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

Shan Yu[1], Guannan Wang[2], Li Wang[3] and Lijian Yang[4]

[1] *University of Virginia, Charlottesville, VA 22904, USA.*

[2] *College of William & Mary, Williamsburg, VA 23187, USA.*

[3] *Iowa State University, Ames, IA 50011, USA.*

[4] *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 $\boldsymbol{a}$, denote $\|\boldsymbol{a}\|$ 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 $\lambda_{\min}(\mathbf{A})$ and $\lambda_{\max}(\mathbf{A})$ be the smallest and largest eigenvalues of $\mathbf{A}$. For a vector valued function $\boldsymbol{g} = (g_0, \ldots, g_p)^\top$, denote $\|\boldsymbol{g}\|_{L_2(\Omega)} = \{\sum_{\ell=0}^p \|g_\ell\|_{L_2(\Omega)}^2\}^{1/2}$ and $\|\boldsymbol{g}\|_{\infty,\Omega} = \max_{0\le\ell\le p} \|g_\ell\|_{\infty,\Omega}$, where $\|g_\ell\|_{L_2,\Omega}$ and $\|g_\ell\|_{\infty,\Omega}$ are the L_2 norm and supremum norm of g_ℓ defined at the beginning of Section 2.3. Further denote $\|\boldsymbol{g}\|_{v,\infty,\Omega} = \max_{0\le\ell\le p} |g_\ell|_{v,\infty,\Omega}$, where $|g_\ell|_{v,\infty,\Omega} = \max_{i+j=v} \|\nabla_{z_1}^i \nabla_{z_2}^j g_\ell(\boldsymbol{z})\|_{\infty,\Omega}$. For notation simplicity, we drop the subscript Ω in the rest of the paper. For $\boldsymbol{g}^{(1)}(\boldsymbol{z}) = (g_0^{(1)}(\boldsymbol{z}), \ldots, g_p^{(1)}(\boldsymbol{z}))^\top$ and $\boldsymbol{g}^{(2)}(\boldsymbol{z}) = (g_0^{(2)}(\boldsymbol{z}), \ldots, g_p^{(2)}(\boldsymbol{z}))^\top$, define the empirical inner product as

$$\langle \boldsymbol{g}^{(1)}, \boldsymbol{g}^{(2)} \rangle_{n,N} = \frac{1}{nN} \sum_{\ell,\ell'=0}^{p} \sum_{i=1}^{n} \sum_{j=1}^{N} X_{i\ell} X_{i\ell'} g_\ell^{(1)}(\boldsymbol{z}_j) g_{\ell'}^{(2)}(\boldsymbol{z}_j), \tag{S1.1}$$

and the theoretical inner product as

$$\langle \boldsymbol{g}^{(1)}, \boldsymbol{g}^{(2)} \rangle = \sum_{\ell,\ell'=0}^{p} E(X_\ell X_{\ell'}) \int_\Omega g_\ell^{(1)}(\boldsymbol{z}) g_{\ell'}^{(2)}(\boldsymbol{z}) d\boldsymbol{z}, \tag{S1.2}$$

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

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

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

Let $A(\Omega)$ be the area of the domain Ω, 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: $\mathbf{B}_0(\boldsymbol{z}) = \mathbf{B}_1(\boldsymbol{z}) = \cdots = \mathbf{B}_p(\boldsymbol{z}) = \mathbf{B}(\boldsymbol{z})$, and $\beta_\ell(\boldsymbol{z}_j) = \mathbf{B}^\top(\boldsymbol{z}_j)\boldsymbol{\gamma}_\ell$.

S1.1 Properties of bivariate splines

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

Lemma S1.1 (Theorem 2.7, Lai and Schumaker (2007))**.** *Let $\{B_m\}_{m\in\mathcal{M}}$ be the Bernstein polynomial basis for spline space $\mathcal{S}_d^r(\triangle)$ defined over a π-quasi-uniform triangulation $\triangle$. Then there exist positive constants c, C depending on the smoothness r, d, and the shape parameter π such that $c|\triangle|^2\sum_{m\in\mathcal{M}}\gamma_m^2 \leq \left\|\sum_{m\in\mathcal{M}}\gamma_m B_m\right\|_{L_2}^2 \leq C|\triangle|^2\sum_{m\in\mathcal{M}}\gamma_m^2$.*

Lemma S1.2 (Theorems 10.2 and 10.10, Lai and Schumaker (2007))**.** *Suppose that $|\triangle|$ is a π-quasi-uniform triangulation of a polygonal domian Ω, and $g(\cdot) \in \mathcal{W}^{d+1,\infty}(\Omega)$.*

(i) For bi-integer (a_1, a_2) with $0 \leq a_1 + a_2 \leq d$, there exists a spline $g^*(\cdot) \in \mathcal{S}_d^0(\triangle)$ such that $\|\nabla_{z_1}^{a_1}\nabla_{z_2}^{a_2}(g - g^*)\|_\infty \leq C|\triangle|^{d+1-a_1-a_2}|g|_{d+1,\infty}$, where C is a constant depending on d, and the shape parameter π.

(ii) For bi-integer (a_1, a_2) with $0 \leq a_1 + a_2 \leq d$, there exists a spline $g^{**}(\cdot) \in \mathcal{S}_d^r(\triangle)$ ($d \geq 3r + 2$) such that $\|\nabla_{z_1}^{a_1}\nabla_{z_2}^{a_2}(g - g^{**})\|_\infty \leq C|\triangle|^{d+1-a_1-a_2}|g|_{d+1,\infty}$, where C is a constant depending on d, r, and the shape parameter π.

Lemma S1.2 shows that $\mathcal{S}_d^0(\triangle)$ has full approximation power, and $\mathcal{S}_d^r(\triangle)$ also has full approximation power if $d \geq 3r + 2$. For any $g(\cdot)$ in Sobolev space $\mathcal{C}^{(0)}(\Omega)$, there exists a spline $g^*(\cdot) \in \mathcal{PC}(\triangle)$ such that $\|g - g^*\|_\infty \leq C|\triangle|\|g\|_\infty$.

Lemma S1.3. *Let $\boldsymbol{g}(\boldsymbol{z}) = (g_0(\boldsymbol{z}), \ldots, g_p(\boldsymbol{z}))^\top$, where $g_\ell(\boldsymbol{z}) = \sum_{m\in\mathcal{M}}\gamma_{\ell m}B_m(\boldsymbol{z})$. Then, under Assumptions (A3) and (A5), $\|\boldsymbol{g}\| \asymp \sum_{\ell=0}^p \|g_\ell\|_{L_2}$.*

Proof. By (S1.1), $\|\mathbf{g}\|^2 = \sum_{\ell,\ell'=0}^{p} E(X_\ell X_{\ell'}) \int_\Omega g_\ell(\boldsymbol{z}) g_{\ell'}(\boldsymbol{z}) d\boldsymbol{z} = \int_\Omega \mathbf{g}^\top(\boldsymbol{z}) \boldsymbol{\Sigma}_X \mathbf{g}(\boldsymbol{z}) d\boldsymbol{z}$. According to Assumptions (A3) and (A5), $\|\mathbf{g}\|^2 \asymp \int_\Omega \mathbf{g}^\top(\boldsymbol{z}) \mathbf{g}(\boldsymbol{z}) d\boldsymbol{z} \asymp \sum_{\ell=0}^{p} \|g_\ell\|_{L_2}$. □

Lemma S1.4. *Under Assumptions (A4) and (A5), for any Bernstein basis polynomials $B_m(\boldsymbol{z})$, $m \in \mathcal{M}$, of degree $d \geq 0$, we have*

$$\max_{m\in\mathcal{M}} \left| \frac{1}{N} \sum_{j=1}^{N} B_m^k(\boldsymbol{z}_j) - \int_\Omega B_m^k(\boldsymbol{z}) d\boldsymbol{z} \right| = O\left(|\triangle| N^{-1/2}\right), \ 1 \leq k < \infty, \tag{S1.3}$$

$$\max_{m,m'\in\mathcal{M}} \left| \frac{1}{N} \sum_{j=1}^{N} B_m(\boldsymbol{z}_j) B_{m'}(\boldsymbol{z}_j) - \int_\Omega B_m(\boldsymbol{z}) B_{m'}(\boldsymbol{z}) d\boldsymbol{z} \right| = O\left(|\triangle| N^{-1/2}\right), \tag{S1.4}$$

$$\max_{m,m'\in\mathcal{M}} \left| \frac{1}{N^2} \sum_{j,j'=1}^{N} G_\eta(\boldsymbol{z}_j, \boldsymbol{z}_{j'}) B_m(\boldsymbol{z}_j) B_{m'}(\boldsymbol{z}_{j'}) - \int_{\Omega^2} G_\eta(\boldsymbol{z}, \boldsymbol{z}') B_m(\boldsymbol{z}) B_{m'}(\boldsymbol{z}') d\boldsymbol{z} d\boldsymbol{z}' \right|$$

$$= O\left(N^{-1/2} |\triangle|^3\right), \tag{S1.5}$$

$$\max_{m\in\mathcal{M}} \left| \|\sigma B_m\|_{N,L_2}^2 - \|\sigma B_m\|_{L_2}^2 \right| = \max_{m\in\mathcal{M}} \left| \frac{1}{N} \sum_{j=1}^{N} B_m^2(\boldsymbol{z}_j) \sigma^2(\boldsymbol{z}_j) - \int_\Omega \sigma^2(\boldsymbol{z}) B_m^2(\boldsymbol{z}) d\boldsymbol{z} \right|$$

$$= O\left(N^{-1/2} |\triangle|\right). \tag{S1.6}$$

Proof. Note that there are $d^* = (d+1)(d+2)/2$ Bernstein basis polynomials on each triangle and $\int_\Omega B_m^k(\boldsymbol{z}) d\boldsymbol{z} = \int_{T_{\lceil m/d^* \rceil}} B_m^k(\boldsymbol{z}) d\boldsymbol{z}$, for any $k \geq 1$.

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

$$\left| \frac{1}{N} \sum_{j=1}^{N} B_m^k(\boldsymbol{z}_j) - \int_\Omega B_m^k(\boldsymbol{z}) d\boldsymbol{z} \right| = \left| \frac{1}{N} \sum_{j=1}^{N} I(\boldsymbol{z}_j \in T_m) - A(T_m) \right|.$$

According to Assumption (A5),

$$\max_{m\in\mathcal{M}} \left| \frac{1}{N} \sum_{j=1}^{N} B_m^k(\boldsymbol{z}_j) - \int_\Omega B_m^k(\boldsymbol{z}) d\boldsymbol{z} \right| \leq C N^{-1/2} |\triangle|.$$

For any $j = 1, \ldots, N$, let $\mathcal{V}_j$ be the jth pixel, and it is clear that

$$\left| \frac{1}{N} \sum_{j=1}^{N} B_m^k(\boldsymbol{z}_j) - \int_\Omega B_m^k(\boldsymbol{z}) d\boldsymbol{z} \right| \leq \left| \sum_{j=1}^{N} \int_{\mathcal{V}_j} \{B_m^k(\boldsymbol{z}_j) - B_m^k(\boldsymbol{z})\} d\boldsymbol{z} \right| + \int_{\Omega \setminus \cup \mathcal{V}_j} B_m^k(\boldsymbol{z}) d\boldsymbol{z}.$$

If $d \geq 1$, by the properties of bivariate spline basis functions in Lai and Schumaker (2007), $\int_{\Omega\backslash\cup\mathcal{V}_j} B_m^k(\boldsymbol{z})d\boldsymbol{z} = O(N^{-1/2}|\triangle|)$, and

$$\left|\sum_{j=1}^N \int_{\mathcal{V}_j} \{B_m^k(\boldsymbol{z}_j) - B_m^k(\boldsymbol{z})\}d\boldsymbol{z}\right| \leq \sum_{\{j:\boldsymbol{z}_j\in T_{\lceil m/d^*\rceil}\}} \int_{\mathcal{V}_j} |B_m^k(\boldsymbol{z}_j) - B_m^k(\boldsymbol{z})|d\boldsymbol{z}$$
$$\leq C(N|\triangle|^2) \times N^{-1} \times (N^{-1/2}|\triangle|^{-1}) \leq CN^{-1/2}|\triangle|.$$

Thus, (S1.3) holds. The proof of (S1.4) is similar to the proof (S1.3), thus omitted.

Next, for any $m, m' \in \mathcal{M}$,

$$\begin{aligned}
&\frac{1}{N^2}\sum_{j=1}^N\sum_{j'=1}^N G_\eta(\boldsymbol{z}_j, \boldsymbol{z}_{j'})B_m(\boldsymbol{z}_j)B_{m'}(\boldsymbol{z}_{j'}) - \int_{\Omega^2} G_\eta(\boldsymbol{z}, \boldsymbol{z}')B_m(\boldsymbol{z})B_{m'}(\boldsymbol{z}')d\boldsymbol{z}d\boldsymbol{z}' \\
&= \sum_{j=1}^N\sum_{j'=1}^N \int_{\mathcal{V}_j\times\mathcal{V}_{j'}} \left\{G_\eta(\boldsymbol{z}_j, \boldsymbol{z}_{j'})B_m(\boldsymbol{z}_j)B_{m'}(\boldsymbol{z}_{j'}) - G_\eta(\boldsymbol{z}, \boldsymbol{z}')B_m(\boldsymbol{z})B_{m'}(\boldsymbol{z}')\right\} d\boldsymbol{z}d\boldsymbol{z}' \\
&\quad + \int_{\Omega^2\backslash\cup_{j,j'}\mathcal{V}_j\times\mathcal{V}_{j'}} \left\{G_\eta(\boldsymbol{z}_j, \boldsymbol{z}_{j'})B_m(\boldsymbol{z}_j)B_{m'}(\boldsymbol{z}_{j'}) - G_\eta(\boldsymbol{z}, \boldsymbol{z}')B_m(\boldsymbol{z})B_{m'}(\boldsymbol{z}')\right\} d\boldsymbol{z}d\boldsymbol{z}'.
\end{aligned}$$

As $N \to \infty$,

$$\int_{\Omega^2\backslash\cup_{j,j'}\mathcal{V}_j\times\mathcal{V}_{j'}} \left\{G_\eta(\boldsymbol{z}_j, \boldsymbol{z}_{j'})B_m(\boldsymbol{z}_j)B_{m'}(\boldsymbol{z}_{j'}) - G_\eta(\boldsymbol{z}, \boldsymbol{z}')B_m(\boldsymbol{z})B_{m'}(\boldsymbol{z}')\right\} d\boldsymbol{z}d\boldsymbol{z}' = O\left(\frac{|\triangle|^3}{\sqrt{N}}\right).$$

Notice that

$$\begin{aligned}
&\left|\sum_{j=1}^N\sum_{j'=1}^N \int_{\mathcal{V}_j\times\mathcal{V}_{j'}} \left\{G_\eta(\boldsymbol{z}_j, \boldsymbol{z}_{j'})B_m(\boldsymbol{z}_j)B_{m'}(\boldsymbol{z}_{j'}) - G_\eta(\boldsymbol{z}, \boldsymbol{z}')B_m(\boldsymbol{z})B_{m'}(\boldsymbol{z}')\right\} d\boldsymbol{z}d\boldsymbol{z}'\right| \\
&\quad \leq \sum_{\{(j,j'):\boldsymbol{z}_j\in T_{\lceil m/d^*\rceil}, \boldsymbol{z}_{j'}\in T_{\lceil m/d^*\rceil}\}} \int_{\mathcal{V}_j\times\mathcal{V}_{j'}} \omega_{jj'}(G_\eta K_m, 2N^{-1/2})d\boldsymbol{z}d\boldsymbol{z}',
\end{aligned}$$

where $K_m(\boldsymbol{z}, \boldsymbol{z}') = B_m(\boldsymbol{z})B_m(\boldsymbol{z}')$ and

$$\omega_{jj'}(g, \varrho) = \sup_{\substack{(\boldsymbol{z}_1,\boldsymbol{z}_1'),(\boldsymbol{z}_2,\boldsymbol{z}_2')\in\mathcal{V}_j\times\mathcal{V}_{j'}, \\ \|\boldsymbol{z}_1-\boldsymbol{z}_2\|^2+\|\boldsymbol{z}_1'-\boldsymbol{z}_2'\|^2=\varrho^2}} |g(\boldsymbol{z}_1, \boldsymbol{z}_1') - g(\boldsymbol{z}_2, \boldsymbol{z}_2')|$$

is the modulus of continuity of g on $\mathcal{V}_j \times \mathcal{V}_{j'}$. Therefore, by Assumption (A4), we have

$$\left|\sum_{j=1}^{N}\sum_{j'=1}^{N}\int_{\mathcal{V}_j\times\mathcal{V}_{j'}}\left\{G_\eta(\boldsymbol{z}_j,\boldsymbol{z}_{j'})B_m(\boldsymbol{z}_j)B_{m'}(\boldsymbol{z}_{j'})-G_\eta(\boldsymbol{z},\boldsymbol{z}')B_m(\boldsymbol{z})B_{m'}(\boldsymbol{z}')\right\}d\boldsymbol{z}d\boldsymbol{z}'\right|$$
$$\leq (N|\triangle|^2)^2 \times N^{-2} \times (N^{-1/2}|\triangle|^{-1}) = O(N^{-1/2}|\triangle|^3).$$

Thus, (S1.5) follows.

Finally, note that

$$\left|\frac{1}{N}\sum_{j=1}^{N}B_m^2(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j)-\int_\Omega\sigma^2(\boldsymbol{z})d\boldsymbol{z}\right| \leq \left|\sum_{j=1}^{N}\int_{\mathcal{V}_j}\{B_m^2(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j)-B_m^2(\boldsymbol{z})\sigma^2(\boldsymbol{z})\}d\boldsymbol{z}\right|$$
$$+\int_{\Omega\setminus\cup\mathcal{V}_j}|B_m^2(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j)-B_m^2(\boldsymbol{z})\sigma^2(\boldsymbol{z})|d\boldsymbol{z}.$$

It is easy to see that $\int_{\Omega\setminus\cup\mathcal{V}_j}|B_m^2(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j)-B_m^2(\boldsymbol{z})\sigma^2(\boldsymbol{z})|d\boldsymbol{z} = O(N^{-1/2}|\triangle|)$. Denote $\omega_j(g,\varrho)=\sup_{\boldsymbol{z},\boldsymbol{z}'\in\mathcal{V}_j,\|\boldsymbol{z}-\boldsymbol{z}'\|=\varrho}|g(\boldsymbol{z})-g(\boldsymbol{z}')|$ is the modulus of continuity of g on the jth pixel $\mathcal{V}_j$, then by Assumption (A4), we have

$$\left|\sum_{j=1}^{N}\int_{\mathcal{V}_j}\{B_m^2(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j)-B_m^2(\boldsymbol{z})\sigma^2(\boldsymbol{z})\}d\boldsymbol{z}\right| \leq \sum_{\{j:\boldsymbol{z}_j\in T_{\lceil m/d^*\rceil}\}}\int_{\mathcal{V}_j}\omega_j(B_m^2\sigma^2,2N^{-1/2})d\boldsymbol{z}$$
$$\leq C(N|\triangle|^2)\times N^{-1}\times(N^{-1/2}|\triangle|^{-1}) \leq CN^{-1/2}|\triangle|.$$

We obtain (S1.6). $\square$

Lemma S1.5. *For any $m\in\mathcal{M}$, $0\leq\ell,\ell'\leq p$, let $\Phi_{m,\ell,\ell'}=E(X_\ell X_{\ell'})\int_\Omega B_m^2(\boldsymbol{z})d\boldsymbol{z}$. Suppose Assumptions (A3) and (A5) hold, and $N^{1/2}|\triangle|\to\infty$ as $N\to\infty$, then with probability 1, one has*

$$\max_{m\in\mathcal{M}}\max_{0\leq\ell,\ell'\leq p}\left|\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}B_m^2(\boldsymbol{z}_j)X_{i\ell}X_{i\ell'}-\Phi_{m,\ell,\ell'}\right| = O\left\{n^{-1/2}|\triangle|^2(\log n)^{1/2}+N^{-1/2}|\triangle|\right\}.$$

Proof. Let $\varsigma_{i,m} \equiv \varsigma_{i,m,\ell,\ell'} = \frac{1}{N}\sum_{j=1}^{N} B_m^2(\boldsymbol{z}_j)X_{i\ell}X_{i\ell'}$. If $N^{1/2}|\triangle| \to \infty$ as $N \to \infty$, then by (S1.3), we can show that $E(\varsigma_{i,m}) = \frac{1}{N}\sum_{j=1}^{N} B_m^2(\boldsymbol{z}_j)E(X_\ell X_{\ell'}) \asymp |\triangle|^2$, and $E\left(\varsigma_{i,m}\right)^2 = \left\{\frac{1}{N}\sum_{j=1}^{N} B_m^2(\boldsymbol{z}_j)\right\}^2 E\left(X_\ell X_{\ell'}\right)^2 \asymp |\triangle|^4$.

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 $X_{i\ell\ell'} = X_{i\ell}X_{i\ell'} = X_{i\ell\ell',1}^{D_n} + X_{i\ell\ell',2}^{D_n}$, where $X_{i\ell\ell',1}^{D_n} = X_{i\ell}X_{i\ell'}I\left\{|E(X_{i\ell}X_{i\ell'}| > D_n\right\}$, $X_{i\ell\ell',2}^{D_n} = X_{i\ell}X_{i\ell'}I\left\{|X_{i\ell}X_{i\ell'}| \le D_n\right\}$. Correspondingly the truncated and tail parts of $\varsigma_{i,m}$ are $\varsigma_{i,m,v} \equiv \varsigma_{i,m,v,\ell,\ell'} = \frac{1}{N}\sum_{j=1}^{N} B_m^2(\boldsymbol{z}_j)X_{i\ell\ell',v}^{D_n}$, $v = 1, 2$. According to Assumption (A3), for any $\ell, \ell' = 0, \ldots, p$,

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

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

$$\begin{aligned}|E(\varsigma_{i,m,1})| &= |E(X_{i\ell\ell',1}^{D_n})|\left\{\frac{1}{N}\sum_{j=1}^{N} B_m^2(\boldsymbol{z}_j)\right\}\\ &\le D_n^{-2}E\left|X_{i\ell}X_{i\ell'}\right|^3\left\{\int_\Omega B_m^2(\boldsymbol{z})d\boldsymbol{z} + O(N^{-1/2}|\triangle|)\right\} \le CD_n^{-2}|\triangle|^2.\end{aligned}$$

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

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

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

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

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

$$
\begin{aligned}
E\left|\varsigma_{i,m,2}^{*}\right|^k &\leq 2^{k-1}\left\{D_n^{k-2}E\left(X_{i\ell\ell'}\right)^2 + O(1)\right\}\left\{\frac{1}{N}\sum_{j=1}^{N} B_m^2(\boldsymbol{z}_j)\right\}^k \\
&\leq 2^k D_n^{k-2} E\left(\varsigma_{i,m,2}^{*}\right)^2\left\{\frac{1}{N}\sum_{j=1}^{N} B_m^2(\boldsymbol{z}_j)\right\}^{k-2} \leq \left(C_\varsigma D_n|\triangle|^2\right)^{k-2} k! E\left(\varsigma_{i,m,2}^{*}\right)^2,
\end{aligned}
$$

which implies that $\left\{\varsigma_{i,m,2}^{*}\right\}_{i=1}^{n}$ satisfies Cramér's condition with constant $C_\varsigma D_n|\triangle|^2$. Applying Bernstein's inequality to $\sum_{i=1}^{n}\varsigma_{i,m,2}^{*}$, for $k>2$ and any large enough $\delta>0$,

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

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

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

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

$$
\begin{aligned}
&\max_{m,\ell,\ell'}\left|n^{-1}\sum_{i=1}^{n}\varsigma_{i,m} - E\varsigma_{i,m}\right| \leq \max_{m,\ell,\ell'}\left|n^{-1}\sum_{i=1}^{n}\varsigma_{i,m,1}\right| + \max_{m,\ell,\ell'}\left|n^{-1}\sum_{i=1}^{n}\varsigma_{i,m,2}^{*}\right| + \max_{m,\ell,\ell'}\left|E\varsigma_{i,m,1}\right| \\
&= O_{a.s.}(n^{-k}) + O_{a.s.}\left\{n^{-1/2}|\triangle|^2(\log n)^{1/2}\right\} + O\left(D_n^{-2}|\triangle|^2\right) = O_{a.s.}\left\{n^{-1/2}|\triangle|^2(\log n)^{1/2}\right\}.
\end{aligned}
$$

Finally, we notice that

$$
\begin{aligned}
&\max_{\substack{m\in\mathcal{M}\\ 0\le\ell,\ell'\le p}}\left|\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}B_m^2(\boldsymbol{z}_j)X_{i\ell}X_{i\ell'}-\Phi_{m,\ell,\ell'}\right|\\
&\quad=\max_{\substack{m\in\mathcal{M}\\ 0\le\ell,\ell'\le p}}\left|n^{-1}\sum_{i=1}^{n}\varsigma_{i,m}-E\varsigma_{i,m}\right|+|EX_{i\ell}X_{i\ell'}|\max_{m\in\mathcal{M}}\left|\frac{1}{N}\sum_{j=1}^{N}B_m^2(\boldsymbol{z}_j)-\int_{\Omega}B_m^2(\boldsymbol{z})d\boldsymbol{z}\right|\\
&\quad=O_{a.s.}\left\{n^{-1/2}|\triangle|^2(\log n)^{1/2}\right\}+O(N^{-1/2}|\triangle|).
\end{aligned}
$$

We obtain the desired result. □

The following lemma provide the uniform convergence rate at which the empirical inner product in (S1.1) approximates the theoretical inner product in (S1.2).

Lemma S1.6. *Let* $g_\ell^{(1)}(\boldsymbol{z})=\sum_{m\in\mathcal{M}}c_{\ell m}^{(1)}B_m(\boldsymbol{z})$, $g_\ell^{(2)}(\boldsymbol{z})=\sum_{m\in\mathcal{M}}c_{\ell m}^{(2)}B_m(\boldsymbol{z})$ *be any spline functions in* $\mathcal{S}_d^r(\triangle)$. *Denote* $\boldsymbol{g}(\boldsymbol{z})=(g_0(\boldsymbol{z}),\ldots,g_p(\boldsymbol{z}))^\top$ *with* $g_\ell\in\mathcal{S}_d^r(\triangle)$, $\ell=0,\ldots,p$. *Suppose Assumptions (A3) and (A5) hold, and* $N^{1/2}|\triangle|\to\infty$ *as* $N\to\infty$, *then*

$$
R_{n,N}=\sup_{\boldsymbol{g}^{(1)},\boldsymbol{g}^{(2)}\in\mathcal{S}_d^r(\triangle)}\left|\frac{\left\langle\boldsymbol{g}^{(1)},\boldsymbol{g}^{(2)}\right\rangle_{n,N}-\left\langle\boldsymbol{g}^{(1)},\boldsymbol{g}^{(2)}\right\rangle}{\|\boldsymbol{g}^{(1)}\|\,\|\boldsymbol{g}^{(2)}\|}\right|=O_P\{n^{-1/2}(\log n)^{1/2}+N^{-1/2}|\triangle|^{-1}\}.
$$

Proof. It is easy to see

$$
\begin{aligned}
\langle\boldsymbol{g}^{(1)},\boldsymbol{g}^{(2)}\rangle_{n,N}&=\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}\left\{\sum_{\ell=0}^{p}\sum_{m\in\mathcal{M}}c_{\ell m}^{(1)}X_{i\ell}B_m(\boldsymbol{z}_j)\right\}\left\{\sum_{\ell'=0}^{p}\sum_{m'\in\mathcal{M}}c_{\ell'm'}^{(2)}X_{i\ell'}B_{m'}(\boldsymbol{z}_j)\right\}\\
&=\sum_{\ell,m}\sum_{\ell',m'}c_{\ell m}^{(1)}c_{\ell'm'}^{(2)}\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}X_{i\ell}X_{i\ell'}B_m(\boldsymbol{z}_j)B_{m'}(\boldsymbol{z}_j).
\end{aligned}
$$

Note that $\|\boldsymbol{g}^{(r)}\|^2=\sum_{\ell,m}\sum_{\ell',m'}c_{\ell m}^{(r)}c_{\ell'm'}^{(r)}E(X_\ell X_{\ell'})\int_\Omega B_m(\boldsymbol{z})B_{m'}(\boldsymbol{z})d\boldsymbol{z}$, $r=1,2$. It follows from Assumptions (A1), (A2), Lemmas S1.1 and S1.3 that,

$$c_v|\triangle|^2 \sum_{\ell,m}\{c_{\ell m}^{(v)}\}^2 \le \|\boldsymbol{g}^{(v)}\|^2 \le C_v|\triangle|^2 \sum_{\ell,m}\{c_{\ell m}^{(v)}\}^2,$$

$$C_1|\triangle|^2 \left[\sum_{\ell,m}\{c_{\ell m}^{(1)}\}^2 \sum_{\ell',m'}\{c_{\ell' m'}^{(2)}\}^2\right]^{1/2} \le \|\boldsymbol{g}^{(1)}\|\|\boldsymbol{g}^{(2)}\| \le C_2|\triangle|^2 \left[\sum_{\ell,m}\{c_{\ell m}^{(1)}\}^2 \sum_{\ell',m'}\{c_{\ell' m'}^{(2)}\}^2\right]^{1/2}.$$

With the above preparation, we have

$$\begin{aligned} R_{n,N} \le & \frac{\sum_{\ell,\ell',|m-m'|\le(d+2)(d+1)/2}|c_{\ell m}^{(1)}c_{\ell' m'}^{(2)}|}{C_1|\triangle|^2\left[\sum_{\ell,m}\{c_{\ell m}^{(1)}\}^2\sum_{\ell',m'}\{c_{\ell' m'}^{(2)}\}^2\right]^{1/2}} \\ & \times \max_{\substack{m,m'\in\mathcal{M} \\ 0\le\ell,\ell'\le p}} \left|\frac{1}{nN}\sum_{i=1}^n\sum_{j=1}^N B_m(\boldsymbol{z}_j)B_{m'}(\boldsymbol{z}_j)X_{i\ell}X_{i\ell'} - E(X_\ell X_{\ell'})\int_\Omega B_m(\boldsymbol{z})B_{m'}(\boldsymbol{z})d\boldsymbol{z}\right| \\ \le & \frac{C}{|\triangle|^2} \max_{\substack{m,m'\in\mathcal{M} \\ 0\le\ell,\ell'\le p}} \left|\frac{1}{nN}\sum_{i=1}^n\sum_{j=1}^N B_m(\boldsymbol{z}_j)B_{m'}(\boldsymbol{z}_j)X_{i\ell}X_{i\ell'} - E(X_\ell X_{\ell'})\int_\Omega B_m(\boldsymbol{z})B_{m'}(\boldsymbol{z})d\boldsymbol{z}\right|. \end{aligned} \tag{S1.7}$$

The desired result follows from (S1.7) and Lemma S1.5. □

As a direct result of Lemma S1.6, we can see that

$$\sup_{\boldsymbol{g}\in\mathcal{S}_d^r(\triangle)} \left| \|\boldsymbol{g}\|_{n,N}^2 \Big/ \|\boldsymbol{g}\|^2 - 1 \right| = O_P\{n^{-1/2}(\log n)^{1/2} + N^{-1/2}|\triangle|^{-1}\}. \tag{S1.8}$$

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 $\boldsymbol{\beta}^o = (\beta_0^o, \ldots, \beta_p^o)^\top$ is defined as

$$\widetilde{\boldsymbol{\beta}} = (\widetilde{\beta}_0, \ldots, \widetilde{\beta}_p)^\top = \underset{\boldsymbol{\beta}\in\mathcal{G}^{(p+1)}}{\arg\min} \sum_{i=1}^n\sum_{j=1}^N \left\{Y_i(\boldsymbol{z}_j) - \sum_{\ell=0}^p X_{i\ell}\beta_\ell(\boldsymbol{z}_j)\right\}^2. \tag{S1.9}$$

Denote

$$\widetilde{\boldsymbol{\theta}}_{\mu} = (\widetilde{\boldsymbol{\theta}}_{\mu,0}^{\top}, \ldots, \widetilde{\boldsymbol{\theta}}_{\mu,p}^{\top})^{\top} = \boldsymbol{\Gamma}_{n,0}^{-1} \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} \left\{ \widetilde{\mathbf{X}}_i \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_j) \right\} \widetilde{\mathbf{X}}_i^{\top} \boldsymbol{\beta}^o(\boldsymbol{z}_j)$$

$$\widetilde{\boldsymbol{\theta}}_{\eta} = (\widetilde{\boldsymbol{\theta}}_{\eta,0}^{\top}, \ldots, \widetilde{\boldsymbol{\theta}}_{\eta,p}^{\top})^{\top} = \boldsymbol{\Gamma}_{n,0}^{-1} \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} \left\{ \widetilde{\mathbf{X}}_i \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_j) \right\} \eta_i(\boldsymbol{z}_j),$$

$$\widetilde{\boldsymbol{\theta}}_{\varepsilon} = (\widetilde{\boldsymbol{\theta}}_{\varepsilon,0}^{\top}, \ldots, \widetilde{\boldsymbol{\theta}}_{\varepsilon,p}^{\top})^{\top} = \boldsymbol{\Gamma}_{n,0}^{-1} \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} \left\{ \widetilde{\mathbf{X}}_i \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_j) \right\} \sigma(\boldsymbol{z}_j) \varepsilon_{ij},$$

where

$$\boldsymbol{\Gamma}_{n,0} = \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} (\widetilde{\mathbf{X}}_i \widetilde{\mathbf{X}}_i^{\top}) \otimes \{\widetilde{\mathbf{B}}(\boldsymbol{z}_j) \widetilde{\mathbf{B}}^{\top}(\boldsymbol{z}_j)\}. \tag{S1.10}$$

Lemma S1.7. *Under Assumptions (A3) and (A5), if $N^{1/2}|\triangle| \to \infty$ as $N \to \infty$, then there exist constants $0 < c_\Gamma < C_\Gamma < \infty$, such that with probability approaching 1, as $N \to \infty$, $n \to \infty$, $c_\Gamma |\triangle|^2 \leq \lambda_{\min}(\boldsymbol{\Gamma}_{n,0}) \leq \lambda_{\max}(\boldsymbol{\Gamma}_{n,0}) \leq C_\Gamma |\triangle|^2$, where $\boldsymbol{\Gamma}_{n,0}$ is in (S1.10).*

Proof. Note that for any vector $\boldsymbol{\theta} = (\boldsymbol{\theta}_0^{\top}, \cdots, \boldsymbol{\theta}_p^{\top})^{\top}$ with $\boldsymbol{\gamma}_\ell = (\gamma_{\ell m}, m \in \mathcal{M})^{\top}$,

$$\boldsymbol{\theta}^{\top} \boldsymbol{\Gamma}_{n,0} \boldsymbol{\theta} = \frac{1}{nN} \boldsymbol{\gamma}^{\top} \sum_{i=1}^{n} \sum_{j=1}^{N} (\widetilde{\mathbf{X}}_i \widetilde{\mathbf{X}}_i^{\top}) \otimes \{\mathbf{B}(\boldsymbol{z}_j) \mathbf{B}^{\top}(\boldsymbol{z}_j)\} \boldsymbol{\gamma} = \|\boldsymbol{g}_{\boldsymbol{\gamma}}\|_{n,N}^2, \tag{S1.11}$$

where $\boldsymbol{\gamma} = \mathbf{Q}_2 \boldsymbol{\theta}$, and $\boldsymbol{g}_{\boldsymbol{\gamma}} = (g_{\gamma_0}, \ldots, g_{\gamma_p})^{\top}$ with $g_{\gamma_\ell} = \sum_{m \in \mathcal{M}} \gamma_{\ell m} B_m$. By (S1.8), we have

$$c(1-R_{n,N})|\triangle|^2 \|\boldsymbol{\gamma}\|^2 \leq (1-R_{n,N}) \|\boldsymbol{g}_{\boldsymbol{\gamma}}\|^2 \leq \|\boldsymbol{g}_{\boldsymbol{\gamma}}\|_{n,N}^2 = (1+R_{n,N}) \|\boldsymbol{g}_{\boldsymbol{\gamma}}\|^2 \leq C(1+R_{n,N})|\triangle|^2 \|\boldsymbol{\gamma}\|^2,$$

in which we have used the stability conditions in Lemma S1.1. $\square$

Next, we consider the following decomposition $\widetilde{\boldsymbol{\beta}}(\boldsymbol{z}) = \widetilde{\boldsymbol{\beta}}_{\mu}(\boldsymbol{z}) + \widetilde{\boldsymbol{\eta}}(\boldsymbol{z}) + \widetilde{\boldsymbol{\varepsilon}}(\boldsymbol{z})$, where

$$\widetilde{\boldsymbol{\beta}}_{\mu}(\boldsymbol{z}) = (\widetilde{\beta}_{\mu,0}(\boldsymbol{z}), \ldots, \widetilde{\beta}_{\mu,p}(\boldsymbol{z}))^{\top} = \{\mathbf{I} \otimes \widetilde{\mathbf{B}}(\boldsymbol{z})\}^{\top} \widetilde{\boldsymbol{\theta}}_{\mu}, \tag{S1.12}$$

$$\widetilde{\boldsymbol{\eta}}(\boldsymbol{z}) = (\widetilde{\eta}_0(\boldsymbol{z}), \ldots, \widetilde{\eta}_p(\boldsymbol{z}))^{\top} = \{\mathbf{I} \otimes \widetilde{\mathbf{B}}(\boldsymbol{z})\}^{\top} \widetilde{\boldsymbol{\theta}}_{\eta}, \tag{S1.13}$$

$$\widetilde{\boldsymbol{\varepsilon}}(\boldsymbol{z}) = (\widetilde{\varepsilon}_0(\boldsymbol{z}), \ldots, \widetilde{\varepsilon}_p(\boldsymbol{z}))^{\top} = \{\mathbf{I} \otimes \widetilde{\mathbf{B}}(\boldsymbol{z})\}^{\top} \widetilde{\boldsymbol{\theta}}_{\varepsilon}. \tag{S1.14}$$

Lemma S1.8. *Under Assumptions (A2)–(A5) and (C1), if $N^{1/2}|\triangle| \to \infty$ as $N \to \infty$, $\|\sum_{k=1}^{\infty} \lambda_k^{1/2} \psi_k\|_{\infty} < \infty$ and $n^{1/(4+\delta_2)} \ll n^{1/2} N^{-1/2} |\triangle|^{-1}$ for some δ_2, then for $\widetilde{\boldsymbol{\eta}}$ and $\widetilde{\boldsymbol{\varepsilon}}$ in (S1.13) and (S1.14), $\|\widetilde{\boldsymbol{\eta}}\|_{\infty} = O_P\{n^{-1/2}(\log n)^{1/2}\}$ and $\|\widetilde{\boldsymbol{\varepsilon}}\|_{\infty} = O_P\{(nN)^{-1/2}(\log n)^{1/2}|\triangle|^{-1}\}$.*

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

$$\|\widetilde{\boldsymbol{\eta}}\|_{\infty} = \max_{0 \le \ell \le p} \|\widetilde{\eta}_{\ell}\|_{\infty} \le C_{\eta} \|\widetilde{\boldsymbol{\theta}}_{\eta,\ell}\|_{\infty} = \left\| (\mathbf{e}_{\ell} \otimes \mathbf{1})^{\top} \boldsymbol{\Gamma}_{n,0}^{-1} \left[\frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} \left\{ \widetilde{\mathbf{X}}_i \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_j) \right\} \eta_i(\boldsymbol{z}_j) \right] \right\|_{\infty},$$

where $\widetilde{\boldsymbol{\theta}}_{\eta} = (\widetilde{\theta}_{\eta,\ell,m})_{m \in \widetilde{\mathcal{M}}}$ with $\widetilde{\mathcal{M}}$ being an index set of the transformed Bernstein basis polynomials $\widetilde{B}_m(\boldsymbol{z})$ and $\boldsymbol{\Gamma}_{n,0}$ is the symmetric positive definite matrix defined in (S1.10). Thus, by Lemma S1.7,

$$\|\widetilde{\boldsymbol{\eta}}\|_{\infty} \le C |\triangle|^{-2} \max_{0 \le \ell \le p} \max_{m \in \widetilde{\mathcal{M}}} \left| \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} X_{i\ell} \eta_i(\boldsymbol{z}_j) \widetilde{B}_m(\boldsymbol{z}_j) \right|,$$

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

$$\max_{0 \le \ell \le p} \max_{m \in \widetilde{\mathcal{M}}} \left| \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} X_{i\ell} \eta_i(\boldsymbol{z}_j) \widetilde{B}_m(\boldsymbol{z}_j) \right| = O\left\{ n^{-1/2} |\triangle|^2 (\log n)^{1/2} \right\}. \tag{S1.15}$$

To prove (S1.15), let $\varpi_i = \varpi_{i,m} = \sum_{k=1}^{\infty} \lambda_k^{1/2} X_{i\ell} \xi_{ik} \frac{1}{N} \sum_{j=1}^{N} \widetilde{B}_m(\boldsymbol{z}_j) \psi_k(\boldsymbol{z}_j)$, where $E(\varpi_i) =$

0 and

$$E(\varpi_i^2) = \frac{E(X_{i\ell}^2)}{N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\widetilde{B}_m(\boldsymbol{z}_{j'})G_\eta(\boldsymbol{z}_j,\boldsymbol{z}_{j'})$$
$$\asymp \int_{\Omega^2} G_\eta(\boldsymbol{z},\boldsymbol{z}')B_m(\boldsymbol{z})B_m(\boldsymbol{z}')d\boldsymbol{z}d\boldsymbol{z}' \asymp |\triangle|^4.$$

We decompose the random variable ϖ_i into a tail part and a truncated part,

$$\varpi_{i,1}^{D_n} = \sum_{k=1}^{\infty}\lambda_k^{1/2}\left\{\frac{1}{N}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\psi_k(\boldsymbol{z}_j)\right\}X_{i\ell}\xi_{ik}I\left\{|X_{i\ell}\xi_{ik}|>D_n\right\},$$
$$\varpi_{i,2}^{D_n} = \sum_{k=1}^{\infty}\lambda_k^{1/2}\left\{\frac{1}{N}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\psi_k(\boldsymbol{z}_j)\right\}X_{i\ell}\xi_{ik}I\left\{|X_{i\ell}\xi_{ik}|\le D_n\right\}-\mu_i^{D_n},$$
$$\mu_i^{D_n} = \sum_{k=1}^{\infty}\lambda_k^{1/2}\left\{\frac{1}{N}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\psi_k(\boldsymbol{z}_j)\right\}E\left[X_{i\ell}\xi_{ik}I\left\{|X_{i\ell}\xi_{ik}|\le D_n\right\}\right],$$

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 $k\ge 1$,

$$\sum_{n=1}^{\infty}P\left\{|X_{n\ell}\xi_{nk}|>D_n\right\}\le\sum_{n=1}^{\infty}\frac{E\left|X_{n\ell}\xi_{nk}\right|^{4+\delta_1}}{D_n^{4+\delta_1}}\le\upsilon_{\delta_1}\sum_{n=1}^{\infty}D_n^{-(4+\delta_1)}<\infty.$$

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

Next, notice that $E(\varpi_{i,2}^{D_n})=0$. Then, $\mathrm{Var}(\varpi_{i,2}^{D_n}) = E(\varpi_i^2)-E(\varpi_{i,1}^{D_n})^2-(\mu^{D_n})^2\asymp|\triangle|^4$. Also, we have, for any $r\ge 3$,

$$E|\varpi_{i,2}^{D_n}|^r = E\left|\sum_{k=1}^{\infty}\lambda_k^{1/2}\frac{1}{N}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\psi_k(\boldsymbol{z}_j)\left[X_{i\ell}\xi_{ik}I\left\{|X_{i\ell}\xi_{ik}|\le D_n\right\}\right]-\mu_i^{D_n}\right|^r$$
$$\le 2^{r-1}\left[E\left|\sum_{k=1}^{\infty}\lambda_r^{1/2}\frac{1}{N}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\psi_k(\boldsymbol{z}_j)X_{i\ell}\xi_{ik}I\left\{|X_{i\ell}\xi_{ik}|\le D_n\right\}\right|^r+\left(\mu_i^{D_n}\right)^r\right]$$
$$\le\left\{2D_n\frac{1}{N}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\sum_{k=1}^{\infty}\lambda_k^{1/2}\psi_k(\boldsymbol{z}_j)\right\}^{r-2}E|\varpi_{i,2}^{D_n}|^2\le(CD_n|\triangle|^2)^{r-2}E|\varpi_{i,2}^{D_n}|^2.$$

Thus, $E\left|\varpi_{i,2}/n\right|^r \leq \{Cn^{-1}D_n|\triangle|^2\}^{r-2}r!E(\varpi_{i,2}^2/n^2) < \infty$ with the Cramer constant $c^* = Cn^{-1}D_n|\triangle|^2$. By the Bernstein inequality, for any large enough $\delta > 0$,

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\varpi_i^{D_n}\right| \geq \delta n^{-1/2}|\triangle|^2(\log n)^{1/2}\right\} \leq 2\exp\left\{\frac{-\delta^2\log n}{4c + 2\delta C D_n(\log n)^{1/2}n^{-1/2}}\right\} \leq 2n^{-3}.$$

Hence,

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

for such $\delta > 0$. Thus, Borel-Cantelli's lemma implies that $\|\widetilde{\boldsymbol{\eta}}\|_\infty = O_P\{n^{-1/2}(\log n)^{1/2}\}$. The result of $\|\widetilde{\boldsymbol{\varepsilon}}\|_\infty = O_P\{(nN)^{-1/2}(\log n)^{1/2}|\triangle|^{-1}\}$ can be established similarly, thus omitted. □

For $\widetilde{\boldsymbol{\beta}}(\boldsymbol{z})$ defined in (S1.9), Theorem S1.1 below provides its uniform convergence rate to $\boldsymbol{\beta}^o$.

Theorem S1.1. *Under Assumptions (A1)–(A6), for $\widetilde{\boldsymbol{\beta}}(\boldsymbol{z})$ defined in (S1.9), $\|\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^o\|_\infty = O_P\{|\triangle|^{d+1}\|\boldsymbol{\beta}^o\|_{d+1,\infty} + n^{-1/2}(\log n)^{1/2}\}$.*

Proof. Note that $\|\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^o\|_\infty \leq \|\widetilde{\boldsymbol{\beta}}_\mu - \boldsymbol{\beta}^o\|_\infty + \|\widetilde{\boldsymbol{\eta}}\|_\infty + \|\widetilde{\boldsymbol{\varepsilon}}\|_\infty$, where

$$\widetilde{\boldsymbol{\beta}}_\mu = \arg\min_{\boldsymbol{g}\in\mathcal{G}^{(p+1)}}\sum_{i=1}^{n}\sum_{j=1}^{N}\left\{\sum_{\ell=0}^{p}X_{i\ell}(\beta_\ell^o - g_\ell)(\boldsymbol{z}_j)\right\}^2.$$

Let $\boldsymbol{\beta}^* = (\beta_0^*,\ldots,\beta_p^*)^* \in \mathcal{G}^{(p+1)}$, where β_ℓ^*'s are the best approximation to β_ℓ^o's with the approximation rate $\|\beta_\ell^* - \beta_\ell^o\|_\infty \leq C|\triangle|^{d+1}\|\boldsymbol{\beta}^o\|_{d+1,\infty}$ for any $\ell = 0,\ldots,p$. By Lai and Wang (2013),

$$\|\widetilde{\boldsymbol{\beta}}_\mu - \boldsymbol{\beta}^o\|_\infty \leq \|\widetilde{\boldsymbol{\beta}}_\mu - \boldsymbol{\beta}^*\|_\infty + \|\boldsymbol{\beta}^* - \boldsymbol{\beta}^o\|_\infty \leq C|\triangle|^{d+1}\|\boldsymbol{\beta}^o\|_{d+1,\infty}. \tag{S1.16}$$

The desired result follows from Lemma S1.8. □

S1.3 Asymptotic properties of penalized spline estimators

Let $\widetilde{\mathbf{B}}(\boldsymbol{z}) = \mathbf{Q}_2^\top \mathbf{B}(\boldsymbol{z})$, then for $\mathbb{U} = \mathbf{X} \otimes (\mathbf{B}\mathbf{Q}_2)$ defined in Section 2.2, we have

$$\mathbb{U}^\top = (\widetilde{\mathbf{X}}_1 \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_1), \ldots, \widetilde{\mathbf{X}}_1 \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_N), \ldots, \widetilde{\mathbf{X}}_n \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_1), \ldots, \widetilde{\mathbf{X}}_n \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_N)),$$

and $\mathbb{U}^\top\mathbb{U} = \sum_{i=1}^n \sum_{j=1}^N (\widetilde{\mathbf{X}}_i \widetilde{\mathbf{X}}_i^\top) \otimes \{\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}^\top(\boldsymbol{z}_j)\}$, $\mathbb{U}^\top\mathbb{Y} = \sum_{i=1}^n \sum_{j=1}^N \{\widetilde{\mathbf{X}}_i \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_j)\} Y_{ij}$. Let

$$\boldsymbol{\Gamma}_{n,\rho} = \frac{1}{nN}\sum_{i=1}^n \sum_{j=1}^N (\widetilde{\mathbf{X}}_i \widetilde{\mathbf{X}}_i^\top) \otimes \{\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}^\top(\boldsymbol{z}_j)\} + \frac{\rho_n}{nN}\mathbf{I}_p \otimes \mathbf{Q}_2^\top [\langle B_m, B_{m'}\rangle_{\mathcal{E}}]_{m,m' \in \mathcal{M}} \mathbf{Q}_2, \tag{S1.17}$$

which is a symmetric positive definite matrix.

Next, we define

$$\widehat{\boldsymbol{\theta}}_\mu = (\widehat{\boldsymbol{\theta}}_{\mu,0}^\top, \ldots, \widehat{\boldsymbol{\theta}}_{\mu,p}^\top)^\top = \boldsymbol{\Gamma}_{n,\rho}^{-1} \frac{1}{nN}\sum_{i=1}^n \sum_{j=1}^N \left\{\widetilde{\mathbf{X}}_i \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\} \widetilde{\mathbf{X}}_i^\top \boldsymbol{\beta}^o(\boldsymbol{z}_j),$$

$$\widehat{\boldsymbol{\theta}}_\eta = (\widehat{\boldsymbol{\theta}}_{\eta,0}^\top, \ldots, \widehat{\boldsymbol{\theta}}_{\eta,p}^\top)^\top = \boldsymbol{\Gamma}_{n,\rho}^{-1} \frac{1}{nN}\sum_{i=1}^n \sum_{j=1}^N \left\{\widetilde{\mathbf{X}}_i \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\} \sum_{k=1}^\infty \lambda_k^{1/2} \xi_{ik}\psi_k(\boldsymbol{z}_j),$$

$$\widehat{\boldsymbol{\theta}}_\varepsilon = (\widehat{\boldsymbol{\theta}}_{\varepsilon,0}^\top, \ldots, \widehat{\boldsymbol{\theta}}_{\varepsilon,p}^\top)^\top = \boldsymbol{\Gamma}_{n,\rho}^{-1} \frac{1}{nN}\sum_{i=1}^n \sum_{j=1}^N \left\{\widetilde{\mathbf{X}}_i \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\} \sigma(\boldsymbol{z}_j)\varepsilon_{ij}.$$

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

$$\widehat{\beta}_\ell(\boldsymbol{z}) = \widehat{\beta}_{\mu,\ell}(\boldsymbol{z}) + \widehat{\eta}_\ell(\boldsymbol{z}) + \widehat{\varepsilon}_\ell(\boldsymbol{z}), \tag{S1.18}$$

where

$$\widehat{\beta}_{\mu,\ell}(\boldsymbol{z}) = \widetilde{\mathbf{B}}(\boldsymbol{z})^\top \widehat{\boldsymbol{\theta}}_{\mu,\ell}, \quad \widehat{\eta}_\ell(\boldsymbol{z}) = \widetilde{\mathbf{B}}(\boldsymbol{z})^\top \widehat{\boldsymbol{\theta}}_{\eta,\ell}, \quad \widehat{\varepsilon}_\ell(\boldsymbol{z}) = \widetilde{\mathbf{B}}(\boldsymbol{z})^\top \widehat{\boldsymbol{\theta}}_{\varepsilon,\ell},$$

Therefore, we have

$$\widehat{\beta}_\ell(\boldsymbol{z}) - \beta_\ell^o(\boldsymbol{z}) = \widehat{\beta}_{\mu,\ell}(\boldsymbol{z}) - \beta_\ell^o(\boldsymbol{z}) + \widehat{\eta}_\ell(\boldsymbol{z}) + \widehat{\varepsilon}_\ell(\boldsymbol{z}). \tag{S1.19}$$

Lemma S1.9. *Under Assumptions (A3)–(A5), if $N^{1/2}|\triangle| \to \infty$ 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 $n \to \infty$, $c_\Gamma|\triangle|^2 \le \lambda_{\min}(\mathbf{\Gamma}_{n,\rho}) \le \lambda_{\max}(\mathbf{\Gamma}_{n,\rho}) \le C_\Gamma\left(|\triangle|^2 + \frac{\rho_n}{nN|\triangle|^2}\right)$.*

Proof. By (S1.11), it is easy to see that, for any vector $\boldsymbol{\theta} = (\boldsymbol{\theta}_0^\top, \cdots, \boldsymbol{\theta}_p^\top)^\top$,

$$\boldsymbol{\theta}^\top \mathbf{\Gamma}_{n,\rho}\boldsymbol{\theta} = \|\boldsymbol{g}_{\boldsymbol{\gamma}}\|_{n,N}^2 + \frac{\rho_n}{nN}\sum_{\ell=0}^{p}\boldsymbol{\gamma}_\ell^\top[\langle B_m, B_{m'}\rangle_{\mathcal{E}}]_{m,m'\in\mathcal{M}}\boldsymbol{\gamma}_\ell,$$

where $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_0, \ldots, \boldsymbol{\gamma}_p)^\top = \mathbf{Q}_2\boldsymbol{\theta}$ with $\boldsymbol{\gamma}_\ell = (\gamma_{\ell m}, m \in \mathcal{M})^\top$. Using the Markov's inequality in the supplement of Lai and Wang (2013) and Lemma S1.1, we have

$$\sum_{\ell=0}^{p}\left\|\sum_{m\in\mathcal{M}}\gamma_{\ell m}B_m\right\|_{\mathcal{E}}^2 \le \frac{C}{|\triangle|^4}\sum_{\ell=0}^{p}\left\|\sum_{m\in\mathcal{M}}\gamma_{\ell m}B_m\right\|_{L_2}^2 \le \frac{C}{|\triangle|^2}\|\boldsymbol{\gamma}\|^2.$$

Thus, the largest eigenvalue of the matrix $\mathbf{\Gamma}_{n,\rho}$ in (S1.17) satisfies that $\lambda_{\max}(\mathbf{\Gamma}_{n,\rho}) \le C\left\{(1 + R_{n,N})|\triangle|^2 + (nN|\triangle|^2)^{-1}\rho_n\right\}$. Thus, we have with probability approaching 1, $\lambda_{\max}(\mathbf{\Gamma}_{n,\rho}) \le C_\Gamma\left\{|\triangle|^2 + (nN|\triangle|^2)^{-1}\rho_n\right\}$ for some positive constant C_Γ. On the other hand, we use Lemma S1.1 and equation (S1.8) to have $\|\boldsymbol{g}_{\boldsymbol{\gamma}}\|_{n,N}^2 = (1 - R_{n,N})\|\boldsymbol{g}_{\boldsymbol{\gamma}}\|^2 \ge c(1 - R_{n,N})|\triangle|^2\|\boldsymbol{\gamma}\|^2$.

Therefore, $\lambda_{\min}(\mathbf{\Gamma}_{n,\rho}) \ge c(1 - R_{n,N})|\triangle|^2 = c_\Gamma|\triangle|^2$. □

Lemma S1.10. *Under Assumptions (A1), (A3) and (A5), if $N^{1/2}|\triangle| \to \infty$, one has $\|\widehat{\boldsymbol{\beta}}_\mu - \boldsymbol{\beta}^o\|_\infty = O_P\left\{\frac{\rho_n}{nN|\triangle|^3}\|\boldsymbol{\beta}^o\|_{2,\infty} + \left(1 + \frac{\rho_n}{nN|\triangle|^5}\right)|\triangle|^{d+1}\|\boldsymbol{\beta}^o\|_{d+1,\infty}\right\}$.*

Proof. Define

$$A_n = \sup_{\boldsymbol{g}\in\mathcal{G}^{(p+1)}}\left\{\frac{\|\boldsymbol{g}\|_\infty}{\|\boldsymbol{g}\|_{n,N}}, \|\boldsymbol{g}\|_{n,N} \ne 0\right\}, \quad \overline{A}_n = \sup_{\boldsymbol{g}\in\mathcal{G}^{(p+1)}}\left\{\frac{\|\boldsymbol{g}\|_{\mathcal{E}}}{\|\boldsymbol{g}\|_{n,N}}, \|\boldsymbol{g}\|_{n,N} \ne 0\right\}, \quad \text{(S1.20)}$$

where random variables A_n and $\overline{A}_n$ depend on the collection of $X_{i\ell}$'s, $i = 1, \ldots, n$,

$\ell = 0, \dots, p$. It is clear that $\|\boldsymbol{\beta}^o - \widehat{\boldsymbol{\beta}}_\mu\|_\infty \le \|\boldsymbol{\beta}^o - \widetilde{\boldsymbol{\beta}}_\mu\|_\infty + \|\widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu\|_\infty$, where $\widetilde{\boldsymbol{\beta}}_\mu$ is given in (S1.12), and $\|\widetilde{\boldsymbol{\beta}}_\mu - \boldsymbol{\beta}^o\|_\infty \le C|\triangle|^{d+1}\|\boldsymbol{\beta}^o\|_{d+1,\infty}$ according to (S1.16).

By the definition of A_n in (S1.20), we have

$$\|\widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu\|_\infty \le A_n\|\widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu\|_{n,N}. \tag{S1.21}$$

Note that the penalized spline $\widehat{\boldsymbol{\beta}}_\mu$ of $\boldsymbol{\beta}^o$ is characterized by the orthogonality relations

$$nN\langle \boldsymbol{\beta}^o - \widehat{\boldsymbol{\beta}}_\mu, \boldsymbol{g}\rangle_{n,N} = \rho_n\langle \widehat{\boldsymbol{\beta}}_\mu, \boldsymbol{g}\rangle_{\mathcal{E}}, \quad \text{for all } \boldsymbol{g} \in \mathcal{G}^{(p+1)}, \tag{S1.22}$$

while $\widetilde{\boldsymbol{\beta}}_\mu$ is characterized by

$$\langle \boldsymbol{\beta}^o - \widetilde{\boldsymbol{\beta}}_\mu, \boldsymbol{g}\rangle_{n,N} = 0, \quad \text{for all } \boldsymbol{g} \in \mathcal{G}^{(p+1)}. \tag{S1.23}$$

By (S1.22) and (S1.23), we have $nN\langle \widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu, \boldsymbol{g}\rangle_{n,N} = \rho_n\langle \widehat{\boldsymbol{\beta}}_\mu, \boldsymbol{g}\rangle_{\mathcal{E}}$, for all $\boldsymbol{g} \in \mathcal{G}^{(p+1)}$. Inserting $\boldsymbol{g} = \widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu$ yields that

$$nN\|\widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu\|_{n,N}^2 = \rho_n\langle \widehat{\boldsymbol{\beta}}_\mu, \widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu\rangle_{\mathcal{E}}. \tag{S1.24}$$

Thus, by Cauchy-Schwarz inequality and the definition of $\overline{A}_n$.

$$nN\|\widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu\|_{n,N}^2 \le \rho_n\|\widehat{\boldsymbol{\beta}}_\mu\|_{\mathcal{E}}\|\widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu\|_{\mathcal{E}} \le \rho_n\overline{A}_n\|\widehat{\boldsymbol{\beta}}_\mu\|_{\mathcal{E}}\|\widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu\|_{n,N}.$$

Similarly, using (S1.24), $nN\|\widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu\|_{n,N}^2 = \rho_n\{\langle \widehat{\boldsymbol{\beta}}_\mu, \widetilde{\boldsymbol{\beta}}_\mu\rangle_{\mathcal{E}} - \langle \widehat{\boldsymbol{\beta}}_\mu, \widehat{\boldsymbol{\beta}}_\mu\rangle_{\mathcal{E}}\} \ge 0$. Thus, by Cauchy-Schwarz inequality, $\|\widehat{\boldsymbol{\beta}}_\mu\|_{\mathcal{E}}^2 \le \langle \widehat{\boldsymbol{\beta}}_\mu, \widetilde{\boldsymbol{\beta}}_\mu\rangle_{\mathcal{E}} \le \|\widehat{\boldsymbol{\beta}}_\mu\|_{\mathcal{E}}\|\widetilde{\boldsymbol{\beta}}_\mu\|_{\mathcal{E}}$, which implies that $\|\widehat{\boldsymbol{\beta}}_\mu\|_{\mathcal{E}} \le \|\widetilde{\boldsymbol{\beta}}_\mu\|_{\mathcal{E}}$. Therefore,

$$\|\widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu\|_{n,N} \le \rho_n(nN)^{-1}\overline{A}_n\|\widetilde{\boldsymbol{\beta}}_\mu\|_{\mathcal{E}}. \tag{S1.25}$$

Combining (S1.21) and (S1.25) yields that

$$\|\widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu\|_\infty \le A_n\|\widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu\|_{n,N} \le \rho_n(nN)^{-1}A_n\overline{A}_n\|\widetilde{\boldsymbol{\beta}}_\mu\|_{\mathcal{E}}.$$

By Lemma S1.2, we have

$$\|\widetilde{\boldsymbol{\beta}}_\mu\|_{\mathcal{E}} = C_1\{\|\boldsymbol{\beta}^o\|_{2,\infty}+\sum_{a_1+a_2=2}\|\nabla_{z_1}^{a_1}\nabla_{z_2}^{a_2}(\boldsymbol{\beta}^o-\widetilde{\boldsymbol{\beta}}_\mu)\|_\infty\} \le C_2(\|\boldsymbol{\beta}^o\|_{2,\infty}+|\triangle|^{d-1}\|\boldsymbol{\beta}^o\|_{d+1,\infty}).$$

It follows

$$\|\widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu\|_\infty = \rho_n(nN)^{-1}A_n\overline{A}_nC_2(\|\boldsymbol{\beta}^o\|_{2,\infty} + |\triangle|^{d-1}\|\boldsymbol{\beta}^o\|_{d+1,\infty}). \tag{S1.26}$$

Next we derive the order of A_n and $\overline{A}_n$. By Markov's inequality, for any $\boldsymbol{g} \in \mathcal{G}^{(p+1)}$, $\|\boldsymbol{g}\|_\infty \le C|\triangle|^{-1}\|\boldsymbol{g}\|$, $\|\boldsymbol{g}\|_{\mathcal{E}} \le C|\triangle|^{-2}\|\boldsymbol{g}\|$. Equation (S1.8) implies that

$$\sup_{\boldsymbol{g}\in\mathcal{G}(\triangle)}\{\|\boldsymbol{g}\|_{n,N}/\,\|\boldsymbol{g}\|\} \ge \left[1 - O_P\left\{(\log n)^{1/2}n^{-1/2} + N^{-1/2}|\triangle|^{-1}\right\}\right]^{1/2}.$$

Thus, we have

$$A_n \le C|\triangle|^{-1}\left[1 - O_P\left\{(\log n)^{1/2}n^{-1/2} + N^{-1/2}|\triangle|^{-1}\right\}\right]^{-1/2} = O_P\left(|\triangle|^{-1}\right),$$

$$\overline{A}_n \le C|\triangle|^{-2}\left[1 - O_P\left\{(\log n)^{1/2}n^{-1/2} + N^{-1/2}|\triangle|^{-1}\right\}\right]^{-1/2} = O_P\left(|\triangle|^{-2}\right).$$

Plugging the order of A_n and $\overline{A}_n$ into (S1.26) yields that

$$\|\widetilde{\boldsymbol{\beta}}_\mu - \widehat{\boldsymbol{\beta}}_\mu\|_\infty = O_P\left\{\frac{C_2\rho_n}{nN|\triangle|^3}(\|\boldsymbol{\beta}^o\|_{2,\infty} + |\triangle|^{d-1}\|\boldsymbol{\beta}^o\|_{d+1,\infty})\right\}.$$

Hence,

$$\|\widehat{\boldsymbol{\beta}}_\mu - \boldsymbol{\beta}^o\|_\infty \le C_1|\triangle|^{d+1}\|\boldsymbol{\beta}^o\|_{d+1,\infty} + O_P\left\{\frac{C_2\rho_n}{nN|\triangle|^3}\left(\|\boldsymbol{\beta}^o\|_{2,\infty} + |\triangle|^{d-1}\|\boldsymbol{\beta}^o\|_{d+1,\infty}\right)\right\}.$$

Therefore, Lemma S1.10 is established. □

Lemma S1.11. *Suppose Assumptions (A2)–(A5) hold and $N^{1/2}|\triangle| \to \infty$ as $N \to \infty$, then $\|\widehat{\boldsymbol{\theta}}_\eta\|^2 = O_P(n^{-1}|\triangle|^{-2})$.*

Proof. Note that $\widehat{\boldsymbol{\theta}}_\eta = \boldsymbol{\Gamma}_{n,\rho}^{-1}\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}\left\{\widetilde{\mathbf{X}}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\}\sum_{k=1}^{\infty}\lambda_k^{1/2}\xi_{ik}\psi_k(\boldsymbol{z}_j)$. According to Lemma S1.9,

$$\|\widehat{\boldsymbol{\theta}}_\eta\|^2 \asymp \frac{1}{n^2N^2|\triangle|^4}\sum_{i,i'=1}^{n}\sum_{j,j'=1}^{N}\left\{\widetilde{\mathbf{X}}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\}^\top \times\sum_{k=1}^{\infty}\lambda_k^{1/2}\xi_{ik}\psi_k(\boldsymbol{z}_j)\left\{\mathbf{X}_{i'}\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\right\}\sum_{k=1}^{\infty}\lambda_k^{1/2}\xi_{i'k}\psi_k(\boldsymbol{z}_{j'}).$$

Note that

$$\widetilde{\mathbf{X}}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\sum_{k=1}^{\infty}\lambda_k^{1/2}\xi_{ik}\psi_k(\boldsymbol{z}_j) = \left(X_{i0}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^\top\sum_{k=1}^{\infty}\lambda_k^{1/2}\xi_{ik}\psi_k(\boldsymbol{z}_j),\ldots,X_{ip}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^\top\sum_{k=1}^{\infty}\lambda_k^{1/2}\xi_{ik}\psi_k(\boldsymbol{z}_j)\right)^\top,$$

so one has

$$\|\widehat{\boldsymbol{\theta}}_\eta\|^2 \asymp \frac{1}{n^2N^2|\triangle|^4}\sum_{\ell=0}^{p}\sum_{i,i'=1}^{n}\sum_{j,j'=1}^{N}X_{i\ell}X_{i'\ell}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^\top\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\sum_{k,k'=1}^{\infty}(\lambda_k\lambda_{k'})^{1/2}\xi_{ik}\psi_k(\boldsymbol{z}_j)\xi_{i'k'}\psi_{k'}(\boldsymbol{z}_{j'}).$$

Because the eigenvalues of $\mathbf{Q}_2\mathbf{Q}_2^\top$ are either 0 or 1, under Assumptions (A2) and (A3), for any ℓ, i,

$$\frac{1}{N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}E\left\{X_{i\ell}^2\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^\top\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\sum_{k,k'=1}^{\infty}(\lambda_k\lambda_{k'})^{1/2}\xi_{ik}\psi_k(\boldsymbol{z}_j)\xi_{ik'}\psi_{k'}(\boldsymbol{z}_{j'})\right\} \leq C\sum_{m\in\mathcal{M}}\frac{1}{N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}B_m(\boldsymbol{z}_j)B_m(\boldsymbol{z}_{j'})G_\eta(\boldsymbol{z}_j,\boldsymbol{z}_{j'}).$$

Assumption (A4) and (S1.5) imply that

$$\frac{1}{N^2}\sum_{j\neq j'}B_m(\boldsymbol{z}_j)B_m(\boldsymbol{z}_{j'})G_\eta(\boldsymbol{z}_j,\boldsymbol{z}_{j'}) = \int_{T_m\times T_m}G_\eta(\boldsymbol{z},\boldsymbol{z}')B_m(\boldsymbol{z})B_m(\boldsymbol{z}')d\boldsymbol{z}d\boldsymbol{z}' \times\{1+O(N^{-1/2}|\triangle|^3)\} = O(|\triangle|^4).$$

Thus,

$$\frac{1}{N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N} EX_{i\ell}^2\mathbf{B}(\boldsymbol{z}_j)^\top\mathbf{B}(\boldsymbol{z}_{j'})\sum_{k,k'=1}^{\infty}(\lambda_k\lambda_{k'})^{1/2}\xi_{ik}\psi_k(\boldsymbol{z}_j)\xi_{ik'}\psi_{k'}(\boldsymbol{z}_{j'}) \le C|\triangle|^2.$$

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

$$\begin{aligned}
&E\left\{X_{i\ell}X_{i'\ell}\mathbf{B}(\boldsymbol{z}_j)^\top\mathbf{B}(\boldsymbol{z}_{j'})\sum_{k,k'=1}^{\infty}(\lambda_k\lambda_{k'})^{1/2}\xi_{ik}\psi_k(\boldsymbol{z}_j)\xi_{ik'}\psi_{k'}(\boldsymbol{z}_{j'})\right\}\\
&\quad= E(X_{i\ell}X_{i'\ell})\sum_{m\in\mathcal{M}}B_m^2(\boldsymbol{z}_j)B_m^2(\boldsymbol{z}_{j'})\sum_{k,k'}E\left\{(\lambda_k\lambda_{k'})^{1/2}\xi_{ik}\xi_{i'k'}\psi_k(\boldsymbol{z}_j)\psi_{k'}(\boldsymbol{z}_{j'})\right\} = 0.
\end{aligned}$$

Therefore, $E\|\widehat{\boldsymbol{\theta}}_\eta\|^2 \le Cp(n^{-1}|\triangle|^{-2})$. The conclusion of the lemma follows. □

Lemma S1.12. *Suppose Assumptions (A2)–(A5) hold and $N^{1/2}|\triangle| \to \infty$ as $N \to \infty$, then $\|\widehat{\boldsymbol{\theta}}_\varepsilon\|^2 = O_P(n^{-1}N^{-1}|\triangle|^{-4})$.*

Proof. By the definition of $\widehat{\boldsymbol{\theta}}_\varepsilon$ in (S1.37), we have

$$\begin{aligned}
\|\widehat{\boldsymbol{\theta}}_\varepsilon\|^2 = &\frac{1}{n^2N^2|\triangle|^4}\left(|\triangle|^{-2}\boldsymbol{\Gamma}_{n,\rho}\right)^{-1}\sum_{i=1}^{n}\sum_{j=1}^{N}\left\{\widetilde{\mathbf{X}}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\}^\top\sigma(\boldsymbol{z}_j)\varepsilon_{ij}\\
&\times\left(|\triangle|^{-2}\boldsymbol{\Gamma}_{n,\rho}\right)^{-1}\sum_{i=1}^{n}\sum_{j'=1}^{N}\left\{\widetilde{\mathbf{X}}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\right\}\sigma(\boldsymbol{z}_{j'})\varepsilon_{ij'}.
\end{aligned}$$

By Lemma S1.9,

$$\|\widehat{\boldsymbol{\theta}}_\varepsilon\|^2 \asymp \frac{1}{n^2N^2|\triangle|^4}\sum_{i,i'=1}^{n}\sum_{j,j'=1}^{N}\left\{\widetilde{\mathbf{X}}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\}^\top\sigma(\boldsymbol{z}_j)\varepsilon_{ij}\left\{\mathbf{X}_{i'}\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\right\}\sigma(\boldsymbol{z}_{j'})\varepsilon_{i'j'}.$$

Note that

$$\widetilde{\mathbf{X}}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij} = \left(X_{i0}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^\top\sigma(\boldsymbol{z}_j)\varepsilon_{ij}, X_{i1}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^\top\sigma(\boldsymbol{z}_j)\varepsilon_{ij},\ldots,X_{ip}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^\top\sigma(\boldsymbol{z}_j)\varepsilon_{ij}\right)^\top,$$

so one has

$$\|\widetilde{\boldsymbol{\theta}}_\varepsilon\|^2 \asymp \frac{1}{n^2N^2|\triangle|^4}\sum_{\ell=0}^{p}\sum_{i,i'=1}^{n}\sum_{j,j'=1}^{N}X_{i\ell}X_{i'\ell}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^\top\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\sigma(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_{j'})\varepsilon_{ij}\varepsilon_{i'j}.$$

Because the eigenvalues of $\mathbf{Q}_2\mathbf{Q}_2^\top$ are either 0 or 1, under Assumption (A2), for any ℓ, i, by (S1.6),

$$\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^\top\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j) = \mathbf{B}(\boldsymbol{z}_j)^\top\mathbf{Q}_2\mathbf{Q}_2^\top\mathbf{B}(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j) \leq C\sum_{m\in\mathcal{M}}\frac{1}{N}\sum_{j=1}^{N}B_m^2(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j)$$
$$\leq C\sum_{m\in\mathcal{M}}\int_{T_{\lceil m/d^*\rceil}}\sigma^2(\boldsymbol{z})B_m^2(\boldsymbol{z})d\boldsymbol{z}\{1+O(N^{-1/2}|\triangle|^{-1})\} \leq C.$$

Next note that for any $\ell, i, j \neq j'$, $E\{X_{i\ell}^2\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^\top\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\varepsilon_{ij}\varepsilon_{ij'}\} = 0$, and for any ℓ, $i \neq i'$, j, j', $E\{X_{i\ell}X_{i'\ell}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^\top\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\sigma(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_{j'})\varepsilon_{ij}\varepsilon_{i'j'}\} = 0$. Therefore,

$$E\|\widehat{\boldsymbol{\theta}}_\varepsilon\|^2 \asymp \frac{1}{nN|\triangle|^4}\sum_{\ell=0}^{p}E(X_{i\ell}^2)\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^\top\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j) \leq Cp(nN)^{-1}|\triangle|^{-4}.$$

The conclusion of the lemma follows. □

Proof of Theorem 1. By Lemma S1.11, Lemma S1.12, and the properties of the bivariate spline basis functions in Lemma S1.1, $\|\widehat{\eta}_\ell\|_{L_2}^2 \asymp |\triangle|^2\|\widehat{\boldsymbol{\theta}}_{\eta,\ell}\|^2 = O_P(n^{-1})$ and $\|\widehat{\varepsilon}_\ell\|_{L_2}^2 \asymp |\triangle|^2\|\widehat{\boldsymbol{\theta}}_{\varepsilon,\ell}\|^2 = O_P(n^{-1}N^{-1}|\triangle|^{-2})$, for any $\ell = 0,1,\ldots,p$. It is clear that $\|\widehat{\beta}_\ell - \beta_\ell^o\|_{L_2}^2 \leq \|\widehat{\beta}_{\mu,\ell} - \beta_\ell^o\|_{L_2}^2 + \|\widehat{\eta}_\ell\|_{L_2}^2 + \|\widehat{\varepsilon}_\ell\|_{L_2}^2$, where the asymptotic order of $\|\widehat{\beta}_{\mu,\ell} - \beta_\ell^o\|_{L_2}$ is the same as $\|\widehat{\beta}_{\mu,\ell} - \beta_\ell^o\|_\infty$. The desired result follows from Lemma S1.10. □

Lemma S1.13. *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 $\mathbf{a} = (\mathbf{a}_0^\top,\ldots,\mathbf{a}_p^\top)^\top$ with $\mathbf{a}^\top\mathbf{a} = 1$, $[\mathrm{Var}\{\mathbf{a}^\top(\widehat{\boldsymbol{\theta}}_\eta + \widehat{\boldsymbol{\theta}}_\varepsilon)\}]^{-1/2}\{\mathbf{a}^\top(\widehat{\boldsymbol{\theta}}_\eta + \widehat{\boldsymbol{\theta}}_\varepsilon)\} \xrightarrow{\mathcal{L}} N(0,1)$, where $\widehat{\boldsymbol{\theta}}_\eta$ and $\widehat{\boldsymbol{\theta}}_\varepsilon$ are given in (S1.37).*

Proof. For coefficient vectors $\widehat{\boldsymbol{\theta}}_\eta$, $\widehat{\boldsymbol{\theta}}_\varepsilon$ and the matrix $\boldsymbol{\Gamma}_{n,\rho}$ defined in (S1.17), $\mathrm{Var}\{\mathbf{a}^\top(\widehat{\boldsymbol{\theta}}_\eta+$

$\widehat{\boldsymbol{\theta}}_\varepsilon)\} = \mathbf{a}^\top\{E(\widehat{\boldsymbol{\theta}}_\eta\widehat{\boldsymbol{\theta}}_\eta^\top) + E(\widehat{\boldsymbol{\theta}}_\varepsilon\widehat{\boldsymbol{\theta}}_\varepsilon^\top)\}\mathbf{a}$. Denote $\boldsymbol{\Psi}_\eta = (\boldsymbol{\Psi}_{\eta,\ell,\ell'})_{\ell,\ell'}$ and $\boldsymbol{\Psi}_\varepsilon = (\boldsymbol{\Psi}_{\epsilon,\ell,\ell'})_{\ell,\ell'}$, with

$$\boldsymbol{\Psi}_{\eta,\ell,\ell'} = \frac{1}{n^2N^2}\sum_{i=1}^n\sum_{j=1}^N\sum_{j'=1}^N X_{i\ell}X_{i\ell'}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}^\top(\boldsymbol{z}_{j'})G_\eta(\boldsymbol{z}_j,\boldsymbol{z}_{j'}),$$
$$\boldsymbol{\Psi}_{\varepsilon,\ell,\ell'} = \frac{1}{n^2N^2}\sum_{i=1}^n\sum_{j=1}^N X_{i\ell}X_{i\ell'}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}^\top(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j),$$

then, we have

$$\begin{aligned}\mathbf{a}^\top E(\widehat{\boldsymbol{\theta}}_\eta\widehat{\boldsymbol{\theta}}_\eta^\top)\mathbf{a} =& E\mathbf{a}^\top\boldsymbol{\Gamma}_{n,\rho}^{-1}\frac{1}{n^2N^2}\sum_{i=1}^n\sum_{j,j'=1}^N\left\{\widetilde{\mathbf{X}}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\}\left\{\widetilde{\mathbf{X}}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\right\}^\top G_\eta(\boldsymbol{z}_j,\boldsymbol{z}_{j'})\boldsymbol{\Gamma}_{n,\rho}^{-1}\mathbf{a}\\ =& E\mathbf{a}^\top\boldsymbol{\Gamma}_{n,\rho}^{-1}\boldsymbol{\Psi}_\eta\boldsymbol{\Gamma}_{n,\rho}^{-1}\mathbf{a},\\ \mathbf{a}^\top E(\widehat{\boldsymbol{\theta}}_\varepsilon\widehat{\boldsymbol{\theta}}_\varepsilon^\top)\mathbf{a} =& E\mathbf{a}^\top\boldsymbol{\Gamma}_{n,\rho}^{-1}\frac{1}{n^2N^2}\sum_{i=1}^n\sum_{j=1}^N\left\{\widetilde{\mathbf{X}}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\}\left\{\widetilde{\mathbf{X}}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\}^\top\sigma^2(\boldsymbol{z}_j)\boldsymbol{\Gamma}_{n,\rho}^{-1}\mathbf{a}\\ =& E\mathbf{a}^\top\boldsymbol{\Gamma}_{n,\rho}^{-1}\boldsymbol{\Psi}_\varepsilon\boldsymbol{\Gamma}_{n,\rho}^{-1}\mathbf{a}.\end{aligned}$$

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

$$\begin{aligned}(a_i^{\eta+\varepsilon})^2 =& \mathbf{a}^\top\boldsymbol{\Gamma}_{n,\rho}^{-1}\frac{1}{n^2N^2}\sum_{j=1}^N\sum_{j'=1}^N\left\{\widetilde{\mathbf{X}}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\}\left\{\widetilde{\mathbf{X}}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\right\}^\top G_\eta(\boldsymbol{z}_j,\boldsymbol{z}_{j'})\boldsymbol{\Gamma}_{n,\rho}^{-1}\mathbf{a}\\ &+\mathbf{a}^\top\boldsymbol{\Gamma}_{n,\rho}^{-1}\frac{1}{n^2N^2}\sum_{j=1}^N\left\{\widetilde{\mathbf{X}}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\}\left\{\widetilde{\mathbf{X}}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\}^\top\sigma^2(\boldsymbol{z}_j)\boldsymbol{\Gamma}_{n,\rho}^{-1}\mathbf{a} = (a_i^\eta)^2 + (a_i^\varepsilon)^2,\end{aligned}$$

and conditional on $\{\widetilde{\mathbf{X}}_i, i = 1,\ldots,n\}$, $\mathfrak{z}_i$ are independent with mean zero and variance one. Thus, $\sum_{i=1}^n(a_i^\eta)^2 = \mathbf{a}^\top\boldsymbol{\Gamma}_{n,\rho}^{-1}\boldsymbol{\Psi}_\eta\boldsymbol{\Gamma}_{n,\rho}^{-1}\mathbf{a}$ and $\sum_{i=1}^n(a_i^\varepsilon)^2 = \mathbf{a}^\top\boldsymbol{\Gamma}_{n,\rho}^{-1}\boldsymbol{\Psi}_\varepsilon\boldsymbol{\Gamma}_{n,\rho}^{-1}\mathbf{a}$.

According to Lemma S1.9, Assumptions (A2) and (A4),

$$E\mathbf{a}^\top\boldsymbol{\Gamma}_{n,\rho}^{-1}\boldsymbol{\Psi}_\eta\boldsymbol{\Gamma}_{n,\rho}^{-1}\mathbf{a} \geq c_\Gamma^{-2}\left(|\triangle|^2 + \frac{\rho_n}{nN|\triangle|^2}\right)^{-2}E\mathbf{a}^\top\boldsymbol{\Psi}_\eta\mathbf{a},$$

where

$$\mathbf{a}^\top \boldsymbol{\Psi}_\eta \mathbf{a} = \frac{1}{n^2 N^2} \sum_{\ell,\ell'=0}^{p} \sum_{i=1}^{n} \sum_{j=1}^{N} \sum_{j'=1}^{N} X_{i\ell} X_{i\ell'} \boldsymbol{a}_\ell^\top \widetilde{\mathbf{B}}(\boldsymbol{z}_j) \widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})^\top \boldsymbol{a}_{\ell'} G_\eta(\boldsymbol{z}_j, \boldsymbol{z}_{j'})$$
$$= \frac{1}{n^2} \sum_{k=1}^{\infty} \sum_{i=1}^{n} \left\{ \frac{1}{N} \sum_{\ell=0}^{p} \sum_{j=1}^{N} \lambda_k^{1/2} X_{i\ell} g_\ell(\boldsymbol{z}_j) \psi_k(\boldsymbol{z}_j) \right\}^2$$

with $g_\ell(\boldsymbol{z}) = \boldsymbol{a}_\ell^\top \widetilde{\mathbf{B}}(\boldsymbol{z})$. Therefore, by Assumption (A3), we have

$$E\mathbf{a}^\top \boldsymbol{\Psi}_\eta \mathbf{a} = \frac{1}{n} \sum_{k=1}^{\infty} \sum_{\ell=0}^{p} \left\{ \frac{1}{N} \sum_{j=1}^{N} \lambda_k^{1/2} g_\ell(\boldsymbol{z}_j) \psi_k(\boldsymbol{z}_j) \right\}^2$$
$$\geq \frac{c}{nN^2} \sum_{\ell=0}^{p} \sum_{j=1}^{N} \sum_{j'=1}^{N} g_\ell(\boldsymbol{z}_j) g_\ell(\boldsymbol{z}_{j'}) G_\eta(\boldsymbol{z}_j, \boldsymbol{z}_{j'})$$
$$\asymp \frac{1}{n} \sum_{\ell=0}^{p} \int_{\Omega^2} g_\ell(\boldsymbol{z}) g_\ell(\boldsymbol{z}') G_\eta(\boldsymbol{z}, \boldsymbol{z}') d\boldsymbol{z} d\boldsymbol{z}'.$$

Noting that the eigenvalues of G_η are strictly positive, we have

$$E\mathbf{a}^\top \boldsymbol{\Psi}_\eta \mathbf{a} \geq c_1 n^{-1} \sum_{\ell=0}^{p} \int_\Omega g_\ell^2(\boldsymbol{z}) d\boldsymbol{z} \geq c_2 n^{-1} |\triangle|^2 \|\mathbf{a}\|^2.$$

Therefore, we have $E\mathbf{a}^\top \boldsymbol{\Gamma}_{n,\rho}^{-1} \boldsymbol{\Psi}_\eta \boldsymbol{\Gamma}_{n,\rho}^{-1} \mathbf{a} \geq cn^{-1} \left(1 + \frac{\rho_n}{nN|\triangle|^4}\right)^{-2} |\triangle|^{-2}$. Similarly, one can show that $E\mathbf{a}^\top \boldsymbol{\Gamma}_{n,\rho}^{-1} \boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{n,\rho}^{-1} \mathbf{a} \geq c(nN)^{-1} \left(1 + \frac{\rho_n}{nN|\triangle|^4}\right)^{-2} |\triangle|^{-2}$. In addition,

$$\max(a_i^\eta)^2 \leq \frac{C}{|\triangle|^4} \mathbf{a}^\top \frac{1}{n^2 N^2} \sum_{j=1}^{N} \sum_{j'=1}^{N} \left\{ (\widetilde{\mathbf{X}}_i \widetilde{\mathbf{X}}_i^\top) \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_j) \widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})^\top \right\} G_\eta(\boldsymbol{z}_j, \boldsymbol{z}_{j'}) \mathbf{a}$$
$$\leq \frac{C}{|\triangle|^4} \sum_{\ell,\ell'=1}^{p} \frac{1}{n^2 N^2} \max_i |X_{i\ell} X_{i\ell'}| \sum_{j=1}^{N} \sum_{j'=1}^{N} g_\ell(\boldsymbol{z}_j) g_{\ell'}(\boldsymbol{z}_{j'}) G_\eta(\boldsymbol{z}_j, \boldsymbol{z}_{j'}) \leq Cn^{-2} |\triangle|^{-2},$$

$$\max(a_i^\varepsilon)^2 \leq \frac{C}{|\triangle|^4} \mathbf{a}^\top \frac{1}{n^2 N^2} \sum_{j=1}^{N} \left\{ (\widetilde{\mathbf{X}}_i \widetilde{\mathbf{X}}_i^\top) \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_j) \widetilde{\mathbf{B}}(\boldsymbol{z}_j)^\top \right\} \sigma^2(\boldsymbol{z}_j) \mathbf{a}$$
$$\leq \frac{C}{|\triangle|^4} \sum_{\ell,\ell'=1}^{p} \frac{1}{n^2 N^2} \max_i |X_{i\ell} X_{i\ell'}| \sum_{j=1}^{N} g_\ell(\boldsymbol{z}_j) g_{\ell'}(\boldsymbol{z}_{j'}) \sigma^2(\boldsymbol{z}_j) \leq Cn^{-2} N^{-1} |\triangle|^{-2}.$$

Thus, if $\rho_n n^{-1} N^{-1} |\triangle|^{-4} \to 0$, we have

$$\frac{\max_{1 \leq i \leq n} (a_i^\eta + a_i^\varepsilon)^2}{\sum_{i=1}^{n} (a_i^\eta + a_i^\varepsilon)^2} \leq Cn^{-1} \left(1 + \frac{\rho_n}{nN|\triangle|^4}\right)^2 \to 0,$$

which satisfies the Lindeberg condition. $\square$

Theorem S1.2. *Under Assumptions (A1)–(A6), if for any $\ell = 0, 1, \ldots, p$, $|X_{i\ell}| \leq C_\ell < \infty$, $\sup_{\boldsymbol{z}\in\Omega}[\mathrm{Var}\{\widehat{\beta}_\ell(\boldsymbol{z})\}]^{-1/2}(\widehat{\beta}_{\mu,\ell}(\boldsymbol{z}) - \beta_\ell^o(\boldsymbol{z})) = o_P(1)$, for $\ell = 0, \ldots, p$.*

Proof. Using similar arguments as in the proof of Lemma S1.13 and the result of Lemma S1.9, we have for any $\|\mathbf{a}\| = 1$, $E\mathbf{a}^\top \boldsymbol{\Gamma}_{n,\rho}^{-1}\boldsymbol{\Psi}_\eta \boldsymbol{\Gamma}_{n,\rho}^{-1}\mathbf{a} \leq C_\Gamma^{-2}|\triangle|^{-4}E\mathbf{a}^\top\boldsymbol{\Psi}_\eta\mathbf{a} \leq Cn^{-1}|\triangle|^{-2}$, and $E\mathbf{a}^\top \boldsymbol{\Gamma}_{n,\rho}^{-1}\boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{n,\rho}^{-1}\mathbf{a} \leq C_\Gamma^{-2}|\triangle|^{-4}E\mathbf{a}^\top\boldsymbol{\Psi}_\varepsilon\mathbf{a} \leq C(nN)^{-1}|\triangle|^{-2}$. Therefore, based on the proof of Lemma S1.13, for any $\|\mathbf{a}\| = 1$,

$$cn^{-1}|\triangle|^{-4}\left(1 + \frac{\rho_n}{nN|\triangle|^4}\right)^{-2} \leq E\mathbf{a}^\top \boldsymbol{\Gamma}_{n,\rho}^{-1}\boldsymbol{\Psi}_\eta \boldsymbol{\Gamma}_{n,\rho}^{-1}\mathbf{a} \leq Cn^{-1}|\triangle|^{-2},$$
$$c(nN)^{-1}\left(1 + \frac{\rho_n}{nN|\triangle|^4}\right)^{-2}|\triangle|^{-2} \leq E\mathbf{a}^\top \boldsymbol{\Gamma}_{n,\rho}^{-1}\boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{n,\rho}^{-1}\mathbf{a} \leq C(nN)^{-1}|\triangle|^{-2}.$$

Thus,

$$\begin{aligned}
&\mathrm{Var}(\widehat{\beta}_\ell) = \{\mathbf{e}_\ell \otimes \widetilde{\mathbf{B}}(\boldsymbol{z})\}^\top E\{\boldsymbol{\Gamma}_{n,\rho}^{-1}(\boldsymbol{\Psi}_\eta + \boldsymbol{\Psi}_\varepsilon)\boldsymbol{\Gamma}_{n,\rho}^{-1}\}\{\mathbf{e}_\ell \otimes \widetilde{\mathbf{B}}(\boldsymbol{z})\} \\
&\asymp \{\mathbf{e}_\ell \otimes \widetilde{\mathbf{B}}(\boldsymbol{z})\}^\top E(\boldsymbol{\Gamma}_{n,\rho}^{-1}\boldsymbol{\Psi}_\eta\boldsymbol{\Gamma}_{n,\rho}^{-1})\{\mathbf{e}_\ell \otimes \widetilde{\mathbf{B}}(\boldsymbol{z})\} \asymp \{\mathbf{e}_\ell \otimes \widetilde{\mathbf{B}}(\boldsymbol{z})\}^\top \\
&\quad \times E\left[\boldsymbol{\Gamma}_{n,\rho}^{-1}\frac{1}{n^2N^2}\sum_{i=1}^n\sum_{j,j'=1}^N\left\{\widetilde{\mathbf{X}}_i \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\}\left\{\widetilde{\mathbf{X}}_i \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\right\}^\top G_\eta(\boldsymbol{z}_j, \boldsymbol{z}_{j'})\boldsymbol{\Gamma}_{n,\rho}^{-1}\right]\{\mathbf{e}_\ell \otimes \widetilde{\mathbf{B}}(\boldsymbol{z})\}.
\end{aligned}$$

By Lemma S1.9, we have

$$\begin{aligned}
\mathrm{Var}(\widehat{\beta}_\ell) &\lesssim \frac{1}{nN^2|\triangle|^2}\sum_{j=1}^N\sum_{j'=1}^N\{\mathbf{e}_\ell \otimes \widetilde{\mathbf{B}}(\boldsymbol{z})\}^\top\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})^\top\{\mathbf{e}_\ell \otimes \widetilde{\mathbf{B}}(\boldsymbol{z})\}G_\eta(\boldsymbol{z}_j, \boldsymbol{z}_{j'}), \\
\mathrm{Var}(\widehat{\beta}_\ell) &\gtrsim \frac{1}{nN^2|\triangle|^2}\sum_{j=1}^N\sum_{j'=1}^N\{\mathbf{e}_\ell \otimes \widetilde{\mathbf{B}}(\boldsymbol{z})\}^\top\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})^\top\{\mathbf{e}_\ell \otimes \widetilde{\mathbf{B}}(\boldsymbol{z})\}G_\eta(\boldsymbol{z}_j, \boldsymbol{z}_{j'}) \\
&\quad \times \left(1 + \frac{\rho_n}{nN|\triangle|^4}\right)^{-2},
\end{aligned}$$

and according to Lemmas S1.1 and S1.4, we have $cn^{-1}\left(1 + \frac{\rho_n}{nN|\triangle|^4}\right)^{-2} \leq \mathrm{Var}(\widehat{\beta}_\ell) \leq$

Cn^{-1}. According to Lemma S1.10, 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 (S1.19) is negligible compared to the order of $[\mathrm{Var}\{\widehat{\beta}_\ell(\boldsymbol{z})\}]^{1/2}$. □

Proof of Theorem 2. Theorem 2 follows from (S1.19), Lemma S1.13 and Theorem S1.2. □

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

$$\widehat{\boldsymbol{\beta}}_\mu^{\mathrm{c}}(\boldsymbol{z}) = (\widehat{\beta}_{\mu,0}^{\mathrm{c}}(\boldsymbol{z}), \ldots, \widehat{\beta}_{\mu,p}^{\mathrm{c}}(\boldsymbol{z}))^\top = \widehat{\mathbf{V}}_{m(\boldsymbol{z})}^{-1} \left\{ \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} B_{m(\boldsymbol{z})}(\boldsymbol{z}_j) X_{i\ell} \sum_{\ell'=0}^{p} \beta_{\ell'}^{o}(\boldsymbol{z}_j) X_{i\ell'} \right\}_{\ell=0}^{p}, \tag{S1.27}$$

$$\widehat{\boldsymbol{\eta}}(\boldsymbol{z}) = (\widehat{\eta}_0(\boldsymbol{z}), \ldots, \widehat{\eta}_p(\boldsymbol{z}))^\top = \widehat{\mathbf{V}}_{m(\boldsymbol{z})}^{-1} \left\{ \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} B_{m(\boldsymbol{z})}(\boldsymbol{z}_j) X_{i\ell} \sum_{k=1}^{\infty} \xi_{ik} \psi_k(\boldsymbol{z}_j) \right\}_{\ell=0}^{p}, \tag{S1.28}$$

$$\widehat{\boldsymbol{\varepsilon}}(\boldsymbol{z}) = (\widehat{\varepsilon}_0(\boldsymbol{z}), \ldots, \widehat{\varepsilon}_p(\boldsymbol{z}))^\top = \widehat{\mathbf{V}}_{m(\boldsymbol{z})}^{-1} \left\{ \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} B_{m(\boldsymbol{z})}(\boldsymbol{z}_j) X_{i\ell} \varepsilon_{ij} \right\}_{\ell=0}^{p}, \tag{S1.29}$$

where $\widehat{\mathbf{V}}_{m(\boldsymbol{z})}$ is defined in (2.7).

The next two theorems concern the functions $\widehat{\beta}_{\mu,\ell}^{\mathrm{c}}(\boldsymbol{z})$, $\widehat{\eta}_\ell(\boldsymbol{z})$, $\widehat{\varepsilon}_\ell(\boldsymbol{z})$, $\ell = 0, \ldots, p$, given in (S1.27), (S1.28) and (S1.29). Theorem S1.3 gives the uniform convergence rate of $\widehat{\beta}_{\mu,\ell}(\boldsymbol{z})$ to $\beta_\ell^o(\boldsymbol{z})$.

Theorem S1.3. *Under Assumptions (A1′), (A2)–(A6), the constant spline functions* $\widehat{\beta}_{\mu,\ell}^{\mathrm{c}}(\boldsymbol{z})$, $\ell = 0, \ldots, p$, *satisfy* $\sup_{\boldsymbol{z}\in\Omega} \sup_{0\le\ell\le p} \left| \widehat{\beta}_{\mu,\ell}^{\mathrm{c}}(\boldsymbol{z}) - \beta_\ell^o(\boldsymbol{z}) \right| = O_P(|\triangle|)$.

In the following, we provide detailed proofs of Theorems S1.3. For the random matrix $\widehat{\mathbf{V}}_m$ defined in (2.7), the lemma below shows that its inverse can be approximated by the inverse of a deterministic matrix $A_m^{-1}\mathbf{\Sigma}_X^{-1}$, where $A_m = \int_\Omega B_m(\boldsymbol{z})d\boldsymbol{z}$.

Lemma S1.14. *Under Assumptions (A3) and (A5), for any $m \in \mathcal{M}$, we have*

$$\widehat{\mathbf{V}}_m^{-1} = A_m^{-1}\mathbf{\Sigma}_X^{-1} + O_P\left\{n^{-1/2}|\triangle|^2(\log n)^{1/2} + N^{-1/2}|\triangle|\right)\right\}. \tag{S1.30}$$

Proof. By Lemma S1.5, $\left\|\widehat{\mathbf{V}}_m - A_m\mathbf{\Sigma}_X\right\|_\infty = O_P\left\{n^{-1/2}|\triangle|^2(\log n)^{1/2} + N^{-1/2}|\triangle|\right)\right\}$. Using the fact that for any matrices $\mathbf{A}$ and $\mathbf{B}$, $(\mathbf{A}+\delta\mathbf{B})^{-1} = \mathbf{A}^{-1} - \delta\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1} + O(\delta^2)$, we obtain (S1.30). □

Proof of Theorem S1.3. According to Lemma S1.2, there exist functions $\beta_\ell^* \in \mathcal{PC}(\triangle)$ that satisfies $\|\beta_\ell^* - \beta_\ell^o\|_\infty = O(|\triangle|)$ for $\ell = 0, 1, \ldots, p$. By the definition of $\widehat{\beta}_{\mu,\ell}(\boldsymbol{z})$ in (S1.27), $\widehat{\boldsymbol{\beta}}_\mu^{\mathrm{c}}(\boldsymbol{z}) = \left(\widehat{\beta}_{\mu,0}^{\mathrm{c}}(\boldsymbol{z}), \widehat{\beta}_{\mu,1}^{\mathrm{c}}(\boldsymbol{z}), \ldots, \widehat{\beta}_{\mu,p}^{\mathrm{c}}(\boldsymbol{z})\right)^\top = \left(\widetilde{\gamma}_{m(\boldsymbol{z}),0}, \ldots, \widetilde{\gamma}_{m(\boldsymbol{z}),p}\right)^\top = \widetilde{\boldsymbol{\gamma}}_{m(\boldsymbol{z})}$, where $\widetilde{\boldsymbol{\gamma}}_m = \widehat{\mathbf{V}}_m^{-1}\left\{(nN)^{-1}\sum_{i=1}^n\sum_{j=1}^N B_m(\boldsymbol{z}_j)X_{i\ell}\sum_{\ell'=0}^p \beta_{\ell'}^o(\boldsymbol{z}_j)X_{i\ell'}\right\}_{\ell=0}^p$ for $\widehat{\mathbf{V}}_m$ defined in (2.7).

Let

$$\widetilde{\boldsymbol{\beta}}(\boldsymbol{z}) = (\widetilde{\beta}_0(\boldsymbol{z}), \widetilde{\beta}_1(\boldsymbol{z}), \ldots, \widetilde{\beta}_p(\boldsymbol{z}))^\top = \widehat{\mathbf{V}}_{m(\boldsymbol{z})}^{-1}\left[\frac{1}{nN}\sum_{i=1}^n\sum_{j=1}^N B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)X_{i\ell}\sum_{\ell'=0}^p \beta_{\ell'}^*(\boldsymbol{z}_j)X_{i\ell'}\right]_{\ell=0}^p,$$

then

$$\widehat{\boldsymbol{\beta}}_\mu^{\mathrm{c}}(\boldsymbol{z}) - \widetilde{\boldsymbol{\beta}}(\boldsymbol{z}) = \widehat{\mathbf{V}}_{m(\boldsymbol{z})}^{-1}\left[\frac{1}{nN}\sum_{i=1}^n\sum_{j=1}^N B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)X_{i\ell}\sum_{\ell'=0}^p \left\{\beta_{\ell'}^o(\boldsymbol{z}_j) - \beta_{\ell'}^*(\boldsymbol{z}_j)\right\}X_{i\ell'}\right]_{\ell=0}^p.$$

Observing that $\widetilde{\beta}_\ell \equiv \beta_\ell^*$ as $\beta_\ell^* \in \mathcal{PC}(\triangle)$, $\widehat{\beta}_{\mu,\ell}^{\mathrm{c}}(\boldsymbol{z}) = \widehat{\beta}_{\mu,\ell}^{\mathrm{c}}(\boldsymbol{z}) - \widetilde{\beta}_\ell(\boldsymbol{z}) + \beta_\ell^*(\boldsymbol{z})$, $\ell = 0, 1, \ldots, p$.

It is easy to see $\|\widehat{\beta}_{\mu,\ell}^{\mathrm{c}} - \widetilde{\beta}_\ell\|_\infty = O_P(|\triangle|)$. Hence, for $\ell = 0, 1, \ldots, p$, $\|\widehat{\beta}_{\mu,\ell}^{\mathrm{c}} - \beta_\ell^o\|_\infty \leq \|\widehat{\beta}_{\mu,\ell}^{\mathrm{c}} - \widetilde{\beta}_\ell\|_\infty + \|\beta_\ell^o - \beta_\ell^*\|_\infty = O_P(|\triangle|)$, which completes the proof. □

By Lemma S1.14, the inverse of the random matrix $\widehat{\mathbf{V}}_m$ can be approximated by that of a deterministic matrix $A_m\mathbf{\Sigma}_X$. Substituting $\widehat{\mathbf{V}}_m$ with $A_m\mathbf{\Sigma}_X$ in (S1.28) and (S1.29), we define the random vectors

$$\widehat{\boldsymbol{\eta}}^*(\boldsymbol{z}) = (\widehat{\eta}_0^*(\boldsymbol{z}),\ldots,\widehat{\eta}_p^*(\boldsymbol{z}))^\top = A_{m(\boldsymbol{z})}^{-1}\mathbf{\Sigma}_X^{-1}\left\{\frac{1}{nN}\sum_{i=1}^n\sum_{j=1}^N B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)X_{i\ell}\sum_{k=1}^\infty \xi_{ik}\psi_k(\boldsymbol{z}_j)\right\}_{\ell=0}^p, \tag{S1.31}$$

$$\widehat{\boldsymbol{\varepsilon}}^*(\boldsymbol{z}) = (\widehat{\varepsilon}_0^*(\boldsymbol{z}),\ldots,\widehat{\varepsilon}_p^*(\boldsymbol{z}))^\top = A_{m(\boldsymbol{z})}^{-1}\mathbf{\Sigma}_X^{-1}\left\{\frac{1}{nN}\sum_{i=1}^n\sum_{j=1}^N B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)X_{i\ell}\varepsilon_{ij}\right\}_{\ell=0}^p. \tag{S1.32}$$

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

Lemma S1.15. *Under Assumptions (A2)–(A5) and (C1), if $N^{1/2}|\triangle|\to\infty$ as $N\to\infty$, $\|\sum_{k=1}^\infty \lambda_k^{1/2}\psi_k\|_\infty<\infty$ and $n^{1/(4+\delta_2)}\ll n^{1/2}N^{-1/2}|\triangle|^{-1}$ for some δ_2, then for $\widehat{\boldsymbol{\eta}}(\boldsymbol{z})$, $\widehat{\boldsymbol{\varepsilon}}(\boldsymbol{z})$ given in (S1.28), (S1.29) and $\widehat{\boldsymbol{\eta}}^*(\boldsymbol{z})$, $\widehat{\boldsymbol{\varepsilon}}^*(\boldsymbol{z})$ given in (S1.31), (S1.32), as $N\to\infty$ and $n\to\infty$, we have*

$$\sup_{\boldsymbol{z}\in\Omega}\|\widehat{\boldsymbol{\eta}}(\boldsymbol{z})-\widehat{\boldsymbol{\eta}}^*(\boldsymbol{z})\|_\infty = O_P\left\{n^{-1}|\triangle|^4\log(n)+n^{-1/2}N^{-1/2}|\triangle|^3(\log n)^{1/2}\right\}, \tag{S1.33}$$

$$\sup_{\boldsymbol{z}\in\Omega}\|\widehat{\boldsymbol{\varepsilon}}(\boldsymbol{z})-\widehat{\boldsymbol{\varepsilon}}^*(\boldsymbol{z})\|_\infty = O_P\left\{n^{-1}N^{-1/2}|\triangle|^3\log(n)+n^{-1/2}N^{-1}|\triangle|^2(\log n)^{1/2}\right\}. \tag{S1.34}$$

Proof. Comparing $\widehat{\boldsymbol{\eta}}(\boldsymbol{z})$ and $\widehat{\boldsymbol{\eta}}^*(\boldsymbol{z})$ given in (S1.28) and (S1.31), we have

$$\widehat{\boldsymbol{\eta}}(\boldsymbol{z})-\widehat{\boldsymbol{\eta}}^*(\boldsymbol{z}) = \left\{\widehat{\mathbf{V}}_m^{-1}-A_{m(\boldsymbol{z})}^{-1}\mathbf{\Sigma}_X^{-1}\right\}\left\{\frac{1}{nN}\sum_{i=1}^n\sum_{j=1}^N B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)X_{i\ell}\sum_{k=1}^\infty \xi_{ik}\psi_k(\boldsymbol{z}_j)\right\}_{\ell=0}^p.$$

Now let $\zeta_{i,m,\ell}\equiv\zeta_i = n^{-1}\left[X_{i\ell}\sum_{k=1}^\infty\left\{\frac{1}{N}\sum_{j=1}^N B_m(\boldsymbol{z}_j)\psi_k(\boldsymbol{z}_j)\right\}\xi_{ik}\right]$, then it is easy to see that $\frac{1}{nN}\sum_{i=1}^n\sum_{j=1}^N B_m(\boldsymbol{z}_j)X_{i\ell}\sum_{k=1}^\infty \xi_{ik}\psi_k(\boldsymbol{z}_j) = \frac{1}{N}\sum_{i=1}^n \zeta_{i,m,\ell}$. It is easy to see that

$E(\zeta_i) = 0$, and

$$\sigma^2_{\zeta_i,n} = E\left(\zeta_i^2\right) = n^{-2}E(X_\ell^2)\int_{T_m\times T_m} G_\eta\left(\boldsymbol{u},\boldsymbol{v}\right) d\boldsymbol{u}d\boldsymbol{v}\{1+O(N^{-1/2}|\triangle|^{-1})\}.$$

Note that $\left\{\sigma^{-1}_{\zeta_i,n}\zeta_i\right\}_{i=1}^n$ are uncorrelated random variables with mean 0. Assume that $|\triangle|^{-2} \asymp n^\tau$ for some $0<\tau<\infty$, we can show that for any large enough $\delta>0$, $P\left[|\sum_{i=1}^n \zeta_i| \geq \delta\left\{C\log(n)n^{-1}|\triangle|^4 E(X_{i\ell}^2)\right\}^{1/2}\right] \leq 2n^{-2-\tau}$. Therefore,

$$\sum_{n=1}^\infty P\left\{\sup_{m\in\mathcal{M},0\leq\ell\leq p}\left|\sum_{i=1}^n \zeta_{i,m,\ell}\right| \geq \delta n^{-1/2}|\triangle|^2(\log n)^{1/2}\right\} < \infty.$$

Thus, $\sup_{m,\ell}|\sum_{i=1}^n \zeta_{i,m,\ell}| = O_P\left\{n^{-1/2}|\triangle|^2(\log n)^{1/2}\right\}$ as $n\to\infty$ by Borel-Cantelli Lemma. It follows that $\sup_{m,\ell}|n^{-1}\sum_{i=1}^n \zeta_{i,m,\ell}| = O_P\left\{n^{-1/2}|\triangle|^2(\log n)^{1/2}\right\}$. Finally, according to (S1.30), we obtain (S1.33). The result in (S1.34) can be proved similarly. □

Lemma S1.16. *For any $\boldsymbol{z}\in\Omega$, the covariance matrices of $\widehat{\boldsymbol{\eta}}^*(\boldsymbol{z})$ and $\widehat{\boldsymbol{\varepsilon}}^*(\boldsymbol{z})$ are*

$$\boldsymbol{\Sigma}_\eta(\boldsymbol{z}) = E\left\{\widehat{\boldsymbol{\eta}}^*(\boldsymbol{z})\widehat{\boldsymbol{\eta}}^{*\top}(\boldsymbol{z})\right\} = A^{-2}_{m(\boldsymbol{z})}\boldsymbol{\Sigma}_X^{-1}\frac{1}{nN^2}\sum_{k=1}^\infty \lambda_k\left\{\sum_{j=1}^N B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)\psi_k(\boldsymbol{z}_j)\right\}^2,$$

$$\boldsymbol{\Sigma}_\varepsilon(\boldsymbol{z}) = E\left\{\widehat{\boldsymbol{\varepsilon}}^*(\boldsymbol{z})\widehat{\boldsymbol{\varepsilon}}^{*\top}(\boldsymbol{z})\right\} = A^{-2}_{m(\boldsymbol{z})}\boldsymbol{\Sigma}_X^{-1}\frac{1}{nN^2}\sum_{j=1}^N B^2_{m(\boldsymbol{z})}(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j),$$

in addition,

$$\sup_{\boldsymbol{z}\in\Omega}\|\boldsymbol{\Sigma}_\eta(\boldsymbol{z})+\boldsymbol{\Sigma}_\varepsilon(\boldsymbol{z})-\boldsymbol{\Sigma}_n(\boldsymbol{z})\|_\infty = O(n^{-1}N^{-1/2}|\triangle|^{-1}), \tag{S1.35}$$

where $\boldsymbol{\Sigma}_n(\boldsymbol{z})$ is given in (2.9).

Proof. Note that $A^2_{m(\boldsymbol{z})}\widehat{\boldsymbol{\eta}}^*(\boldsymbol{z})\widehat{\boldsymbol{\eta}}^{*\top}(\boldsymbol{z})$ is equal to

$$\boldsymbol{\Sigma}_X^{-1}\left\{\frac{1}{n^2N^2}\sum_{i=1}^n\sum_{j=1}^N B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)X_{i\ell}\sum_{k=1}^\infty \xi_{ik}\psi_k(\boldsymbol{z}_j)\sum_{i'=1}^n\sum_{j'=1}^N B_{m(\boldsymbol{z})}(\boldsymbol{z}_{j'})X_{i'\ell'}\sum_{k'=1}^\infty \xi_{i'k'}\psi_{k'}(\boldsymbol{z}_{j'})\right\}_{\ell,\ell'=0}^p\boldsymbol{\Sigma}_X^{-1}.$$

Thus,

$$\mathbf{\Sigma}_{\eta}(\boldsymbol{z}) = E\left\{\widehat{\boldsymbol{\eta}}^{*}(\boldsymbol{z})\widetilde{\boldsymbol{\eta}}^{\top}(\boldsymbol{z})\right\} = A_{m(\boldsymbol{z})}^{-2}\mathbf{\Sigma}_X^{-1}\frac{1}{nN^2}\sum_{k=1}^{\infty}\lambda_k\left\{\sum_{j=1}^{N}B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)\psi_k(\boldsymbol{z}_j)\right\}^2.$$

Similarly, we can derive the covariance of $\widehat{\boldsymbol{\varepsilon}}^{*}(\boldsymbol{z})$: $\mathbf{\Sigma}_{\varepsilon}(\boldsymbol{z}) = A_{m(\boldsymbol{z})}^{-2}\mathbf{\Sigma}_X^{-1}\frac{1}{nN^2}\sum_{j=1}^{N}B_{m(\boldsymbol{z})}^2(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j)$.

Observe that

$$\sum_{k=1}^{\infty}\lambda_k\left\{\frac{1}{N}\sum_{j=1}^{N}B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)\psi_k(\boldsymbol{z}_j)\right\}^2 = \frac{1}{N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}G_{\eta}(\boldsymbol{z}_j,\boldsymbol{z}_{j'})B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)B_{m(\boldsymbol{z})}(\boldsymbol{z}_{j'}).$$

Hence, by (S1.5) and (S1.6) in Lemma S1.4, (S1.35) holds. Therefore,

$$\begin{aligned}\mathbf{\Sigma}_{\eta}(\boldsymbol{z}) + \mathbf{\Sigma}_{\varepsilon}(\boldsymbol{z}) =& (nA_{m(\boldsymbol{z})}^2)^{-1}\mathbf{\Sigma}_X^{-1}\int_{T_{m(\boldsymbol{z})}\times T_{m(\boldsymbol{z})}} G_{\eta}\left(\boldsymbol{u},\boldsymbol{v}\right)d\boldsymbol{u}d\boldsymbol{v}\{1+O(N^{-1/2}|\triangle|^{-1})\}\\ &+ (nNA_{m(\boldsymbol{z})}^2)^{-1}\mathbf{\Sigma}_X^{-1}\int_{T_{m(\boldsymbol{z})}}\sigma^2(\boldsymbol{u})d\boldsymbol{u}\{1+O(N^{-1/2}|\triangle|^{-1})\}\\ =& n^{-1}\mathbf{\Sigma}_X^{-1}G_{\eta}\left(\boldsymbol{z},\boldsymbol{z}\right)\{1+O(N^{-1/2}|\triangle|^{-1})\}.\end{aligned}$$

Therefore, $\sup_{\boldsymbol{z}\in\Omega}\left\|\mathbf{\Sigma}_{\eta}(\boldsymbol{z}) + \mathbf{\Sigma}_{\varepsilon}(\boldsymbol{z}) - n^{-1}\mathbf{\Sigma}_X^{-1}G_{\eta}\left(\boldsymbol{z},\boldsymbol{z}\right)\right\|_{\infty} = O(n^{-1}N^{-1/2}|\triangle|^{-1})$. The desired result in (S1.35) follows. □

Proof of Theorem 3. Note that, for any vector $\mathbf{a} = (a_0,\ldots,a_p)^{\top} \in \mathcal{R}^{(p+1)}$, we have $E\left[\sum_{\ell=0}^{p}a_{\ell}\left\{\widehat{\eta}_{\ell}^{*}(\boldsymbol{z}) + \widehat{\varepsilon}_{\ell}^{*}(\boldsymbol{z})\right\}\right] = 0$, and

$$\begin{aligned}\sum_{\ell=0}^{p}a_{\ell}\widehat{\eta}_{\ell}^{*}(\boldsymbol{z}) &= \boldsymbol{a}^{\top}\frac{A_{m(\boldsymbol{z})}^{-1}\mathbf{\Sigma}_X^{-1}}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)\sum_{k=1}^{\infty}\xi_{ik}\psi_k(\boldsymbol{z}_j)\mathbf{X}_i = \sum_{i=1}^{n}\boldsymbol{a}^{\top}A_{m(\boldsymbol{z})}^{-1}\mathbf{\Sigma}_X^{-1}\mathfrak{z}_i^{\eta},\\ \sum_{\ell=0}^{p}a_{\ell}\widehat{\varepsilon}_{\ell}^{*}(\boldsymbol{z}) &= \boldsymbol{a}^{\top}\frac{A_{m(\boldsymbol{z})}^{-1}\mathbf{\Sigma}_X^{-1}}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)\varepsilon_{ij}\mathbf{X}_i = \sum_{i=1}^{n}\boldsymbol{a}^{\top}A_{m(\boldsymbol{z})}^{-1}\mathbf{\Sigma}_X^{-1}\mathfrak{z}_i^{\varepsilon},\end{aligned}$$

where $\mathfrak{z}_i^{\eta} = \frac{1}{nN}\sum_{j=1}^{N}B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)\sum_{k=1}^{\infty}\xi_{ik}\psi_k(\boldsymbol{z}_j)\mathbf{X}_i$ and $\mathfrak{z}_i^{\varepsilon} = \frac{1}{nN}\sum_{j=1}^{N}B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)\varepsilon_{ij}\mathbf{X}_i$ are independent sequences with variances $\mathrm{Var}(\mathfrak{z}_i^{\eta}) = \frac{1}{n^2N^2}\sum_{j,j'}B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)B_{m(\boldsymbol{z})}(\boldsymbol{z}_j')G_{\eta}(\boldsymbol{z}_j,\boldsymbol{z}_{j'})\mathbf{\Sigma}_X$ and $\mathrm{Var}(\mathfrak{z}_i^{\varepsilon}) = \frac{1}{n^2N^2}\sum_{j=1}^{N}B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j)\mathbf{\Sigma}_X$, respectively. Therefore, we have

$$\mathrm{Var}\left(\boldsymbol{a}^\top A_{m(\boldsymbol{z})}^{-1}\boldsymbol{\Sigma}_X^{-1}\boldsymbol{\mathfrak{z}}_i^{\eta}\right) = \frac{1}{n}\boldsymbol{a}^\top\boldsymbol{\Sigma}_\eta(\boldsymbol{z})\boldsymbol{a} = \frac{A_{m(\boldsymbol{z})}^{-2}}{n^2N^2}\sum_{j,j'=1}^{N} B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)B_{m(\boldsymbol{z})}(\boldsymbol{z}_{j'})G_\eta(\boldsymbol{z}_j,\boldsymbol{z}_{j'})\boldsymbol{a}^\top\boldsymbol{\Sigma}_X^{-1}\boldsymbol{a},$$

$$\mathrm{Var}\left(\boldsymbol{a}^\top A_{m(\boldsymbol{z})}^{-1}\boldsymbol{\Sigma}_X^{-1}\boldsymbol{\mathfrak{z}}_i^{\varepsilon}\right) = \frac{1}{n}\boldsymbol{a}^\top\boldsymbol{\Sigma}_\varepsilon(\boldsymbol{z})\boldsymbol{a} = \frac{A_{m(\boldsymbol{z})}^{-2}}{n^2N^2}\sum_{j=1}^{N} B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)B_{m(\boldsymbol{z})}(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j)\boldsymbol{a}^\top\boldsymbol{\Sigma}_X^{-1}\boldsymbol{a}.$$

Using central limit theorem, we have

$$\left[\boldsymbol{a}^\top\left\{\boldsymbol{\Sigma}_\eta(\boldsymbol{z})+\boldsymbol{\Sigma}_\varepsilon(\boldsymbol{z})\right\}\boldsymbol{a}\right]^{-1/2}\sum_{\ell=0}^{p} a_\ell\left\{\widehat{\eta}_\ell^*(\boldsymbol{z})+\widehat{\varepsilon}_\ell^*(\boldsymbol{z})\right\}\xrightarrow{\mathcal{L}} N(0,1).$$

By (S1.35), as $N\to\infty$ and $n\to\infty$, $\{\mathbf{a}^\top\boldsymbol{\Sigma}_n(\boldsymbol{z})\mathbf{a}\}^{-1/2}\sum_{\ell=0}^{p}a_\ell\left\{\widehat{\eta}_\ell^*(\boldsymbol{z})+\widehat{\varepsilon}_\ell^*(\boldsymbol{z})\right\}\xrightarrow{\mathcal{L}} N(0,1)$. Therefore, $\{\mathbf{a}^\top\boldsymbol{\Sigma}_n(\boldsymbol{z})\mathbf{a}\}^{-1/2}\sum_{\ell=0}^{p}a_\ell\{\widehat{\beta}_\ell^{\mathrm{c}}(\boldsymbol{z})-\beta_\ell^o(\boldsymbol{z})\}\xrightarrow{\mathcal{L}} N(0,1)$ follows from (S1.18), Theorem S1.3, Lemma S1.15 and Slutsky's Theorem. Applying Cramér-Wold's device, we obtain $\boldsymbol{\Sigma}_n^{-1/2}(\boldsymbol{z})\{\widehat{\beta}_\ell^{\mathrm{c}}(\boldsymbol{z})-\beta_\ell^o(\boldsymbol{z})\}_{\ell=0}^{p}\xrightarrow{\mathcal{L}} N\left(\mathbf{0},\mathbf{I}_{(p+1)\times(p+1)}\right)$, as $N\to\infty$ and $n\to\infty$, and consequently, $\sigma_{n,\ell\ell}^{-1}(\boldsymbol{z})\{\widehat{\beta}_\ell^{\mathrm{c}}(\boldsymbol{z})-\beta_\ell^o(\boldsymbol{z})\}\xrightarrow{\mathcal{L}} N(0,1)$, for any $\boldsymbol{z}\in\Omega$ and $\ell=0,\ldots,p$. □

S1.5 Convergence of the covariance estimator

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

$$\boldsymbol{\Upsilon}_n=\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)\widetilde{\mathbf{B}}_\eta^\top(\boldsymbol{z}_j),$$

then we have

$$\begin{aligned}\widehat{\boldsymbol{\vartheta}}_i &= \boldsymbol{\Upsilon}_n^{-1}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)\widehat{R}_{ij} \\ &= \boldsymbol{\Upsilon}_n^{-1}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)\left[\sum_{\ell=0}^{p}X_{i\ell}\{\beta_\ell^o(\boldsymbol{z}_j)-\widehat{\beta}_\ell(\boldsymbol{z}_j)\}+\eta_i(\boldsymbol{z}_j)+\sigma(\boldsymbol{z}_j)\varepsilon_{ij}\right]. \quad \text{(S1.36)}\end{aligned}$$

Lemma S1.17. *Under Assumptions (A3)–(A5), if* $(N^{1/2}|\triangle_\eta|)/\log(|\triangle_\eta|^{-1})\to\infty$ *as* $N\to\infty$*, then there exist constants* $0<c_\Upsilon<C_\Upsilon<\infty$*, such that with probability approaching 1 as* $N\to\infty$*,* $n\to\infty$*,* $c_\Upsilon|\triangle_\eta|^2\le\lambda_{\min}(\boldsymbol{\Upsilon}_n)\le\lambda_{\max}(\boldsymbol{\Upsilon}_n)\le C_\Upsilon|\triangle_\eta|^2$.

The proof is similar to the proof of S1.7, thus omitted.

Next we define

$$\widetilde{b}_i(\boldsymbol{z}) = \widetilde{\mathbf{B}}_\eta(\boldsymbol{z})^\top\boldsymbol{\Upsilon}_n^{-1}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)\sum_{\ell=0}^{p}X_{i\ell}\{\beta_\ell^o(\boldsymbol{z}_j)-\widehat{\beta}_\ell(\boldsymbol{z}_j)\}, \quad \text{(S1.37)}$$

$$\widetilde{\eta}_i(\boldsymbol{z}) = \widetilde{\mathbf{B}}_\eta(\boldsymbol{z})^\top\boldsymbol{\Upsilon}_n^{-1}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)\eta_i(\boldsymbol{z}_j),\ \widetilde{\varepsilon}_i(\boldsymbol{z}) = \widetilde{\mathbf{B}}_\eta(\boldsymbol{z})^\top\boldsymbol{\Upsilon}_n^{-1}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}.$$

Then, the estimation error $D_i(\boldsymbol{z})=\widehat{\eta}_i(\boldsymbol{z})-\eta_i(\boldsymbol{z})$ in (3.1) can be decomposed as the following:

$$D_i(\boldsymbol{z}) = \widetilde{b}_i(\boldsymbol{z})+\nabla\eta_i(\boldsymbol{z})+\widetilde{\varepsilon}_i(\boldsymbol{z}).$$

For any $\boldsymbol{z}$, $\boldsymbol{z}'\in\Omega$, denote

$$\widetilde{G}_\eta(\boldsymbol{z},\boldsymbol{z}') = n^{-1}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\eta_i(\boldsymbol{z}').$$

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

Lemma S1.18. *Under Assumptions (A1)–(A5) and (C1)–(C3),* $\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}|\widetilde{G}_\eta(\boldsymbol{z},\boldsymbol{z}')-G_\eta(\boldsymbol{z},\boldsymbol{z}')|=O_P\{n^{-1/2}(\log n)^{1/2}\}$.

Proof. Let $\bar{\xi}_{\cdot kk'} = n^{-1}\sum_{i=1}^{n}\xi_{ik}\xi_{ik'}$, then

$$\widetilde{G}_\eta(\boldsymbol{z},\boldsymbol{z}') - G_\eta(\boldsymbol{z},\boldsymbol{z}') = \sum_{k=1}^{\infty}\lambda_k\psi_k(\boldsymbol{z})\psi_k(\boldsymbol{z}')\left(\bar{\xi}_{\cdot kk}-1\right) + \sum_{k\neq k'}\bar{\xi}_{\cdot kk'}(\lambda_k\lambda_{k'})^{1/2}\psi_k(\boldsymbol{z})\psi_{k'}(\boldsymbol{z}').$$

As $E\left[\sum_{k=1}^{\infty}\lambda_k\psi_k(\boldsymbol{z})\psi_k(\boldsymbol{z}')\left(\bar{\xi}_{\cdot kk}-1\right)\right] = 0$, then $E\{\widetilde{G}_\eta(\boldsymbol{z},\boldsymbol{z}') - G_\eta(\boldsymbol{z},\boldsymbol{z}')\} = 0$. Note that $E\{\eta^2(\boldsymbol{z})\eta^2(\boldsymbol{z}')\} = G_\eta(\boldsymbol{z},\boldsymbol{z})G_\eta(\boldsymbol{z}',\boldsymbol{z}')+2G_\eta^2(\boldsymbol{z},\boldsymbol{z}')+\sum_{k=1}^{\infty}\lambda_k^2E(\xi_{1k}^4-3)\psi_k^2(\boldsymbol{z})\psi_k^2(\boldsymbol{z}')$. Next,

$$\begin{aligned}E\left\{\widetilde{G}_\eta(\boldsymbol{z},\boldsymbol{z}') - G_\eta(\boldsymbol{z},\boldsymbol{z}')\right\}^2 &= E\left\{\frac{1}{n}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\eta_i(\boldsymbol{z}') - G_\eta(\boldsymbol{z},\boldsymbol{z}')\right\}^2\\ &= \frac{1}{n}\left\{G_\eta(\boldsymbol{z},\boldsymbol{z})G_\eta(\boldsymbol{z}',\boldsymbol{z}') + G_\eta^2(\boldsymbol{z},\boldsymbol{z}') + \sum_{k=1}^{\infty}\lambda_k^2E(\xi_{1k}^4-3)\psi_k^2(\boldsymbol{z})\psi_k^2(\boldsymbol{z}')\right\}.\end{aligned}$$

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

Proof of Theorem 4. Note that

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}|\widehat{G}_\eta(\boldsymbol{z},\boldsymbol{z}')-G_\eta(\boldsymbol{z},\boldsymbol{z}')| \leq \sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\{|\widehat{G}_\eta(\boldsymbol{z},\boldsymbol{z}')-\widetilde{G}_\eta(\boldsymbol{z},\boldsymbol{z}')|+|\widetilde{G}_\eta(\boldsymbol{z},\boldsymbol{z}')-G_\eta(\boldsymbol{z},\boldsymbol{z}')|\},$$

where $\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}|\widetilde{G}_\eta(\boldsymbol{z},\boldsymbol{z}') - G_\eta(\boldsymbol{z},\boldsymbol{z}')| = o_P(1)$ according to Lemma S1.18, and

$$\begin{aligned}\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}|\widehat{G}_\eta(\boldsymbol{z},\boldsymbol{z}') - \widetilde{G}_\eta(\boldsymbol{z},\boldsymbol{z}')| &\leq \sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|n^{-1}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})D_i(\boldsymbol{z}')\right|\\ &+ \sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|n^{-1}\sum_{i=1}^{n}\eta_i(\boldsymbol{z}')D_i(\boldsymbol{z})\right| + \sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|n^{-1}\sum_{i=1}^{n}D_i(\boldsymbol{z})D_i(\boldsymbol{z}')\right|.\end{aligned}$$

With some simple calculations, we have

$$\sum_{i=1}^{n}\eta_i(\boldsymbol{z})D_i(\boldsymbol{z}') = \sum_{i=1}^{n}\eta_i(\boldsymbol{z})\widetilde{b}_i(\boldsymbol{z}') + \sum_{i=1}^{n}\eta_i(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}') + \sum_{i=1}^{n}\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}'),$$

where $\nabla\eta_i = \widetilde{\eta}_i - \eta_i$. According to (S1.39), (S1.42) and (S1.47), we have

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

Note that

$$\begin{aligned}\sum_{i=1}^{n}D_i(\boldsymbol{z})D_i(\boldsymbol{z}') = &\sum_{i=1}^{n}\widetilde{b}_i(\boldsymbol{z})\widetilde{b}_i(\boldsymbol{z}') + \sum_{i=1}^{n}\nabla\eta_i(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}') + \sum_{i=1}^{n}\widetilde{b}_i(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}') + \sum_{i=1}^{n}\widetilde{\varepsilon}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\\ &+ \sum_{i=1}^{n}\nabla\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}') + \sum_{i=1}^{n}\widetilde{b}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}').\end{aligned}$$

It follows from (S1.38), (S1.41), (S1.43)–(S1.46) that $\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}|n^{-1}\sum_{i=1}^{n}D_i(\boldsymbol{z})D_i(\boldsymbol{z}')| = o_P(1)$. The desired result is established. □

Lemma S1.19. *Under Assumptions (A1)–(A5), (C1)–(C3), we have*

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|n^{-1}\sum_{i=1}^{n}\widetilde{b}_i(\boldsymbol{z})\widetilde{b}_i(\boldsymbol{z}')\right| = O_P\{n^{-1}|\triangle_\eta|^{-2}(\log n)^{1/2}\}, \tag{S1.38}$$

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\widetilde{b}_i(\boldsymbol{z}')\right| = O_P\{n^{-1}(\log n)^{1/2}\}. \tag{S1.39}$$

Proof. According to (S1.19) and (S1.37), we have

$$\begin{aligned}\widetilde{b}_i(\boldsymbol{z}) &= \widetilde{\mathbf{B}}_\eta(\boldsymbol{z})^\top\boldsymbol{\Upsilon}_n^{-1}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\sum_{\ell=0}^{p}X_{i\ell}\{\beta_\ell^o(\boldsymbol{z}_j) - \widehat{\beta}_\ell(\boldsymbol{z}_j)\}\\ &= \widetilde{\mathbf{B}}_\eta(\boldsymbol{z})^\top\boldsymbol{\Upsilon}_n^{-1}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\sum_{\ell=0}^{p}X_{i\ell}\{\beta_\ell^o(\boldsymbol{z}_j) - \widehat{\beta}_{\mu,\ell}(\boldsymbol{z}_j) - \widehat{\eta}_\ell(\boldsymbol{z}_j) - \widehat{\varepsilon}_\ell(\boldsymbol{z}_j)\}.\end{aligned} \tag{S1.40}$$

Thus,

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\widetilde{b}_i(\boldsymbol{z})\widetilde{b}_i(\boldsymbol{z}') &= \frac{1}{n}\sum_{i=1}^{n}\widetilde{\mathbf{B}}_{\eta}(\boldsymbol{z})^{\top}\boldsymbol{\Upsilon}_n^{-1}\left[\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\sum_{\ell=0}^{p}X_{i\ell}\{\beta_{\ell}^{o}(\boldsymbol{z}_j)-\widehat{\beta}_{\ell}(\boldsymbol{z}_j)\}\right] \\
&\quad\times\left[\frac{1}{N}\sum_{j'=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})^{\top}\sum_{\ell'=0}^{p}X_{i\ell'}\{\beta_{\ell'}^{o}(\boldsymbol{z}_{j'})-\widehat{\beta}_{\ell'}(\boldsymbol{z}_{j'})\}\right]\boldsymbol{\Upsilon}_n^{-1}\widetilde{\mathbf{B}}_{\eta}(\boldsymbol{z}') \\
&\asymp \frac{1}{n|\triangle_{\eta}|^{4}}\sum_{i=1}^{n}\widetilde{\mathbf{B}}_{\eta}(\boldsymbol{z})^{\top}\left[\frac{1}{N^2}\sum_{j,j'=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})^{\top}\sum_{\ell,\ell'=0}^{p}X_{i\ell}X_{i\ell'}\{\beta_{\ell}^{o}(\boldsymbol{z}_j)-\widehat{\beta}_{\ell}(\boldsymbol{z}_j)\}\right. \\
&\quad\left.\times\{\beta_{\ell'}^{o}(\boldsymbol{z}_{j'})-\widehat{\beta}_{\ell'}(\boldsymbol{z}_{j'})\}\right]\widetilde{\mathbf{B}}_{\eta}(\boldsymbol{z}').
\end{aligned}
$$

Therefore, by Theorem 1, we have

$$
E\left\{\frac{1}{n}\sum_{i=1}^{n}\widetilde{b}_i(\boldsymbol{z})\widetilde{b}_i(\boldsymbol{z}')\right\} \asymp \sum_{\ell=0}^{p}\sum_{\ell'=0}^{p}|\triangle_{\eta}|^{-2}\|\beta_{\ell}^{o}-\widehat{\beta}_{\ell}\|\|\beta_{\ell'}^{o}-\widehat{\beta}_{\ell'}\| \asymp n^{-1}|\triangle_{\eta}|^{-2}.
$$

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

$$
\begin{aligned}
&E\left\{\widetilde{b}_i(\boldsymbol{z})\widetilde{b}_i(\boldsymbol{z}')\widetilde{b}_{i'}(\boldsymbol{z})\widetilde{b}_{i'}(\boldsymbol{z}')\right\} \asymp |\triangle_{\eta}|^{-8} \\
&\times E\widetilde{\mathbf{B}}_{\eta}(\boldsymbol{z})^{\top}\left[\frac{1}{N^2}\sum_{j,j'=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})^{\top}\sum_{\ell,\ell'=0}^{p}X_{i\ell}X_{i\ell'}\{\beta_{\ell}^{o}(\boldsymbol{z}_j)-\widehat{\beta}_{\ell}(\boldsymbol{z}_j)\}\{\beta_{\ell'}^{o}(\boldsymbol{z}_{j'})-\widehat{\beta}_{\ell'}(\boldsymbol{z}_{j'})\}\right]\widetilde{\mathbf{B}}_{\eta}(\boldsymbol{z}') \\
&\times\widetilde{\mathbf{B}}_{\eta}(\boldsymbol{z})^{\top}\left[\frac{1}{N^2}\sum_{j,j'=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})^{\top}\sum_{\ell,\ell'=0}^{p}X_{i'\ell}X_{i'\ell'}\{\beta_{\ell}^{o}(\boldsymbol{z}_j)-\widehat{\beta}_{\ell}(\boldsymbol{z}_j)\}\{\beta_{\ell'}^{o}(\boldsymbol{z}_{j'})-\widehat{\beta}_{\ell'}(\boldsymbol{z}_{j'})\}\right]\widetilde{\mathbf{B}}_{\eta}(\boldsymbol{z}') \\
&\asymp n^{-2}|\triangle_{\eta}|^{-4}.
\end{aligned}
$$

Thus, (S1.38) follows from the Bernstein inequality after the discretization.

Following from (S1.40), we have, for any $i, i' = 1, \ldots, n$,

$$
\begin{aligned}
& E\left\{\widetilde{b}_i(\boldsymbol{z}')\widetilde{b}_{i'}(\boldsymbol{z}')\eta_i(\boldsymbol{z})\eta_{i'}(\boldsymbol{z})\right\} \\
& \asymp |\triangle_\eta|^{-4}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}')^\top \frac{1}{N^2}\sum_{j,j'=1}^N \widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})^\top E\left\{\sum_{\ell,\ell'=0}^p X_{i\ell}X_{i\ell'}\widehat{\eta}_\ell(\boldsymbol{z}_j)\widehat{\eta}_{\ell'}(\boldsymbol{z}_{j'})\eta_i(\boldsymbol{z})\eta_{i'}(\boldsymbol{z})\right\}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}'), \\
& E\left\{\sum_{\ell,\ell'=0}^p X_{i\ell}X_{i\ell'}\widehat{\eta}_\ell(\boldsymbol{z}_j)\widehat{\eta}_{\ell'}(\boldsymbol{z}_{j'})\eta_i(\boldsymbol{z})\eta_{i'}(\boldsymbol{z})\right\} = \frac{1}{n^2N^2}\sum_{i'',i'''=1}^n E\left[\left\{\mathbf{X}_i \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\}^\top \boldsymbol{\Gamma}_{n,\rho}^{-1}\right. \\
& \quad \left. \times \sum_{j'',j'''=1}^N \mathbf{X}_{i''}\mathbf{X}_{i'''}^\top \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_{j''})\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'''})^\top \boldsymbol{\Gamma}_{n,\rho}^{-1}\mathbf{X}_{i'} \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\right] E\left\{\eta_i(\boldsymbol{z})\eta_{i'}(\boldsymbol{z})\eta_{i''}(\boldsymbol{z}_{j''})\eta_{i'''}(\boldsymbol{z}_{j'''})\right\} \\
& = \frac{1}{n^2N^2}E\left[\left\{\mathbf{X}_i \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\}^\top \boldsymbol{\Gamma}_{n,\rho}^{-1}\sum_{j'',j'''=1}^N \mathbf{X}_{i''}\mathbf{X}_{i'''}^\top \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_{j''})\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'''})^\top \boldsymbol{\Gamma}_{n,\rho}^{-1}\left\{\mathbf{X}_{i'} \otimes \widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\right\}\right] \\
& \quad \times E\left\{\eta_i(\boldsymbol{z})\eta_{i'}(\boldsymbol{z})\eta_i(\boldsymbol{z}_{j''})\eta_{i'}(\boldsymbol{z}_{j'''}) + \eta_i(\boldsymbol{z})\eta_{i'}(\boldsymbol{z})\eta_{i'}(\boldsymbol{z}_{j''})\eta_i(\boldsymbol{z}_{j'''})\right\} \asymp n^{-2}.
\end{aligned}
$$

Therefore, $E\left\{\frac{1}{n}\sum_{i=1}^n \eta_i(\boldsymbol{z})\widetilde{b}_i(\boldsymbol{z}')\right\}^2 = \frac{1}{n^2}\sum_{i,i'=1}^n E\{\widetilde{b}_i(\boldsymbol{z}')\widetilde{b}_{i'}(\boldsymbol{z}')\eta_i(\boldsymbol{z})\eta_{i'}(\boldsymbol{z})\} = O(n^{-2})$.

□

Lemma S1.20. *Under Assumptions (A1)–(A5), (C1)–(C3), we have*

$$
\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|n^{-1}\sum_{i=1}^n \nabla\eta_i(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}')\right| = O_P\left\{|\triangle_\eta|^{2(s+1)}\sum_{k=1}^{K_n}\lambda_k\|\psi_k\|_{s+1,\infty}^2 + \sum_{k=K_n+1}^\infty \lambda_k\|\psi_k\|_\infty^2\right\}, \tag{S1.41}
$$

$$
\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|n^{-1}\sum_{i=1}^n \eta_i(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}')\right| = O_P\left\{|\triangle_\eta|^{s+1}\sum_{k=1}^{K_n}\lambda_k\|\psi_k\|_{s+1,\infty}\|\psi_k\|_\infty + \sum_{k=K_n+1}^\infty \lambda_k\|\psi_k\|_\infty^2\right\}, \tag{S1.42}
$$

$$
\begin{aligned}
\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\sum_{i=1}^n \nabla\eta_i(\boldsymbol{z})\widetilde{b}_i(\boldsymbol{z}')\right| &= O_P\left\{(\log n)^{1/2}n^{-1}|\triangle_\eta|^{s+1}\sum_{k=1}^{K_n}\lambda_k\|\psi_k\|_{s+1,\infty}\|\psi_k\|_\infty\right\} \\
&\quad + O_P\left\{(\log n)^{1/2}n^{-1}\sum_{k=K_n+1}^\infty \lambda_k\|\psi_k\|_\infty^2\right\}.
\end{aligned} \tag{S1.43}
$$

Proof. For any $k \geq 1$, denote $\widetilde{\psi}_k(\boldsymbol{z}) = \widetilde{\mathbf{B}}(\boldsymbol{z})^\top \boldsymbol{\Upsilon}_n^{-1} \frac{1}{N}\sum_{j=1}^N \widetilde{\mathbf{B}}(\boldsymbol{z}_j)\psi_k(\boldsymbol{z}_j)$, and $\nabla\psi_k = \widetilde{\psi}_k - \psi_k$. According to Assumption (C2), we hav(C3)e, for any $k \geq 1$, $\|\nabla\psi_k\|_\infty \leq C|\triangle_\eta|^{s+1}\|\psi_k\|_{s+1,\infty}$ and $\|\widetilde{\psi}_k\|_\infty \leq \|\psi_k\|_\infty + \|\nabla\psi_k\|_\infty \leq 2\|\psi_k\|_\infty$, as $n \to \infty$. It is easy to see that $\nabla\eta_i(\boldsymbol{z}') = \sum_{k=1}^\infty \lambda_k^{1/2}\xi_{ik}\nabla\psi_k(\boldsymbol{z}')$.

We first show (S1.41). Let $\bar{\xi}_{\cdot kk'} = n^{-1}\sum_{i=1}^n \xi_{ik}\xi_{ik'}$, where $E(\bar{\xi}_{\cdot kk'}) = I(k = k')$ and $E(\bar{\xi}_{\cdot kk'})^2 \leq (E\xi_{ik}^4 E\xi_{ik'}^4)^{1/2} \leq C$. Simple calculation yields that $\frac{1}{n}\sum_{i=1}^n \nabla\eta_i(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}') = \sum_{k,k'=1}^\infty \bar{\xi}_{\cdot kk'}(\lambda_k\lambda_{k'})^{1/2}\nabla\psi_k(\boldsymbol{z})\nabla\psi_{k'}(\boldsymbol{z}')$. Thus, by Assumption (C2), we have

$$
\begin{aligned}
\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|E\left\{\frac{1}{n}\sum_{i=1}^n \nabla\eta_i(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}')\right\}\right| &= \sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\sum_{k=1}^\infty \lambda_k\nabla\psi_k(\boldsymbol{z})\nabla\psi_k(\boldsymbol{z}')\right| \\
&\leq |\triangle_\eta|^{2(s+1)}\sum_{k=1}^{K_n}\lambda_k\|\psi_k\|_{s+1,\infty}^2 + C_\psi\sum_{k=K_n+1}^\infty \lambda_k\|\psi_k\|_\infty^2.
\end{aligned}
$$

In addition, we have

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

Thus,

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

Therefore, using the discretization method and Bernstein inequality

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

Next we derive (S1.42). Noting that $n^{-1}\sum_{i=1}^n \eta_i(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}') = \xi_{\cdot kk'}(\lambda_k\lambda_{k'})^{1/2}\psi_k(\boldsymbol{z}')(\nabla\psi_{k'})(\boldsymbol{z}')$, we have

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

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

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

$$\sup_{\boldsymbol{z},\boldsymbol{z}'\in\Omega} E\left\{\eta_i^2(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}')^2\right\} = \sup_{\boldsymbol{z},\boldsymbol{z}'\in\Omega}\left\{\sum_{k=1}^{\infty}E\xi_{ik}^4\lambda_k^2\psi_k^2(\boldsymbol{z})(\nabla\psi_k)^2(\boldsymbol{z}') + \sum_{k\neq k'}\lambda_k\lambda_{k'}\psi_k^2(\boldsymbol{z})(\nabla\psi_k)^2(\boldsymbol{z}')\right\}$$

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

and

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

Therefore,

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

Hence,

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

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

using the discretization method and Bernstein inequality.

Finally, we provide the proof of (S1.43). Note that

$$E\left\{\widetilde{b}_i(\boldsymbol{z}')\widetilde{b}_{i'}(\boldsymbol{z}')\nabla\eta_i(\boldsymbol{z})\nabla\eta_{i'}(\boldsymbol{z})\right\} \asymp |\triangle_\eta|^{-4}$$

$$\times\,\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}')^\top\frac{1}{N^2}\sum_{j,j'=1}^{N}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_{j'})^\top E\left\{\sum_{\ell,\ell'=0}^{p}X_{i\ell}X_{i\ell'}\widehat{\eta}_\ell(\boldsymbol{z}_j)\widehat{\eta}_{\ell'}(\boldsymbol{z}_{j'})\nabla\eta_i(\boldsymbol{z})\nabla\eta_{i'}(\boldsymbol{z})\right\}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}'),$$

and by (S1.42),

$$E\left\{\sum_{\ell,\ell'=0}^{p} X_{i\ell}X_{i\ell'}\widehat{\eta}_\ell(\boldsymbol{z}_j)\widehat{\eta}_{\ell'}(\boldsymbol{z}_{j'})\nabla\eta_i(\boldsymbol{z})\nabla\eta_{i'}(\boldsymbol{z})\right\} = \frac{1}{n^2N^2}\sum_{i'',i'''=1}^{n} E\left[\left\{\mathbf{X}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\}^\top \boldsymbol{\Gamma}_{n,\rho}^{-1}\right.$$
$$\left.\times\sum_{j'',j'''=1}^{N}\mathbf{X}_{i''}\mathbf{X}_{i'''}^\top\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_{j''})\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'''})^\top\boldsymbol{\Gamma}_{n,\rho}^{-1}\mathbf{X}_{i'}\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\right]E\left\{\eta_{i''}(\boldsymbol{z}_{j''})\eta_{i'''}(\boldsymbol{z}_{j'''})\nabla\eta_i(\boldsymbol{z})\nabla\eta_{i'}(\boldsymbol{z})\right\}$$
$$=\frac{1}{n^2N^2}E\left[\left\{\mathbf{X}_i\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\right\}^\top\boldsymbol{\Gamma}_{n,\rho}^{-1}\sum_{j'',j'''=1}^{N}\mathbf{X}_{i''}\mathbf{X}_{i'''}^\top\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_{j''})\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'''})^\top\boldsymbol{\Gamma}_{n,\rho}^{-1}\left\{\mathbf{X}_{i'}\otimes\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\right\}\right]$$
$$\times E\left\{\eta_i(\boldsymbol{z}_{j''})\eta_{i'}(\boldsymbol{z}_{j'''})\nabla\eta_i(\boldsymbol{z})\nabla\eta_{i'}(\boldsymbol{z})+\eta_{i'}(\boldsymbol{z}_{j''})\eta_i(\boldsymbol{z}_{j'''})\nabla\eta_i(\boldsymbol{z})\nabla\eta_{i'}(\boldsymbol{z})\right\}.$$

If $i\neq i'$, we have

$$E\left\{\eta_i(\boldsymbol{z}_{j''})\eta_{i'}(\boldsymbol{z}_{j'''})\nabla\eta_i(\boldsymbol{z})\nabla\eta_{i'}(\boldsymbol{z})+\eta_{i'}(\boldsymbol{z}_{j''})\eta_i(\boldsymbol{z}_{j'''})\nabla\eta_i(\boldsymbol{z})\nabla\eta_{i'}(\boldsymbol{z})\right\}$$
$$\asymp\left\{\sum_{k=1}^{K_n}\lambda_k|\triangle|_\eta^{s+1}\|\psi_k\|_{s+1,\infty}\|\psi_k\|_\infty+\sum_{k=K_n+1}^{\infty}\lambda_k\|\psi_k\|_\infty^2\right\}^2.$$

If $i=i'$, then we have

$$E\left\{\eta_i(\boldsymbol{z}_{j''})\eta_i(\boldsymbol{z}_{j'''})\nabla\eta_i(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z})\right\}=\sum_{k=1}^{\infty}\lambda_k^2E\xi_{ik}^4\psi_k(\boldsymbol{z}'')\psi_k(\boldsymbol{z}''')\nabla\psi_k(\boldsymbol{z})\nabla\psi_k(\boldsymbol{z})$$
$$\leq\sum_{k=1}^{K_n}\lambda_k^2|\triangle|_\eta^{2s+2}\|\psi_k\|_{s+1,\infty}^2\|\psi_k\|_\infty^2+\sum_{k=K_n+1}^{\infty}\lambda_k^2\|\psi_k\|_\infty^4.$$

Thus,

$$E\left\{\sum_{\ell,\ell'=0}^{p}X_{i\ell}X_{i\ell'}\widehat{\eta}_\ell(\boldsymbol{z}_j)\widehat{\eta}_{\ell'}(\boldsymbol{z}_{j'})\nabla\eta_i(\boldsymbol{z})\nabla\eta_{i'}(\boldsymbol{z})\right\}$$
$$\asymp\left\{\sum_{k=1}^{K_n}\lambda_k|\triangle|_\eta^{s+1}\|\psi_k\|_{s+1,\infty}\|\psi_k\|_\infty+\sum_{k=K_n+1}^{\infty}\lambda_k\|\psi_k\|_\infty^2\right\}^2.$$

Therefore,

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

Thus, (S1.43) is obtained. □

Lemma S1.21. *Under Assumptions (A1)–(A5), (C1)–(C3), we have*

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|n^{-1}\sum_{i=1}^{n}\widetilde{\varepsilon}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right| = O_P(N^{-1}|\triangle_\eta|^{-2}), \tag{S1.44}$$

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|n^{-1}\sum_{i=1}^{n}\nabla\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right| = O_P\left\{n^{-1/2}N^{-1/2}(\log n)^{1/2}|\triangle_\eta|^s\sum_{k=1}^{K_n}\lambda_k^{1/2}\|\psi_k\|_{s+1,\infty}\right\}$$
$$+ O_P\left\{n^{-1/2}N^{-1/2}|\triangle_\eta|^{-1}(\log n)^{1/2}\sum_{k=K_n+1}^{\infty}\lambda_k^{1/2}\|\psi_k\|_\infty\right\}, \tag{S1.45}$$

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|n^{-1}\sum_{i=1}^{n}\widetilde{b}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right| = O_P\{n^{-1}N^{-1}|\triangle|^{-2}(\log n)^{1/2}\}, \tag{S1.46}$$

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|n^{-1}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right| = O_P\{n^{-1/2}N^{-1/2}|\triangle_\eta|^{-1}(\log n)^{1/2}\}. \tag{S1.47}$$

Proof. We first show (S1.44). Let $\bar{\varepsilon}_{\cdot jj'} = n^{-1}\sum_{i=1}^n \varepsilon_{ij}\varepsilon_{ij'}$, where $E(\bar{\varepsilon}_{\cdot jj'}) = I(j = j')$. Note that

$$\frac{1}{n}\sum_{i=1}^{n}\widetilde{\varepsilon}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}') = \widetilde{\mathbf{B}}(\boldsymbol{z})^\top\boldsymbol{\Upsilon}_n^{-1}\left\{\frac{1}{N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})^\top\sigma(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_{j'})\bar{\varepsilon}_{\cdot jj'}\right\}\boldsymbol{\Upsilon}_n^{-1}\widetilde{\mathbf{B}}(\boldsymbol{z}').$$

It is easy to see that,

$$E\left\{\frac{1}{n}\sum_{i=1}^{n}\widetilde{\varepsilon}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right\} = \widetilde{\mathbf{B}}(\boldsymbol{z})^\top\boldsymbol{\Upsilon}_n^{-1}\left\{\frac{1}{N^2}\sum_{j=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})^\top\sigma^2(\boldsymbol{z}_j)\right\}\boldsymbol{\Upsilon}_n^{-1}\widetilde{\mathbf{B}}(\boldsymbol{z}').$$

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

$$E\left\{\widetilde{\varepsilon}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right\} = \widetilde{\mathbf{B}}(\boldsymbol{z})^\top \boldsymbol{\Upsilon}_n^{-1}\left\{\frac{1}{N^2}\sum_{j=1}^N \widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})^\top \sigma^2(\boldsymbol{z}_j)\right\}\boldsymbol{\Upsilon}_n^{-1}\widetilde{\mathbf{B}}(\boldsymbol{z}') = O(N^{-1}|\triangle_\eta|^{-2}),$$

$$\begin{aligned} E\left\{\widetilde{\varepsilon}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right\}^2 =& E\left[\widetilde{\mathbf{B}}_\eta(\boldsymbol{z})^\top \boldsymbol{\Upsilon}_n^{-1}\left\{\frac{1}{N^2}\sum_{j=1}^N\sum_{j'=1}^N \widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_{j'})^\top \sigma(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_{j'})\varepsilon_{ij}\varepsilon_{ij'}\right\}\boldsymbol{\Upsilon}_n^{-1}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}')\right]^2 \\ \asymp& \frac{|\triangle_\eta|^{-8}}{N^4}\sum_{j,j',j'',j'''=1}^N \widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_{j'})^\top \widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_{j''})\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_{j'''})^\top \\ &\times \sigma(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_{j'})\sigma(\boldsymbol{z}_{j''})\sigma(\boldsymbol{z}_{j'''})\varepsilon_{ij}\varepsilon_{ij'}\varepsilon_{ij''}\varepsilon_{ij'''} \asymp N^{-2}|\triangle_\eta|^{-4}. \end{aligned}$$

Thus, $\mathrm{var}\left\{\frac{1}{n}\sum_{i=1}^n \widetilde{\varepsilon}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right\} = \frac{1}{n^2}\sum_{i=1}^n \mathrm{var}\left\{\widetilde{\varepsilon}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right\} \asymp n^{-1}N^{-2}|\triangle_\eta|^{-4}$. Therefore,

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|n^{-1}\sum_{i=1}^n \widetilde{\varepsilon}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}') - E\left\{\widetilde{\varepsilon}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right\}\right| = O_P\{n^{-1/2}N^{-1}(\log n)^{1/2}|\triangle_\eta|^{-2}\}$$

using the discretization method and Bernstein inequality.

Next we derive (S1.45). Note that

$$\frac{1}{n}\sum_{i=1}^n \nabla\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}') = \frac{1}{n}\sum_{i=1}^n\sum_{k=1}^\infty \xi_{ik}\lambda_k^{1/2}\nabla\psi_k(\boldsymbol{z})\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}')^\top \boldsymbol{\Upsilon}_n^{-1}\left\{\frac{1}{N}\sum_{j=1}^N \widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}\right\},$$

$$\begin{aligned} \left\{\frac{1}{n}\sum_{i=1}^n \nabla\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right\}^2 &= \frac{1}{n^2}\sum_{i=1}^n\sum_{i'=1}^n\sum_{k=1}^\infty\sum_{k'=1}^\infty \xi_{ik}\xi_{i'k'}(\lambda_k\lambda_{k'})^{1/2}\nabla\psi_k(\boldsymbol{z}) \\ \times \nabla\psi_{k'}(\boldsymbol{z})\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}')^\top \boldsymbol{\Upsilon}_n^{-1}&\left\{\frac{1}{N^2}\sum_{j=1}^N\sum_{j'=1}^N \widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_{j'})^\top \sigma(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_{j'})\varepsilon_{ij}\varepsilon_{i'j'}\right\}\boldsymbol{\Upsilon}_n^{-1}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}'). \end{aligned}$$

Next observe that $E\left[\frac{1}{n}\sum_{i=1}^n \nabla\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right] = 0$ and

$$\begin{aligned} E\left\{\frac{1}{n}\sum_{i=1}^n \nabla\eta_i(\boldsymbol{z})(\nabla\psi_k)^2(\boldsymbol{z})\right\} &= \frac{1}{n^2}\sum_{i=1}^n\sum_{k=1}^\infty \lambda_k(\nabla\psi_k)^2(\boldsymbol{z}) \\ &\times \widetilde{\mathbf{B}}_\eta(\boldsymbol{z}')^\top \boldsymbol{\Upsilon}_n^{-1}\left\{\frac{1}{N^2}\sum_{j=1}^N \widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)^\top \sigma^2(\boldsymbol{z}_j)\right\}\boldsymbol{\Upsilon}_n^{-1}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}') \end{aligned}$$

So,

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

Thirdly, we prove (S1.46). Note that for any i, i', j, j', we have

$$\begin{aligned}
&E\left\{\widetilde{b}_i(\boldsymbol{z})\varepsilon_{ij}\widetilde{b}_{i'}(\boldsymbol{z})\varepsilon_{i'j'}\right\}\\
&=E\left[\mathbf{B}_\eta(\boldsymbol{z})^\top\boldsymbol{\Upsilon}_n^{-1}\frac{1}{N^2}\sum_{j'',j'''=1}^{N}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_{j''})\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_{j'''})^\top\sum_{\ell,\ell'=0}^{p}X_{i\ell}\widehat{\varepsilon}_\ell(\boldsymbol{z}_{j''})X_{i\ell'}\widehat{\varepsilon}_{\ell'}(\boldsymbol{z}_{j'''})\varepsilon_{ij}\varepsilon_{i'j'}\boldsymbol{\Upsilon}_n^{-1}\mathbf{B}_\eta(\boldsymbol{z})\right]\\
&=O(n^{-2}N^{-2}|\triangle|^{-4}).
\end{aligned}$$

Therefore,

$$\begin{aligned}
E\left\{\widetilde{b}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\widetilde{b}_{i'}(\boldsymbol{z})\widetilde{\varepsilon}_{i'}(\boldsymbol{z}')\right\}&=\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}')^\top\boldsymbol{\Upsilon}_n^{-1}\frac{1}{N^2}\sum_{j,j'=1}^{N}E\left\{\widetilde{b}_i(\boldsymbol{z})\varepsilon_{ij}\widetilde{b}_{i'}(\boldsymbol{z})\varepsilon_{i'j'}\right\}\boldsymbol{\Upsilon}_n^{-1}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}')\\
&=O(n^{-2}N^{-2}|\triangle|^{-4}),\\
E\left[n^{-1}\sum_{i=1}^{n}\widetilde{b}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right]^2&=\frac{1}{n^2}\sum_{i,i'=1}^{n}E\left\{\widetilde{b}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\widetilde{b}_{i'}(\boldsymbol{z})\widetilde{\varepsilon}_{i'}(\boldsymbol{z}')\right\}=O(n^{-2}N^{-2}|\triangle|^{-4}).
\end{aligned}$$

Finally, we show (S1.47). Note that

$$\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')=\sum_{i=1}^{n}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}')^\top\boldsymbol{\Upsilon}_n^{-1}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}\sum_{k=1}^{\infty}\xi_{ik}\lambda_k^{1/2}\psi_k(\boldsymbol{z}),$$

where $E\left\{n^{-1}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right\}=0$, and

$$\begin{aligned}
E\left\{n^{-1}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right\}^2&=n^{-1}E\{\eta_i(\boldsymbol{z})^2\}E\{\widetilde{\varepsilon}_i(\boldsymbol{z}')^2\}=n^{-1}G_\eta(\boldsymbol{z},\boldsymbol{z}')\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}')^\top\\
&\times\boldsymbol{\Upsilon}_n^{-1}\frac{1}{N^2}\sum_{j=1}^{N}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}_j)^\top\sigma^2(\boldsymbol{z}_j)\boldsymbol{\Upsilon}_n^{-1}\widetilde{\mathbf{B}}_\eta(\boldsymbol{z}')=O(n^{-1}N^{-1}|\triangle_\eta|^{-2}).
\end{aligned}$$

Thus, (S1.47) is obtained. $\square$

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 S2.1 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 S2.2 and S2.3, respectively. The plots are obtained based on the setting: $n = 50$, $\lambda_1 = 0.2$, $\lambda_2 = 0.05$, $\sigma = 1.0$. Table S2.1 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 S2.2 and S2.3, 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 closed 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.

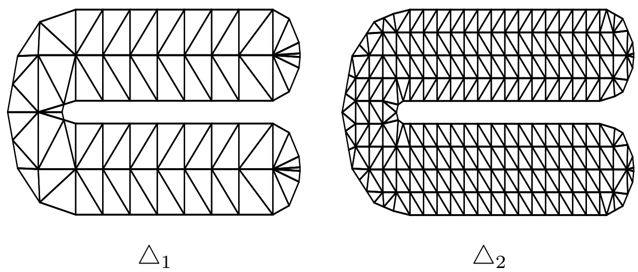

Figure S2.1: Triangulations for the horseshoe domain.

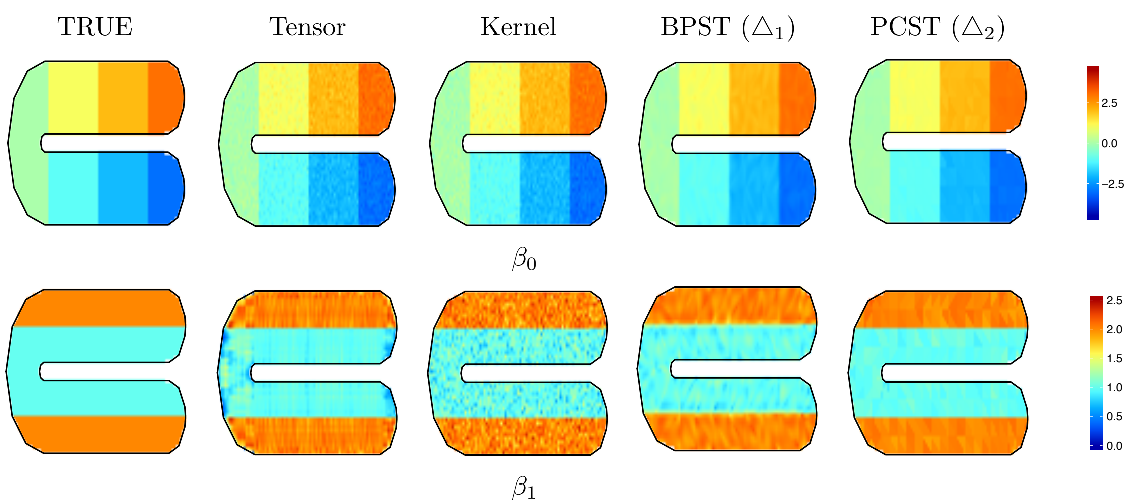

Figure S2.2: True coefficient functions and their different estimators for Case I in Example 1.

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 S2.2 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

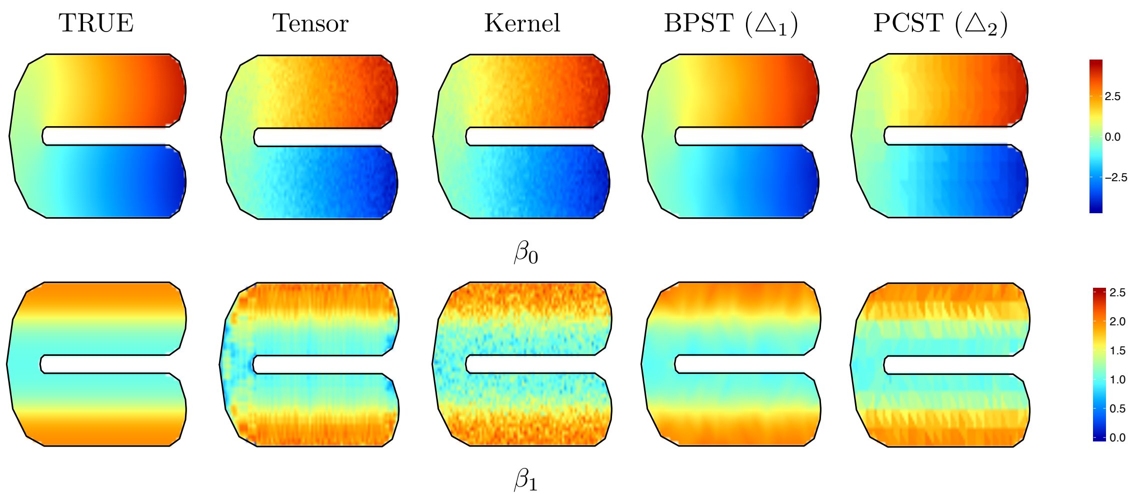

Figure S2.3: True coefficient functions and their different estimators for Case II in Example 1.

triangulations: $\triangle_5$ and $\triangle_6$; see Figure S2.4. Table S2.3 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 S2.4 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 S2.5 and S2.6 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 S2.5 below summarizes the distribution of patients by diagnosis status and sex. Next, Figure S2.7 displays the triangulations of slices used for the BPST method in the model fitting and

Table S2.1: Estimation errors of the coefficient estimators, $\sigma = 1.0$.

Function Type	n	Method	$\lambda_1 = 0.03,\ \lambda_2 = 0.006$		$\lambda_1 = 0.2,\ \lambda_2 = 0.05$	
			β_0	β_1	β_0	β_1
Jump	50	BPST	0.0059	0.0075	0.0066	0.0082
		PCST	0.0023	0.0023	0.0028	0.0030
		Kernel	0.0201	0.0206	0.0207	0.0213
		Tensor	0.0201	0.0132	0.0206	0.0142
	100	BPST	0.0038	0.0050	0.0042	0.0054
		PCST	0.0011	0.0011	0.0014	0.0015
		Kernel	0.0100	0.0102	0.0104	0.0105
		Tensor	0.0099	0.0112	0.0103	0.0120
Smooth	50	BPST	0.0010	0.0012	0.0016	0.0019
		PCST	0.0049	0.0065	0.0054	0.0072
		Kernel	0.0201	0.0206	0.0207	0.0213
		Tensor	0.0189	0.0132	0.0207	0.0153
	100	BPST	0.0006	0.0007	0.0009	0.0010
		PCST	0.0037	0.0054	0.0040	0.0057
		Kernel	0.0100	0.0102	0.0104	0.0105
		Tensor	0.0100	0.0113	0.0103	0.0128

constructing the SCCs. Finally, Figures S2.8 and S2.9 provide the image maps of the estimated coefficient functions for eighth, 15th, 35th, 55th, 62nd, and 65th slices, and Figures S2.10 and S2.11 show the corresponding significance maps. The significance maps in the eighth and 15th slice show that the increase of age increases the brain activities in the cerebellum and temporal lobe, and people with the 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

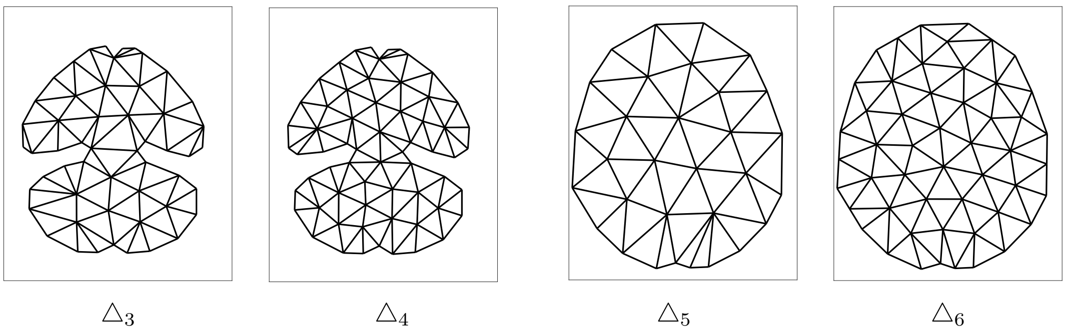

Figure S2.4: Triangulations for the fifth slice ($\triangle_3$, $\triangle_4$) and 35th slice ($\triangle_5$, $\triangle_6$) of the brain image in Simulation Example 2.

Table S2.2: Estimation errors of the coefficient function estimators, $\sigma = 0.5$.

n	Method	$\lambda_1 = 0.1,\ \lambda_2 = 0.02$			$\lambda_1 = 0.2,\ \lambda_2 = 0.05$		
		β_0	β_1	β_2	β_0	β_3	β_2
50	BPST($\triangle_3$)	0.003	0.005	0.005	0.007	0.011	0.010
	BPST($\triangle_4$)	0.003	0.005	0.005	0.006	0.009	0.009
	Kernel	0.008	0.011	0.011	0.011	0.016	0.016
	Tensor	0.008	0.007	0.010	0.011	0.012	0.014
100	BPST($\triangle_3$)	0.002	0.002	0.002	0.003	0.005	0.005
	BPST($\triangle_4$)	0.002	0.002	0.002	0.003	0.004	0.004
	Kernel	0.004	0.005	0.005	0.005	0.008	0.007
	Tensor	0.004	0.005	0.005	0.005	0.007	0.009

increase of brain 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

Andersen, K., Andersen, B. B., and Pakkenberg, B. (2012), "Stereological quantification of the cerebellum in patients with Alzheimer's disease," *Neurobiology of aging*, 33, 197–e11.

Table S2.3: Estimation errors of the coefficient function estimators in the 35th slice.

n	σ	Method	$\lambda_1 = 0.1,\ \lambda_2 = 0.02$			$\lambda_1 = 0.2,\ \lambda_2 = 0.05$		
			β_0	β_1	β_2	β_0	β_1	β_2
50	0.5	BPST($\triangle_5$)	0.003	0.005	0.005	0.007	0.011	0.011
		BPST($\triangle_6$)	0.003	0.005	0.005	0.007	0.011	0.010
		Kernel	0.008	0.012	0.012	0.018	0.018	0.017
		Tensor	0.008	0.009	0.011	0.012	0.015	0.015
	1.0	BPST($\triangle_5$)	0.003	0.005	0.005	0.007	0.011	0.011
		BPST($\triangle_6$)	0.003	0.005	0.005	0.007	0.011	0.011
		Kernel	0.023	0.033	0.033	0.027	0.039	0.038
		Tensor	0.023	0.012	0.019	0.027	0.017	0.023
100	0.5	BPST($\triangle_5$)	0.002	0.002	0.002	0.003	0.005	0.005
		BPST($\triangle_6$)	0.002	0.002	0.002	0.003	0.005	0.005
		Kernel	0.004	0.006	0.006	0.006	0.008	0.008
		Tensor	0.004	0.006	0.007	0.006	0.010	0.009
	1.0	BPST($\triangle_5$)	0.002	0.002	0.002	0.003	0.005	0.005
		BPST($\triangle_6$)	0.002	0.002	0.002	0.003	0.005	0.005
		Kernel	0.012	0.016	0.016	0.013	0.018	0.018
		Tensor	0.013	0.010	0.013	0.011	0.007	0.011

Bernard, J. A. and Seidler, R. D. (2014), "Moving forward: age effects on the cerebellum underlie cognitive and motor declines," *Neuroscience & Biobehavioral Reviews*, 42, 193–207.

Dubb, A., Gur, R., Avants, B., and Gee, J. (2003), "Characterization of sexual dimorphism in the human corpus callosum," *Neuroimage*, 20, 512–519.

Lai, M.-J. and Schumaker, L. L. (2007), *Spline functions on triangulations*, Cambridge University Press.

Lai, M. J. and Wang, L. (2013), "Bivariate penalized splines for regression." *Statistica Sinica*, 23, 1399–1417.

Sangalli, L. M., Ramsay, J. O., and Ramsay, T. O. (2013), "Spatial spline regression models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75, 681–703.

Table S2.4: The coverage rate of the 95% SCCs for the coefficient functions defined over the 35th slice.

n	λ	σ	Coverage			Width		
			β_0	β_1	β_2	β_0	β_1	β_2
50	(0.1,0.02)	0.5	0.962	0.916	0.934	0.307	0.344	0.347
		1.0	0.964	0.926	0.940	0.331	0.368	0.371
	(0.2,0.05)	0.5	0.952	0.920	0.930	0.426	0.490	0.492
		1.0	0.96	0.920	0.934	0.449	0.512	0.512
100	(0.1,0.02)	0.5	0.956	0.952	0.940	0.214	0.240	0.244
		1.0	0.962	0.952	0.948	0.239	0.262	0.265
	(0.2,0.05)	0.5	0.946	0.954	0.932	0.298	0.340	0.346
		1.0	0.952	0.954	0.938	0.317	0.359	0.365

Table S2.5: Distribution of patients by diagnosis status and gender.

	CN	MCI	AD	All
Male	70	136	72	278
Female	42	77	50	169
All	112	213	122	447

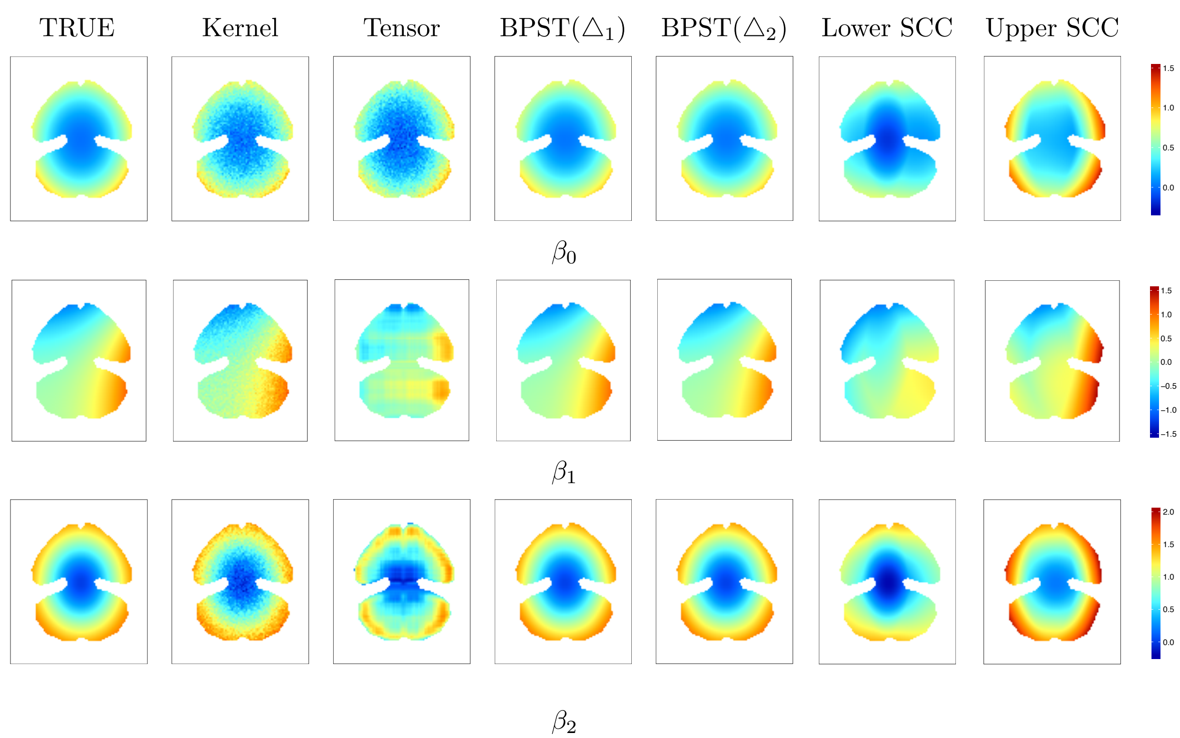

Figure S2.5: True coefficient functions and their estimators and 95% SCCs based on the fifth slice.

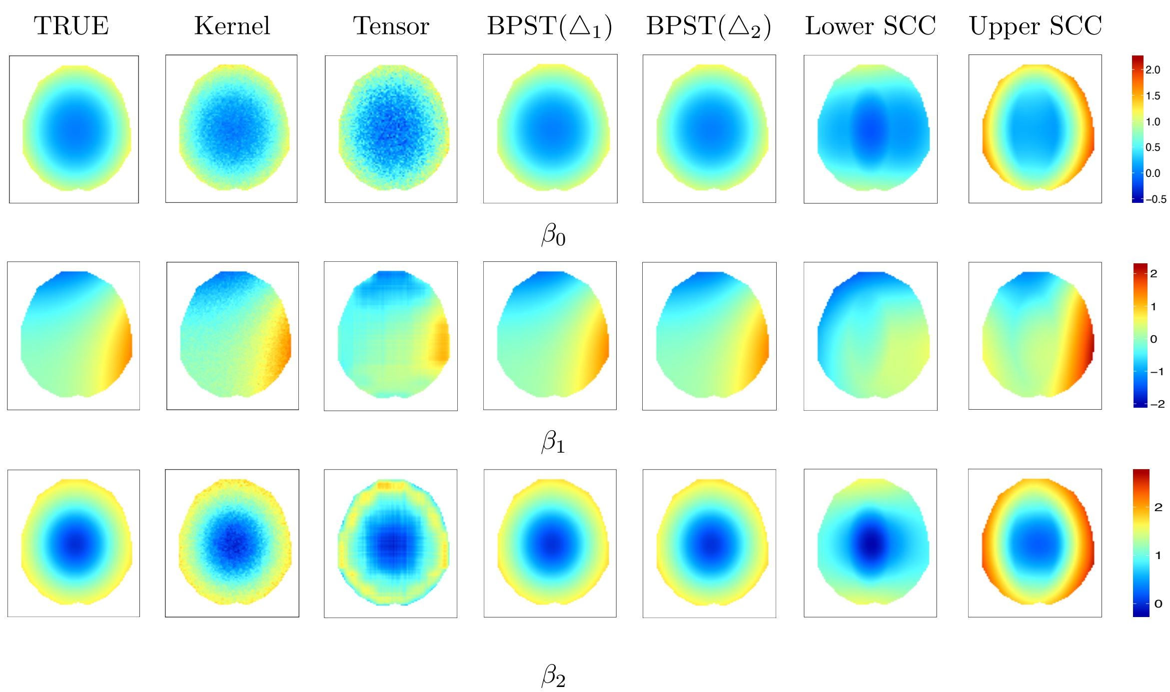

Figure S2.6: True coefficient functions and their estimators and 95% SCCs based on the 35th slice.

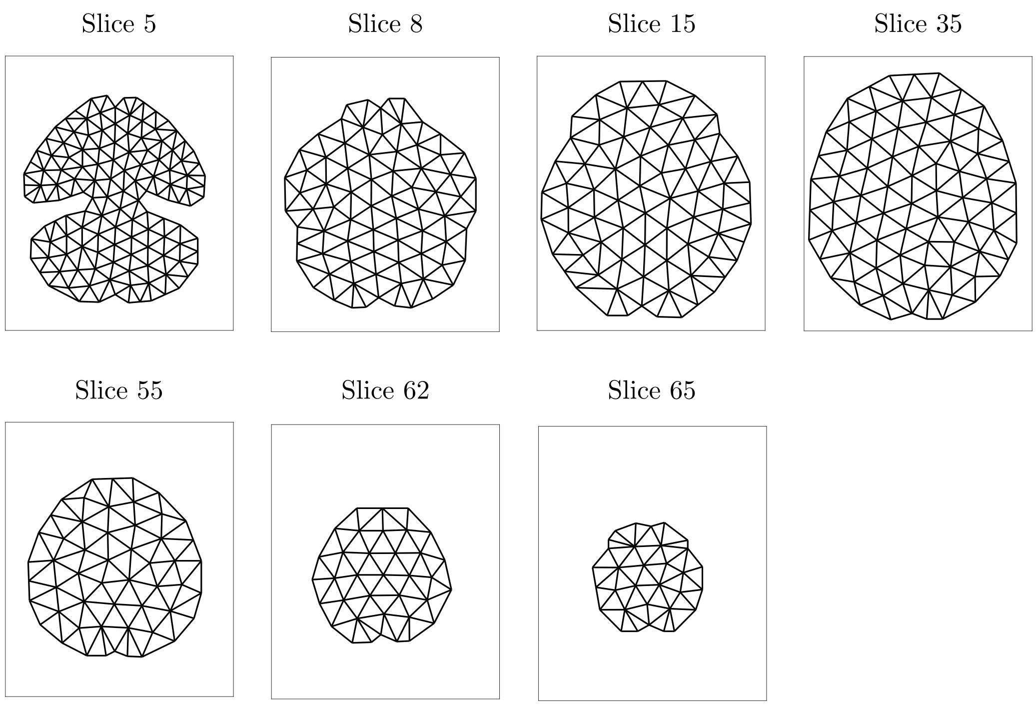

Figure S2.7: Triangulation sets used in the ADNI data analysis.

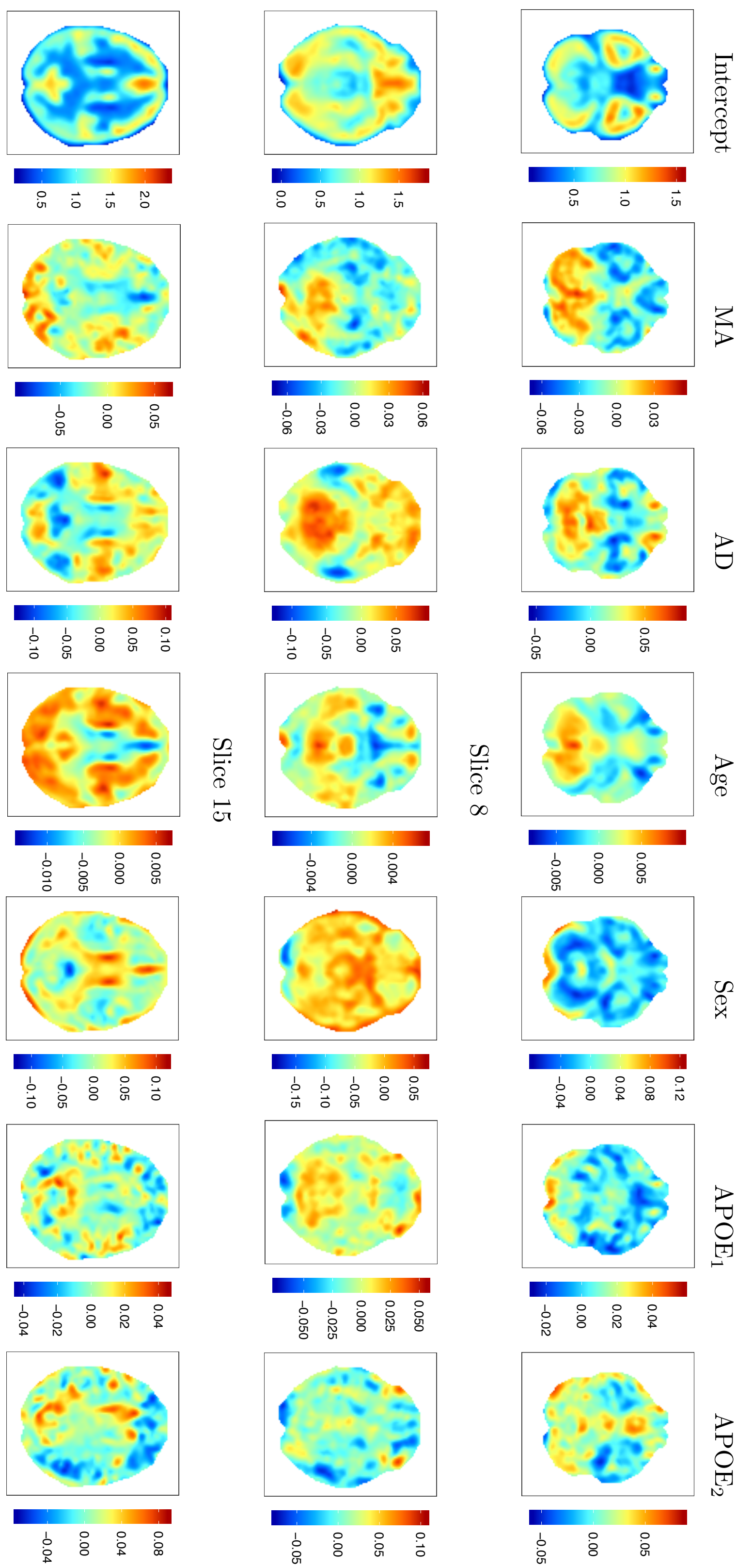

Figure S2.8: The BPST estimates of the coefficient functions for the ADNI data based on the eighth, 15th and 35th slices, respectively.

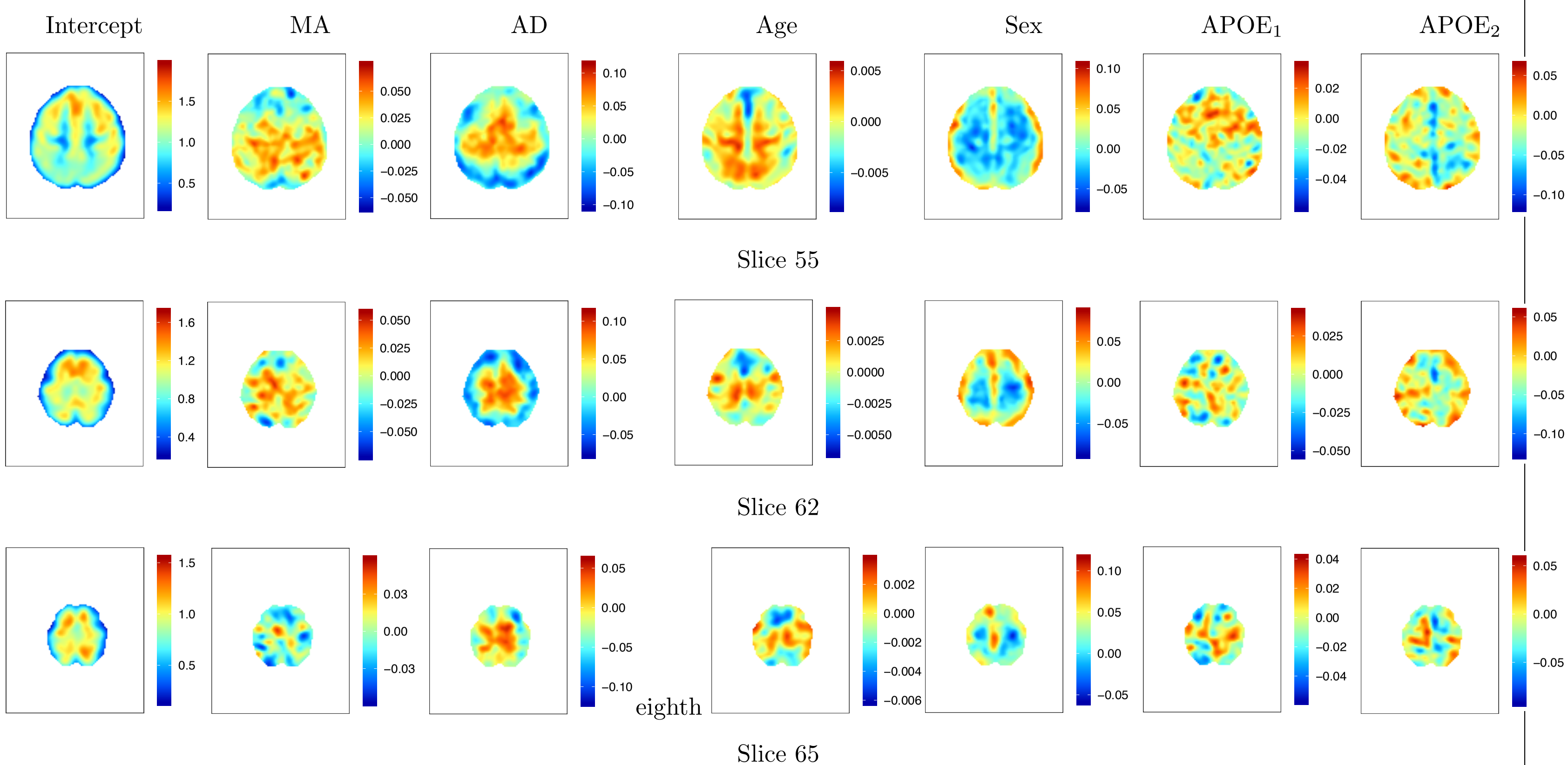

Figure S2.9: The BPST estimates of the coefficient functions for the ADNI data based on the 55th, 62nd and 65th slices, respectively.

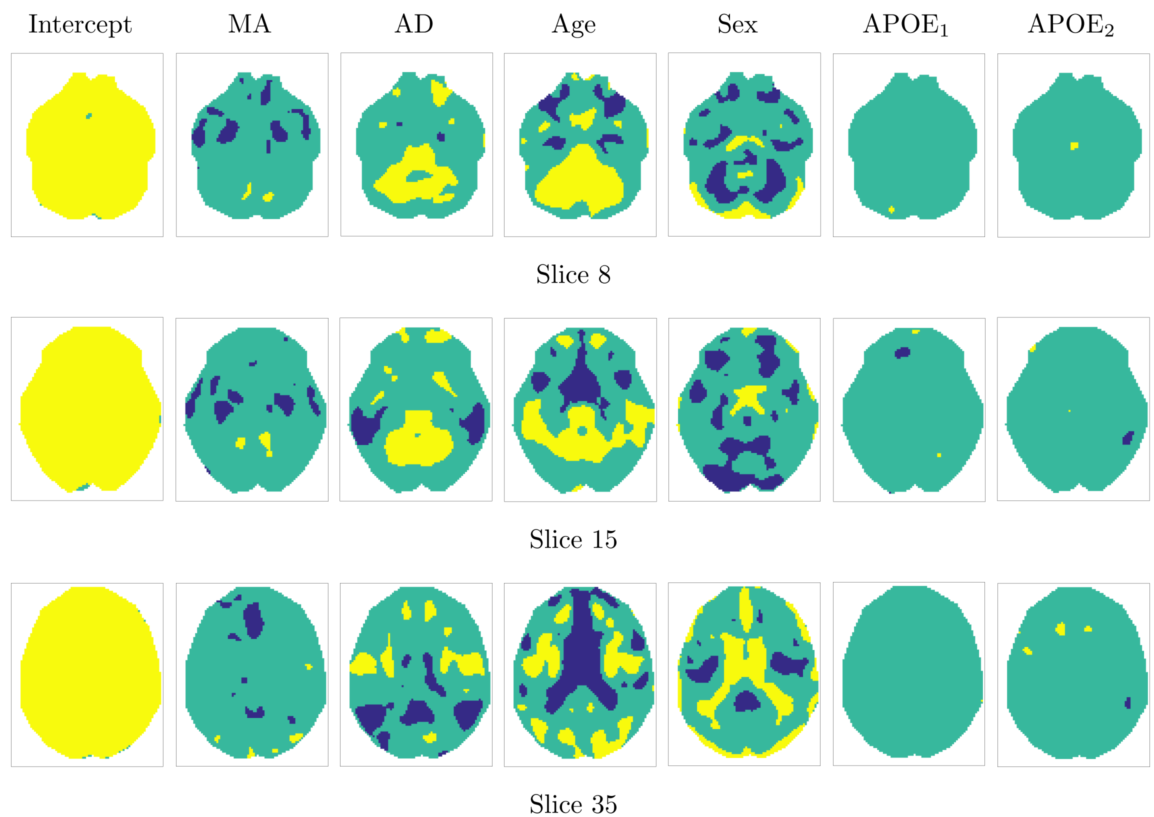

Figure S2.10: The "significance" map (based on the 95% SCC) for the coefficient functions for the ADNI data. The yellow color and blue color on the map indicate the regions that zero is below the lower SCC or above the upper SCC, respectively.

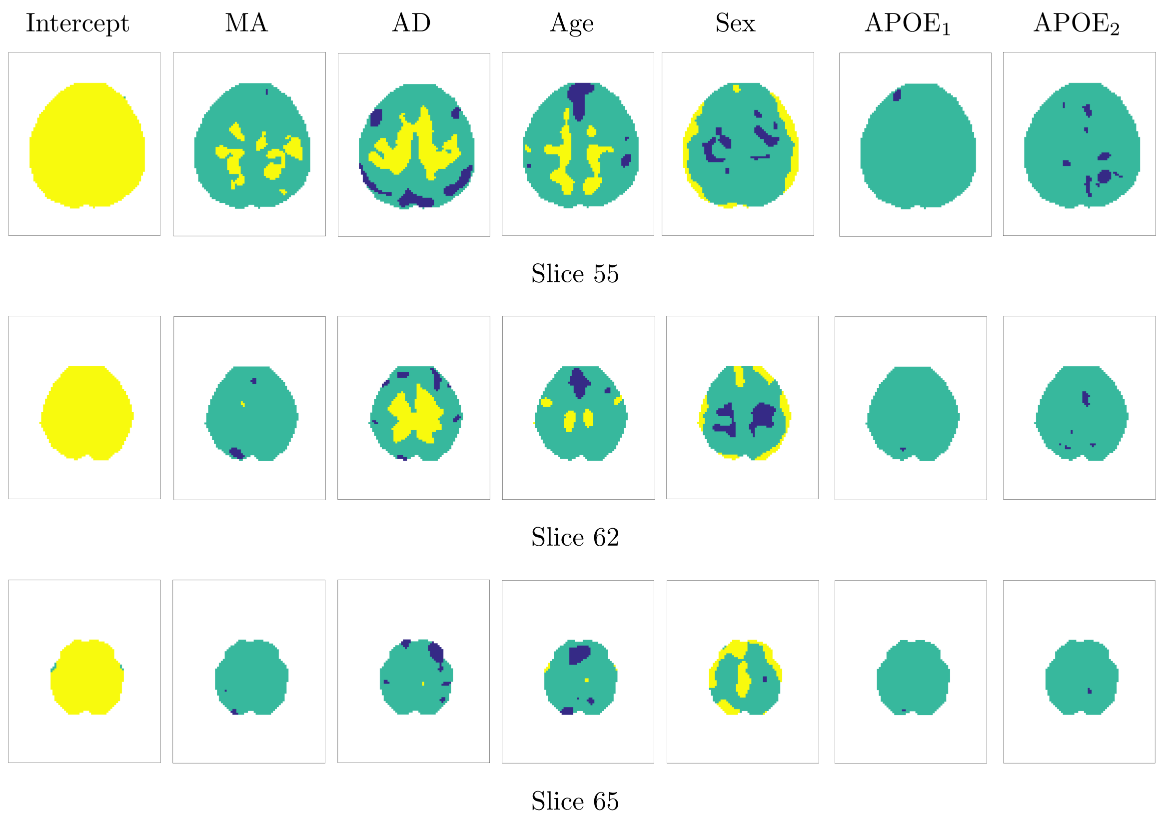

Figure S2.11: The "significance" map (based on the 95% SCC) for the coefficient functions for the ADNI data. The yellow color and blue color on the map indicate the regions that zero is below the lower SCC or above the upper SCC, respectively.