eta_i(z) \nabla \eta_i(z') + \sum_{i=1}^{n} \eta_i(z) \tilde{\epsilon}_i(z'),
\]

\[32\]
where $\nabla \eta_i = \tilde{\eta}_i - \eta_i$. According to (S39), (S42) and (S47), we have

$$
\sup_{(z, z') \in \Omega} \left| n^{-1} \sum_{i=1}^{n} \eta_i(z) D_i(z') + n^{-1} \sum_{i=1}^{n} \eta_i(z') D_i(z) \right| = o_P(1).
$$

Note that

$$
\sum_{i=1}^{n} D_i(z) D_i(z') = \sum_{i=1}^{n} \tilde{b}_i(z) \tilde{b}_i(z') + \sum_{i=1}^{n} \nabla \eta_i(z) \nabla \eta_i(z') + \sum_{i=1}^{n} \tilde{b}_i(z) \nabla \eta_i(z') + \sum_{i=1}^{n} \tilde{e}_i(z) \tilde{e}_i(z')
$$

$$
+ \sum_{i=1}^{n} \nabla \eta_i(z) \tilde{e}_i(z') + \sum_{i=1}^{n} \tilde{b}_i(z) \tilde{e}_i(z').
$$

It follows from (S38), (S41), (S43)–(S46) that $\sup_{(z, z') \in \Omega} |n^{-1} \sum_{i=1}^{n} D_i(z) D_i(z')| = o_P(1)$. The desired result is established.

**Lemma S19.** Under Assumptions (A1)–(A5), (C1)–(C3), we have

$$
\sup_{(z, z') \in \Omega} \left| n^{-1} \sum_{i=1}^{n} \tilde{b}_i(z) \tilde{b}_i(z') \right| = O_P \left\{ n^{-1} |\Delta_\eta|^{-2} (\log n)^{1/2} \right\}, \quad \text{(S38)}
$$

$$
\sup_{(z, z') \in \Omega} \left| \sum_{i=1}^{n} \eta_i(z) \tilde{b}_i(z') \right| = O_P \left\{ n^{-1} (\log n)^{1/2} \right\}. \quad \text{(S39)}
$$

**Proof.** According to (S19) and (S37), we have

$$
\tilde{b}_i(z) = \tilde{B}_\eta(z)^\top Y_n^{-1} \frac{1}{N} \sum_{j=1}^{N} \tilde{B}(z_j) \sum_{\ell=0}^{p} X_{i\ell} \{ \beta_\ell^0(z_j) - \hat{\beta}_\ell(z_j) \}
$$

$$
= \tilde{B}_\eta(z)^\top Y_n^{-1} \frac{1}{N} \sum_{j=1}^{N} \tilde{B}(z_j) \sum_{\ell=0}^{p} X_{i\ell} \{ \beta_\ell^0(z_j) - \hat{\beta}_\ell(z_j) - \hat{\eta}_\ell(z_j) - \tilde{e}_\ell(z_j) \}. \quad \text{(S40)}
$$
Thus,

\[
\frac{1}{n} \sum_{i=1}^{n} \tilde{b}_i(z) \tilde{b}_i(z') = \frac{1}{n} \sum_{i=1}^{n} \tilde{B}_\eta(z)^\top Y_n^{-1} \left[ \frac{1}{N} \sum_{j=1}^{N} \tilde{B}(z_j) \sum_{\ell=0}^{p} X_{i\ell} \{ \beta_\ell^o(z_j) - \hat{\beta}_\ell(z_j) \} \right] \\
\times \left[ \frac{1}{N} \sum_{j'=1}^{N} \tilde{B}(z_{j'})^\top \sum_{\ell'=0}^{p} X_{i\ell'} \{ \beta_{\ell'}^o(z_{j'}) - \hat{\beta}_{\ell'}(z_{j'}) \} \right] Y_n^{-1} \tilde{B}_\eta(z') \\
\simeq \frac{1}{n |\Delta_\eta|^4} \sum_{i=1}^{n} \tilde{B}_\eta(z)^\top \left[ \frac{1}{N^2} \sum_{j,j'=1}^{N} \tilde{B}(z_j) \tilde{B}(z_{j'})^\top \sum_{\ell,\ell'=0}^{p} X_{i\ell} X_{i\ell'} \{ \beta_\ell^o(z_j) - \hat{\beta}_\ell(z_j) \} \right] \\
\times \{ \beta_{\ell'}^o(z_{j'}) - \hat{\beta}_{\ell'}(z_{j'}) \} \tilde{B}_\eta(z').
\]

Therefore, by Theorem [1] we have

\[
E \left\{ \frac{1}{n} \sum_{i=1}^{n} \tilde{b}_i(z) \tilde{b}_i(z') \right\} \simeq \frac{1}{n^2} \sum_{\ell=0}^{p} \sum_{\ell'=0}^{p} |\Delta_\eta|^{-2} |\beta_\ell^o - \hat{\beta}_\ell|||\beta_{\ell'}^o - \hat{\beta}_{\ell'}|| \simeq n^{-1} |\Delta_\eta|^{-2}.
\]

We have \( E \left\{ n^{-1} \sum_{i=1}^{n} \tilde{b}_i(z) \tilde{b}_i(z') \right\}^2 = \frac{1}{n^2} \sum_{i,i'=1}^{n} E \left\{ \tilde{b}_i(z) \tilde{b}_i(z') \tilde{b}_{i'}(z') \tilde{b}_{i'}(z') \right\} \), where

\[
E \left\{ \tilde{b}_i(z) \tilde{b}_i(z') \tilde{b}_{i'}(z) \tilde{b}_{i'}(z') \right\} \simeq |\Delta_\eta|^{-8} \\
\times E \tilde{B}_\eta(z)^\top \left[ \frac{1}{N^2} \sum_{j,j'=1}^{N} \tilde{B}(z_j) \tilde{B}(z_{j'})^\top \sum_{\ell,\ell'=0}^{p} X_{i\ell} X_{i\ell'} \{ \beta_\ell^o(z_j) - \hat{\beta}_\ell(z_j) \} \{ \beta_{\ell'}^o(z_{j'}) - \hat{\beta}_{\ell'}(z_{j'}) \} \right] \tilde{B}_\eta(z') \\
\times \tilde{B}_\eta(z)^\top \left[ \frac{1}{N^2} \sum_{j,j'=1}^{N} \tilde{B}(z_j) \tilde{B}(z_{j'})^\top \sum_{\ell,\ell'=0}^{p} X_{i\ell} X_{i\ell'} \{ \beta_\ell^o(z_j) - \hat{\beta}_\ell(z_j) \} \{ \beta_{\ell'}^o(z_{j'}) - \hat{\beta}_{\ell'}(z_{j'}) \} \right] \tilde{B}_\eta(z') \\
\simeq n^{-2} |\Delta_\eta|^{-4}.
\]

Thus, (S38) follows from the Bernstein inequality after the discretization.
Following from [S40], we have, for any $i, i' = 1, \ldots, n$,

\[
E \left\{ \tilde{b}_i(z') \tilde{b}_{i'}(z') \eta_i(z) \eta_{i'}(z) \right\} 
\lesssim |\Delta_\eta|^4 \left( \frac{1}{N^2} \right) \sum_{j, j'=1}^N \tilde{B}(z_j) \tilde{B}(z_{j'}) ^\top E \left\{ \sum_{i, i'=0}^p X_{ii} X_{i'i'} \tilde{\eta}_i(z_j) \tilde{\eta}_{i'}(z_{j'}) \eta_i(z) \eta_{i'}(z) \right\} \tilde{B}_\eta(z'),
\]

\[
E \left\{ \sum_{i, i'=0}^p X_{ii} X_{i'i'} \tilde{\eta}_i(z_j) \tilde{\eta}_{i'}(z_{j'}) \eta_i(z) \eta_{i'}(z) \right\} = \frac{1}{n^2 N^2} \sum_{i', i''=1}^n E \left\{ \{ X_i \otimes \tilde{B}(z_j) \} \Gamma^{-1}_{n, \rho} \right\} \sum_{j, j'=1}^N X_{i'i'} X_{i''i''} ^\top \tilde{B}(z_{j'}) ^\top \Gamma^{-1}_{n, \rho} \tilde{B}(z_{j''}) \Gamma^{-1}_{n, \rho} \left\{ \{ X_{i'i'} \otimes \tilde{B}(z_j) \} \right\}
\]

\[
= \frac{1}{n^2 N^2} E \left\{ \{ X_i \otimes \tilde{B}(z_j) \} \Gamma^{-1}_{n, \rho} \right\} \sum_{j, j'=1}^N X_{i'i'} X_{i''i''} ^\top \tilde{B}(z_{j'}) ^\top \Gamma^{-1}_{n, \rho} \tilde{B}(z_{j''}) \Gamma^{-1}_{n, \rho} \left\{ \{ X_{i'i'} \otimes \tilde{B}(z_j) \} \right\}
\]

\[
\times E \left\{ \eta_i(z) \eta_{i'}(z) \eta_i(z') \eta_{i'}(z') + \eta_i(z) \eta_{i'}(z) \eta_i(z'') \eta_{i'}(z'') + \eta_i(z) \eta_{i'}(z) \eta_i(z'''') \eta_{i'}(z''') \right\} \simeq n^{-2}.
\]

Therefore, 

\[
E \left\{ \frac{1}{n} \sum_{i=1}^n \eta_i(z) \tilde{b}_i(z') \right\}^2 = \frac{1}{n^2} \sum_{i, i'=1}^n E \left\{ \tilde{b}_i(z') \tilde{b}_{i'}(z') \eta_i(z) \eta_{i'}(z) \right\} = O(n^{-2}).
\]

\[\square\]

**Lemma S20.** Under Assumptions (A1)–(A5), (C1)–(C3), we have

\[
\sup_{(z, z') \in \Omega^2} \left| n^{-1} \sum_{i=1}^n \nabla_\eta_i(z) \nabla_\eta_i(z') \right| = O_P \left\{ |\Delta_\eta|^{2(s+1)} \sum_{k=1}^{K_n} \lambda_k \| \psi_k \|^2_{s+1, \infty} + \sum_{k=K_n+1}^{\infty} \lambda_k \| \psi_k \|^2_{\infty} \right\}, 
\]

\[\text{(S41)}\]

\[
\sup_{(z, z') \in \Omega^2} \left| n^{-1} \sum_{i=1}^n \eta_i(z) \nabla_\eta_i(z') \right| = O_P \left\{ |\Delta_\eta|^{s+1} \sum_{k=1}^{K_n} \lambda_k \| \psi_k \|^2_{s+1, \infty} \| \psi_k \|_{\infty} + \sum_{k=K_n+1}^{\infty} \lambda_k \| \psi_k \|^2_{\infty} \right\}, 
\]

\[\text{(S42)}\]

\[
\sup_{(z, z') \in \Omega^2} \left| \sum_{i=1}^n \nabla_\eta_i(z) \tilde{b}_i(z') \right| = O_P \left\{ (\log n)^{1/2} n^{-1} |\Delta_\eta|^{s+1} \sum_{k=1}^{K_n} \lambda_k \| \psi_k \|^2_{s+1, \infty} \| \psi_k \|_{\infty} \right\} 
\]

\[+ O_P \left\{ (\log n)^{1/2} n^{-1} \sum_{k=K_n+1}^{\infty} \lambda_k \| \psi_k \|^2_{\infty} \right\}.\]

\[\text{(S43)}\]
Proof. For any $k \geq 1$, denote $\tilde{\psi}_k(z) = \tilde{B}(z)^T \Upsilon_{n^{-1}\sum_{j=1}^N} \tilde{B}(z_j) \psi(z_j)$, and $\nabla \psi_k = \tilde{\psi}_k - \psi_k$. According to Assumption (C2), we have $\nabla \psi_k \leq C |\Delta_{n}|^{s+1}\|\psi_k\|_{s+1,\infty}$ and $\|\tilde{\psi}_k\|_{\infty} \leq \|\psi_k\|_{\infty} + \|\nabla \psi_k\|_{\infty} \leq 2\|\psi_k\|_{\infty}$, as $n \to \infty$. It is easy to see that $\nabla \eta_i(z') = \sum_{k=1}^\infty \lambda_k^{1/2} \xi_k \nabla \psi_k(z')$.

We first show (S41). Let $\bar{k} = n^{-1} \sum_{i=1}^n \xi_{ik} \xi_{ik'}$, where $E(\bar{k}i) = I(k = k')$ and $E(\bar{k}i)^2 \leq (E \xi_{ik}^2) E(\xi_{ik'})^2 \leq C$. Simple calculation yields that $\frac{1}{n} \sum_{i=1}^n \nabla \eta_i(z) \nabla \eta_i(z') = \sum_{k,k'=1}^\infty \bar{k}kk' \lambda_k \lambda_{k'} \frac{1}{2} \nabla \psi_k(z) \nabla \psi_k'(z')$. Thus, by Assumption (C2), we have

$$\sup_{(z,z') \in \Omega} \left| E \left\{ \frac{1}{n} \sum_{i=1}^n \nabla \eta_i(z) \nabla \eta_i(z') \right\} \right| = \sup_{(z,z') \in \Omega} \left| \sum_{k=1}^\infty \lambda_k \nabla \psi_k(z) \nabla \psi_k(z') \right| \leq |\Delta_\eta|^{2(s+1)} \sum_{k=1}^\infty \lambda_k \|\psi_k\|_{s+1,\infty}^2 + C \sum_{k=K_n+1}^\infty \lambda_k \|\psi_k\|_{\infty}^2.$$

In addition, we have

$$\sup_{z, z' \in \Omega} \left[ E \left\{ \nabla \eta_i(z) \nabla \eta_i(z') \right\} \right]^2 = \sup_{z, z' \in \Omega} \left[ \sum_{k=1}^\infty \xi_{ik}^2 \lambda_k \nabla \psi_k(z) \sum_{k'=1}^\infty \xi_{ik'}^2 \lambda_{k'} \nabla \psi_k'(z') \right]^2 \asymp n^{-1} \left[ |\Delta_\eta|^{2(s+1)} \sum_{k=1}^\infty \lambda_k \|\psi_k\|_{s+1,\infty}^2 + \sum_{k=K_n+1}^\infty \lambda_k \|\psi_k\|_{\infty}^2 \right]^2.$$

Thus,

$$\sup_{z, z' \in \Omega} \text{Var} \left\{ \frac{1}{n} \sum_{i=1}^n \nabla \eta_i(z) \nabla \eta_i(z') \right\} \asymp \left[ |\Delta_\eta|^{2(s+1)} \sum_{k=1}^\infty \lambda_k \|\psi_k\|_{s+1,\infty}^2 + \sum_{k=K_n+1}^\infty \lambda_k \|\psi_k\|_{\infty}^2 \right]^2.$$

Therefore, using the discretization method and Bernstein inequality

$$\sup_{(z,z') \in \Omega} \left| \frac{1}{n} \sum_{i=1}^n \nabla \eta_i(z) \nabla \eta_i(z') - E \left\{ \nabla \eta_i(z) \nabla \eta_i(z') \right\} \right| = O_P \left( (\log n)^{1/2} n^{-1/2} |\Delta_\eta|^{2(s+1)} \sum_{k=1}^\infty \lambda_k \|\psi_k\|_{s+1,\infty}^2 (\log n)^{1/2} n^{-1/2} \sum_{k=K_n+1}^\infty \lambda_k \|\psi_k\|_{\infty}^2 \right).$$

Next we derive (S42). Noting that $n^{-1} \sum_{i=1}^n \eta_i(z) \nabla \eta_i(z') = \xi_{kk'} (\lambda_k \lambda_{k'})^{1/2} \psi_k(z') (\nabla \psi_k')(z')$,
we have

$$\sup_{(z, z') \in \Omega^2} \left| E \left\{ \frac{1}{n} \sum_{i=1}^{n} \eta_i(z) \nabla \eta_i(z') \right\} \right| \leq \sum_{k=1}^{\infty} \lambda_k \| \psi_k \|_{\infty} \| \nabla \psi_{k'} \|_{\infty}$$

$$\leq C |\Delta_n|^{s+1} \sum_{k=1}^{K_n} \lambda_k \| \psi_k \|_{s+1, \infty} \| \psi_k \|_{\infty} + \sum_{k=K_n+1}^{\infty} \lambda_k \| \psi_k \|_{\infty}^2,$$

$$\text{var} \left\{ n^{-1} \sum_{i=1}^{n} \eta_i(z) \nabla \eta_i(z') \right\} = n^{-1} \left[ E \left\{ \eta_i^2(z) \nabla \eta_i(z') \right\} - \left\{ E \eta_i(z) \nabla \eta_i(z') \right\}^2 \right],$$

$$\sup_{z, z' \in \Omega} E \left\{ \eta_i^2(z) \nabla \eta_i(z') \right\} = \sup_{z, z' \in \Omega} \left\{ \sum_{k=1}^{\infty} E \xi_{i k} \lambda_k \psi_k^2(z) (\nabla \psi_k)(z') + \sum_{k \neq k'} \lambda_k \lambda_{k'} \psi_k^2(z) (\nabla \psi_k)(z') \right\}$$

$$\leq C \left\{ |\Delta_n|^{2(s+1)} \sum_{k=1}^{K_n} \lambda_k \| \psi_k \|_{s+1, \infty}^2 + \sum_{k=K_n+1}^{\infty} \lambda_k \| \psi_k \|_{\infty}^2 \right\},$$

and

$$\sup_{z, z' \in \Omega} \left| E \left\{ \eta_i(z) \nabla \eta_i(z') \right\} \right| \leq C \left\{ |\Delta_n|^{s+1} \sum_{k=1}^{K_n} \lambda_k \| \psi_k \|_{s+1, \infty} \| \psi_k \|_{\infty} + \sum_{k=K_n+1}^{\infty} \lambda_k \| \psi_k \|_{\infty}^2 \right\}.$$

Therefore,

$$\sup_{z, z' \in \Omega} E \left\{ n^{-1} \sum_{i=1}^{n} \eta_i(z) \nabla \eta_i(z') \right\}^2 = O \left\{ \left| \Delta_n \right|^{s+1} \sum_{k=1}^{K_n} \lambda_k \| \psi_k \|_{s+1, \infty} \| \psi_k \|_{\infty} + \sum_{k=K_n+1}^{\infty} \lambda_k \| \psi_k \|_{\infty}^2 \right\}^2.$$

Hence,

$$\sup_{(z, z') \in \Omega^2} \left| n^{-1} \sum_{i=1}^{n} \nabla \eta_i(z) \bar{\varepsilon}_i(z') - E \left\{ \nabla \eta_i(z) \bar{\varepsilon}_i(z') \right\} \right|$$

$$= O_P \left\{ (\log n)^{1/2} n^{-1/2} |\Delta_n|^{2(s+1)} \sum_{k=1}^{K_n} \lambda_k \| \psi_k \|_{s+1, \infty}^2 + (\log n)^{1/2} n^{-1/2} \sum_{k=K_n+1}^{\infty} \lambda_k \| \psi_k \|_{\infty}^2 \right\}$$

using the discretization method and Bernstein inequality.

Finally, we provide the proof of (S43). Note that

$$E \left\{ b_i(z') \tilde{b}_{i'}(z') \nabla \eta_i(z) \nabla \eta_{i'}(z) \right\} \propto |\Delta_n|^{-4}$$

$$\times \tilde{B}_\eta(z')^\top \frac{1}{N^2} \sum_{j, j'=1}^{N} \tilde{B}_\eta(z_j) \tilde{B}_\eta(z_{j'})^\top E \left\{ \sum_{\ell, \ell'=0}^{p} X_{i\ell} X_{i'\ell'} (z_j) \tilde{\eta}_\ell(z_{j'}) \nabla \eta_i(z) \nabla \eta_{i'}(z) \right\} \tilde{B}_\eta(z'),$$
and by (S42),

\[
E \left\{ \sum_{i,i'=0}^p X_{ii'}X_{i'i'}^T \eta_i(z_j) \eta_{i'}(z_{j'}) \nabla \eta_i(z) \nabla \eta_{i'}(z) \right\} = \frac{1}{n^2N^2} \sum_{i,i'=1}^n E \left[ \left\{ X_i \otimes \tilde{B}(z_j) \right\}^T \Gamma_{n,\rho}^{-1} \right.
\]
\[
\times \sum_{j''}^N X_{i''}X_{i'''}^T \otimes \tilde{B}(z_{j''}) \tilde{B}(z_{j''})^T \Gamma_{n,\rho}^{-1} X_{i'} \otimes \tilde{B}(z_{j'}) \right] E \left\{ \eta_i(z_{j''}) \eta_{i'}(z_{j''}) \nabla \eta_i(z) \nabla \eta_{i'}(z) \right\}
\]
\[
= \frac{1}{n^2N^2} E \left[ \left\{ X_i \otimes \tilde{B}(z_j) \right\}^T \Gamma_{n,\rho}^{-1} \sum_{j''}^N X_{i''}X_{i'''}^T \otimes \tilde{B}(z_{j''}) \tilde{B}(z_{j''})^T \Gamma_{n,\rho}^{-1} \left\{ X_{i'} \otimes \tilde{B}(z_{j'}) \right\} \right]
\]
\[
\times E \left\{ \eta_i(z_{j''}) \eta_{i'}(z_{j''}) \nabla \eta_i(z) \nabla \eta_{i'}(z) + \eta_i(z_{j''}) \eta_{i'}(z_{j''}) \nabla \eta_i(z) \nabla \eta_{i'}(z) \right\}.
\]

If \( i \neq i' \), we have

\[
E \left\{ \eta_i(z_{j''}) \eta_{i'}(z_{j''}) \nabla \eta_i(z) \nabla \eta_{i'}(z) + \eta_i(z_{j''}) \eta_{i'}(z_{j''}) \nabla \eta_i(z) \nabla \eta_{i'}(z) \right\}
\]
\[
\leq \sum_{k=1}^{K_n} \lambda_k \| \Delta \|_{\eta}^{s+1} \| \psi_k \|_{s+1,\infty} \| \psi_k \|_{\infty} + \sum_{k=K_n+1}^\infty \lambda_k \| \psi_k \|_{\infty}^2.
\]

If \( i = i' \), then we have

\[
E \left\{ \eta_i(z_{j''}) \eta_i(z_{j''}) \nabla \eta_i(z) \nabla \eta_{i}(z) \right\} = \sum_{k=1}^\infty \lambda_k^2 E \xi_{ik}^4 \psi_k(z_{j''}) \psi_k(z_{j''}) \nabla \psi_k(z) \nabla \psi_k(z)
\]
\[
\leq \sum_{k=1}^{K_n} \lambda_k^2 \| \Delta \|_{\eta}^{2s+2} \| \psi_k \|_{s+1,\infty}^2 \| \psi_k \|_{\infty}^2 + \sum_{k=K_n+1}^\infty \lambda_k^2 \| \psi_k \|_{\infty}^4.
\]

Thus,

\[
E \left\{ \sum_{i,i'=0}^p X_{ii'}X_{i'i'}^T \eta_i(z_j) \eta_{i'}(z_{j'}) \nabla \eta_i(z) \nabla \eta_{i'}(z) \right\}
\]
\[
\leq \left\{ \sum_{k=1}^{K_n} \lambda_k \| \Delta \|_{\eta}^{s+1} \| \psi_k \|_{s+1,\infty} \| \psi_k \|_{\infty} + \sum_{k=K_n+1}^\infty \lambda_k \| \psi_k \|_{\infty}^2 \right\}^2.
\]
Therefore,

\[
E \left\{ \frac{1}{n} \sum_{i=1}^{n} \nabla \eta_i(z) \tilde{b}_i(z') \right\}^2 = \frac{1}{n^2} \sum_{i,i'=1}^{n} E \tilde{b}_i(z') \tilde{b}_{i'}(z') \nabla \eta_i(z) \nabla \eta_{i'}(z) \\
= O \left[ n^{-2} \sum_{k=1}^{K_n} \lambda_k^2 \| \Delta \|_{\psi,2}^2 \| \psi_k \|_{s+1,\infty}^2 + n^{-2} \sum_{k=K_n+1}^{\infty} \lambda_k^2 \| \psi_k \|_{4,\infty}^4 \right].
\]

Thus, (S43) is obtained. \(\square\)

**Lemma S21.** Under Assumptions (A1)–(A5), (C1)–(C3), we have

\[
\begin{align*}
\sup_{(z,z') \in \Omega^2} & \left| n^{-1} \sum_{i=1}^{n} \tilde{\epsilon}_i(z) \tilde{\epsilon}_i(z') \right| = O_P(N^{-1} |\Delta\eta|^{-2}), & (S44) \\
\sup_{(z,z') \in \Omega^2} & \left| n^{-1} \sum_{i=1}^{n} \nabla \eta_i(z) \tilde{\epsilon}_i(z') \right| = O_P \left\{ n^{-1/2} N^{-1/2} (\log n)^{1/2} |\Delta\eta|^{-1} \sum_{k=1}^{K_n} \lambda_k^{1/2} \| \psi_k \|_{s+1,\infty}^1 \right\} \\
& \quad + O_P \left\{ n^{-1/2} N^{-1/2} |\Delta\eta|^{-1} (\log n)^{1/2} \sum_{k=K_n+1}^{\infty} \lambda_k^{1/2} \| \psi_k \|_{\infty} \right\}, & (S45) \\
\sup_{(z,z') \in \Omega^2} & \left| n^{-1} \sum_{i=1}^{n} \tilde{b}_i(z) \tilde{\epsilon}_i(z') \right| = O_P \left\{ n^{-1} N^{-1} |\Delta|^{-2} (\log n)^{1/2} \right\}, & (S46) \\
\sup_{(z,z') \in \Omega^2} & \left| n^{-1} \sum_{i=1}^{n} \eta_i(z) \tilde{\epsilon}_i(z') \right| = O_P \left\{ n^{-1/2} N^{-1/2} |\Delta\eta|^{-1} (\log n)^{1/2} \right\}. & (S47)
\end{align*}
\]

**Proof.** We first show (S44). Let \( \tilde{\epsilon}_{jj'} = n^{-1} \sum_{i=1}^{n} \epsilon_{ij} \epsilon_{ij'} \), where \( E(\tilde{\epsilon}_{jj'}) = I(j = j') \).

Note that

\[
\frac{1}{n} \sum_{i=1}^{n} \tilde{\epsilon}_i(z) \tilde{\epsilon}_i(z') = \tilde{B}(z)^\top \Psi_n^{-1} \left\{ \frac{1}{N^2} \sum_{j=1}^{N} \sum_{j'=1}^{N} \tilde{B}(z_j) \tilde{B}(z_{j'})^\top \sigma(z_j) \sigma(z_{j'}) \tilde{\epsilon}_{jj'} \right\} \Psi_n^{-1} \tilde{B}(z').
\]

It is easy to see that,

\[
E \left\{ \frac{1}{n} \sum_{i=1}^{n} \tilde{\epsilon}_i(z) \tilde{\epsilon}_i(z') \right\} = \tilde{B}(z)^\top \Psi_n^{-1} \left\{ \frac{1}{N^2} \sum_{j=1}^{N} \tilde{B}(z_j) \tilde{B}(z_j)^\top \sigma^2(z_j) \right\} \Psi_n^{-1} \tilde{B}(z').
\]

Therefore, \( \sup_{(z,z') \in \Omega^2} | E \left\{ \frac{1}{n} \sum_{i=1}^{n} \tilde{\epsilon}_i(z) \tilde{\epsilon}_i(z') \right\} | = O(N^{-1} |\Delta\eta|^{-2}) \). In addition, note
that

\[ E \{ \tilde{\epsilon}_i(z) \tilde{\epsilon}_i(z') \} = \tilde{\mathbf{B}}(z) \mathbf{Y}_n^{-1} \left\{ \frac{1}{N^2} \sum_{j=1}^N \tilde{\mathbf{B}}(z_j) \tilde{\mathbf{B}}(z_{j'})^\top \sigma^2(z_j) \right\} \mathbf{Y}_n^{-1} \tilde{\mathbf{B}}(z') = O(N^{-1}|\Delta \eta|^{-2}), \]

\[ E \{ \tilde{\epsilon}_i(z) \tilde{\epsilon}_i(z') \}^2 = E \left[ \tilde{\mathbf{B}}_\eta(z)^\top \mathbf{Y}_n^{-1} \left\{ \frac{1}{N^2} \sum_{j=1}^N \sum_{j'=1}^N \tilde{\mathbf{B}}_\eta(z_j) \tilde{\mathbf{B}}_\eta(z_{j'})^\top \sigma(z_j) \sigma(z_{j'}) \varepsilon_{ij} \varepsilon_{ij'} \right\} \mathbf{Y}_n^{-1} \tilde{\mathbf{B}}_\eta(z') \right]^2 \approx \frac{|\Delta \eta|^{-8}}{N^4} \sum_{j,j',j''=1}^N \tilde{\mathbf{B}}_\eta(z_j) \tilde{\mathbf{B}}_\eta(z_{j'})^\top \tilde{\mathbf{B}}_\eta(z_{j''}) \tilde{\mathbf{B}}_\eta(z_{j'''})^\top \sigma(z_j) \sigma(z_{j'}) \sigma(z_{j''}) \varepsilon_{ij} \varepsilon_{ij'} \varepsilon_{ij''} \varepsilon_{ij'''} \approx N^{-2}|\Delta \eta|^{-4}. \]

Thus, \( \text{var} \left\{ \frac{1}{n} \sum_{i=1}^n \tilde{\epsilon}_i(z) \tilde{\epsilon}_i(z') \right\} = \frac{1}{n^2} \sum_{i=1}^n \text{var} \left\{ \tilde{\epsilon}_i(z) \tilde{\epsilon}_i(z') \right\} \approx n^{-1} N^{-2} |\Delta \eta|^{-4} \). Therefore,

\[ \sup_{(z,z') \in \Omega} \left| n^{-1} \sum_{i=1}^n \tilde{\epsilon}_i(z) \tilde{\epsilon}_i(z') - E \{ \tilde{\epsilon}_i(z) \tilde{\epsilon}_i(z') \} \right| = O_P \left\{ n^{-1/2} N^{-1} (\log n)^{1/2} |\Delta \eta|^{-2} \right\}
\]

using the discretization method and Bernstein inequality.

Next we derive (S45). Note that

\[ \frac{1}{n} \sum_{i=1}^n \nabla \eta_i(z) \tilde{\epsilon}_i(z') = \frac{1}{n^2} \sum_{i=1}^n \sum_{k=1}^\infty \xi_{ik} \lambda_i^{1/2} \nabla \psi_k(z) \tilde{\mathbf{B}}_\eta(z') \mathbf{Y}_n^{-1} \left\{ \frac{1}{N} \sum_{j=1}^N \tilde{\mathbf{B}}_\eta(z_j) \sigma(z_j) \varepsilon_{ij} \right\}, \]

\[ \left\{ \frac{1}{n} \sum_{i=1}^n \nabla \eta_i(z) \tilde{\epsilon}_i(z') \right\}^2 = \frac{1}{n^2} \sum_{i=1}^n \sum_{j'=1}^\infty \sum_{k=1}^\infty \sum_{k'=1}^\infty \xi_{ik} \xi_{i'k'} (\lambda_i \lambda_{k'})^{1/2} \nabla \psi_k(z) \nabla \psi_{k'}(z) \tilde{\mathbf{B}}_\eta(z') \mathbf{Y}_n^{-1} \left\{ \frac{1}{N^2} \sum_{j=1}^N \sum_{j'=1}^N \tilde{\mathbf{B}}_\eta(z_j) \tilde{\mathbf{B}}_\eta(z_{j'})^\top \sigma(z_j) \sigma(z_{j'}) \varepsilon_{ij} \varepsilon_{ij'} \right\} \mathbf{Y}_n^{-1} \tilde{\mathbf{B}}_\eta(z'). \]

Next observe that \( E \left\{ \frac{1}{n} \sum_{i=1}^n \nabla \eta_i(z) \tilde{\epsilon}_i(z') \right\} = 0 \) and

\[ E \left\{ \frac{1}{n} \sum_{i=1}^n \nabla \eta_i(z) (\nabla \psi_k)^2(z) \right\} = \frac{1}{n^2} \sum_{i=1}^n \sum_{k=1}^\infty \lambda_i (\nabla \psi_k)^2(z) \tilde{\mathbf{B}}_\eta(z') \mathbf{Y}_n^{-1} \left\{ \frac{1}{N^2} \sum_{j=1}^N \tilde{\mathbf{B}}_\eta(z_j) \tilde{\mathbf{B}}_\eta(z_j)^\top \sigma^2(z_j) \right\} \mathbf{Y}_n^{-1} \tilde{\mathbf{B}}_\eta(z'). \]
So,

\[
E \left\{ \frac{1}{n} \sum_{i=1}^{n} \nabla \eta_i(z)(\nabla \psi_k)^2(z) \right\} \leq C_1 \frac{\Delta \eta}{nN} \left( |\Delta \eta|^2 (s+1) \sum_{k=1}^{K_n} \lambda_k \| \psi_k \|_{s+1, \infty}^2 + \sum_{k=K_n+1}^{\infty} \lambda_k \| \psi_k \|^2_{\infty} \right).
\]

Thirdly, we prove (S46). Note that for any \( i, i', j, j' \), we have

\[
E \left\{ \tilde{b}_i(z) \tilde{e}_{ij} \tilde{b}_{i'}(z) \tilde{e}_{i'j'} \right\} = E \left[ B_{\eta}^T(z) Y_n^{-1} \frac{1}{N^2} \sum_{j''} B_{\eta}(z_{j''}) B_{\eta}(z_{j''})^T \sum_{\ell, \ell'=0}^{P} X_{i\ell} \tilde{e}_{\ell}(z_{j''}) X_{i\ell'} \tilde{e}_{\ell'}(z_{j''}) \tilde{e}_{ij} \tilde{e}_{i'j'} Y_n^{-1} B_{\eta}(z) \right]
\]

\[
= O(n^{-2} N^{-2} |\Delta|^{-4}).
\]

Therefore,

\[
E \left\{ \tilde{b}_i(z) \tilde{e}_{i}(z') \tilde{b}_{i'}(z) \tilde{e}_{i'}(z') \right\} = \tilde{B}_{\eta}(z')^T Y_n^{-1} \frac{1}{N^2} \sum_{j''} E \left\{ \tilde{b}_i(z) \tilde{e}_{ij} \tilde{b}_{i'}(z) \tilde{e}_{i'j'} \right\} Y_n^{-1} \tilde{B}_{\eta}(z')
\]

\[
= O(n^{-2} N^{-2} |\Delta|^{-4}),
\]

\[
E \left[ n^{-1} \sum_{i=1}^{n} \tilde{b}_i(z) \tilde{e}_{i}(z') \right] = \frac{1}{n^2} \sum_{i, i'=1}^{n} E \left\{ \tilde{b}_i(z) \tilde{e}_{i}(z') \tilde{b}_{i'}(z) \tilde{e}_{i'}(z') \right\} = O(n^{-2} N^{-2} |\Delta|^{-4}).
\]

Finally, we show (S47). Note that

\[
\sum_{i=1}^{n} \eta_i(z) \tilde{e}_{i}(z') = \sum_{i=1}^{n} \tilde{B}_{\eta}(z')^T Y_n^{-1} \frac{1}{N} \sum_{j=1}^{N} \tilde{B}_{\eta}(z_j) \sigma(z_j) \tilde{e}_{ij} \sum_{k=1}^{K_n} \xi_{ik} \lambda_k^{1/2} \psi_k(z),
\]

where \( E \left\{ n^{-1} \sum_{i=1}^{n} \eta_i(z) \tilde{e}_{i}(z') \right\} = 0 \), and

\[
E \left\{ n^{-1} \sum_{i=1}^{n} \eta_i(z) \tilde{e}_{i}(z') \right\}^2 = n^{-1} E \{ \eta_i(z)^2 \} E \{ \tilde{e}_i(z')^2 \} = n^{-1} G_{\eta}(z, z') \tilde{B}_{\eta}(z')^T
\]

\[
\times Y_n^{-1} \frac{1}{N^2} \sum_{j=1}^{N} \tilde{B}_{\eta}(z_j) \tilde{B}_{\eta}(z_j)^T \sigma^2(z_j) Y_n^{-1} \tilde{B}_{\eta}(z') = O(n^{-1} N^{-1} |\Delta\eta|^2).
\]

Thus, (S47) is obtained.\qquad \square
S2. Appendix 2

In this section, we provide some additional results from simulation studies and real application analysis.

S2.1 More Results of Simulation Studies

In Section 5.1 of the main paper, we illustrated the advantage of the proposed method over the complex horseshoe domain in Sangalli et al. (2013). Figure S1 shows the two triangulations used for the horseshoe domain in this example. For implementation, the BPST method is conducted over triangulation, $\triangle_1$, while triangulation, $\triangle_2$, is used for PCST method. To visually compare different methods, we display the estimated coefficient functions for Case I (jump function) and Case II (smooth function) in Figures S2 and S3, respectively. The plots are obtained based on the setting: $n = 50$, $\lambda_1 = 0.2$, $\lambda_2 = 0.05$, $\sigma = 1.0$. Table S1 summarizes the estimation results based on the noise level $\sigma = 1.0$.

From these figures, one sees that the BPST and PCST estimates are both very close to the true coefficient functions. When the true coefficient functions are smooth, BPST provides the best estimation, while when the true coefficient function contains jumps, PCST provides a better estimation. The performance of the Tensor method will be affected by the design of the coefficient function. Moreover, from Figure S2 and S3, one can see that even when the coefficient function is smooth across the boundary, the estimation accuracy is also affected by the domain of the true signal, especially the pixels which are close to the boundary. The performance of the Kernel method is not
affected by the design of the coefficient functions, instead, it heavily depends on the noise level due to the three-stage structure. As the noise level increases, the Kernel estimates are getting more blurred.

Figure S1: Triangulations for the horseshoe domain.

Figure S2: True coefficient functions and their different estimators for Case I in Example 1.

In Section 5.2 of the main paper, we conduct a simulation study based on the domain of the 5th slice of the brain images illustrated in Section 6. Table S2 demonstrates the estimation results for $\sigma = 0.5$. In this example, we focus on the domain of the 35th slices of the brain image. Based on this domain, we consider two types of
Figure S3: True coefficient functions and their different estimators for Case II in Example 1.

triangulations: $\Delta_5$ and $\Delta_6$; see Figure S4. Table S3 summarizes the MSE results of the BPST, Kernel and Tensor methods. The findings are similar to those described in Section 5.2. Tables 3 and S4 summarize the ECRs of the 95% SCCs for the 5th and 35th slices, respectively, and they are all close to 95%. As the sample size increases, the ECRs are getting closer to 95%. Figures S5 and S6 show the true coefficient functions and an example of their estimators and 95% SCCs based on the 5th and 35th slices, respectively. The plots are generated based on the setting: $n = 50$, $\lambda_1 = 0.1$, $\lambda_2 = 0.02$ and $\sigma = 0.5$.

S2.2 Additional ADNI Data Analysis Results

For the ADNI data described in Section 6 in the main paper, Table S5 below summarizes the distribution of patients by diagnosis status and sex. Next, Figure S7 displays the triangulations of slices used for the BPST method in the model fitting and con-
Table S1: Estimation errors of the coefficient estimators, $\sigma = 1.0$.

<table>
<thead>
<tr>
<th>Function Type</th>
<th>$n$</th>
<th>Method</th>
<th>$\lambda_1 = 0.03, \lambda_2 = 0.006$</th>
<th>$\lambda_1 = 0.2, \lambda_2 = 0.05$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\beta_0$</td>
<td>$\beta_1$</td>
</tr>
<tr>
<td>Jump</td>
<td>50</td>
<td>BPST</td>
<td>0.0059</td>
<td>0.0075</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PCST</td>
<td>0.0023</td>
<td>0.0023</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kernel</td>
<td>0.0201</td>
<td>0.0206</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tensor</td>
<td>0.0201</td>
<td>0.0132</td>
</tr>
<tr>
<td>100</td>
<td>BPST</td>
<td>0.0038</td>
<td>0.0050</td>
<td>0.0042</td>
</tr>
<tr>
<td></td>
<td>PCST</td>
<td>0.0011</td>
<td>0.0011</td>
<td>0.0014</td>
</tr>
<tr>
<td>Smooth</td>
<td>50</td>
<td>Kernel</td>
<td>0.0100</td>
<td>0.0102</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tensor</td>
<td>0.0099</td>
<td>0.0112</td>
</tr>
<tr>
<td>100</td>
<td>BPST</td>
<td>0.0010</td>
<td>0.0012</td>
<td>0.0016</td>
</tr>
<tr>
<td></td>
<td>PCST</td>
<td>0.0049</td>
<td>0.0065</td>
<td>0.0054</td>
</tr>
<tr>
<td></td>
<td>Kernel</td>
<td>0.0201</td>
<td>0.0206</td>
<td>0.0207</td>
</tr>
<tr>
<td></td>
<td>Tensor</td>
<td>0.0189</td>
<td>0.0132</td>
<td>0.0207</td>
</tr>
<tr>
<td>50</td>
<td>BPST</td>
<td>0.0006</td>
<td>0.0007</td>
<td>0.0009</td>
</tr>
<tr>
<td></td>
<td>PCST</td>
<td>0.0037</td>
<td>0.0054</td>
<td>0.0040</td>
</tr>
<tr>
<td>Smooth</td>
<td>100</td>
<td>Kernel</td>
<td>0.0100</td>
<td>0.0102</td>
</tr>
<tr>
<td></td>
<td>Tensor</td>
<td>0.0100</td>
<td>0.0113</td>
<td>0.0103</td>
</tr>
</tbody>
</table>

Finally, Figures S8 and S9 provide the image maps of the estimated coefficient functions for 8th, 15th, 35th, 55th, 62nd, and 65th slices, and Figures S10 and S11 show the corresponding significance maps. The significance maps in the 8th and 15th slice show that the increase of age increases the brain activities in the cerebellum and temporal lobe, and people with Alzheimer’s disease are more active in the cerebellum, while less active in the temporal lobe. The significance maps of the 35th slide display that the age has a negative effect on the brain activities in the anterior cingulate gyrus, corpus callosum, and part of the cerebral white matter, while the female has a higher level of activities in these regions. These regions connect the left and right cerebral hemispheres and enabling communication between them. From the significance maps of the 55th, 62nd, and 65th slices, we could see an increase of brain...
Figure S4: Triangulations for the 5th slice (Δ₃, Δ₄) and 35th slice (Δ₅, Δ₆) of the brain image in Simulation Example 2.

Table S2: Estimation errors of the coefficient function estimators, σ = 0.5.

<table>
<thead>
<tr>
<th>n</th>
<th>Method</th>
<th>λ₁ = 0.1, λ₂ = 0.02</th>
<th></th>
<th>λ₁ = 0.2, λ₂ = 0.05</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>β₀</td>
<td>β₁</td>
<td>β₂</td>
<td>β₀</td>
</tr>
<tr>
<td>50</td>
<td>BPST(Δ₃)</td>
<td>0.003</td>
<td>0.005</td>
<td>0.005</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>BPST(Δ₄)</td>
<td>0.003</td>
<td>0.005</td>
<td>0.005</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>Kernel</td>
<td>0.008</td>
<td>0.011</td>
<td>0.011</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>Tensor</td>
<td>0.008</td>
<td>0.007</td>
<td>0.010</td>
<td>0.011</td>
</tr>
<tr>
<td>100</td>
<td>BPST(Δ₃)</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>BPST(Δ₄)</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Kernel</td>
<td>0.004</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>Tensor</td>
<td>0.004</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
</tr>
</tbody>
</table>

activities in the frontal gyrus, precentral gyrus and postcentral gyrus for people with Alzheimer’s disease. Our findings are consistent with the findings in the literature, see Andersen et al. (2012), Bernard and Seidler (2014), and Dubb et al. (2003).

Bibliography


Table S4: The coverage rate of the 95% SCCs for the coefficient functions defined over the 35th slice.

<table>
<thead>
<tr>
<th>n</th>
<th>λ</th>
<th>σ</th>
<th>Coverage Width</th>
<th>β₀</th>
<th>β₁</th>
<th>β₂</th>
<th>β₀</th>
<th>β₁</th>
<th>β₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>(0.1,0.02)</td>
<td>0.5</td>
<td>0.962 0.916 0.934</td>
<td>0.307</td>
<td>0.344</td>
<td>0.347</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>0.964 0.926 0.940</td>
<td>0.331</td>
<td>0.368</td>
<td>0.371</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>(0.2,0.05)</td>
<td>0.5</td>
<td>0.952 0.920 0.930</td>
<td>0.426</td>
<td>0.490</td>
<td>0.492</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>0.96 0.920 0.934</td>
<td>0.449</td>
<td>0.512</td>
<td>0.512</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>(0.1,0.02)</td>
<td>0.5</td>
<td>0.956 0.952 0.940</td>
<td>0.214</td>
<td>0.240</td>
<td>0.244</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>0.962 0.952 0.948</td>
<td>0.239</td>
<td>0.262</td>
<td>0.265</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>(0.2,0.05)</td>
<td>0.5</td>
<td>0.946 0.954 0.932</td>
<td>0.298</td>
<td>0.340</td>
<td>0.346</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>0.952 0.954 0.938</td>
<td>0.317</td>
<td>0.359</td>
<td>0.365</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table S5: Distribution of patients by diagnosis status and gender.

<table>
<thead>
<tr>
<th></th>
<th>CN</th>
<th>MCI</th>
<th>AD</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>70</td>
<td>136</td>
<td>72</td>
<td>278</td>
</tr>
<tr>
<td>Female</td>
<td>42</td>
<td>77</td>
<td>50</td>
<td>169</td>
</tr>
<tr>
<td>All</td>
<td>112</td>
<td>213</td>
<td>122</td>
<td>447</td>
</tr>
</tbody>
</table>
Figure S5: True coefficient functions and their estimators and 95% SCCs based on the 5th slice.
Figure S6: True coefficient functions and their estimators and 95% SCCs based on the 35th slice.
Figure S7: Triangulation sets used in the ADNI data analysis.
Figure S8: The BPSF estimates of the coefficient functions for the ADNI data based on the 8th, 15th and 35th slices, respectively.
Figure S9: The BPST estimates of the coefficient functions for the ADNI data based on the 55th, 62nd and 65th slices, respectively.
Figure S10: The "significance" map (based on the 95% SCC) for the coefficient functions for the ADNI data. The yellow color and blue color on the map indicate the regions that zero is below the lower SCC or above the upper SCC, respectively.
Figure S11. The “significance” map (based on the 95% SCC) for the coefficient functions for the ADNI data. The yellow color and blue color on the map indicate the regions that zero is below the lower SCC or above the upper SCC, respectively.