
STATISTICAL INFERENCE FOR MEAN FUNCTIONS OF COMPLEX 3D OBJECTS

Yueying Wang[1], Guannan Wang[2], Brandon Klinedinst[3]
Auriel Willette[4], and Li Wang[5]

[1]*Amazon.com, Inc., Bellevue, WA 98170, USA.*
[2]*William & Mary, Williamsburg, VA 23185, USA.*
[3]*University of Washington, Seattle, WA 98195, USA.*
[4]*Iowa State University, Ames, IA 50011, USA.*
[5]*George Mason University, Fairfax, VA 22030, USA.*

Supplementary Material

In this supplement, Section S1 presents additional simulation studies, and S2 includes detailed proofs of the theoretical results in the main article.

S1 More Results from Simulation Studies

In this section, we conduct a series of Monte Carlo simulation studies to illustrate the finite-sample performance of the proposed method in estimating the mean function. We evaluate the performance of the proposed TPST method by comparing it to several existing methods, including Tensor-Product Splines (TPS), the MASS procedure by Zhu et al. (2014), and the Fréchet regression approach in Petersen et al. (2019). The MASS procedure uses a tensor-product kernel smoothing method with adaptive weights. The Fréchet regression is a least-square estimate for the mean functions and serves as a baseline comparison in our simulation studies.

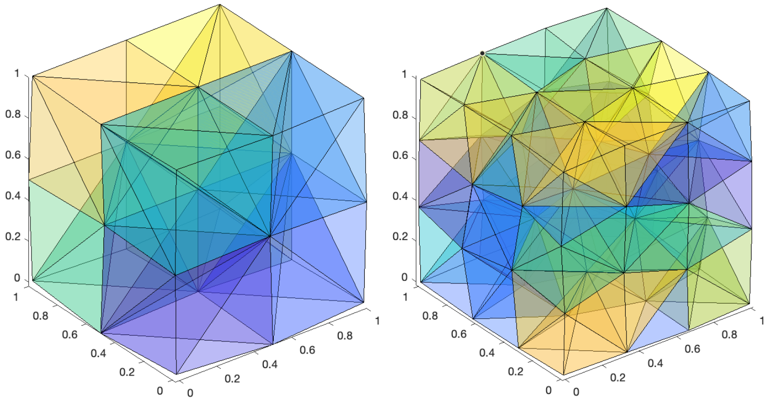
We generate data from the following model: for $i = 1, \ldots, n$ and $j = 1, \ldots, N$,

$$Y_{ij} = \mu(\boldsymbol{z}_j) + \sum_{k=1}^{\kappa} \sqrt{\lambda_k} \xi_{ik} \psi_k(\boldsymbol{z}_j) + \sigma(\boldsymbol{z}_j)\varepsilon_{ij},\ \boldsymbol{z}_j = (z_{1j}, z_{2j}, z_{3j}) \in \Omega.$$

We consider two types of domains: a ball with domain $\Omega_2 = \{\boldsymbol{z} \in [0,1]^3 : (z_1 - 0.5)^2 + (z_2 - 0.5)^2 + (z_3 - 0.5)^2 \leq 0.5^2\}$, as described in the main paper, and a regular cube with the domain $\Omega_3 = [0,1]^3$. The inclusion of the regular cube domain enables comparison with other existing methods, while the ball-shaped domain showcases our method's ability to handle irregular domains. To examine the performance under different scenarios, we adopt the same mean functions and standard deviation function, as well as a similar data generation process as described in the main paper. The eigenvalues and eigenfunctions are $\lambda_1 = 0.5, \lambda_2 = 0.2, \lambda_3 = 0.1$ and $\psi_1(\boldsymbol{z}) = c_{11}\sin(\pi z_1) + c_{12}$, $\psi_2(\boldsymbol{z}) = c_2\cos(\pi z_2)$, $\psi_3(\boldsymbol{z}) = c_3(z_3 - 1/2)$. The constants c_{11}, c_{12}, c_2 and c_3 are chosen separately for different domains to ensure that $\int_{\Omega_i} \psi_k^2(\boldsymbol{z})\mathrm{d}\boldsymbol{z} = 1$ and $\int_{\Omega_i} \psi_k(\boldsymbol{z})\psi_{k'}(\boldsymbol{z})\mathrm{d}\boldsymbol{z} = 0$ for $i = 2, 3,\ k, k' = 1, \ldots, \kappa$ and $k' \neq k$.

We generate data on M^3 equally spaced grid points in the cube domain $\Omega_3 = [0,1]^3$, with M set to 25 and 35. We remove points outside the domain for the ball domain Ω_2. To assess performance with varying sample sizes, we use sample sizes of 50, 150, 300, and 600.

For TPST estimation, we set the smoothness parameter $r = 1$ and consider degrees $d = 2, 3, \ldots, 6$. We employ two triangulations for each domain. For the ball domain Ω_2, we employ triangulations $\triangle_{2,1}$ and $\triangle_{2,2}$ consisting of 176 and 332 tetrahedra, respectively, as illustrated in the main paper. In the cube domain Ω_3, we utilize triangulations $\triangle_{3,1}$ and $\triangle_{3,2}$ with 48 and 172 tetrahedra, respectively, as depicted in



(a) $\triangle_{3,1}$ with 48 tetrahedra (b) $\triangle_{3,2}$ with 172 tetrahedra

Figure S1: Triangulation used in rectangle domain Ω_3.

Figure S1. We calculate TPS using quadratic and cubic B-spline basis functions with five equally spaced knots in each coordinate. We set the kernel bandwidth for the MASS procedure to $h = 5$.

Firstly, we evaluate the influence of the spline basis degrees (d) on estimation accuracy. To assess the performance, we calculate the average mean squared errors (AMSE) defined as AMSE $= N^{-1}\sum_{j=1}^{N}\{\widehat{\mu}(\boldsymbol{z}_j)-\mu(\boldsymbol{z}_j)\}^2$. The simulations are replicated 100 times, and the AMSEs of the TPST method with different degrees of spline basis are presented in Table S1. From this table, one can see that a higher degree of spline basis provides a more flexible estimator, allowing for a finer representation of the underlying patterns in the data. However, it is crucial to consider the risk of overfitting the data when using a higher degree. Overfitting occurs when the method captures not only the actual underlying signal but also the noise or individual-level fluctuations, resulting in a less reliable estimation. For example, when the underlying function is quadratic (as in μ_1), $d = 6$ usually results in the highest AMSE and provides the least accurate result. Conversely, using a lower degree of spline basis may lead to underfitting, where

Table S1: Average mean squared errors (AMSEs$\times 10^{-3}$) of the TPST method with different degrees of spline basis for cube domain Ω_3.

Sample		Number of Grid Points (M^3)					
Size	Method	$25 \times 25 \times 25$			$35 \times 35 \times 35$		
n		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
50	TPST($d = 2,\ \triangle_{1,1}$)	17.3525	22.0265	18.2578	17.3528	22.0057	18.2512
	TPST($d = 3,\ \triangle_{1,1}$)	17.3555	17.3554	17.3829	17.3558	17.3557	17.3832
	TPST($d = 4,\ \triangle_{1,1}$)	17.3481	17.3479	17.3481	17.3535	17.3534	17.3535
	TPST($d = 5,\ \triangle_{1,1}$)	17.3588	17.3554	17.3563	17.3574	17.3566	17.3569
	TPST($d = 6,\ \triangle_{1,1}$)	17.3873	17.3639	17.3639	17.3667	17.3609	17.3615
150	TPST($d = 2,\ \triangle_{1,1}$)	5.0306	9.7157	5.9432	5.0286	9.6847	5.9297
	TPST($d = 3,\ \triangle_{1,1}$)	5.0319	5.0319	5.0601	5.0295	5.0295	5.0573
	TPST($d = 4,\ \triangle_{1,1}$)	5.0294	5.0294	5.0295	5.0288	5.0288	5.0289
	TPST($d = 5,\ \triangle_{1,1}$)	5.0329	5.0325	5.0327	5.0301	5.0301	5.0301
	TPST($d = 6,\ \triangle_{1,1}$)	5.0431	5.0394	5.0393	5.0335	5.0327	5.0329
300	TPST($d = 2,\ \triangle_{1,1}$)	2.3751	7.0551	3.2862	2.3747	7.0292	3.2752
	TPST($d = 3,\ \triangle_{1,1}$)	2.3760	2.3760	2.4040	2.3752	2.3752	2.4029
	TPST($d = 4,\ \triangle_{1,1}$)	2.3797	2.3797	2.3797	2.3764	2.3764	2.3765
	TPST($d = 5,\ \triangle_{1,1}$)	2.3814	2.3813	2.3814	2.3771	2.3771	2.3771
	TPST($d = 6,\ \triangle_{1,1}$)	2.3817	2.3798	2.3798	2.3772	2.3768	2.3770
600	TPST($d = 2,\ \triangle_{1,1}$)	1.3199	5.9997	2.2323	1.3202	5.9746	2.2212
	TPST($d = 3,\ \triangle_{1,1}$)	1.3204	1.3204	1.3485	1.3205	1.3205	1.3482
	TPST($d = 4,\ \triangle_{1,1}$)	1.3220	1.3220	1.3221	1.3210	1.3210	1.3211
	TPST($d = 5,\ \triangle_{1,1}$)	1.3229	1.3228	1.3229	1.3213	1.3213	1.3214
	TPST($d = 6,\ \triangle_{1,1}$)	1.3233	1.3225	1.3231	1.3215	1.3214	1.3214

the model fails to capture the complexity and nuances present in the data. Underfitting can result in a loss of important information and may yield less accurate estimates. For example, when the underlying function is cubic (as in μ_2), $d = 2$ is not enough for a good estimation, and a higher d is usually preferred. It is worth noting that the choice of degree (d) should also be considered with the number of observations per object (N). As the degree increases, the number of basis functions and the complexity of the model also increase. Therefore, when using a higher degree, it is generally recommended to have a larger N to ensure sufficient data support for the estimation.

In the following simulation studies, we chose $d = 4$ as a moderate degree to demonstrate the general performance of our method and compare the performance of our proposed method with other existing methods. The average AMSEs for different methods are presented in Tables S2 and S3. As seen from Tables S2 and S3, it is evident that the estimation errors for all methods decrease as the sample size increases. As the number of grid points increases, the accuracy of all four methods has improved. Based on the AMSEs, the proposed TPST method outperforms all other methods in estimation accuracy in most simulation settings. Specifically, in the cube domain Ω_3, TPST exhibits slightly better estimation performance compared to TPS and Fréchet. In the ball domain Ω_2, the AMSEs of TPST are around $15 \sim 500$ times smaller than those of TPS and MASS. The superior performance of TPST illustrates its efficiency in estimation, as well as its capability in dealing with irregular domains. Meanwhile, it is worth noticing that the choice of triangulation has only a minimal effect on the performance as long as the triangulation is constructed properly.

S2 Technical Proofs

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 $\mathbf{A}$. 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 paper. Let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ be a vector of nonnegative integers. Denote $D^{\boldsymbol{\alpha}} = D_{z_1}^{\alpha_1} D_{z_2}^{\alpha_2} D_{z_3}^{\alpha_3}$ and $|\boldsymbol{\alpha}| = \alpha_1 + \alpha_2 + \alpha_3$. For any function g over the closure of domain Ω, let $|g|_{v,\infty,\Omega} =$

Table S2: Average mean squared errors (AMSEs$\times 10^{-3}$) of the TPST, TPS, MASS and Fréchet regression methods for cube domain Ω_3.

Sample Size	Method	Number of Grid Points (M^3)					
		$25 \times 25 \times 25$			$35 \times 35 \times 35$		
n		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
50	TPST($d = 4,\ \triangle_{3,1}$)	17.3481	17.3479	17.3481	17.3535	17.3534	17.3535
	TPST($d = 4,\ \triangle_{3,2}$)	17.3483	17.3473	17.3481	17.3535	17.3533	17.3535
	TPS(Quadratic)	17.3577	17.3578	17.3579	17.3570	17.3570	17.3572
	TPS(Cubic)	17.3641	17.3641	17.3641	17.3593	17.3593	17.3593
	MASS	33.4096	21.8239	18.3113	24.7088	19.1995	17.6987
	Fréchet	17.8363	17.8363	17.8363	17.8538	17.8538	17.8538
150	TPST($d = 4,\ \triangle_{3,1}$)	5.0294	5.0294	5.0295	5.0288	5.0288	5.0289
	TPST($d = 4,\ \triangle_{3,2}$)	5.0294	5.0293	5.0294	5.0288	5.0288	5.0288
	TPS(Quadratic)	5.0326	5.0327	5.0328	5.0300	5.0300	5.0302
	TPS(Cubic)	5.0347	5.0347	5.0347	5.0308	5.0308	5.0308
	MASS	16.6515	8.8635	6.1068	11.4963	6.7204	5.4236
	Fréchet	5.1921	5.1921	5.1921	5.1955	5.1955	5.1955
300	TPST($d = 4,\ \triangle_{3,1}$)	2.3797	2.3797	2.3797	2.3764	2.3764	2.3765
	TPST($d = 4,\ \triangle_{3,2}$)	2.3797	2.3797	2.3797	2.3764	2.3764	2.3764
	TPS(Quadratic)	2.3813	2.3814	2.3815	2.3770	2.3771	2.3772
	TPS(Cubic)	2.3824	2.3824	2.3824	2.3774	2.3774	2.3774
	MASS	12.1217	6.1830	3.4692	8.4402	4.1060	2.7742
	Fréchet	2.4609	2.4609	2.4609	2.4596	2.4596	2.4596
600	TPST($d = 4,\ \triangle_{3,1}$)	1.3220	1.3220	1.3221	1.3210	1.3210	1.3211
	TPST($d = 4,\ \triangle_{3,2}$)	1.3220	1.3220	1.3220	1.3210	1.3210	1.3210
	TPS(Quadratic)	1.3228	1.3229	1.3230	1.3213	1.3214	1.3215
	TPS(Cubic)	1.3233	1.3233	1.3233	1.3215	1.3215	1.3215
	MASS	10.8086	5.1865	2.4279	7.4036	3.0899	1.7247
	Fréchet	1.3627	1.3627	1.3627	1.3627	1.3627	1.3627

Table S3: Average mean squared errors (AMSEs$\times 10^{-3}$) of the TPST, TPS, MASS and Fréchet regression methods for ball domain Ω_2.

Sample		Number of Grid Points (M^3)					
Size	Method	$25 \times 25 \times 25$			$35 \times 35 \times 35$		
n		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
50	TPST($d = 4,\ \triangle_{2,1}$)	17.0357	17.0342	17.0342	17.0046	17.0044	17.0044
	TPST($d = 4,\ \triangle_{2,2}$)	17.0377	17.0362	17.0360	17.0053	17.0050	17.0050
	TPS (Quadratic)	1267.4574	1149.1364	572.3607	1248.657	1125.3152	563.5462
	TPS (Cubic)	1134.8375	1012.8265	511.4966	1124.127	1000.4489	506.441
	MASS	393.7273	1078.052	260.6591	388.2174	1549.54	291.0762
	Fréchet	17.6111	17.6111	17.6111	17.5948	17.5948	17.5948
150	TPST($d = 4,\ \triangle_{2,1}$)	4.9404	4.9402	4.9402	4.9319	4.9318	4.9318
	TPST($d = 4,\ \triangle_{2,2}$)	4.9411	4.9408	4.9409	4.9321	4.9321	4.9321
	TPS (Quadratic)	1253.344	1135.2666	559.1026	1234.6315	1111.5447	550.3515
	TPS (Cubic)	1121.0293	999.2215	498.4112	1110.3232	986.8671	493.3675
	MASS	378.1072	1106.247	304.1948	343.2141	1245.853	207.0073
	Fréchet	5.1320	5.1320	5.1320	5.1284	5.1284	5.1284
300	TPST($d = 4,\ \triangle_{2,1}$)	2.3277	2.3276	2.3276	2.3225	2.3225	2.3225
	TPST($d = 4,\ \triangle_{2,2}$)	2.3280	2.3280	2.3280	2.3226	2.3226	2.3226
	TPS (Quadratic)	1251.6722	1133.8673	557.1999	1232.9401	1110.1195	548.436
	TPS (Cubic)	1119.2052	997.6327	496.3970	1108.5280	985.2979	491.3743
	MASS	434.9718	1310.1950	354.5279	461.0489	1439.208	310.4805
	Fréchet	2.4235	2.4235	2.4235	2.4206	2.4206	2.4206
600	TPST($d = 4,\ \triangle_{2,1}$)	1.2921	1.2921	1.2921	1.2904	1.2904	1.2904
	TPST($d = 4,\ \triangle_{2,2}$)	1.2923	1.2923	1.2923	1.2905	1.2905	1.2905
	TPS (Quadratic)	1250.2103	1132.7704	555.8878	1231.479	1109.0274	547.1256
	TPS (Cubic)	1117.8003	996.5359	495.1206	1107.1138	984.2003	490.0932
	MASS	398.3621	1203.492	354.4963	388.2509	1304.9950	261.9709
	Fréchet	1.3401	1.3401	1.3401	1.3396	1.3396	1.3396

$\max_{|\boldsymbol{\alpha}|=\upsilon} \|D^{\boldsymbol{\alpha}} g\|_{\infty,\Omega}$ be the maximum norms of all the υth order derivatives of g over Ω and $\|f\|_{L^q(\Omega)} = \{\int_\Omega |f(\boldsymbol{v})|^q d\boldsymbol{v}\}^{1/q}$ be the L^q norm.

For $g_1(\boldsymbol{z})$, $g_1(\boldsymbol{z})$, define the theoretical and empirical inner products as

$$\langle g_1, g_2 \rangle = \int_\Omega g_1(\boldsymbol{z}) g_2(\boldsymbol{z}) \mathrm{d}\boldsymbol{z},\ \langle g_1, g_2 \rangle_N = \frac{1}{N} \sum_{j=1}^{N} g_1(\boldsymbol{z}_j) g_2(\boldsymbol{z}_j), \tag{S2.1}$$

and denote the corresponding theoretical and empirical norms $\|\cdot\|$ and $\|\cdot\|_N$. Furthermore, let $\|\cdot\|_{\mathcal{E}}$ be the norm introduced by the inner product $\langle \cdot, \cdot \rangle_{\mathcal{E}}$, where, for $g_1(\boldsymbol{z})$ and $g_2(\boldsymbol{z})$,

$$\langle g_1, g_2 \rangle_{\mathcal{E}} = \sum_{|\boldsymbol{\alpha}|=2} \binom{2}{\alpha_1} \binom{2-\alpha_1}{\alpha_2} \sum_{T \in \triangle} \int_T \{D^{\boldsymbol{\alpha}} g_1(\boldsymbol{z})\}\{D^{\boldsymbol{\alpha}} g_2(\boldsymbol{z})\} \mathrm{d}\boldsymbol{z}.$$

S2.1 Properties of trivariate Splines

We cite two important results from Lai and Schumaker (2007) and Li et al. (2022).

Lemma S1 (Lemma B.5 in Li et al. (2022)). *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*

$$c|\triangle|^3 \sum_{m\in\mathcal{M}} \gamma_m^2 \leq \left\| \sum_{m\in\mathcal{M}} \gamma_m B_m \right\|_{L^2(\Omega)}^2 \leq C|\triangle|^3 \sum_{m\in\mathcal{M}} \gamma_m^2.$$

Proof. This lemma follows directly from Theorem 17.18 in Lai and Schumaker (2007).

□

Lemma S2 (Theorem 3.5.2 in Lai (1989) or Lemma 1 in Li et al. (2022)). *For all $f \in \mathcal{W}^{d+1,q}(\Omega)$ with $1 \leq q \leq \infty$, $r \geq 0$, and $d \geq 6r+3$, there exists a spline $s_f \in \mathcal{S}_d^r(\triangle)$ such that*

$$\|D^{\boldsymbol{\alpha}}(f - s_f)\|_{L^q(\Omega)} \leq K|\triangle|^{d+1-|\boldsymbol{\alpha}|} |f|_{d+1,q,\Omega},$$

for all $0 \leq |\boldsymbol{\alpha}| \leq m$*, where* $K > 0$ *is a constant independent of* f *and* $|\triangle|$ *but is dependent on the geometry of* $\triangle$*.*

Lemma S2 shows that when $d \geq 6r+3$, the space $\mathcal{S}_d^r(\triangle)$ can attain the full approximation power (approximation with optimal convergence rate).

Lemma S3. *Under Assumptions (A3) and (A4), for any Bernstein basis polynomials* $B_m(\boldsymbol{z})$*,* $m \in \mathcal{M}$*, of degree* $d \geq 0$*, one has*

$$\max_{m\in\mathcal{M}} \left| \frac{1}{N}\sum_{j=1}^{N} B_m^k(\boldsymbol{z}_j) - \int_\Omega B_m^k(\boldsymbol{z})\mathrm{d}\boldsymbol{z} \right| = O(N^{-1/3}|\triangle|^2), \quad 1 \leq k < \infty, \tag{S2.2}$$

$$\max_{m,m'\in\mathcal{M}} \left| \frac{1}{N}\sum_{j=1}^{N} B_m(\boldsymbol{z}_j)B_{m'}(\boldsymbol{z}_j) - \int_\Omega B_m(\boldsymbol{z})B_{m'}(\boldsymbol{z})\mathrm{d}\boldsymbol{z} \right| = O(N^{-1/3}|\triangle|^2), \quad 1 \leq k < \infty, \tag{S2.3}$$

$$\max_{m,m'\in\mathcal{M}} \left| \frac{1}{N^2}\sum_{j,j'=1}^{N} G_\eta(\boldsymbol{z}_j,\boldsymbol{z}_{j'})B_m(\boldsymbol{z}_j)B_{m'}(\boldsymbol{z}_{j'}) - \int_{\Omega^2} G_\eta(\boldsymbol{z},\boldsymbol{z}')B_m(\boldsymbol{z})B_{m'}(\boldsymbol{z}')d\boldsymbol{z}\mathrm{d}\boldsymbol{z}' \right|$$
$$= O(N^{-1/3}|\triangle|^5), \tag{S2.4}$$

$$\max_{m\in\mathcal{M}} \left| \|\sigma B_m\|_N^2 - \|\sigma B_m\|^2 \right| = \max_{m\in\mathcal{M}} \left| \frac{1}{N}\sum_{j=1}^{N} B_m^2(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j) - \int_\Omega \sigma^2(\boldsymbol{z})B_m^2(\boldsymbol{z})\mathrm{d}\boldsymbol{z} \right|$$
$$= O(N^{-1/3}|\triangle|^2). \tag{S2.5}$$

Proof. For $m \in \mathcal{M}$, denote T_m be the tetrahedron that basis function $B_m|_{T_m} > 0$. Thus, based on the property of Bernstein basis polynomials, we have $\int_\Omega B_m^k(\boldsymbol{z})\mathrm{d}\boldsymbol{z} = \int_{T_m} B_m^k(\boldsymbol{z})\mathrm{d}\boldsymbol{z}$ for any $k \geq 1$. We first show equation (S2.2). When $d = 0$, i.e. piecewise constant basis functions, we have $B_m(\boldsymbol{z}) = I(\boldsymbol{z} \in T_m)$. Then for all $k \geq 1$,

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

where $V(T_m)$ is the volume of tetrahedron T_m. According to Assumption (A4),

$$\max_{m\in\mathcal{M}}\left|\frac{1}{N}\sum_{j=1}^{N} I(\boldsymbol{z}_j\in T_m)-V(T_m)\right|\leq CN^{-1/3}|\triangle|^2.$$

When $d\geq 1$, denote $\mathcal{V}_j$ to be the jth voxel containing $\boldsymbol{z}_j$ for $j=1,\ldots,N$, we have

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

By the properties of trivariate spline basis described in Theorem 15.3 of Lai and Schumaker (2007),

$$\int_{\Omega\backslash\cup\mathcal{V}_j} B_m^k(\boldsymbol{z})\mathrm{d}\boldsymbol{z}=O(N^{-1/3}|\triangle|^2),$$

$$\left|\sum_{j=1}^{N}\int_{\mathcal{V}_j}\left\{B_m^k(\boldsymbol{z}_j)-B_m^k(\boldsymbol{z})\right\}\mathrm{d}\boldsymbol{z}\right|\leq CN|\triangle|^3\times N^{-1}\times N^{-1/3}|\triangle|^{-1}\leq CN^{-1/3}|\triangle|^2.$$

Equation (S2.3) can be proved following a similar procedure.

Next, note that for all $m,m'\in\mathcal{M}$,

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

As $N\to\infty$, we have

$$\int_{\Omega^2\backslash\cup_{j,j'}\mathcal{V}_j\times\mathcal{V}_{j'}} G_\eta(\boldsymbol{z},\boldsymbol{z}')B_m(\boldsymbol{z})B_{m'}(\boldsymbol{z}')d\boldsymbol{z}\mathrm{d}\boldsymbol{z}'=O(N^{-1/3}|\triangle|^5).$$

Meanwhile, for $j,j'=1,\ldots,N$, denote

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

and then

$$
\begin{aligned}
&\left|\sum_{j,j'=1}^{N}\int_{\mathcal{V}_j\times\mathcal{V}_{j'}}\left\{G_\eta(\boldsymbol{z}_j,\boldsymbol{z}_{j'})B_m(\boldsymbol{z}_j)B_{m'}(\boldsymbol{z}_{j'}) - G_\eta(\boldsymbol{z},\boldsymbol{z}')B_m(\boldsymbol{z})B_{m'}(\boldsymbol{z}')\right\}d\boldsymbol{z}\mathrm{d}\boldsymbol{z}'\right| \\
&\leq \sum_{\{(j,j'):\boldsymbol{z}_j\in T_m,\boldsymbol{z}_{j'}\in T_{m'}\}}\int_{\mathcal{V}_j\times\mathcal{V}_{j'}} d_{j,j'}(G_\eta B_m B_{m'},\sqrt{6}N^{-1/3}) \\
&\leq C(N|\triangle|^3)^2\times N^{-2}\times N^{-1/3}\times|\triangle|^{-1} = O(N^{-1/3}|\triangle|^5).
\end{aligned}
$$

Thus, equation (S2.4) is proved.

Finally, using similar decomposition procedure, we have

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

According to Assumptions (A3)-(A4), $\int_{\Omega\setminus\cup_j\mathcal{V}_j}\sigma^2(\boldsymbol{z})B_m^2(\boldsymbol{z})\mathrm{d}\boldsymbol{z} = O(N^{-1/3}|\triangle|^2)$, as $n\to\infty$. Similar to the proof of (S2.4), we denote $d_j(g,\rho) = \sup_{\boldsymbol{z},\boldsymbol{z}'\in\mathcal{V}_j,\|\boldsymbol{z}-\boldsymbol{z}'\|\leq\rho}|g(\boldsymbol{z}) - g(\boldsymbol{z}')|$ and obtain

$$
\begin{aligned}
&\left|\sum_{j=1}^{N}\int_{\mathcal{V}_j}\left\{B_m^2(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j) - B_m^2(\boldsymbol{z})\sigma^2(\boldsymbol{z})\right\}\mathrm{d}\boldsymbol{z}\right| \leq \sum_{\{j:\boldsymbol{z}_j\in T_m\}}\int_{\mathcal{V}_j} d_j(B_m^2\sigma^2,\sqrt{3}N^{-1/3})\mathrm{d}\boldsymbol{z} \\
&\quad\leq C(N|\triangle|^3)\times N^{-1}\times N^{-1/3}|\triangle|^{-1}\leq CN^{-1/3}|\triangle|^2.
\end{aligned}
$$

Hence, equation (S2.5) follows. □

The following lemma provides the uniform convergence rate at which the empirical inner product approximates the theoretical inner product defined in (S2.1).

Lemma S4. *Let* $g_1(\boldsymbol{z}) = \sum_{m\in\mathcal{M}}\gamma_{1,m}B_m(\boldsymbol{z})$, $g_2(\boldsymbol{z}) = \sum_{m\in\mathcal{M}}\gamma_{2,m}B_m(\boldsymbol{z})$ *be any spline functions in* $\mathcal{S}_d^r(\triangle)$. *Suppose Assumption (A4)hold, and* $N^{1/3}|\triangle|\to\infty$ *as* $N\to\infty$, *then*

$$\omega_N = \sup_{g_1, g_2 \in \mathcal{S}_d^r(\triangle)} \left| \frac{\langle g_1, g_2\rangle_N - \langle g_1, g_2\rangle}{\|g_1\| \, \|g_2\|} \right| = O\left(N^{-1/3}|\triangle|^{-1}\right) = o(1).$$

Proof. It is easy to see

$$\begin{aligned}\langle g_1, g_2\rangle_N &= \frac{1}{N}\sum_{j=1}^{N}\left\{\sum_{m\in\mathcal{M}} \gamma_{1,m} B_m(\boldsymbol{z}_j)\right\}\left\{\sum_{m'\in\mathcal{M}} \gamma_{2,m'} B_{m'}(\boldsymbol{z}_j)\right\} \\ &= \sum_{m}\sum_{m'} \gamma_{1,m}\gamma_{2,m'} \frac{1}{N}\sum_{j=1}^{N} B_m(\boldsymbol{z}_j) B_{m'}(\boldsymbol{z}_j).\end{aligned}$$

Note that $\langle g_1, g_2\rangle = \sum_m \sum_{m'} \gamma_{1,m}\gamma_{2,m'} \int_\Omega B_m(\boldsymbol{z}) B_{m'}(\boldsymbol{z}) d\boldsymbol{z}$. It follows from Assumptions (A1), (A2) and Lemma S1 that, for any $l = 1, 2$, $\widetilde{c}_l |\triangle|^3 \sum_m \gamma_{l,m}^2 \leq \|g_l\|^2 \leq \widetilde{C}_l |\triangle|^3 \sum_m \gamma_{l,m}^2$, and

$$C_1|\triangle|^3\left[\sum_m \gamma_{1,m}^2 \sum_{m'} \gamma_{2,m'}^2\right]^{1/2} \leq \|g_1\|\|g_2\| \leq C_2|\triangle|^3\left[\sum_m \gamma_{1,m}^2 \sum_{m'} \gamma_{2,m'}^2\right]^{1/2}.$$

Therefore, one has

$$\begin{aligned}\omega_N \leq & \frac{\sum_{|m'-m|\leq (d+3)(d+2)(d+1)/6} |\gamma_{1,m}\gamma_{2,m'}|}{C_1|\triangle|^3\left[\sum_m \gamma_{1,m}^2 \sum_{m'} \gamma_{2,m'}^2\right]^{1/2}} \\ & \times \max_{m,m'\in\mathcal{M}} \left| \frac{1}{N}\sum_{j=1}^{N} B_m(\boldsymbol{z}_j) B_{m'}(\boldsymbol{z}_j) - \int_\Omega B_m(\boldsymbol{z}) B_{m'}(\boldsymbol{z}) d\boldsymbol{z} \right| \\ \leq & C|\triangle|^{-3} \max_{m,m'\in\mathcal{M}} \left| \frac{1}{N}\sum_{j=1}^{N} B_m(\boldsymbol{z}_j) B_{m'}(\boldsymbol{z}_j) - \int_\Omega B_m(\boldsymbol{z}) B_{m'}(\boldsymbol{z}) d\boldsymbol{z} \right|.\end{aligned}$$

The desired result follows from Lemma S3. □

As a direct result of Lemma S4, we can see that

$$\sup_{g\in\mathcal{S}_d^r(\triangle)} \left| \|g\|_N^2 / \|g\|^2 - 1 \right| = O\left(N^{-1/3}|\triangle|^{-1}\right) = o(1). \tag{S2.6}$$

Lemma S5. *Suppose Assumption (A4) hold, and $N^{1/3}|\triangle| \to \infty$ as $N \to \infty$, then*

$$S_N = \sup_{g\in\mathcal{S}_d^r(\triangle)} \left\{ \frac{\|g\|_\infty}{\|g\|_N}, \|g\|_N \neq 0 \right\} = O(|\triangle|^{-3/2}), \tag{S2.7}$$

$$\overline{S}_N = \sup_{g \in \mathcal{S}_d^r(\triangle)} \left\{ \frac{\|g\|_{\mathcal{E}}}{\|g\|_N}, \|g\|_N \neq 0 \right\} = O(|\triangle|^{-2}). \tag{S2.8}$$

Proof. By Markov's inequality, for any $g \in \mathcal{S}_d^r(\triangle)$, $\|g\|_{\mathcal{E}} \leq C|\triangle|^{-2}\|g\|$ (Lemma 15.2 and Theorem 15.28 in Lai and Schumaker (2007)). Moreover, for some other constant C_1, $\|g\|_\infty \leq C_1|\triangle|^{-3/2}\|g\|$. By (S2.6), $\|g\|_N/\|g\| \geq \left[1 - O_P\left\{N^{-1/3}|\triangle|^{-1}\right\}\right]^{1/2}$. Thus, one has

$$S_N \leq C|\triangle|^{-3/2}\left[1 - O_P\left\{N^{-1/3}|\triangle|^{-1}\right\}\right]^{-1/2} = O\left(|\triangle|^{-3/2}\right),$$
$$\overline{S}_N \leq C|\triangle|^{-2}\left[1 - O_P\left\{N^{-1/3}|\triangle|^{-1}\right\}\right]^{-1/2} = O\left(|\triangle|^{-2}\right).$$

Lemma S5 is established. □

S2.2 Convergence of Penalized Spline Estimators

Let $\left\{\widetilde{B}_m(\boldsymbol{z}), m \in \widetilde{\mathcal{M}}\right\}$ be a set of transformed Bernstein basis polynomials and $\widetilde{\mathbf{B}}(\boldsymbol{z}) = \mathbf{Q}_2^\top \mathbf{B}(\boldsymbol{z})$, then $\mathbf{U} = \mathbf{B}\mathbf{Q}_2$, $\mathbf{U}^\top\mathbf{U} = \sum_{j=1}^N \widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}^\top(\boldsymbol{z}_j)$ and $\mathbf{U}^\top\bar{\mathbf{Y}} = \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^N \widetilde{\mathbf{B}}(\boldsymbol{z}_j)Y_{ij}$. Denote by

$$\boldsymbol{\Gamma}_{N,\rho} = \frac{1}{N}\sum_{j=1}^N\{\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}^\top(\boldsymbol{z}_j)\} + \frac{\rho_n}{nN}\mathbf{Q}_2^\top[\langle B_m, B_{m'}\rangle_{\mathcal{E}}]_{m,m'\in\mathcal{M}}\mathbf{Q}_2 \tag{S2.9}$$

a symmetric positive definite matrix.

According to Li et al. (2022), the matrix $\boldsymbol{\Gamma}_{N,\rho}$ possess the following asymptotic properties.

Lemma S6 (Lemma B.7 in Li et al. (2022))**.** *Under Assumption (A4), $d \geq 6r+3$, if $N^{1/3}|\triangle| \to \infty$ as $N \to \infty$, then we have the following properties:*

(i) There exist constants $0 < c_\rho < C_\rho < \infty$, such that with probability approaching 1 as $n \to \infty$,

$$c_\rho|\triangle|^3 \leq \lambda_{\min}(\boldsymbol{\Gamma}_{N,\rho}) \leq \lambda_{\max}(\boldsymbol{\Gamma}_{N,\rho}) \leq C_\rho\left(|\triangle|^3 + \frac{\rho_n}{nN|\triangle|}\right).$$

Specifically, when $\rho_n = 0$, *one has* $c_0|\triangle|^3 \leq \lambda_{\min}(\mathbf{\Gamma}_{N,0}) \leq \lambda_{\max}(\mathbf{\Gamma}_{N,0}) \leq C_0|\triangle|^3$.

(ii) There exists constant $M_d > 0$ *such that* $\|\widetilde{\mathbf{\Gamma}}_{N,\rho}^{-1}\|_\infty \leq M_d|\triangle|^{-3}$.

(iii) For every vector $\boldsymbol{v} = (v_1, \ldots, v_N)^\top$, *there exist a constant* $C_d > 0$ *such that*

$$\left\|\widetilde{\mathbf{B}}(\boldsymbol{z})^\top \ \mathbf{\Gamma}_{N,\rho}^{-1} \frac{1}{N} \sum_{j=1}^{N} \widetilde{\mathbf{B}}(\boldsymbol{z}_j) v_j\right\|_\infty \leq C_d \|\boldsymbol{v}\|_\infty.$$

Using $\mathbf{\Gamma}_{N,\rho}$ defined in (S2.9), the solution of the penalized regression problem (2.4) is given by

$$\widehat{\boldsymbol{\theta}} = \mathbf{\Gamma}_{N,\rho}^{-1} \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} \widetilde{\mathbf{B}}(\boldsymbol{z}_j) Y_{ij}.$$

Next we define

$$\widehat{\boldsymbol{\theta}}_\mu = \mathbf{\Gamma}_{N,\rho}^{-1} \frac{1}{N} \sum_{j=1}^{N} \widetilde{\mathbf{B}}(\boldsymbol{z}_j)\mu(\boldsymbol{z}_j), \quad \widehat{\boldsymbol{\theta}}_\eta = \mathbf{\Gamma}_{N,\rho}^{-1} \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} \widetilde{\mathbf{B}}(\boldsymbol{z}_j) \sum_{k=1}^{\infty} \xi_{ik}\phi_k(\boldsymbol{z}_j),$$
$$\widehat{\boldsymbol{\theta}}_\varepsilon = \mathbf{\Gamma}_{N,\rho}^{-1} \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} \widetilde{\mathbf{B}}(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}. \tag{S2.10}$$

Note that, the BPS estimator $\widehat{\mu}$ in Section 2.3 can be written as $\widehat{\mu}(\boldsymbol{z}) = \widehat{\mu}^o(\boldsymbol{z}) + \widehat{\eta}(\boldsymbol{z}) + \widehat{\varepsilon}(\boldsymbol{z})$, where

$$\widehat{\mu}^o(\boldsymbol{z}) = \widetilde{\mathbf{B}}(\boldsymbol{z})^\top \widehat{\boldsymbol{\theta}}_\mu, \quad \widehat{\eta}(\boldsymbol{z}) = \widetilde{\mathbf{B}}(\boldsymbol{z})^\top \widehat{\boldsymbol{\theta}}_\eta, \quad \widehat{\varepsilon}(\boldsymbol{z}) = \widetilde{\mathbf{B}}(\boldsymbol{z})^\top \widehat{\boldsymbol{\theta}}_\varepsilon, \tag{S2.11}$$

Therefore,

$$\widehat{\mu}(\boldsymbol{z}) - \mu(\boldsymbol{z}) = \widehat{\mu}^o(\boldsymbol{z}) - \mu(\boldsymbol{z}) + \widehat{\eta}(\boldsymbol{z}) + \widehat{\varepsilon}(\boldsymbol{z}). \tag{S2.12}$$

Lemma S7. *Suppose Assumptions (A2)–(A4) hold and* $N^{1/3}|\triangle| \to \infty$ *as* $N \to \infty$, *then* $\|\widehat{\boldsymbol{\theta}}_\eta\|^2 = O_P(n^{-1}|\triangle|^{-3})$ *and* $\|\widehat{\boldsymbol{\theta}}_\varepsilon\|^2 = O_P(n^{-1}N^{-1}|\triangle|^{-6})$.

Proof. Note that

$$\widehat{\boldsymbol{\theta}}_\eta = \mathbf{\Gamma}_{n,\rho}^{-1} \frac{1}{nN} \sum_{i=1}^{n} \sum_{j=1}^{N} \widetilde{\mathbf{B}}(\boldsymbol{z}_j) \sum_{k=1}^{\infty} \xi_{ik}\phi_k(\boldsymbol{z}_j).$$

By Lemma S6, one has

$$\|\widehat{\boldsymbol{\theta}}_{\eta}\|^2 \asymp \frac{1}{n^2N^2|\triangle|^6}\sum_{i,i'=1}^{n}\sum_{j,j'=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^{\top}\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\sum_{k,k'=1}^{\infty}\xi_{ik}\phi_k(\boldsymbol{z}_j)\xi_{i'k'}\phi_{k'}(\boldsymbol{z}_{j'}).$$

Note that by Assumption (A2), for any $i \neq i'$, j, j', one has

$$\begin{aligned}&\mathrm{E}\Big\{\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^{\top}\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\sum_{k,k'=1}^{\infty}\xi_{ik}\phi_k(\boldsymbol{z}_j)\xi_{i'k'}\phi_{k'}(\boldsymbol{z}_{j'})\Big\}\\ =&\sum_{m\in\widetilde{\mathcal{M}}}\widetilde{B}_m(\boldsymbol{z}_j)\widetilde{B}_m(\boldsymbol{z}_{j'})\sum_{k,k'}\mathrm{E}\xi_{ik}\xi_{i'k'}\phi_k(\boldsymbol{z}_j)\phi_{k'}(\boldsymbol{z}_{j'})=0.\end{aligned}$$

Next, for any i, because $\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^{\top}\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'}) = \mathbf{B}(\boldsymbol{z}_j)^{\top}\mathbf{Q}_2\mathbf{Q}_2^{\top}\mathbf{B}(\boldsymbol{z}_{j'})$ and the eigenvalues of $\mathbf{Q}_2\mathbf{Q}_2^{\top}$ are either 0 or 1,

$$\begin{aligned}&\frac{1}{N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}\mathrm{E}\left\{\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^{\top}\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\sum_{k,k'=1}^{\infty}\xi_{ik}\phi_k(\boldsymbol{z}_j)\xi_{ik'}\phi_{k'}(\boldsymbol{z}_{j'})\right\}\\ &\quad\leq\frac{1}{N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}\mathrm{E}\left\{\mathbf{B}(\boldsymbol{z}_j)^{\top}\mathbf{B}(\boldsymbol{z}_{j'})\sum_{k=1}^{\infty}\xi_{ik}^2\phi_k(\boldsymbol{z}_j)\phi_k(\boldsymbol{z}_{j'})\right\}\\ &\quad=\sum_{m\in\mathcal{M}}\frac{1}{N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}B_m(\boldsymbol{z}_j)B_m(\boldsymbol{z}_{j'})G_{\eta}(\boldsymbol{z}_j,\boldsymbol{z}_{j'}).\end{aligned}$$

Assumption (A4) and Lemma S3 imply that

$$\begin{aligned}\frac{1}{N^2}\sum_{j\neq j'}B_m(\boldsymbol{z}_j)B_m(\boldsymbol{z}_{j'})G_{\eta}(\boldsymbol{z}_j,\boldsymbol{z}_{j'}) &= \int_{\Omega^2}G_{\eta}(\boldsymbol{z},\boldsymbol{z}')B_m(\boldsymbol{z})B_m(\boldsymbol{z}')d\boldsymbol{z}d\boldsymbol{z}'\\ &\qquad\times\{1+O(N^{-1/3}|\triangle|^5)\}=O(|\triangle|^6).\end{aligned}$$

Thus,

$$\frac{1}{N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}\mathrm{E}\left\{\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^{\top}\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\sum_{k,k'=1}^{\infty}\xi_{ik}\phi_k(\boldsymbol{z}_j)\xi_{ik'}\phi_{k'}(\boldsymbol{z}_{j'})\right\}\leq C|\triangle|^3.$$

Therefore, $\mathrm{E}\|\widehat{\boldsymbol{\theta}}_{\eta}\|^2 \leq C(n^{-1}|\triangle|^{-3})$.

Similarly, by the definition of $\widehat{\boldsymbol{\theta}}_{\varepsilon}$ in (S2.10) and Lemma S6, one has

$$\|\widehat{\boldsymbol{\theta}}_{\varepsilon}\|^2 \asymp \frac{1}{n^2N^2|\triangle|^6}\sum_{i,i'=1}^{n}\sum_{j,j'=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^{\top}\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\sigma(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_{j'})\varepsilon_{ij}\varepsilon_{i'j}.$$

Note that for any $i \neq i'$, j, j', $\mathrm{E}\{\varepsilon_{ij}\varepsilon_{i'j'}\} = 0$ and for any $i, j \neq j'$, $\mathrm{E}\{\varepsilon_{ij}\varepsilon_{ij'}\} = 0$. Because the eigenvalues of $\mathbf{Q}_2\mathbf{Q}_2^\top$ are either 0 or 1, by Assumption (A2) and Lemma S3, for any i,

$$\mathrm{E}\Big\{\frac{1}{N}\sum_{j,j'=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^\top\widetilde{\mathbf{B}}(\boldsymbol{z}_{j'})\sigma(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_{j'})\varepsilon_{ij}\varepsilon_{i'j}\Big\} = \frac{1}{N}\sum_{j=1}^{N}\mathbf{B}(\boldsymbol{z}_j)^\top\mathbf{Q}_2\mathbf{Q}_2^\top\mathbf{B}(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j)$$

$$\leq C\sum_{m\in\mathcal{M}}\frac{1}{N}\sum_{j=1}^{N}B_m^2(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j) \leq C\sum_{m\in\mathcal{M}}\int_\Omega \sigma^2(\boldsymbol{z})B_m^2(\boldsymbol{z})d\boldsymbol{z}\{1+O(N^{-1/3}|\triangle|^2)\} = O(1).$$

Therefore,

$$\mathrm{E}\|\widehat{\boldsymbol{\theta}}_\varepsilon\|^2 \asymp \frac{1}{nN|\triangle|^6}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)^\top\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j) \leq C(nN)^{-1}|\triangle|^{-6}.$$

The conclusion of the lemma follows. □

Next, the following lemmas give the uniform convergence rate of $\widehat{\mu}(\boldsymbol{z})$ to $\mu(\boldsymbol{z})$. We start by introducing some notations for the specific situation when there is no penalty in the regression problem, i.e., $\rho_n = 0$. Denote $\boldsymbol{\Gamma}_{N,0} = \frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}^\top(\boldsymbol{z}_j)$. Let $\bar{\xi}_{\cdot k} = \frac{1}{n}\sum_{i=1}^{n}\xi_{ik}$, for any $k \geq 1$, and $\bar{\varepsilon}_{\cdot j} = \frac{1}{n}\sum_{i=1}^{n}\varepsilon_{ij}$ for any $j = 1,\ldots,N$, and denote

$$\begin{aligned}
\widetilde{\boldsymbol{\theta}}_\mu &= \boldsymbol{\Gamma}_{N,0}^{-1}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\mu(\boldsymbol{z}_j),\\
\widetilde{\boldsymbol{\theta}}_\eta &= \boldsymbol{\Gamma}_{N,0}^{-1}\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\eta_i(\boldsymbol{z}_j) = \boldsymbol{\Gamma}_{N,0}^{-1}\frac{1}{N}\sum_{j=1}^{N}\sum_{k=1}^{\kappa}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\bar{\xi}_{\cdot k}\phi_k(\boldsymbol{z_j}),\\
\widetilde{\boldsymbol{\theta}}_\varepsilon &= \boldsymbol{\Gamma}_{N,0}^{-1}\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij} = \boldsymbol{\Gamma}_{N,0}^{-1}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\bar{\varepsilon}_{\cdot j},
\end{aligned}$$

Then we can have the following decomposition $\widetilde{\mu}(\boldsymbol{z}) = \widetilde{\mu}^o(\boldsymbol{z}) + \widetilde{\eta}(\boldsymbol{z}) + \widetilde{\varepsilon}(\boldsymbol{z})$, where

$$\widetilde{\mu}^o(\boldsymbol{z}) = \widetilde{\mathbf{B}}(\boldsymbol{z})^\top\widetilde{\boldsymbol{\theta}}_\mu,\ \widetilde{\eta}(\boldsymbol{z}) = \widetilde{\mathbf{B}}(\boldsymbol{z})^\top\widetilde{\boldsymbol{\theta}}_\eta,\ \widetilde{\varepsilon}(\boldsymbol{z}) = \widetilde{\mathbf{B}}(\boldsymbol{z})^\top\widetilde{\boldsymbol{\theta}}_\varepsilon. \tag{S2.13}$$

Lemma S8. *Under Assumptions (A1) and (A4), if $N^{1/3}|\triangle| \to \infty$ as $N \to \infty$, the functions $\widehat{\mu}^o(\boldsymbol{z})$ satisfy*

$$\|\widehat{\mu}^o - \mu\|_\infty = O_P\left\{\frac{\rho_n}{nN|\triangle|^{7/2}}\|\mu\|_{2,\infty} + \left(1 + \frac{\rho_n}{nN|\triangle|^{11/2}}\right)|\triangle|^{\ell+1}\|\mu\|_{\ell+1,\infty}\right\}.$$

Proof. Note that $\|\mu - \widehat{\mu}^o\|_\infty \leq \|\mu - \widetilde{\mu}^o\|_\infty + \|\widetilde{\mu}^o - \widehat{\mu}^o\|_\infty$, where $\widetilde{\mu}^o$ is given in (S2.13). According to Proposition 1 in Li et al. (2022), $\|\widetilde{\mu}^o - \mu\|_\infty \leq C|\triangle|^{\ell+1}\|\mu\|_{\ell+1,\infty}$. Thus we only need to show the order of $\|\widetilde{\mu}^o - \widehat{\mu}^o\|_\infty$.

By the definition of S_N in (S2.7), one has

$$\|\widetilde{\mu}^o - \widehat{\mu}^o\|_\infty \leq S_N\|\widetilde{\mu}^o - \widehat{\mu}^o\|_N. \tag{S2.14}$$

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

$$nN\langle \mu - \widehat{\mu}^o, g\rangle_N = \rho_n\langle\widehat{\mu}^o, g\rangle_{\mathcal{E}}, \quad \text{for all } g \in \mathcal{S}_d^r(\triangle),$$

while $\widetilde{\mu}^o$ is characterized by $\langle \mu - \widetilde{\mu}^o, g\rangle_N = 0$, for all $g \in \mathcal{S}_d^r(\triangle)$.

Combining the two orthogonality relations, one has $nN\langle\widetilde{\mu}^o - \widehat{\mu}^o, g\rangle_N = \rho_n\langle\widehat{\mu}^o, g\rangle_{\mathcal{E}}$, for all $g \in \mathcal{S}_d^r(\triangle)$. Inserting $g = \widetilde{\mu}^o - \widehat{\mu}^o$ yields that

$$nN\|\widetilde{\mu}^o - \widehat{\mu}^o\|_N^2 = \rho_n\langle\widehat{\mu}^o, \widetilde{\mu}^o - \widehat{\mu}^o\rangle_{\mathcal{E}} = \rho_n\left\{\langle\widehat{\mu}^o, \widetilde{\mu}^o\rangle_{\mathcal{E}} - \|\widehat{\mu}^o\|_{\mathcal{E}}^2\right\} \geq 0.$$

Thus, by Cauchy-Schwarz inequality, $\|\widehat{\mu}^o\|_{\mathcal{E}}^2 \leq \langle\widehat{\mu}^o, \widetilde{\mu}^o\rangle_{\mathcal{E}} \leq \|\widehat{\mu}^o\|_{\mathcal{E}}\|\widetilde{\mu}^o\|_{\mathcal{E}}$, which implies that $\|\widehat{\mu}^o\|_{\mathcal{E}} \leq \|\widetilde{\mu}^o\|_{\mathcal{E}}$. Meanwhile, by the definition of $\overline{S}_N$,

$nN\|\widetilde{\mu}^o - \widehat{\mu}^o\|_N^2 \leq \rho_n\|\widehat{\mu}^o\|_{\mathcal{E}}\|\widetilde{\mu}^o - \widehat{\mu}^o\|_{\mathcal{E}} \leq \rho_n\overline{S}_N\|\widehat{\mu}^o\|_{\mathcal{E}}\|\widetilde{\mu}^o - \widehat{\mu}^o\|_N \leq \rho_n\overline{S}_N\|\widetilde{\mu}^o\|_{\mathcal{E}}\|\widetilde{\mu}^o - \widehat{\mu}^o\|_N$.
Therefore,

$$\|\widetilde{\mu}^o - \widehat{\mu}^o\|_N \leq \rho_n(nN)^{-1}\overline{S}_N\|\widetilde{\mu}^o\|_{\mathcal{E}}. \tag{S2.15}$$

Combining (S2.14) and (S2.15) yields that

$$\|\widetilde{\mu}^o - \widehat{\mu}^o\|_\infty \leq S_N\|\widetilde{\mu}^o - \widehat{\mu}^o\|_N \leq \rho_n(nN)^{-1}S_N\overline{S}_N\|\widetilde{\mu}^o\|_{\mathcal{E}}.$$

By Lemma S2, one has

$$\|\widetilde{\mu}^o\|_{\mathcal{E}} = C_1\{\|\mu\|_{2,\infty} + \sum_{|\boldsymbol{\alpha}|=2}\|D^{\boldsymbol{\alpha}}(\mu - \widetilde{\mu}^o)\|_\infty\} \leq C_2(\|\mu\|_{2,\infty} + |\triangle|^{\ell-1}\|\mu\|_{\ell+1,\infty}).$$

It follows $\|\widetilde{\mu}^o - \widehat{\mu}^o\|_\infty = \rho_n (nN)^{-1} S_N \overline{S}_N C_2 (\|\mu\|_{2,\infty} + |\triangle|^{\ell-1}\|\mu\|_{\ell+1,\infty})$. By Lemma S5, one has $S_N = O_P\left(|\triangle|^{-3/2}\right)$ and $\overline{S}_N = O_P\left(|\triangle|^{-2}\right)$. Thus,

$$\|\widetilde{\mu}^o - \widehat{\mu}^o\|_\infty = O_P\left\{\frac{\rho_n}{nN|\triangle|^{7/2}}(\|\mu\|_{2,\infty} + |\triangle|^{\ell-1}\|\mu\|_{\ell+1,\infty})\right\}.$$

Hence, $\|\widehat{\mu}^o - \mu\|_\infty \leq C_1|\triangle|^{\ell+1}\|\mu\|_{\ell+1,\infty} + O_P\left\{\frac{\rho_n}{nN|\triangle|^{7/2}}\left(\|\mu\|_{2,\infty} + |\triangle|^{\ell-1}\|\mu\|_{\ell+1,\infty}\right)\right\}$. Lemma S8 is established. □

Lemma S9. *Under Assumptions (A2)–(A5), if* $N^{1/3}|\triangle| \to \infty$ *as* $N \to \infty$, *then* $\|\widetilde{\varepsilon}\|_\infty = O_P\{(nN)^{-1/2}(\log n)^{1/2}|\triangle|^{-3/2}\}$. *In addition, if Assumption (A6) holds, then* $\|\widetilde{\eta}\|_\infty = O_P\{n^{-1/2}(\log n)^{1/2}\}$.

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

$$\begin{aligned}\|\widetilde{\varepsilon}\|_\infty &= \left\|\widetilde{\mathbf{B}}(\boldsymbol{z})^\top \boldsymbol{\Gamma}_{N,0}^{-1}\left[\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}\right]_{m\in\widetilde{\mathcal{M}}}\right\|_\infty \\ &\leq C|\triangle|^{-3}\max_{m\in\widetilde{\mathcal{M}}}\left|\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}\right|,\end{aligned}$$

almost surely, where $\widetilde{\boldsymbol{\theta}}_\varepsilon = (\widetilde{\theta}_{\varepsilon,m})_{m\in\widetilde{\mathcal{M}}}$ with $\widetilde{\mathcal{M}}$ being an index set of the transformed Bernstein basis polynomials $\widetilde{B}_m(\boldsymbol{z})$. Next we show that with probability 1

$$\max_{m\in\widetilde{\mathcal{M}}}\left|\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}\right| = O\left\{(\log n)^{1/2}|\triangle|^{3/2}(nN)^{-1/2}\right\}. \tag{S2.16}$$

To prove (S2.16), let $\tau_i = \tau_{i,m} = \frac{1}{nN}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}$. We decompose the random variable τ_i into a truncated part and a tail part,

$$\begin{aligned}
\tau_{i,1}^{L_n} &= \frac{1}{nN}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}I\left\{|\varepsilon_{ij}|>L_n\right\},\\
\tau_{i,2}^{L_n} &= \frac{1}{nN}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}I\left\{|\varepsilon_{ij}|\le L_n\right\}-\mu_i^{L_n},\\
\mu_i^{L_n} &= \frac{1}{nN}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\mathrm{E}\left[\varepsilon_{ij}I\left\{|\varepsilon_{ij}|\le L_n\right\}\right],
\end{aligned}$$

where $L_n = n^{\alpha}$, and $n^{1/(4+\delta_2)} \ll n^{\alpha} \ll \sqrt{\frac{n}{N\log n}}|\triangle|^{-3/2}$.

It is straightforward to verify that $\mu_i^{L_n} = O(n^{-1}L_n^{-3}|\triangle|^3)$. Next we show that tail part vanishes almost surely. Note that

$$\sum_{n=1}^{\infty}P\left\{|\varepsilon_{nj}|>L_n\right\}\le\sum_{n=1}^{\infty}\frac{\mathrm{E}\left|\varepsilon_{nj}\right|^{4+\delta_2}}{L_n^{4+\delta_2}}\le\upsilon_\delta\sum_{n=1}^{\infty}L_n^{-(4+\delta_2)}<\infty. \tag{S2.17}$$

By Borel Cantelli lemma, one has $\left|\sum_{i=1}^{n}\tau_{i,1}^{L_n}\right| = O_{a.s.}\left(n^{-k}\right)$, for any $k>0$. Next, note that $\mathrm{E}\left(\tau_{i,2}^{L_n}\right)=0$, one has

$$\begin{aligned}
\mathrm{Var}(\tau_{i,2}^{L_n}) =& \frac{1}{n^2N^2}\sum_{j=1}^{N}\widetilde{B}_m^2(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j)\\
&\times\left\{\mathrm{E}\left(\varepsilon_{ij}^2\right)-\mathrm{E}[\varepsilon_{ij}^2I\{|\varepsilon_{ij}|>L_n\}]-\left(\mathrm{E}\left[\varepsilon_{ij}I\left\{|\varepsilon_{ij}|\le L_n\right\}\right]\right)^2\right\}\\
\asymp& n^{-2}N^{-1}|\triangle|^3.
\end{aligned}$$

Using the independence of $\tau_{i,2}^{L_n}$, $i=1,\ldots,n$, one has $\mathrm{Var}\left(\sum_{i=1}^{n}\tau_{i,2}^{L_n}\right)\asymp(nN)^{-1}|\triangle|^3$.

Now Minkowski's inequality implies that

$$\begin{aligned}
\mathrm{E}\left|\tau_{i,2}^{L_n}\right|^k =& \mathrm{E}\Big|\frac{1}{nN}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}I\left\{|\varepsilon_{ij}|\le L_n\right\}-\mu_i^{L_n}\Big|^k\\
\le& 2^{k-1}\left[\left\{\frac{1}{nN}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)L_n\right\}^{k-2}\right.\\
&\left.\times\mathrm{E}\left|\frac{1}{nN}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}I\left\{|\varepsilon_{ij}|\le L_n\right\}\right|^2+\left(\mu_i^{L_n}\right)^k\right]\\
\le& 2\left(\frac{C|\triangle|^3L_n}{n}\right)^{k-2}\mathrm{E}\left|\tau_{i,2}^{L_n}\right|^2,\quad k\ge3.
\end{aligned}$$

Thus, $\mathrm{E}\left|\tau_{i,2}^{L_n}\right|^k \le \left(\frac{CL_n|\triangle|^3}{n}\right)^{k-2} k!\mathrm{E}|\tau_{i,2}^{L_n}|^2 < \infty$ with the Cramer constant $c^* = Cn^{-1}L_n|\triangle|^3$. By the Bernstein inequality, for any large enough $\delta > 0$,

$$\begin{aligned}
&P\left(\left|\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}\right| \ge \delta|\triangle|^{3/2}\sqrt{\frac{\log n}{nN}}\right)\\
&= P\left(\left|\sum_{i=1}^{n}\tau_{i,2}^{L_n}\right| \ge \delta|\triangle|^{3/2}\sqrt{\frac{\log n}{nN}}\right)\\
&\le 2\exp\left\{\frac{-\delta^2|\triangle|^3\frac{\log n}{nN}}{4\mathrm{Var}\left(\sum_{i=1}^{n}\tau_{i,2}^{L_n}\right)+2c^*\delta|\triangle|^{3/2}\sqrt{\frac{\log n}{nN}}}\right\}\\
&= 2\exp\left\{\frac{-\delta^2|\triangle|^3\frac{\log n}{nN}}{\frac{4c}{nN}|\triangle|^3+2CL_nn^{-1}\delta|\triangle|^{9/2}\sqrt{\frac{\log n}{nN}}}\right\}\\
&= 2\exp\left\{\frac{-\delta^2\log n}{4c+2CL_n\delta|\triangle|^{3/2}\sqrt{\frac{N\log n}{n}}}\right\} \le 2n^{-3},
\end{aligned}$$

given that $L_n = n^{\alpha} = o\{n^{1/2}(N\log n)^{-1/2}|\triangle|^{-3/2}\}$. Hence

$$\sum_{n=1}^{\infty}P\left(\max_{m\in\mathcal{M}}\left|\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}\right| \ge \delta|\triangle|^{3/2}\sqrt{\frac{\log n}{nN}}\right) \le c|\triangle|^{-3}\sum_{n=1}^{\infty}n^{-3} < \infty,$$

for such $\delta > 0$. Thus, Borel-Cantelli's lemma implies (S2.16).

Similarly, for $\widetilde{\eta}(\boldsymbol{z}) = \sum_{m\in\widetilde{\mathcal{M}}}\widetilde{\theta}_{\eta,m}\widetilde{B}_m(\boldsymbol{z})$ one has

$$\|\widetilde{\eta}\|_{\infty} \le C|\triangle|^{-3}\max_{m\in\widetilde{\mathcal{M}}}\left|\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}\eta_i(\boldsymbol{z}_j)\widetilde{B}_m(\boldsymbol{z}_j)\right|,$$

almost surely. Then we can show that with probability 1

$$\max_{m\in\widetilde{\mathcal{M}}}\left|\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}\eta_i(\boldsymbol{z}_j)\widetilde{B}_m(\boldsymbol{z}_j)\right| = O\left\{n^{-1/2}|\triangle|^3(\log n)^{1/2}\right\},$$

by decomposing mean 0 random variable $u_i = u_{i,m} = \frac{1}{N}\sum_{j=1}^{N} \eta_i(\boldsymbol{z}_j)\widetilde{B}_m(\boldsymbol{z}_j)$ into

$$\begin{aligned} u_{i,1}^{T_n} &= \sum_{k=1}^{\infty}\left\{\frac{1}{nN}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\phi_k(\boldsymbol{z}_j)\right\}\xi_{ik}I\left\{|\xi_{ik}| > T_n\right\}, \\ u_{i,2}^{T_n} &= \sum_{k=1}^{\infty}\left\{\frac{1}{nN}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\phi_k(\boldsymbol{z}_j)\right\}\xi_{ik}I\left\{|\xi_{ik}| \le T_n\right\} - \mu_i^{T_n}, \\ \mu_i^{T_n} &= \sum_{k=1}^{\infty}\left\{\frac{1}{nN}\sum_{j=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\phi_k(\boldsymbol{z}_j)\right\}\mathrm{E}\left[\xi_{ik}I\left\{|\xi_{ik}| \le T_n\right\}\right], \end{aligned}$$

where $T_n = n^{\alpha}$ and $n^{1/(4+\delta_1)} \ll n^{\alpha} \ll \sqrt{n/\log n}$.

Using Borel Cantelli lemma and similar method in (S2.17), we can show that tail part vanishes almost surely, i.e., $\left|\sum_{i=1}^{n} u_{i,1}^{T_n}\right| = O_{a.s.}\left(n^{-r}\right),$ for any $r > 0$. As $\mathrm{E}u_i = 0$, then it is straightforward to verify that $\mu_i^{T_n} = -\mathrm{E}u_{i,1}^{T_n} = O(n^{-1}T_n^{-3}|\triangle|^3)$.

Next, notice that $\mathrm{E}u_{i,2}^{T_n} = 0$. Then, one has

$$\begin{aligned} \mathrm{Var}\left(u_{i,2}^{T_n}\right) =& \frac{1}{n^2N^2}\sum_{k=1}^{\infty}\left\{\sum_{j=1}^{N}\sum_{j'=1}^{N}\widetilde{B}_m(\boldsymbol{z}_j)\widetilde{B}_m(\boldsymbol{z}_{j'})\phi_k(\boldsymbol{z}_l)\phi_k(\boldsymbol{z}_{j'})\right\} \\ &\times\left\{\mathrm{E}(\xi_{ik}^2) - \mathrm{E}\Big[\xi_{ik}^2 I\{|\xi_{ik}| > T_n\}\Big] - \left(\mathrm{E}\left[\xi_{ik}I\left\{|\xi_{ik}| \le T_n\right\}\right]\right)^2\right\} = O(n^{-2}|\triangle|^6), \end{aligned}$$

which indicates $\mathrm{Var}\left(\sum_{i=1}^{n} u_{i,2}^{T_n}\right) = n^{-1}|\triangle|^6$.

Similarly, we can show that there exists some constant C for any $r \ge 3$ such that

$$\mathrm{E}|u_{i,2}^{T_n}|^r \le \left(\frac{C|\triangle|^3 T_n}{n}\right)^{r-2} r!\mathrm{E}|u_{i,2}^{T_n}|^2.$$

Using Bernstein inequality, one has

$$P\left\{\left|\sum_{i=1}^{n} u_i\right| \ge \delta n^{-1/2}|\triangle|^3(\log n)^{1/2}\right\} \le 2\exp\left\{\frac{-\delta^2\log n}{4c + 2\delta C T_n(\log n)^{1/2}n^{-1/2}}\right\} \le 2n^{-3}.$$

Hence,

$$\sum_{n=1}^{\infty} P\left\{\max_{m\in\widetilde{\mathcal{M}}}\left|\sum_{i=1}^{n} u_i\right| \ge \delta n^{-1/2}|\triangle|^3(\log n)^{1/2}\right\} \le C|\triangle|^{-3}\sum_{n=1}^{\infty} n^{-3} < \infty$$

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

$\square$

Lemma S10. *Under Assumptions (A2)–(A4), one has*

$$\|\widehat{\varepsilon}\|_\infty = O_P\left\{\frac{(\log n)^{1/2}}{\sqrt{nN}|\triangle|^{3/2}} + \frac{\rho_n}{n^{3/2}N^{3/2}|\triangle|^7}\right\},\ \|\widehat{\eta}\|_\infty = O_P\left\{\frac{(\log n)^{1/2}}{\sqrt{n}} + \frac{\rho_n}{n^{3/2}N|\triangle|^{11/2}}\right\}.$$

Proof. We only show the infinity norm of $\widehat{\varepsilon}$. The conclusion of $\|\widehat{\eta}\|_\infty$ follows similarly. Note that the penalized spline $\widehat{\varepsilon}$ of ε is characterized by the orthogonality relations

$$nN\left\langle \varepsilon - \widehat{\varepsilon}, g\right\rangle_N = \rho_n \left\langle \widehat{\varepsilon}, g\right\rangle_{\mathcal{E}}, \quad \text{for all } g \in \mathcal{S}_d^r(\triangle).$$

In particular, $\widetilde{\varepsilon}$ is characterized by $\left\langle \varepsilon - \widetilde{\varepsilon}, g\right\rangle_N = 0$, for all $g \in \mathcal{S}_d^r(\triangle)$. Inserting $g = \widetilde{\varepsilon} - \widehat{\varepsilon}$ yield that

$$nN\left\|\widetilde{\varepsilon} - \widehat{\varepsilon}\right\|_N^2 = \rho_n \left\langle \widehat{\varepsilon}, \widetilde{\varepsilon} - \widehat{\varepsilon}\right\rangle_{\mathcal{E}} = \rho_n(\langle \widehat{\varepsilon}, \widetilde{\varepsilon}\rangle_{\mathcal{E}} - \|\widehat{\varepsilon}\|_{\mathcal{E}}).$$

It follows, by Cauchy-Schwarz inequality, that $\|\widehat{\varepsilon}\|_{\mathcal{E}}^2 \leq \langle \widehat{\varepsilon}, \widetilde{\varepsilon}\rangle_{\mathcal{E}} \leq \|\widehat{\varepsilon}\|_{\mathcal{E}}\|\widetilde{\varepsilon}\|_{\mathcal{E}}$, which implies that $\|\widehat{\varepsilon}\|_{\mathcal{E}} \leq \|\widetilde{\varepsilon}\|_{\mathcal{E}}$. Thus, by Cauchy-Schwarz inequality and the definition of $\overline{S}_N$ in (S2.8), one has

$$nN\left\|\widetilde{\varepsilon} - \widehat{\varepsilon}\right\|_N^2 \leq \rho_n \|\widehat{\varepsilon}\|_{\mathcal{E}}\|\widetilde{\varepsilon} - \widehat{\varepsilon}\|_{\mathcal{E}} \leq \overline{S}_N \rho_n \|\widehat{\varepsilon}\|_{\mathcal{E}}\|\widetilde{\varepsilon} - \widehat{\varepsilon}\|_N.$$

Hence, $\|\widetilde{\varepsilon} - \widehat{\varepsilon}\|_N \leq (nN)^{-1}\overline{S}_N \rho_n \|\widehat{\varepsilon}\|_{\mathcal{E}}$. Using (S2.7), we obtain

$$\|\widetilde{\varepsilon} - \widehat{\varepsilon}\|_\infty \leq S_N \|\widetilde{\varepsilon} - \widehat{\varepsilon}\|_N \leq (nN)^{-1} S_N \overline{S}_N \rho_n \|\widehat{\varepsilon}\|_{\mathcal{E}}.$$

Finally, we use Markov's inequality to get $\|\widetilde{\varepsilon}\|_{\mathcal{E}} \leq C_1|\triangle|^{-2}\|\widetilde{\varepsilon}\|$. It, therefore, follows that

$$\|\widehat{\varepsilon}\|_\infty \leq \|\widetilde{\varepsilon}\|_\infty + \|\widetilde{\varepsilon} - \widehat{\varepsilon}\|_\infty \leq \|\widetilde{\varepsilon}\|_\infty + \frac{\rho_n}{nN} S_N \overline{S}_N \frac{C_1}{|\triangle|^2}\|\widetilde{\varepsilon}\|_{L_2}.$$

According to Lemmas S5, S7 and S9, one has $\|\widehat{\varepsilon}\|_\infty = O_P\left\{(nN)^{-1/2}(\log n)^{1/2}|\triangle|^{-3/2}\right\}$ and $\|\widehat{\varepsilon}\|_{L_2}^2 \asymp |\triangle|^3\|\widehat{\boldsymbol{\theta}}_\varepsilon\|^2 = O_P(n^{-1}N^{-1}|\triangle|^{-3})$. Hence,

$$\|\widehat{\varepsilon}\|_\infty = O_P\left\{\frac{(\log n)^{1/2}}{\sqrt{nN}|\triangle|^{3/2}} + \frac{\rho_n}{n^{3/2}N^{3/2}|\triangle|^7}\right\}.$$

□

Proof of Theorem 1. Note that $\widehat{\mu} - \mu = \widehat{\mu}^o - \mu + \widehat{\eta} + \widehat{\varepsilon}$. Therefore, $\|\widehat{\mu} - \mu\|_{L_2} \le \|\widehat{\mu}^o - \mu\|_{L_2} + \|\widehat{\eta}\|_{L_2} + \|\widehat{\varepsilon}\|_{L_2}$. By Lemmas S1 and S7, one has

$$\|\widehat{\eta}\|_{L_2}^2 \asymp |\triangle|^3\|\widehat{\boldsymbol{\theta}}_\eta\|^2 = O_P(n^{-1}),\ \|\widehat{\varepsilon}\|_{L_2}^2 \asymp |\triangle|^3\|\widehat{\boldsymbol{\theta}}_\varepsilon\|^2 = O_P(n^{-1}N^{-1}|\triangle|^{-3}),$$

and the asymptotic order of $\|\widehat{\mu}^o - \mu\|_{L_2}$ is the same as $\|\widehat{\mu}^o - \mu\|_\infty$. By Lemmas S8 and S10,

$$\|\widehat{\mu}^o - \mu\|_\infty = O_P\left\{\frac{\rho_n}{nN|\triangle|^{7/2}}\|\mu\|_{2,\infty} + \left(1 + \frac{\rho_n}{nN|\triangle|^{11/2}}\right)|\triangle|^{\ell+1}\|\mu\|_{\ell+1,\infty}\right\},$$
$$\|\widehat{\eta}\|_\infty = O_P\left\{\frac{(\log n)^{1/2}}{\sqrt{n}} + \frac{\rho_n}{n^{3/2}N|\triangle|^{11/2}}\right\},\ \|\widehat{\varepsilon}\|_\infty = O_P\left\{\frac{(\log n)^{1/2}}{\sqrt{nN}|\triangle|^{3/2}} + \frac{\rho_n}{n^{3/2}N^{3/2}|\triangle|^7}\right\}.$$

Thus, by Assumption (A5), $\|\widehat{\mu}-\mu\|_\infty = O_P\{(n^{-1}\log(n))^{1/2}\}$ and $\|\widehat{\mu}-\mu\|_{L_2} = O_P(n^{-1/2})$.

□

S2.3 Simultaneous Confidence Bands

Proof of Theorem 2

Lemma S11 (Lemma A.5, Cao et al. (2012)). *Let* $\bar{\xi}_{\cdot k} = n^{-1}\sum_{i=1}^n \xi_{ik}$ *and* $\bar{\varepsilon}_{\cdot j} = n^{-1}\sum_{i=1}^n \varepsilon_{ij}$. *Under Assumption (A2), there exists some constant* $C_\beta > 0$ *such that* $\max_{1\le k\le\kappa} \mathrm{E}|\bar{\xi}_{\cdot k} - \bar{Z}_{\cdot k,\xi}| \le C_\beta n^{\beta-1}$, and $\max_{1\le j\le N}|\bar{\varepsilon}_{\cdot j} - \bar{Z}_{\cdot j,\varepsilon}| = O_{a.s.}(n^{\beta-1})$, for some $\beta \in (0, 1/2)$, *where* $\{Z_{ik,\xi}\}_{i=1,k=1}^{n,\kappa}$ *and* $\{Z_{ij,\varepsilon}\}_{i=1,j=1}^{n,N}$ *are iid* $N(0,1)$ *variables and* $\bar{Z}_{\cdot k,\xi} = \frac{1}{n}\sum_{i=1}^n Z_{ik,\xi}$, $\bar{Z}_{\cdot j,\varepsilon} = n^{-1}\sum_{i=1}^n Z_{ij,\varepsilon}$, $1 \le j \le N$, $1 \le k \le \kappa$.

Lemma S12. *Let* $\bar{\eta}(\boldsymbol{z}) = n^{-1}\sum_{i=1}^{n}\eta_i(\boldsymbol{z}) = \sum_{k=1}^{\kappa}\bar{\xi}_{\cdot k}\phi_k(\boldsymbol{z})$*, then under Assumptions (A2)–(A6), for* $\widehat{\eta}(\boldsymbol{z})$ *defined in (S2.11), one has* $n^{1/2}\|\widehat{\eta}-\bar{\eta}\|_\infty = o_P(1)$*. In addition, for any* $\boldsymbol{z} \in \Omega$*,*

$$\lim_{n\to\infty} P\left\{\sup_{\boldsymbol{z}\in\Omega} n^{1/2}G_\eta(\boldsymbol{z},\boldsymbol{z})^{-1/2}|\bar{\eta}(\boldsymbol{z})| \le q_{1-\alpha}\right\} = 1-\alpha,$$

$$\lim_{n\to\infty} P\left\{n^{1/2}G_\eta(\boldsymbol{z},\boldsymbol{z})^{-1/2}|\bar{\eta}(\boldsymbol{z})| \le z_{1-\alpha/2}\right\} = 1-\alpha.$$

Proof. Denote $\widetilde{\zeta}_k(\boldsymbol{z}) = \bar{Z}_{\cdot k,\xi}\phi_k(\boldsymbol{z}),\ k=1,\dots,\kappa$, and define

$$\widetilde{\zeta}(\boldsymbol{z}) = n^{1/2}\left\{\sum_{k=1}^{\kappa}\phi_k^2(\boldsymbol{z})\right\}^{-1/2}\sum_{k=1}^{\kappa}\widetilde{\zeta}_k(\boldsymbol{z}) = n^{1/2}G_\eta(\boldsymbol{z},\boldsymbol{z})^{-1/2}\sum_{k=1}^{\kappa}\widetilde{\zeta}_k(\boldsymbol{z}),$$

then $\{\widetilde{\zeta}(\boldsymbol{z}), \boldsymbol{z}\in\Omega\}$ is a Gaussian random field with mean 0, variance 1 and covariance

$$\mathrm{Cov}\left\{\widetilde{\zeta}(\boldsymbol{z}),\widetilde{\zeta}(\boldsymbol{z}')\right\} = G_\eta(\boldsymbol{z},\boldsymbol{z})^{-1/2}G_\eta(\boldsymbol{z},\boldsymbol{z}')G_\eta(\boldsymbol{z}',\boldsymbol{z}')^{-1/2}.$$

Therefore, $\widetilde{\zeta}(\boldsymbol{z})$ has the same distribution as $\zeta(\boldsymbol{z}), \boldsymbol{z}\in\Omega$.

Next, let $\widehat{\phi}_k(\boldsymbol{z}) = \widetilde{\mathbf{B}}(\boldsymbol{z})^\top\boldsymbol{\Gamma}_{N,\rho}^{-1}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\phi_k(\boldsymbol{z}_j)$. Similar as the the proof for Lemma S9, by Lemma S6 one has $\|\widehat{\phi}_k\|_\infty \le C|\triangle|^{-3}\left\|\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\phi_k(\boldsymbol{z}_j)\right\|_\infty \le C_1\|\phi_k\|_\infty$. According to Lemma S8,

$$\|\widehat{\phi}_k-\phi_k\|_\infty = O_P\left\{\frac{\rho_n}{nN|\triangle|^{7/2}}\|\phi_k\|_{2,\infty} + \left(1+\frac{\rho_n}{nN|\triangle|^{11/2}}\right)|\triangle|^{s+1}\|\phi_k\|_{s+1,\infty}\right\}.$$

Therefore, by Assumptions (A4)–(A6), one has

$$\begin{aligned}
&\mathrm{E}\Big\{n^{1/2}\sup_{\boldsymbol{z}\in\Omega}G_\eta(\boldsymbol{z},\boldsymbol{z})^{-1/2}|\widehat{\eta}(\boldsymbol{z})-\bar{\eta}(\boldsymbol{z})|\Big\}\\
=&\mathrm{E}\left[n^{1/2}\sup_{\boldsymbol{z}\in\Omega}G_\eta(\boldsymbol{z},\boldsymbol{z})^{-1/2}\left|\sum_{k=1}^{\kappa}\bar{\xi}_{\cdot k}\{\phi_k(\boldsymbol{z})-\widehat{\phi}_k(\boldsymbol{z})\}\right|\right]\\
\le&n^{1/2}G_\eta(\boldsymbol{z},\boldsymbol{z})^{-1/2}\left(\sum_{k=1}^{\kappa_n}\|\widehat{\phi}_k-\phi_k\|_\infty\mathrm{E}|\bar{\xi}_{\cdot k}| + C\sum_{k=\kappa_n+1}^{\kappa}\|\phi_k\|_\infty\mathrm{E}|\bar{\xi}_{\cdot k}|\right).
\end{aligned}$$

Thus,

$$
\begin{aligned}
&\mathrm{E}\Big\{n^{1/2}\sup_{\boldsymbol{z}\in\Omega} G_\eta(\boldsymbol{z},\boldsymbol{z})^{-1/2}|\widehat{\eta}(\boldsymbol{z})-\bar{\eta}(\boldsymbol{z})|\Big\}\\
\leq& n^{1/2}G_\eta(\boldsymbol{z},\boldsymbol{z})^{-1/2}C_1\left[\sum_{k=1}^{\kappa_n}\left\{\frac{\rho_n}{nN|\triangle|^{7/2}}\|\phi_k\|_{2,\infty}+\left(1+\frac{\rho_n}{nN|\triangle|^{11/2}}\right)|\triangle|^{s+1}\|\phi_k\|_{s+1,\infty}\right\}\right.\\
&\left.\quad\times\mathrm{E}|\bar{\xi}_{\cdot k}|+\sum_{k=\kappa_n+1}^{\kappa}\mathrm{E}|\bar{\xi}_{\cdot k}|\|\phi_k\|_\infty\right]\\
\leq& C_2n^{1/2}\left\{\frac{\rho_n}{nN|\triangle|^{7/2}}\sum_{k=1}^{\kappa_n}\|\phi_k\|_{2,\infty}+\left(1+\frac{\rho_n}{nN|\triangle|^{11/2}}\right)\sum_{k=1}^{\kappa_n}|\triangle|^{s+1}\|\phi_k\|_{s+1,\infty}+\sum_{k=\kappa_n+1}^{\kappa}\|\phi_k\|_\infty\right\}\\
=&o(1).
\end{aligned}
$$

Hence, $n^{1/2}\sup_{\boldsymbol{z}\in\Omega} G_\eta(\boldsymbol{z},\boldsymbol{z})^{-1/2}|\widehat{\eta}(\boldsymbol{z})-\bar{\eta}(\boldsymbol{z})|=o_P(1)$.

Under Assumption (A3), it follows that $\|\widehat{\eta}-\bar{\eta}\|_\infty=o_P(n^{-1/2})$.

By Lemma S11, for some $\beta\in(0,1/2)$,

$$
\begin{aligned}
&\mathrm{E}\Big\{\sup_{\boldsymbol{z}\in\Omega}\left|\widetilde{\zeta}(\boldsymbol{z})-n^{1/2}G_\eta(\boldsymbol{z},\boldsymbol{z})^{-1/2}\bar{\eta}(\boldsymbol{z})\right|\Big\}=\mathrm{E}\left\{n^{1/2}\sup_{\boldsymbol{z}\in\Omega}G_\eta(\boldsymbol{z},\boldsymbol{z})^{-1/2}\left|\sum_{k=1}^{\kappa}(\bar{Z}_{\cdot k,\xi}-\bar{\xi}_{\cdot k})\phi_k(\boldsymbol{z})\right|\right\}\\
&\quad\leq Cn^{1/2}\sum_{k=1}^{\kappa}\|\phi_k\|_\infty\mathrm{E}|\bar{Z}_{\cdot k,\xi}-\bar{\xi}_{\cdot k}|\leq\widetilde{C}n^{\beta-1/2}\sum_{k=1}^{\kappa}\|\phi_k\|_\infty.
\end{aligned}
$$

Thus, by Assumption (A6), $\sup_{\boldsymbol{z}\in\Omega}\left|\widetilde{\zeta}(\boldsymbol{z})-n^{1/2}G_\eta(\boldsymbol{z},\boldsymbol{z})^{-1/2}\bar{\eta}(\boldsymbol{z})\right|=o_P(1)$.
Note that $P\left\{\sup_{\boldsymbol{z}\in\Omega}\left|\widetilde{\zeta}(\boldsymbol{z})\right|\leq q_{1-\alpha}\right\}=1-\alpha$. The lemma is proved. □

Proof of Theorem 2. Note that “oracle” estimator $\bar{\mu}(\boldsymbol{z})=\mu(\boldsymbol{z})+\bar{\eta}(\boldsymbol{z})$ implies that $\widehat{\mu}-\bar{\mu}=\widehat{\mu}^o-\mu+\widehat{\eta}-\bar{\eta}+\widehat{\varepsilon}$. By Lemmas S8, S10, Assumptions (A5) and (A6),

$$
\begin{aligned}
&\|\widehat{\mu}^o-\mu\|_\infty=O_P\left\{\frac{\rho_n}{nN|\triangle|^{7/2}}\|\mu\|_{2,\infty}+\left(1+\frac{\rho_n}{nN|\triangle|^{11/2}}\right)|\triangle|^{\ell+1}\|\mu\|_{\ell+1,\infty}\right\}=o_P(n^{-1/2}),\\
&\|\widehat{\varepsilon}\|_\infty=O_P\left\{\frac{(\log n)^{1/2}}{\sqrt{nN}|\triangle|^{3/2}}+\frac{\rho_n}{n^{3/2}N^{3/2}|\triangle|^7}\right\}=o_P(n^{-1/2}).
\end{aligned}
$$

Thus, according to Lemma S12, the theorem is established. □

Proof of Theorem 4

According to (S2.12), for $H = 1, 2$, we can decompose the unpenalized spline estimator $\widehat{\mu}_H(\cdot)$ as

$$\widehat{\mu}_H(\boldsymbol{z}) = \widehat{\mu}_H^o(\boldsymbol{z}) + \widehat{\eta}_H(\boldsymbol{z}) + \widehat{\varepsilon}_H(\boldsymbol{z}).$$

Therefore, asymptotic error $(\widehat{\mu}_1 - \widehat{\mu}_2) - (\mu_1 - \mu_2)$ can be decomposed into three components:

$$(\widehat{\mu}_1^o - \widehat{\mu}_2^o - \mu_1 + \mu_2) + (\widehat{\eta}_1 - \widehat{\eta}_2) + (\widehat{\varepsilon}_1 - \widehat{\varepsilon}_2).$$

Similar as the proof of Theorem 2, the first and third components of the decomposition can be proved to have $\sqrt{n}$ asymptotic efficiency. Here we focus on the second component.

Lemma S13. *If Assumptions (A1)–(A6) are modified for each group accordingly, then one has*

$$\sup_{\boldsymbol{z}\in\Omega}\left|\widetilde{W}(\boldsymbol{z}) - n_1^{1/2} V\left(\boldsymbol{z},\boldsymbol{z}\right)^{-1/2}\left\{\widehat{\eta}_1(\boldsymbol{z}) - \widehat{\eta}_2(\boldsymbol{z})\right\}\right| = o_P(1).$$

Proof. According to Lemma S11, one can find i.i.d $Z_{Hik,\xi} \sim N(0,1)$, $i = 1, \ldots, n_H$ such that $\max_{1\le k\le\kappa} \mathrm{E}|\bar{\xi}_{H\cdot k} - \bar{Z}_{H\cdot k,\xi}| \le C_0 n^{\beta-1}$ and $\bar{Z}_{H\cdot k,\xi} = n_H^{-1}\sum_{i=1}^{n_H} Z_{Hik,\xi}$. Likewise, for the white noise sequence $\{\varepsilon_{Hij}, i \ge 1\}$, one can also find iid $Z_{Hik,\varepsilon} \sim N(0,1)$, $i = 1, \ldots, n_H$ such that $\max_{1\le j\le N} |\bar{\varepsilon}_{H\cdot j} - \bar{Z}_{H\cdot j,\varepsilon}| = O_{a.s.}(n^{\beta-1})$, where $\beta \in (0, 1/2)$. Let $V(\boldsymbol{z},\boldsymbol{z}') = G_{\eta,1}(\boldsymbol{z},\boldsymbol{z}') + \tau G_{\eta,2}(\boldsymbol{z},\boldsymbol{z}')$, $\widetilde{W}_k(\boldsymbol{z}) = \bar{Z}_{1\cdot k,\xi}\phi_{1k}(\boldsymbol{z}) - \bar{Z}_{2\cdot k,\xi}\phi_{2k}(\boldsymbol{z})$, $k = 1, \ldots, \kappa$, where $\tau = \lim_{n_1\to\infty} n_1/n_2$ and define

$$\widetilde{W}(\boldsymbol{z}) = n_1^{1/2}\left\{\sum_{k=1}^{\kappa}\phi_{1k}^2(\boldsymbol{z}) + \tau\phi_{2k}^2(\boldsymbol{z})\right\}^{-1/2}\sum_{k=1}^{\kappa}\widetilde{W}_k(\boldsymbol{z}) = n_1^{1/2}V\left(\boldsymbol{z},\boldsymbol{z}\right)^{-1/2}\sum_{k=1}^{\kappa}\widetilde{W}_k(\boldsymbol{z}).$$

Then, for any $\boldsymbol{z} \in \Omega$, $\widetilde{W}(\boldsymbol{z})$ is Gaussian with mean 0 and variance 1, and the covariance

$$\mathrm{E}\left\{\widetilde{W}(\boldsymbol{z})\widetilde{W}(\boldsymbol{z}')\right\} = V\left(\boldsymbol{z},\boldsymbol{z}\right)^{-1/2} V\left(\boldsymbol{z},\boldsymbol{z}'\right) V\left(\boldsymbol{z}',\boldsymbol{z}'\right)^{-1/2}.$$

That is, the distribution of $\widetilde{W}(\boldsymbol{z})$, $\boldsymbol{z} \in \Omega$ and the distribution of $W(\boldsymbol{z})$, $\boldsymbol{z} \in \Omega$ are identical. Similarly, for $H = 1, 2$, let $\widehat{\phi}_{Hk}(\boldsymbol{z}) = \widetilde{\mathbf{B}}(\boldsymbol{z})^{\top}\boldsymbol{\Gamma}_{N,\rho}^{-1}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}(\boldsymbol{z}_j)\phi_{Hk}(\boldsymbol{z}_j)$. Note that

$$\begin{aligned}
&\bar{\eta}_H(\boldsymbol{z}) = \frac{1}{n}\sum_{i=1}^{n}\eta_{Hi}(\boldsymbol{z}) = \sum_{k=1}^{\kappa}\bar{\xi}_{H\cdot k}\phi_{Hk}(\boldsymbol{z}),\ \widehat{\eta}_H(\boldsymbol{z}) = \sum_{k=1}^{\kappa}\bar{\xi}_{H\cdot k}\widehat{\phi}_{Hk}(\boldsymbol{z}),\\
&\widetilde{W}(\boldsymbol{z}) = n_1^{1/2}V(\boldsymbol{z},\boldsymbol{z})^{-1/2}\sum_{k=1}^{\kappa}\left\{\bar{Z}_{1\cdot k,\xi}\phi_{1k}(\boldsymbol{z}) - \bar{Z}_{2\cdot k,\xi}\phi_{2k}(\boldsymbol{z})\right\}.
\end{aligned}$$

Note that, by similar discussion in Lemma S12,

$$\begin{aligned}
&\mathrm{E}\left[\sup_{\boldsymbol{z}\in\Omega}\left|\widetilde{W}(\boldsymbol{z}) - n_1^{1/2}V(\boldsymbol{z},\boldsymbol{z})^{-1/2}\left\{\widehat{\eta}_1(\boldsymbol{z}) - \widehat{\eta}_2(\boldsymbol{z})\right\}\right|\right]\\
=&\mathrm{E}\left[n_1^{1/2}\sup_{\boldsymbol{z}\in\Omega}V(\boldsymbol{z},\boldsymbol{z})^{-1/2}\left|\sum_{k=1}^{\kappa}\widetilde{W}_k(\boldsymbol{z}) - \left\{\widehat{\eta}_1(\boldsymbol{z}) - \widehat{\eta}_2(\boldsymbol{z})\right\}\right|\right]\\
\leq& n_1^{1/2}\sup_{\boldsymbol{z}\in\Omega}V(\boldsymbol{z},\boldsymbol{z})^{-1/2}\sum_{k=1}^{\kappa}\Big\{\mathrm{E}\left|\bar{Z}_{1\cdot k,\xi} - \bar{\xi}_{1\cdot k}\right|\left|\phi_{1k}(\boldsymbol{z})\right| + \mathrm{E}\left|\bar{\xi}_{1\cdot k}\right|\left|\phi_{1k}(\boldsymbol{z}) - \widehat{\phi}_{1k}(\boldsymbol{z})\right|\\
&+ \mathrm{E}\left|\bar{Z}_{2\cdot k,\xi} - \bar{\xi}_{2\cdot k}\right|\left|\phi_{2k}(\boldsymbol{z})\right| + \mathrm{E}\left|\bar{\xi}_{2\cdot k}\right|\left|\phi_{2k}(\boldsymbol{z}) - \widehat{\phi}_{2k}(\boldsymbol{z})\right|\Big\} = o(1).
\end{aligned}$$

Therefore, one has $\sup_{\boldsymbol{z}\in\Omega}\left|\widetilde{W}(\boldsymbol{z}) - n_1^{1/2}V(\boldsymbol{z},\boldsymbol{z})^{-1/2}\left\{\widehat{\eta}_1(\boldsymbol{z}) - \widehat{\eta}_2(\boldsymbol{z})\right\}\right| = o_P(1)$. Observe that $P\left\{\sup_{\boldsymbol{z}\in\Omega}|\widetilde{W}(\boldsymbol{z})| \leq q_{12,\alpha}\right\} = 1 - \alpha$, for any $\alpha \in (0, 1)$,

$$\begin{aligned}
&\lim_{n_1\to\infty}P\left\{\sup_{\boldsymbol{z}\in\Omega}n_1^{1/2}V(\boldsymbol{z},\boldsymbol{z})^{-1/2}\left|\widehat{\eta}_1(\boldsymbol{z}) - \widehat{\eta}_2(\boldsymbol{z})\right| \leq q_{12,\alpha}\right\} = 1 - \alpha,\\
&\lim_{n_1\to\infty}P\left\{n_1^{1/2}V(\boldsymbol{z},\boldsymbol{z})^{-1/2}\left|\widehat{\eta}_1(\boldsymbol{z}) - \widehat{\eta}_2(\boldsymbol{z})\right| \leq z_{1-\alpha/2}\right\} = 1 - \alpha, \quad \text{for all } \boldsymbol{z} \in \Omega.
\end{aligned}$$

The conclusion of the lemma is proved. □

S2.4 Convergence of the Covariance Estimator

Without loss of generality, we prove Theorem 3 based on the unpenalized trivariate spline estimator. Using similar arguments in Section S2.2, we can easily extend this proof to the penalized case. Based on the estimated residuals $\widehat{R}_{ij} = Y_{ij} - \widehat{\mu}(\boldsymbol{z}_j)$, $i = 1, \ldots, n$, $j = 1, \ldots, N$, denote

$$\widehat{\boldsymbol{\beta}}_i = \arg\min_{\boldsymbol{\beta}} \sum_{j=1}^{N} \left\{ \widehat{R}_{ij} - \mathbf{B}^*(\boldsymbol{z}_j)^\top \mathbf{Q}_2^* \boldsymbol{\beta} \right\}^2,$$

where $\mathbf{B}^*(\boldsymbol{z})$ is the set of trivariate spline basis functions used to estimate $\eta_i(\boldsymbol{z})$, and the transpose of the smoothness matrix $\mathbf{H}^*$ admits the following QR decomposition: $\mathbf{H}^{*\top} = \mathbf{Q}^*\mathbf{R}^* = (\mathbf{Q}_1^* \ \mathbf{Q}_2^*)\binom{\mathbf{R}_1^*}{\mathbf{R}_2^*}$. Then, the trivariate spline estimator of $\eta_i(\boldsymbol{z})$ can be written as

$$\widehat{\eta}_i(\boldsymbol{z}) = \mathbf{B}^*(\boldsymbol{z})^\top \mathbf{Q}_2^* \widehat{\boldsymbol{\beta}}_i = \widetilde{\mathbf{B}}^*(\boldsymbol{z})^\top \widehat{\boldsymbol{\beta}}_i. \tag{S2.18}$$

Let $\boldsymbol{\Gamma}_N^* = \frac{1}{N}\sum_{j=1}^{N} \widetilde{\mathbf{B}}^*(\boldsymbol{z}_j)\widetilde{\mathbf{B}}^*(\boldsymbol{z}_j)^\top$, then one has

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}_i &= \boldsymbol{\Gamma}_N^{*-1} \frac{1}{N} \sum_{j=1}^{N} \widetilde{\mathbf{B}}^*(\boldsymbol{z}_j) \widehat{R}_{ij} = \boldsymbol{\Gamma}_N^{*-1} \frac{1}{N} \sum_{j=1}^{N} \widetilde{\mathbf{B}}^*(\boldsymbol{z}_j) \{Y_{ij} - \widehat{\mu}(\boldsymbol{z}_j)\} \\
&= \boldsymbol{\Gamma}_N^{*-1} \frac{1}{N} \sum_{j=1}^{N} \widetilde{\mathbf{B}}^*(\boldsymbol{z}_j) \left[\{\mu(\boldsymbol{z}_j) - \widehat{\mu}(\boldsymbol{z}_j)\} + \eta_i(\boldsymbol{z}_j) + \sigma(\boldsymbol{z}_j)\varepsilon_{ij} \right].
\end{aligned}$$

Next we define

$$\begin{aligned}
\widetilde{r}(\boldsymbol{z}) &= \widetilde{\mathbf{B}}^*(\boldsymbol{z})^\top \boldsymbol{\Gamma}_N^{*-1} \frac{1}{N} \sum_{j=1}^{N} \widetilde{\mathbf{B}}^*(\boldsymbol{z}_j) \{\mu(\boldsymbol{z}_j) - \widehat{\mu}(\boldsymbol{z}_j)\}, \\
\widetilde{\eta}_i(\boldsymbol{z}) &= \widetilde{\mathbf{B}}^*(\boldsymbol{z})^\top \boldsymbol{\Gamma}_N^{*-1} \frac{1}{N} \sum_{j=1}^{N} \widetilde{\mathbf{B}}^*(\boldsymbol{z}_j) \eta_i(\boldsymbol{z}_j), \\
\widetilde{\varepsilon}_i(\boldsymbol{z}) &= \widetilde{\mathbf{B}}^*(\boldsymbol{z})^\top \boldsymbol{\Gamma}_N^{*-1} \frac{1}{N} \sum_{j=1}^{N} \widetilde{\mathbf{B}}^*(\boldsymbol{z}_j) \sigma(\boldsymbol{z}_j)\varepsilon_{ij}.
\end{aligned}$$

Then, the estimation error $d_i(\boldsymbol{z}) = \widehat{\eta}_i(\boldsymbol{z}) - \eta_i(\boldsymbol{z})$ in (S2.18) can be decomposed as the following: $d_i(\boldsymbol{z}) = \widetilde{r}(\boldsymbol{z}) + \widetilde{\eta}_i(\boldsymbol{z}) - \eta_i(\boldsymbol{z}) + \widetilde{\varepsilon}_i(\boldsymbol{z})$.

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 S14. *Under Assumptions (A4), (A5) and (A7), then there exist constants* $0 < c_\eta < C_\eta < \infty$ *such that with probability approaching 1 as* $N \to \infty, n \to \infty$*,*

$$c_\eta |\triangle_\eta|^3 \leq \lambda_{min}(\mathbf{\Gamma}^*_{\mathbf{N}}) < \lambda_{max}(\mathbf{\Gamma}^*_{\mathbf{N}}) \leq C_\eta |\triangle_\eta|^3.$$

The lemma can be justified similarly to Lemma S6. Thus the proof is omitted.

Lemma S15. *Under Assumptions (A2), (A3) and (A5)–(A7), one has*

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2} |\widetilde{G}_\eta(\boldsymbol{z},\boldsymbol{z}') - G_\eta(\boldsymbol{z},\boldsymbol{z}')| = O_P\{n^{-1/2}(\log n)^{1/2}\}.$$

Proof. Note that

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

where $u_i(\boldsymbol{z},\boldsymbol{z}') = \eta_i(\boldsymbol{z})\eta_i(\boldsymbol{z}') - G_\eta(\boldsymbol{z},\boldsymbol{z}') = \sum_{k,k'=1}^{\infty} \phi_k(\boldsymbol{z})\phi_{k'}(\boldsymbol{z}')(\xi_{ik}\xi_{ik'} - \delta_{kk'})$ and $\delta_{kk'} = 1$ if $k = k'$ and 0 otherwise. Based on Assumption (A2), we have $\mathrm{E}[u_i(\boldsymbol{z},\boldsymbol{z}')] = 0$ and

$$\begin{aligned}
\mathrm{E}\left\{u_i^2(\boldsymbol{z},\boldsymbol{z}')\right\} &= \mathrm{E}\{\eta_i^2(\boldsymbol{z})\eta_i^2(\boldsymbol{z}')\} - G_\eta^2(\boldsymbol{z},\boldsymbol{z}') \\
&= \mathrm{E}\left[\sum_{k=1}^{\infty}\sum_{k'>k} \xi_{ik}^2\xi_{ik'}^2 \left\{4\phi_k(\boldsymbol{z})\phi_k(\boldsymbol{z}')\phi_{k'}(\boldsymbol{z})\phi_{k'}(\boldsymbol{z}') + \phi_k^2(\boldsymbol{z})\phi_{k'}^2(\boldsymbol{z}) + \phi_k^2(\boldsymbol{z}')\phi_{k'}^2(\boldsymbol{z})\right\}\right. \\
&\quad \left. + \sum_{k=1}^{\infty} \xi_{ik}^4 \phi_k^2(\boldsymbol{z})\phi_k^2(\boldsymbol{z}')\right] - G_\eta^2(\boldsymbol{z},\boldsymbol{z}') \\
&= G_\eta^2(\boldsymbol{z},\boldsymbol{z}') + G_\eta(\boldsymbol{z},\boldsymbol{z})G_\eta(\boldsymbol{z}',\boldsymbol{z}') + \sum_{k=1}^{\infty}(\mathrm{E}\xi_{ik}^4 - 3)\phi_k^2(\boldsymbol{z})\phi_k^2(\boldsymbol{z}')
\end{aligned}$$

The existence of $\mathrm{E}[u_i^2(\boldsymbol{z},\boldsymbol{z}')]$ is guaranteed by Mercer's theorem and Assumption (A3). Therefore, $\mathrm{E}\{\widetilde{G}_\eta(\boldsymbol{z},\boldsymbol{z}') - G_\eta(\boldsymbol{z},\boldsymbol{z}')\}^2 = O(n^{-1})$. Following a similar truncation method in Lemma S9 and applying Bernstein inequality, we have

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2} |\widetilde{G}_\eta(\boldsymbol{z},\boldsymbol{z}') - G_\eta(\boldsymbol{z},\boldsymbol{z}')| = O_P\{n^{-1/2}(\log n)^{1/2}\}.$$

□

Lemma S16. *Under Assumptions (A1)–(A7), denote* $\nabla\eta_i(\boldsymbol{z}) = \widetilde{\eta}_i(\boldsymbol{z}) - \eta_i(\boldsymbol{z})$, *then*

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2} |\widetilde{r}(\boldsymbol{z})\widetilde{r}(\boldsymbol{z}')| = O_P\{n^{-1}\log n\}, \tag{S2.19}$$

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2} \left|\frac{1}{n}\sum_{i=1}^{n}\widetilde{r}(\boldsymbol{z})\eta_i(\boldsymbol{z}')\right| = O_P\{n^{-1}(\log n)^{1/2}\}, \tag{S2.20}$$

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2} \left|\frac{1}{n}\sum_{i=1}^{n}\widetilde{r}(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}')\right| = o_P\{n^{-1}(\log n)^{1/2}\}, \tag{S2.21}$$

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2} \left|\frac{1}{n}\sum_{i=1}^{n}\widetilde{r}(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right| = O_P\{n^{-1}\log n N^{-1/2}|\triangle_\eta|^{-3/2}\}. \tag{S2.22}$$

Proof. Note that by definition, $\widetilde{r}(\boldsymbol{z}) = \widetilde{\mathbf{B}}^*(\boldsymbol{z})^\top\boldsymbol{\Gamma}_N^{*-1}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}^*(\boldsymbol{z}_j)\{\mu(\boldsymbol{z}_j) - \widehat{\mu}(\boldsymbol{z}_j)\}$. According to Lemma S14, there exists some constant C such that $\sup_{\boldsymbol{z}\in\Omega}|\widetilde{r}(\boldsymbol{z})| \leq C\|\mu - \widehat{\mu}\|_\infty$. Then by Theorem 1, we have Equation (S2.19) proved as follows.

$$\sup_{\boldsymbol{z},\boldsymbol{z}'\in\Omega^2} |\widetilde{r}(\boldsymbol{z})\widetilde{r}(\boldsymbol{z}')| \leq \left(\sup_{\boldsymbol{z}\in\Omega}|\widetilde{r}(\boldsymbol{z})|\right)^2 \leq C^2\|\mu - \widehat{\mu}\|_\infty^2 = O_P(n^{-1}\log n).$$

Similarly, denote $\overline{\xi}_{\cdot k} = n^{-1}\sum_{i=1}^{n}\xi_{ik}$,

$$\mathrm{E}\left\{\sup_{\boldsymbol{z}\in\Omega}\left|\frac{1}{n}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\right|\right\} = \mathrm{E}\left\{\sup_{\boldsymbol{z}\in\Omega}\left|\sum_{k=1}^{\infty}\overline{\xi}_{\cdot k}\phi_k(\boldsymbol{z})\right|\right\} \leq \sum_{k=1}^{\infty}\mathrm{E}|\overline{\xi}_{\cdot k}|\|\phi_k\|_\infty = O(n^{-1/2}),$$

which is based on Assumptions (A2) and (A6). Thus,

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\frac{1}{n}\sum_{i=1}^{n}\widetilde{r}(\boldsymbol{z})\eta_i(\boldsymbol{z}')\right| \leq \sup_{\boldsymbol{z}\in\Omega}|\widetilde{r}(\boldsymbol{z})|\sup_{\boldsymbol{z}\in\Omega}\left|\frac{1}{n}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\right| = O_P\{n^{-1}(\log n)^{1/2}\},$$

and Equation (S2.20) holds.

Next, we derive (S2.21). Denote $\widetilde{\phi}_k^*(\boldsymbol{z}) = \widetilde{\mathbf{B}}^*(\boldsymbol{z})^\top\boldsymbol{\Gamma}_N^{*-1}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}^*(\boldsymbol{z}_j)\phi_k(\boldsymbol{z}_j)$, then we have $\|\widetilde{\phi}_k^* - \phi_k\|_\infty \leq C|\triangle_\eta|^{s+1}\|\phi_k\|_{s+1,\infty}$ according to Lemma S2 and

$$\nabla\eta_i(\boldsymbol{z}) = \widetilde{\eta}_i(\boldsymbol{z}) - \eta_i(\boldsymbol{z}) = \sum_{k=1}^{\infty}\xi_{ik}\{\widetilde{\phi}_k^*(\boldsymbol{z}) - \phi_k(\boldsymbol{z})\}. \tag{S2.23}$$

Therefore, following the similar procedure in Lemma S12,

$$\begin{aligned}
&\mathrm{E}\left[\sup_{\boldsymbol{z}\in\Omega}\left|\frac{1}{n}\sum_{i=1}^{n}\nabla\eta_i(\boldsymbol{z})\right|\right] \le \sum_{k=1}^{\infty}\mathrm{E}|\overline{\xi}_{\cdot k}|\|\widetilde{\phi}_k^{*}-\phi_k\|_{\infty}\\
&\le C\Big\{\sum_{k=1}^{\kappa_n}\mathrm{E}|\overline{\xi}_{\cdot k}||\triangle_\eta|^{s+1}\|\phi_k\|_{s+1,\infty}+\sum_{k=\kappa_n+1}^{\infty}\mathrm{E}|\overline{\xi}_{\cdot k}|\|\phi_k\|_{\infty}\Big\}=o(n^{-1/2}).
\end{aligned}$$

Thus,

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\frac{1}{n}\sum_{i=1}^{n}\widetilde{r}(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}')\right| \le \sup_{\boldsymbol{z}\in\Omega}|\widetilde{r}(\boldsymbol{z})|\sup_{\boldsymbol{z}\in\Omega}\left|\frac{1}{n}\sum_{i=1}^{n}\nabla\eta_i(\boldsymbol{z})\right| = o_P\{n^{-1}(\log n)^{1/2}\}.$$

Finally, note that $n^{-1}\sum_{i=1}^{n}\widetilde{\varepsilon}_i(\boldsymbol{z})=\widetilde{\mathbf{B}}^{*}(\boldsymbol{z})^{\top}\mathbf{\Gamma}_N^{*-1}\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}\widetilde{\mathbf{B}}^{*}(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}$. Applying similar truncation technique and Bernstein's inequality as in Lemma S9, it can be shown that

$$\sup_{\boldsymbol{z}\in\Omega}\left|\frac{1}{n}\sum_{i=1}^{n}\widetilde{\varepsilon}_i(\boldsymbol{z})\right| = O_P\left\{(nN)^{-1/2}(\log n)^{1/2}|\triangle_\eta|^{-3/2}\right\}. \tag{S2.24}$$

Hence, $\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\frac{1}{n}\sum_{i=1}^{n}\widetilde{r}(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right| = O_P\{n^{-1}\log nN^{-1/2}|\triangle_\eta|^{-3/2}\}$. □

Lemma S17. *Under Assumptions (A1)–(A7), we have*

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\frac{1}{n}\sum_{i=1}^{n}\widetilde{\varepsilon}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right| = O_P\{N^{-1}|\triangle_\eta|^{-3}\}, \tag{S2.25}$$

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\frac{1}{n}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right| = O_P\{n^{-1/2}(\log n)^{1/2}N^{-1/2}|\triangle_\eta|^{-3/2}\}, \tag{S2.26}$$

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\frac{1}{n}\sum_{i=1}^{n}\nabla\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right| = O_P\{n^{-1/2}(\log n)^{1/2}N^{-1/2}|\triangle_\eta|^{-3/2}\}. \tag{S2.27}$$

Proof. We begin by showing (S2.25). Note that

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\widetilde{\varepsilon}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\\
&=\frac{1}{n}\sum_{i=1}^{n}\widetilde{\mathbf{B}}^{*}(\boldsymbol{z})^{\top}\mathbf{\Gamma}_N^{*-1}\left\{\frac{1}{N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}\widetilde{\mathbf{B}}^{*}(\boldsymbol{z}_j)\widetilde{\mathbf{B}}^{*}(\boldsymbol{z}_{j'})^{\top}\sigma(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_{j'})\varepsilon_{ij}\varepsilon_{ij'}\right\}\mathbf{\Gamma}_N^{*-1}\widetilde{\mathbf{B}}^{*}(\boldsymbol{z}'),
\end{aligned}$$

and

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

Hence, we have

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\Big|\mathrm{E}\Big\{\frac{1}{n}\sum_{i=1}^{n}\widetilde{\varepsilon}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\Big\}\Big| = O(N^{-1}|\triangle_\eta|^{-3}).$$

Meanwhile, let $w^*_{i,mm'} = n^{-1}N^{-2}\sum_{j,j'=1}^{N}\widetilde{B}^*_m(\boldsymbol{z}_j)\widetilde{B}^*_{m'}(\boldsymbol{z}_{j'})\sigma(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_{j'})\varepsilon_{ij}\varepsilon_{ij'}$, we are able to derive that

$$\begin{aligned}\mathrm{E}[w^*_{i,mm'}] =& \frac{1}{nN^2}\sum_{j=1}^{N}\widetilde{B}^*_m(\boldsymbol{z}_j)\widetilde{B}^*_{m'}(\boldsymbol{z}_j)\sigma^2(\boldsymbol{z}_j) = O(n^{-1}N^{-1}|\triangle_\eta|^3),\\ \mathrm{E}[w^{*2}_{i,mm'}] =& \frac{1}{n^{-}N^4}\sum_{j,j',j'',j'''=1}^{N}\widetilde{B}^*_m(\boldsymbol{z}_j)\widetilde{B}^*_m(\boldsymbol{z}_{j'})\widetilde{B}^*_{m'}(\boldsymbol{z}_{j''})\widetilde{B}^*_{m'}(\boldsymbol{z}_{j'''})\sigma(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_{j'})\sigma(\boldsymbol{z}_{j''})\sigma(\boldsymbol{z}_{j'''})\\ &\times\mathrm{E}[\varepsilon_j\varepsilon_{j'}\varepsilon_{j''}\varepsilon_{j'''}] = O(n^{-2}N^{-2}|\triangle_\eta|^6).\end{aligned}$$

Then, let $\tau^*_{i,mm'} = w^*_{i,mm'} - \mathrm{E}[w^*_{i,mm'}]$, we have $\mathrm{Var}(\sum_{i=1}^{n}\tau^*_{i,mm'}) = O(n^{-1}N^{-2}|\triangle_\eta|^6)$. Therefore, using Bernstein's inequality, we can obtain

$$\begin{aligned}&\max_{m\in\widetilde{\mathcal{M}}^*}\sum_{m'\in\widetilde{\mathcal{M}}^*}\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}\widetilde{B}^*_m(\boldsymbol{z}_j)\widetilde{B}^*_{m'}(\boldsymbol{z}_{j'})\sigma(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_{j'})\{\varepsilon_{ij}\varepsilon_{ij'} - \mathrm{E}\varepsilon_{ij}\varepsilon_{ij'}\}\right|\\ &\quad = O_P\{n^{-1}N^{-2}(\log n)^{1/2}\},\end{aligned}$$

and

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\frac{1}{n}\sum_{i=1}^{n}\widetilde{\varepsilon}_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}') - \mathrm{E}\Big\{\frac{1}{n}\sum_{i=1}^{n}\widetilde{\varepsilon}(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\Big\}\right| = O_P\{n^{-1}N^{-2}(\log n)^{1/2}|\triangle_\eta|^{-6}\}.$$

Thus, equation (S2.25) holds.

Next, note that

$$\frac{1}{n}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}') = \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{\infty}\xi_{ik}\phi_k(\boldsymbol{z})\widetilde{\mathbf{B}}^*(\boldsymbol{z})^\top\boldsymbol{\Gamma}_N^{*-1}\frac{1}{N}\sum_{j=1}^{N}\widetilde{\mathbf{B}}^*(\boldsymbol{z}_j)\sigma(\boldsymbol{z}_j)\varepsilon_{ij}.$$

which indicates that $\mathrm{E}\{\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\}=0$ and

$$\begin{aligned}\mathrm{E}\Big[\Big\{\frac{1}{n}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\Big\}^2\Big] &= \frac{1}{n}\mathrm{E}\{\eta_i^2(\boldsymbol{z})\}\mathrm{E}\{\widetilde{\varepsilon}_i^2(\boldsymbol{z}')\}\\ &= \frac{1}{n}G_\eta(\boldsymbol{z},\boldsymbol{z})\widetilde{\mathbf{B}}^*(\boldsymbol{z}')^\top\boldsymbol{\Gamma}_N^{*-1}\left\{\frac{1}{N^2}\sum_{j=1}^{N}\widetilde{\mathbf{B}}^*(\boldsymbol{z}_j)\widetilde{\mathbf{B}}^*(\boldsymbol{z}_j)^\top\sigma^2(\boldsymbol{z}_j)\right\}\boldsymbol{\Gamma}_N^{*-1}\widetilde{\mathbf{B}}^*(\boldsymbol{z}')\\ &= O(n^{-1}N^{-1}|\triangle_\eta|^{-3}).\end{aligned}$$

Then, applying Bernstein inequality, equation (S2.26) can be obtained.

Finally, we derive (S2.27). Denote $\nabla\phi_k(\boldsymbol{z})=\widetilde{\phi}_k^*(\boldsymbol{z})-\phi_k(\boldsymbol{z})$, according to equation (S2.23) we have $\nabla\eta_i(\boldsymbol{z})=\sum_{k=1}^{\infty}\xi_{ik}\nabla\phi_k(\boldsymbol{z})$ and $\|\nabla\phi_k\|_\infty\le|\triangle_\eta|^{s+1}\|\phi_k\|_{s+1,\infty}$. Together with (S2.24), Assumptions (A2) and (A6), we can obtain that

$$\begin{aligned}\mathrm{E}\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\frac{1}{n}\sum_{i=1}^{n}\nabla\eta_i(\boldsymbol{z})\widetilde{\varepsilon}_i(\boldsymbol{z}')\right| &\le\sum_{k=1}^{\kappa}\mathrm{E}|\xi_{1k}|\|\nabla\phi_k\|_\infty\sup_{\boldsymbol{z}\in\Omega}\left|\frac{1}{n}\sum_{i=1}^{n}\widetilde{\varepsilon}_i(\boldsymbol{z})\right|\\ &\le C(nN)^{-1/2}(\log n)^{1/2}|\triangle_\eta|^{-3/2}\\ &\quad\times\left(|\triangle_\eta|^{s+1}\sum_{k=1}^{\kappa_n}\|\phi_k\|_{s+1,\infty}+\sum_{k=\kappa_n+1}^{\kappa}\|\phi_k\|_\infty\right).\end{aligned}$$

Therefore, equation (S2.27) is proved. □

Lemma S18. *Under Assumptions (A1)–(A7), we have*

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\frac{1}{n}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}')\right|=O_P\left\{|\triangle_\eta|^{s+1}\sum_{k=1}^{\kappa_n}\|\phi_k\|_\infty\|\phi_k\|_{s+1,\infty}+\sum_{k=\kappa_n+1}^{\kappa}\|\phi_k\|_\infty^2\right\}, \tag{S2.28}$$

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\frac{1}{n}\sum_{i=1}^{n}\nabla\eta_i(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}')\right|=O_P\left\{|\triangle_\eta|^{2(s+1)}\sum_{k=1}^{\kappa_n}\|\phi_k\|_{s+1,\infty}^2+\sum_{k=\kappa_n+1}^{\kappa}\|\phi_k\|_\infty^2\right\}. \tag{S2.29}$$

Proof. First, we show (S2.28). Note that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\triangledown\eta_i(\boldsymbol{z}') &= \frac{1}{n}\sum_{i=1}^{n}\sum_{k,k'=1}^{\kappa}\xi_{ik}\xi_{ik'}\phi_k(\boldsymbol{z})\triangledown\phi_{k'}(\boldsymbol{z}')\\ &= \sum_{k,k'=1}^{\kappa}\left(\frac{1}{n}\sum_{i=1}^{n}\xi_{ik}\xi_{ik'}\right)\phi_k(\boldsymbol{z})\triangledown\phi_{k'}(\boldsymbol{z}').\end{aligned}$$

Let $\bar{\xi}_{\cdot kk'} = n^{-1}\sum_{i=1}^{n}\xi_{ik}\xi_{ik'}$, then $\mathrm{E}(\bar{\xi}_{\cdot kk'}) = \mathbf{I}$, $\mathrm{E}\{(\xi_{ik}\xi_{ik'})^2\} \leq \{\mathrm{E}(\xi_{ik}^4)\mathrm{E}(\xi_{ik'}^4)\}^{1/2} \leq C$ and $\mathrm{E}(\bar{\xi}_{\cdot kk'}^2) \leq C$. Therefore,

$$\begin{aligned}\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\mathrm{E}\left\{\frac{1}{n}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\triangledown\eta_i(\boldsymbol{z}')\right\}\right| &= \sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\sum_{k=1}^{\kappa}\phi_k(\boldsymbol{z})\triangledown\phi_k(\boldsymbol{z}')\right|\\ &\leq C\left\{|\triangle_\eta|^{s+1}\sum_{k=1}^{\kappa_n}\|\phi_k\|_\infty\|\phi_k\|_{s+1,\infty} + \sum_{k=\kappa_n+1}^{\kappa}\|\phi_k\|_\infty^2\right\}.\end{aligned}$$

Next, let $v_i^*(\boldsymbol{z},\boldsymbol{z}') = \sum_{k,k'=1}^{\kappa}\phi_k(\boldsymbol{z})\triangledown\phi_{k'}(\boldsymbol{z}')\{\xi_{ik}\xi_{ik'} - \mathrm{E}(\xi_{ik}\xi_{ik'})\}$, then

$$\frac{1}{n}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\triangledown\eta_i(\boldsymbol{z}') - \mathrm{E}\left\{\frac{1}{n}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\triangledown\eta_i(\boldsymbol{z}')\right\} = \frac{1}{n}\sum_{i=1}^{n}v_i^*(\boldsymbol{z},\boldsymbol{z}').$$

Thus, $\mathrm{E}[v_i^*(\boldsymbol{z},\boldsymbol{z}')] = 0$ and

$$\mathrm{E}[v_i^{*2}(\boldsymbol{z},\boldsymbol{z}')] \leq C\left(|\triangle_\eta|^{s+1}\sum_{k=1}^{\kappa_n}\|\phi_k\|_\infty\|\phi_k\|_{s+1,\infty} + \sum_{k=\kappa_n+1}^{\kappa}\|\phi_k\|_\infty^2\right)^2.$$

By applying the truncation method and Bernstein inequality,

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

Therefore,

$$\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\frac{1}{n}\sum_{i=1}^{n}\eta_i(\boldsymbol{z})\triangledown\eta_i(\boldsymbol{z}')\right| = O_P\left\{|\triangle_\eta|^{s+1}\sum_{k=1}^{\kappa_n}\|\phi_k\|_\infty\|\phi_k\|_{s+1,\infty} + \sum_{k=\kappa_n+1}^{\kappa}\|\phi_k\|_\infty^2\right\}.$$

Next, equation (S2.29) can be proved following a similar procedure. Note that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\nabla\eta_i(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}') &= \sum_{k,k'=1}^{\kappa}\left(\frac{1}{n}\sum_{i=1}^{n}\xi_{ik}\xi_{ik'}\right)\nabla\phi_k(\boldsymbol{z})\nabla\phi_{k'}(\boldsymbol{z}')\\ &= \sum_{k,k'=1}^{\kappa}\bar{\xi}_{\cdot kk'}\nabla\phi_k(\boldsymbol{z})\nabla\phi_{k'}(\boldsymbol{z}'),\end{aligned}$$

which indicates that

$$\begin{aligned}\sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\mathrm{E}\left\{\frac{1}{n}\sum_{i=1}^{n}\nabla\eta_i(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}')\right\}\right| &= \sup_{(\boldsymbol{z},\boldsymbol{z}')\in\Omega^2}\left|\sum_{k=1}^{\kappa}\nabla\phi_k(\boldsymbol{z})\nabla\phi_k(\boldsymbol{z}')\right|\\ &\le C\left\{|\triangle_\eta|^{2(s+1)}\sum_{k=1}^{\kappa_n}\|\phi_k\|_{s+1,\infty}^2+\sum_{k=\kappa_n+1}^{\kappa}\|\phi_k\|_{\infty}^2\right\}.\end{aligned}$$

Meanwhile,

$$\mathrm{Var}\left\{\frac{1}{n}\sum_{i=1}^{n}\nabla\eta_i(\boldsymbol{z})\nabla\eta_i(\boldsymbol{z}')\right\}\asymp\frac{1}{n}\left(|\triangle_\eta|^{2(s+1)}\sum_{k=1}^{\kappa_n}\|\phi_k\|_{s+1,\infty}^2+\sum_{k=\kappa_n+1}^{\kappa}\|\phi_k\|_{\infty}^2\right)^2.$$

Based on similar derivation using Bernstein inequality,

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

Hence, (S2.29) holds. □

Proof of Theorem 3 (i). Note that

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

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 S15, and

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

Based on Lemmas S16–S18, one can show that

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

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

The desired result is established. □

Proof of Theorem 3 (ii). Denote $\Delta\psi_k(\boldsymbol{z}) = \int(\widehat{G}-G)(\boldsymbol{z},\boldsymbol{z}')\psi_k(\boldsymbol{z}')d\boldsymbol{z}'$. By Theorem 3 (i), $\|\widehat{G}-G\|_\infty = o_P(1)$. Thus, for any $k \geq 1$, $\|\Delta\psi_k\|_\infty = o_P(1)$. According to Hall and Hosseini-Nasab (2006), let $\|\mathit{\Delta}\|_2 = [\iint(\widehat{G}(\boldsymbol{z},\boldsymbol{z}') - G(\boldsymbol{z},\boldsymbol{z}'))^2 d\boldsymbol{z}d\boldsymbol{z}']^{1/2}$, then $\widehat{\psi}_k - \psi_k = \sum_{j\neq k}(\lambda_k-\lambda_j)^{-1}\langle\Delta\psi_k,\psi_j\rangle\psi_j + O(\|\mathit{\Delta}\|_2^2)$. It follows from Bessel's inequality that $\|\widehat{\psi}_k - \psi_k\|_2 \leq C(\|\Delta\psi_k\|_\infty^2 + O(\|\mathit{\Delta}\|_2^2)) = o_P(1)$. By (2.9) in Hall and Hosseini-Nasab (2006),

$$\widehat{\lambda}_k - \lambda_k = \iint(\widehat{G}-G)(\boldsymbol{z},\boldsymbol{z}')\psi_k(\boldsymbol{z})\psi_k(\boldsymbol{z}')d\boldsymbol{z}d\boldsymbol{z}' + O(\|\Delta\psi_k\|_2^2).$$

Thus, using Theorem 3 (i), one has $|\widehat{\lambda}_k - \lambda_k| = o_P(1), \forall k \geq 1$.

Next, note that

$$\begin{aligned}\widehat{\lambda}_k\widehat{\psi}_k(\boldsymbol{z}) - \lambda_k\psi_k(\boldsymbol{z}) &= \int\widehat{G}(\boldsymbol{z},\boldsymbol{z}')\widehat{\psi}_k(\boldsymbol{z}')d\boldsymbol{z}' - \int G(\boldsymbol{z},\boldsymbol{z}')\psi_k(\boldsymbol{z}')d\boldsymbol{z}'\\ &= \int(\widehat{G}-G)(\boldsymbol{z},\boldsymbol{z}')(\widehat{\psi}(\boldsymbol{z}') - \psi_k(\boldsymbol{z}'))d\boldsymbol{z}' + \int(\widehat{G}-G)(\boldsymbol{z},\boldsymbol{z}')\psi_k(\boldsymbol{z}')d\boldsymbol{z}'\\ &\quad + \int G(\boldsymbol{z},\boldsymbol{z}')\{\widehat{\psi}_k(\boldsymbol{z}') - \psi_k(\boldsymbol{z}')\}d\boldsymbol{z}'.\end{aligned}$$

By Cauchy-Schwarz inequality and Theorem 3 (i), for all $\boldsymbol{z} \in \Omega$,

$$\int G(\boldsymbol{z},\boldsymbol{z}')\{\widehat{\psi}_k(\boldsymbol{z}') - \psi_k(\boldsymbol{z}')\}d\boldsymbol{z}' \leq \left(\int G^2(\boldsymbol{z},\boldsymbol{z}')d\boldsymbol{z}'\right)^{1/2}\|\widehat{\psi}_k - \psi_k\|_2 = o_P(1),$$

$$\int(\widehat{G}-G)(\boldsymbol{z},\boldsymbol{z}')(\widehat{\psi}(\boldsymbol{z}') - \psi_k(\boldsymbol{z}'))d\boldsymbol{z}' \leq \|\widehat{G}-G\|_\infty\|\widehat{\psi}_k - \psi_k\|_2 = o_P(1),$$

$$\int(\widehat{G}-G)(\boldsymbol{z},\boldsymbol{z}')\psi_k(\boldsymbol{z}')d\boldsymbol{z}' \leq \|\widehat{G}-G\|_\infty\|\psi_k\|_2 = o_P(1).$$

Therefore, $\|\widehat{\lambda}_k\widehat{\psi}_k - \lambda_k\psi_k\|_\infty = o_P(1)$, and

$$\lambda_k\|\widehat{\psi}_k - \psi_k\|_\infty \leq \|\widehat{\lambda}_k\widehat{\psi}_k - \lambda_k\psi_k\|_\infty + |\widehat{\lambda}_k - \lambda_k|\|\widehat{\psi}_k\|_\infty = o_P(1).$$

It follows that $\|\widehat{\psi}_k - \psi_k\|_\infty = o_P(1)$. □

Bibliography

Cao, G., Yang, L., and Todem, D. (2012), "Simultaneous inference for the mean function based on dense functional data," *Journal of nonparametric statistics*, 24, 359–377.

Hall, P. and Hosseini-Nasab, M. (2006), "On properties of functional principal components analysis," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68, 109–126.

Lai, M. J. (1989), "On construction of bivariate and trivariate vertex splines on arbitrary mixed grid partitions," Phd dissertation, Texas A&M University.

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

Li, X., Yu, S., Wang, Y., Wang, G., and Wang, L. (2022), "Nonparametric regression for 3D point cloud Learning," *arXiv preprint arXiv:2106.04255.*

Petersen, A., Müller, H.-G., et al. (2019), "Fréchet regression for random objects with Euclidean predictors," *The Annals of Statistics*, 47, 691–719.

Zhu, H., Fan, J., and Kong, L. (2014), "Spatially varying coefficient model for neuroimaging data with jump discontinuities," *Journal of the American Statistical Association*, 109, 1084–1098.