Statistica Sinica

SUPPLEMENT TO "STATISTICAL INFERENCE FOR FUNCTIONAL DATA OVER MULTI-DIMENSIONAL DOMAIN"

Qirui Hu[1], Lijian Yang[1]

Tsinghua University

Abstract: This supplement provides tables and figures of simulation results and detailed proofs of the theoretical results with necessarily technical lemmas.

S1. Tables and Figures

Table S.1: Simulation results based on homogeneous errors with $\sigma(x_1, x_2) \equiv 0.1$.

Distribution of ξ						Normal						
Number of grids		$N=50$			$N=100$			$N=200$			$N=400$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
0.8	0.728	0.723	0.72	0.747	0.754	0.748	0.746	0.753	0.759	0.787	0.784	0.789
0.9	0.861	0.856	0.857	0.879	0.879	0.867	0.876	0.878	0.877	0.881	0.888	0.889
0.95	0.925	0.921	0.916	0.933	0.933	0.919	0.941	0.942	0.933	0.948	0.943	0.949
0.975	0.955	0.962	0.963	0.965	0.962	0.959	0.971	0.97	0.972	0.973	0.975	0.971
0.99	0.98	0.979	0.98	0.988	0.985	0.985	0.989	0.985	0.985	0.991	0.992	0.992
Distribution of ξ						Uniform						
Number of grids		$N=50$			$N=100$			$N=200$			$N=400$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
0.8	0.702	0.707	0.709	0.74	0.745	0.73	0.768	0.766	0.758	0.794	0.784	0.796
0.9	0.824	0.843	0.83	0.861	0.857	0.847	0.873	0.877	0.87	0.906	0.906	0.9
0.95	0.911	0.913	0.912	0.924	0.927	0.921	0.932	0.931	0.931	0.942	0.938	0.946
0.975	0.948	0.947	0.943	0.958	0.958	0.96	0.964	0.966	0.966	0.968	0.968	0.974
0.99	0.978	0.979	0.976	0.981	0.979	0.978	0.979	0.983	0.979	0.99	0.994	0.994
Distribution of ξ						Laplace						
Number of grids		$N=50$			$N=100$			$N=200$			$N=400$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
0.8	0.746	0.753	0.746	0.741	0.738	0.748	0.754	0.757	0.752	0.776	0.790	0.774
0.9	0.871	0.879	0.874	0.863	0.872	0.863	0.865	0.869	0.867	0.878	0.880	0.870
0.95	0.941	0.94	0.939	0.937	0.934	0.934	0.925	0.928	0.933	0.932	0.940	0.936
0.975	0.96	0.965	0.965	0.974	0.974	0.976	0.961	0.96	0.961	0.972	0.980	0.972
0.99	0.986	0.986	0.989	0.986	0.987	0.989	0.986	0.985	0.986	0.988	0.990	0.990

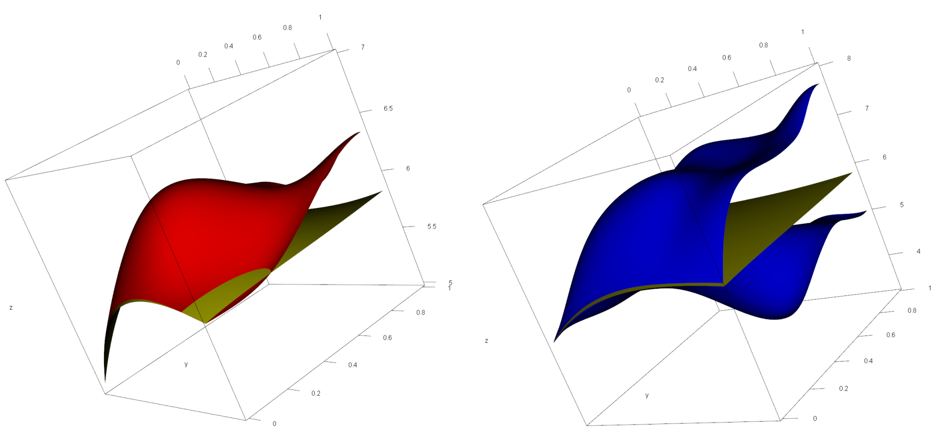

Figure S.1: Plot of true mean function (yellow), tensor product $B-$spline estimator (red) and the 95% simultaneous confidence region (blue) when $N = 100$.

Table S.2: Simulation results based on homogeneous errors with $\sigma(x_1, x_2) \equiv 0.2$.

Distribution of ξ						Normal						
Number of grids		$N = 50$			$N = 100$			$N = 200$			$N = 400$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
0.8	0.699	0.697	0.707	0.735	0.733	0.743	0.735	0.737	0.742	0.786	0.782	0.775
0.9	0.854	0.85	0.852	0.867	0.867	0.869	0.866	0.871	0.871	0.881	0.885	0.883
0.95	0.92	0.917	0.917	0.922	0.922	0.923	0.929	0.939	0.929	0.947	0.944	0.942
0.975	0.959	0.959	0.957	0.957	0.959	0.963	0.966	0.968	0.968	0.965	0.969	0.971
0.99	0.982	0.976	0.976	0.983	0.982	0.984	0.985	0.985	0.985	0.984	0.984	0.989
Distribution of ξ						Uniform						
Number of grids		$N = 50$			$N = 100$			$N = 200$			$N = 400$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
0.8	0.693	0.697	0.682	0.729	0.721	0.72	0.742	0.748	0.733	0.782	0.762	0.774
0.9	0.822	0.824	0.815	0.853	0.846	0.85	0.859	0.86	0.848	0.894	0.888	0.882
0.95	0.908	0.908	0.903	0.925	0.917	0.919	0.923	0.924	0.923	0.940	0.932	0.936
0.975	0.951	0.94	0.946	0.958	0.961	0.956	0.964	0.962	0.958	0.972	0.970	0.970
0.99	0.973	0.976	0.973	0.984	0.977	0.978	0.981	0.974	0.977	0.986	0.986	0.994
Distribution of ξ						Laplace						
Number of grids		$N = 50$			$N = 100$			$N = 200$			$N = 400$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
0.8	0.725	0.728	0.733	0.732	0.73	0.732	0.716	0.726	0.729	0.758	0.758	0.768
0.9	0.867	0.861	0.864	0.859	0.86	0.862	0.853	0.844	0.858	0.870	0.858	0.864
0.95	0.933	0.936	0.93	0.927	0.93	0.931	0.921	0.91	0.92	0.924	0.932	0.924
0.975	0.96	0.961	0.959	0.966	0.97	0.968	0.951	0.957	0.958	0.972	0.978	0.968
0.99	0.986	0.983	0.987	0.985	0.986	0.989	0.981	0.977	0.98	0.986	0.992	0.990

Table S.3: Simulation results based on heteroscedastic errors with $\sigma(x_1, x_2) = 0.15\,(5 - \exp(s+t))\,/\,(5 + \exp(s+t))$.

Distribution of ξ						Normal						
Number of grids		$N = 50$			$N = 100$			$N = 200$			$N = 400$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
0.8	0.725	0.721	0.727	0.753	0.756	0.747	0.749	0.755	0.759	0.786	0.794	0.772
0.9	0.864	0.856	0.858	0.877	0.881	0.869	0.878	0.878	0.876	0.886	0.892	0.890
0.95	0.923	0.924	0.921	0.934	0.932	0.921	0.941	0.942	0.931	0.944	0.940	0.950
0.975	0.959	0.962	0.962	0.964	0.964	0.964	0.97	0.97	0.973	0.968	0.974	0.966
0.99	0.979	0.98	0.981	0.988	0.984	0.988	0.987	0.987	0.984	0.986	0.994	0.994
Distribution of ξ						Uniform						
Number of grids		$N = 50$			$N = 100$			$N = 200$			$N = 400$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
0.8	0.717	0.715	0.713	0.738	0.746	0.729	0.769	0.769	0.76	0.796	0.790	0.790
0.9	0.831	0.834	0.832	0.86	0.858	0.845	0.876	0.879	0.873	0.906	0.900	0.900
0.95	0.901	0.908	0.908	0.93	0.931	0.919	0.934	0.936	0.936	0.940	0.942	0.942
0.975	0.94	0.936	0.945	0.961	0.96	0.964	0.964	0.965	0.966	0.972	0.968	0.978
0.99	0.98	0.972	0.976	0.983	0.979	0.979	0.984	0.982	0.979	0.986	0.994	0.994
Distribution of ξ						Laplace						
Number of grids		$N = 50$			$N = 100$			$N = 200$			$N = 400$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
0.8	0.725	0.721	0.727	0.742	0.745	0.755	0.757	0.744	0.746	0.774	0.790	0.776
0.9	0.864	0.856	0.858	0.872	0.877	0.87	0.891	0.885	0.878	0.878	0.878	0.874
0.95	0.923	0.924	0.921	0.941	0.935	0.934	0.952	0.955	0.952	0.936	0.940	0.934
0.975	0.959	0.962	0.962	0.968	0.97	0.973	0.977	0.976	0.975	0.972	0.980	0.976
0.99	0.979	0.980	0.981	0.985	0.986	0.989	0.986	0.989	0.986	0.986	0.990	0.990

Table S.4: Simulation results based on heteroscedastic errors with $\sigma(x_1, x_2) = 0.3\,(5 - \exp(x_1 + x_2))\,/\,(5 + \exp(x_1 + x_2))$.

Distribution of ξ	Normal											
Number of grids		$N = 50$			$N = 100$			$N = 200$			$N = 400$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
0.8	0.703	0.7	0.708	0.737	0.734	0.746	0.742	0.74	0.745	0.784	0.788	0.788
0.9	0.854	0.85	0.851	0.868	0.869	0.869	0.866	0.872	0.869	0.884	0.892	0.896
0.95	0.925	0.92	0.92	0.923	0.923	0.922	0.932	0.938	0.93	0.944	0.948	0.942
0.975	0.956	0.959	0.959	0.959	0.959	0.964	0.967	0.968	0.966	0.964	0.966	0.972
0.99	0.981	0.979	0.978	0.985	0.98	0.985	0.986	0.985	0.984	0.982	0.982	0.988
Distribution of ξ	Uniform											
Number of grids		$N = 50$			$N = 100$			$N = 200$			$N = 400$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
0.8	0.706	0.707	0.69	0.735	0.725	0.725	0.739	0.751	0.736	0.784	0.764	0.780
0.9	0.83	0.825	0.819	0.852	0.843	0.848	0.864	0.861	0.854	0.894	0.892	0.886
0.95	0.899	0.9	0.891	0.926	0.919	0.92	0.923	0.924	0.923	0.936	0.936	0.938
0.975	0.945	0.937	0.943	0.959	0.962	0.958	0.963	0.958	0.961	0.974	0.970	0.974
0.99	0.977	0.972	0.973	0.985	0.979	0.978	0.983	0.974	0.977	0.986	0.988	0.994
Distribution of ξ	Laplace											
Number of grids		$N = 50$			$N = 100$			$N = 200$			$N = 400$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
0.8	0.713	0.73	0.722	0.737	0.736	0.733	0.728	0.736	0.733	0.752	0.762	0.764
0.9	0.865	0.861	0.861	0.864	0.875	0.865	0.874	0.874	0.876	0.876	0.864	0.864
0.95	0.938	0.939	0.931	0.926	0.931	0.928	0.944	0.932	0.94	0.926	0.936	0.930
0.975	0.963	0.966	0.964	0.968	0.966	0.963	0.971	0.972	0.97	0.968	0.978	0.970
0.99	0.989	0.988	0.988	0.983	0.988	0.987	0.985	0.987	0.985	0.988	0.990	0.992

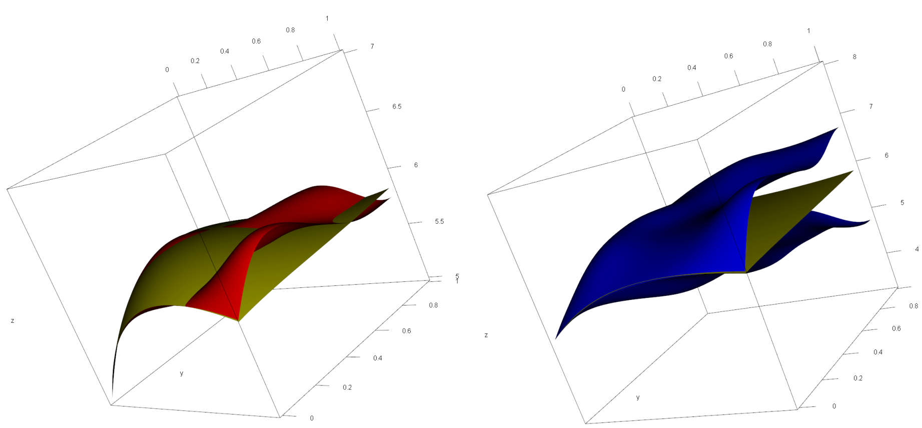

Figure S.2: Plot of true mean function (yellow), tensor product $B-$spline estimator (red) and the 95% simultaneous confidence region (blue) when $N = 200$.

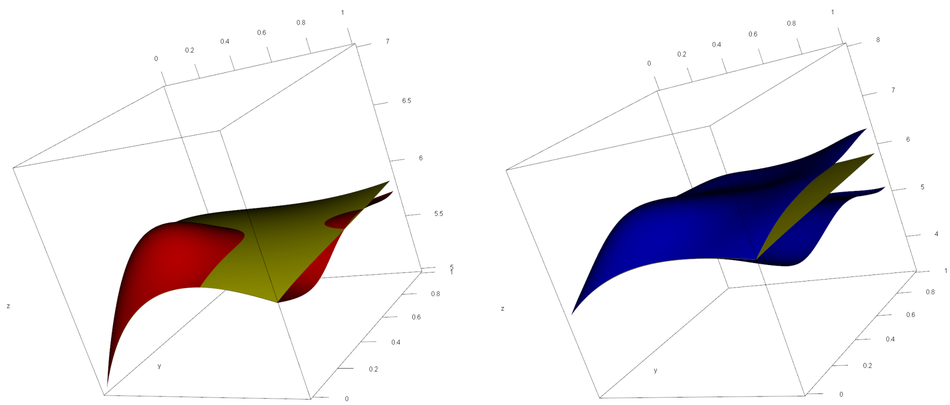

Figure S.3: Plot of true mean function (yellow), tensor product $B-$spline estimator (red) and the 95% simultaneous confidence region (blue) when $N = 400$.

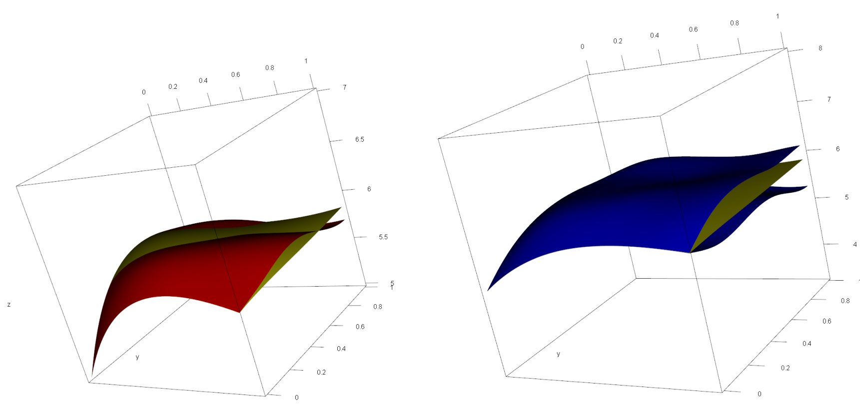

Figure S.4: Plot of true mean function (yellow), tensor product $B-$spline estimator (red) and the 95% simultaneous confidence region (blue) when $N = 800$.

S2. Additional simulation results

S2.1 2D case

Following the reviewer's suggestion, we report the MSE, variance, and bias based on the setting in Section 5. Since these results are similar, we only consider the case

$\sigma(x_1, x_2) \equiv 0.2$.

Table S.5: Additional 2D simulation results based on homogeneous errors with $\sigma(x_1, x_2) \equiv 0.2$.

Distribution of ξ	Normal											
Number of grids		$N=50$			$N=100$			$N=200$			$N=400$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
MSE	0.1693	0.1693	0.1693	0.0825	0.0825	0.0825	0.0444	0.0444	0.0444	0.0249	0.0249	0.0249
Bias	0.0059	0.0059	0.0059	0.0003	0.0003	0.0003	0.0007	0.0007	0.0007	0.0015	0.0015	0.0015
Variance	0.1617	0.1617	0.1617	0.0829	0.0829	0.0829	0.0441	0.0441	0.0441	0.0244	0.0244	0.0244
Distribution of ξ	Uniform											
Number of grids		$N=50$			$N=100$			$N=200$			$N=400$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
MSE	0.1651	0.1651	0.1651	0.0819	0.0819	0.0819	0.0440	0.0440	0.0440	0.0247	0.0247	0.0247
Bias	0.0018	0.0018	0.0018	0.0037	0.0037	0.0037	0.0062	0.0062	0.0062	0.0016	0.0016	0.0016
Variance	0.1610	0.1610	0.1610	0.0826	0.0826	0.0826	0.0442	0.0442	0.0442	0.0243	0.0243	0.0243
Distribution of ξ	Laplace											
Number of grids		$N=50$			$N=100$			$N=200$			$N=400$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
MSE	0.1636	0.1636	0.1637	0.0858	0.0858	0.0858	0.0439	0.0439	0.0439	0.0251	0.0251	0.0251
Bias	0.0003	0.0003	0.0003	-0.0076	-0.0076	-0.0076	-0.0049	-0.0049	-0.0049	0.0004	0.0004	0.0004
Variance	0.1618	0.1618	0.1618	0.0829	0.0829	0.0829	0.0443	0.0443	0.0443	0.0244	0.0244	0.0244

S2.2 3D case

Following the reviewer's suggestion, we report the MSE, variance, and bias of the proposed estimator in the 3D case. For $D = 3$, with $N_1 = N_2 = N_3$, we consider the following data generating process:

$$
\begin{aligned}
m(\boldsymbol{x}) &= 2\sin\{\pi(x_1+x_2)/2\}\, e^{-(x_1+x_2)} + x_1 \sin x_2 + x_1 x_2 x_3, \\
\phi_1(\boldsymbol{x}) &= 2\sqrt{2}\sin(\pi x_1/2)\sin(\pi x_2/2)\sin(\pi x_3/2), \\
\phi_2(\boldsymbol{x}) &= 2\sqrt{2}\sin(3\pi x_1/2)\sin(\pi x_2/2)\sin(3\pi x_3/2), \\
\phi_3(\boldsymbol{x}) &= 2\sin(3\pi x_1/2)\sin(\pi x_2/2)\sin(5\pi x_3/2), \\
\phi_4(\boldsymbol{x}) &= 2\sin(3\pi x_1/2)\sin(3\pi x_2/2)\sin(7\pi x_3/2),
\end{aligned}
$$

$$\phi_5(\boldsymbol{x}) = \sqrt{2}\sin(5\pi x_1/2)\sin(3\pi x_2/2)\sin(9\pi x_3/2),$$

$$\phi_6(\boldsymbol{x}) = \sqrt{2}\sin(5\pi x_1/2)\sin(5\pi x_2/2)\sin(11\pi x_3/2),$$

$$\phi_k(\boldsymbol{x}) = 0 \qquad k \geq 7.$$

For the same reason, we only consider the case $\xi_k \sim N(0,1)$ and $\sigma(x_1, x_2, x_3) \equiv 0.2$. The observations are taken for $N = 50, 75, 100, 125$ (corresponding to $n = 44, 65, 85, 105$).

Table S.6: Additional 3D simulation results based on homogeneous errors with $\sigma(x_1, x_2, x_3) \equiv 0.2$.

Number of grids		$N=50$			$N=75$			$N=100$			$N=125$	
Distribution of ε	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace	normal	uniform	laplace
MSE	0.0790	0.0790	0.0790	0.0534	0.0533	0.0533	0.0372	0.0372	0.0372	0.0315	0.0315	0.0315
Bias	-0.0046	-0.0045	-0.0046	-0.0014	-0.0014	-0.0014	-0.0045	-0.0045	-0.0045	-0.0047	-0.0047	-0.0047
Variance	0.0250	0.0250	0.0250	0.0166	0.0166	0.0166	0.0128	0.0128	0.0128	0.0105	0.0105	0.0105

One finds that the proposed estimator performs well in both 2D and 3D scenarios in terms of MSE, variance, and bias.

S3. Preliminaries

For any vector $\boldsymbol{a} = (a_1, \ldots, a_n) \in \mathbb{R}^n$, take $\|\boldsymbol{a}\|_r = \left(\sum_{i=1}^{n} |a_i|^r\right)^{1/r}, 1 \leq r < +\infty$, $\|\boldsymbol{a}\|_\infty = \max_{1\leq i\leq n} |a_i|$. For any matrix $\boldsymbol{A} = (a_{ij})_{i,j=1}^{m,n}$, denote $\|\boldsymbol{A}\|_r = \max_{\boldsymbol{a}\in\mathbb{R}^n, \boldsymbol{a}\neq\boldsymbol{0}} \|\boldsymbol{A}\boldsymbol{a}\|_r \|\boldsymbol{a}\|_r^{-1}$,for $r < +\infty$ and $\|\boldsymbol{A}\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|$. Denote $\mathrm{Vec}(A)$ as the vectorization of matrix A, i.e., $\mathrm{Vec}(A) = (a_{11}, \ldots, a_{m1}, a_{12}, \ldots, a_{m2}, \ldots, a_{1n}, \ldots, a_{mn})^\top$.

For any two functions $\phi(\cdot), \varphi(\cdot) \in \mathcal{L}^2\left([0,1]^D\right)$, set

$$\langle \phi, \varphi \rangle = \int_{[0,1]^D} \phi(\boldsymbol{x}) \varphi(\boldsymbol{x}) \, d\boldsymbol{x},$$

as their theoretical inner product, with norm $\|\phi\|_2^2 = \langle \phi, \phi \rangle$. We then define the empirical inner product between ϕ and φ

$$\langle \phi, \varphi \rangle_{D,N} = T_D^{-1} \sum_{j_1=1}^{N_1} \cdots \sum_{j_D=1}^{N_D} \phi\left(\boldsymbol{x}_{j_1\ldots j_D}\right) \varphi\left(\boldsymbol{x}_{j_1\ldots j_D}\right),$$

with T_D as given in Assumption (A3), and the empirical norm $\|\phi\|_{D,N}^2 = \langle \phi, \phi \rangle_{D,N}$.

For any positive integer p, the theoretical and empirical inner product matrices of $\left\{B_{J_1\ldots J_D}^{[D]}(\cdot)\right\}_{J_1\ldots J_D=1-p}^{N_{s_1}\ldots N_{s_D}}$ are defined as

$$\boldsymbol{V}_{p,D} = \left(\left\langle B_{j_1\ldots j_D,p}^{[D]}(\cdot), B_{j_1'\ldots j_D',p}^{[D]}(\cdot) \right\rangle\right)_{\substack{j_1\ldots j_D=1-p \\ j_1'\ldots j_D'=1-p}}^{N_{s_1}\ldots N_{s_D}},$$

$$\widehat{\boldsymbol{V}}_{p,D} = \left(\left\langle B^{[D]}_{j_1\ldots j_D,p}(\cdot), B^{[D]}_{j'_1\ldots j'_D,p}(\cdot)\right\rangle_{D,N}\right)^{N_{s_1}\ldots N_{s_D}}_{\substack{j_1\ldots j_D=1-p\\ j'_1\ldots j'_D=1-p}}.$$

It is easy to calculate that $\widehat{\boldsymbol{V}}_{p,D} = T_D^{-1}\left(\mathbf{X}\mathbf{X}^\top\right)$, so we will establish some asymptotic properties of $\widehat{\boldsymbol{V}}_{p,D}$ and closeness between $\widehat{\boldsymbol{V}}_{p,D}$ and $\boldsymbol{V}_{p,D}$.

Lemma S.1 (Lemma A.3 of Cao et al. (2012))**.** *For any $p \in \mathbb{N}_+$ there exists a constant $M_p > 0$ depending only on p, such that for large enough n, $\left\|\boldsymbol{V}_{p,1}^{-1}\boldsymbol{u}\right\|_\infty \leq M_p N_s \|\boldsymbol{u}\|_\infty$ for any vector $\boldsymbol{u} \in \mathbb{R}^{N_s+p}$.*

By basic properties of Kronecker product, one easily obtains the following.

Lemma S.2. *For any $p \in \mathbb{N}_+$ there exists a constant $M_p > 0$ depending only on p, such that for large enough n $\left\|\boldsymbol{V}_{p,D}^{-1}\boldsymbol{u}\right\|_\infty \leq M_p^D N_s^D \|\boldsymbol{u}\|_\infty$. for any vector $\boldsymbol{u} \in \mathbb{R}^{K_n}$.*

The difference between theoretical inner product matrix $\boldsymbol{V}_{p,D}$ and the emprical $\widehat{\boldsymbol{V}}_{p,D}$ of $\left\{B^{[D]}_{J_1\ldots J_D}(\cdot)\right\}^{N_{s_1}\ldots N_{s_D}}_{J_1\ldots J_D=1-p}$ is asymptotically negligible, and $\widehat{\boldsymbol{V}}_{p,D}^{-1}$ behaves similarly as $\boldsymbol{V}_{p,D}^{-1}$.

Lemma S.3. *Under Assumption (A6), $\forall D \in \mathbb{N}^+$, $\left\|\widehat{\boldsymbol{V}}_{p,D} - \boldsymbol{V}_{p,D}\right\|_\infty = \mathcal{O}\left(N^{-D}\right)$ and $\left\|\widehat{\boldsymbol{V}}_{p,D}^{-1}\right\|_\infty = \mathcal{O}\left(N_s^D\right)$*

PROOF. Using the basic properties of B-splines in Lemma S.8, one can obtains

$$\left\|\widehat{\boldsymbol{V}}_{p,D} - \boldsymbol{V}_{p,D}\right\|_\infty = \max_{1-p\leq l_d\leq N_{s_d}, 1\leq d\leq D}\left|\sum_{k_1=1}^{N_{s_1}}\cdots\sum_{k_D=1}^{N_{s_D}} T_D^{-1}\right.$$

$$
\begin{aligned}
& B_{l_1 \ldots l_D}^{[D]}\left(\boldsymbol{x}_{j_1 \ldots j_D}\right) B_{k_1 \ldots k_D}^{[D]}\left(\boldsymbol{x}_{j_1 \ldots j_D}\right)-\int_{[0,1]^D} B_{l_1 \ldots l_D}^{[D]}(\boldsymbol{x}) B_{k_1 \ldots k_D}^{[D]}\left(\boldsymbol{x}^{\prime}\right) d \boldsymbol{x} d \boldsymbol{x}^{\prime} \Bigg| \\
\leq & \max_{1-p \leq l_d \leq N_{s_d}, 1 \leq d \leq D} \sum_{k_1=1}^{N_{s_1}} \cdots \sum_{k_D=1}^{N_{s_D}} \left|T_D^{-1}\right. \\
& B_{l_1 \ldots l_D}^{[D]}\left(\boldsymbol{x}_{j_1 \ldots j_D}\right) B_{k_1 \ldots k_D}^{[D]}\left(\boldsymbol{x}_{j_1 \ldots j_D}\right)-\int_{[0,1]^D} B_{l_1 \ldots l_D}^{[D]}(\boldsymbol{x}) B_{k_1 \ldots k_D}^{[D]}\left(\boldsymbol{x}^{\prime}\right) d \boldsymbol{x} d \boldsymbol{x}^{\prime} \Bigg| \\
= & \max_{1-p \leq l_d \leq N_{s_d}, 1 \leq d \leq D} \sum_{k_1=1}^{N_{s_1}} \cdots \sum_{k_D=1}^{N_{s_D}} \sum_{j_1=1}^{N_1} \cdots \sum_{j_D=1}^{N_D} \\
& \left|\int_{\Delta_{j_1 \ldots j_D}} B_{l_1 \ldots l_D}^{[D]}\left(\boldsymbol{x}_{j_1 \ldots j_D}\right) B_{k_1 \ldots k_D}^{[D]}\left(\boldsymbol{x}_{j_1 \ldots j_D}\right) B_{l_1 \ldots l_D}^{[D]}(\boldsymbol{x}) B_{k_1 \ldots k_D}^{[D]}\left(\boldsymbol{x}^{\prime}\right) d \boldsymbol{x} d \boldsymbol{x}^{\prime}\right| \\
\leq & \sum_{l_1=1}^{N_{s_1}} \cdots \sum_{l_D=1}^{N_{s_D}} \sum_{k_1=1}^{N_{s_1}} \cdots \sum_{k_D=1}^{N_{s_D}} \sum_{j_1=1}^{N_1} \cdots \sum_{j_D=1}^{N_D} \\
& \left|\int_{\Delta_{j_1 \ldots j_D}} B_{l_1 \ldots l_D}^{[D]}\left(\boldsymbol{x}_{j_1 \ldots j_D}\right) B_{k_1 \ldots k_D}^{[D]}\left(\boldsymbol{x}_{j_1 \ldots j_D}\right)-B_{l_1 \ldots l_D}^{[D]}(\boldsymbol{x}) B_{k_1 \ldots k_D}^{[D]}\left(\boldsymbol{x}^{\prime}\right) d \boldsymbol{x} d \boldsymbol{x}^{\prime}\right| \\
= & \mathcal{O}\left(1 \times\left(N N_s^{-1}\right)^D \times\left(N^{-1} N_s\right)^D \times N^{-D}\right)=\mathcal{O}\left(N^{-D}\right),
\end{aligned}
$$

where $\Delta_{j_1 \ldots j_D} = \left[\left(j_1-1\right)/N_1, j_1/N_1\right] \times \cdots \times \left[\left(j_D-1\right)/N_D, j_D/N_D\right]$. For any K_n dimensional vector $\boldsymbol{\tau}$, Lemma S.2 implies $\left\|\boldsymbol{V}_{p,D}^{-1}\boldsymbol{\tau}\right\|_{\infty} \leq M_p^D N_s^D$. Hence $\left\|\boldsymbol{V}_{p,D}^{-1}\boldsymbol{\tau}\right\|_{\infty} \geq M_p^{-D} N_s^{-D} \|\boldsymbol{\tau}\|_{\infty}$. Note next that

$$
\begin{aligned}
\left\|\widehat{\boldsymbol{V}}_{p,D}^{-1}\boldsymbol{\tau}\right\|_{\infty} & \geq \left\|\boldsymbol{V}_{p,D}^{-1}\boldsymbol{\tau}\right\|_{\infty}-\left\|\left(\widehat{\boldsymbol{V}}_{p,D}^{-1}-\boldsymbol{V}_{p,D}^{-1}\right)\boldsymbol{\tau}\right\|_{\infty} \\
& = \mathcal{O}\left(N_s^{-D}\right)\|\boldsymbol{\tau}\|_{\infty}.
\end{aligned}
$$

The proof is complete. □

The next lemma on Gaussian strong approximation is explicit and powerful.

Lemma S.4 (Theorem 4 of Götze and Zaitsev (2010)). *Let ξ be a $\mathbb{R}$-valued random variable with $\mathbb{E}\xi = 0$ and $\mathbb{E}|\xi|^r < \infty$ for some $r > 2$. Then there exists independent random variables $\xi_1, \xi_2, \ldots$ and independent random variables $Z_1, Z_2, \ldots$ defined on a common probability space such that each ξ_i equals ξ in distribution and each Z_i follows $N(0, \mathbb{E}\xi^2)$ and for $x > 0, n \in \mathbb{N}_+$*

$$\mathbb{P}\left\{\max_{1\le\tau\le n}\left|\sum_{i=1}^{\tau}\xi_i - \sum_{i=1}^{\tau}Z_i\right| > x\right\} \le cn\mathbb{E}|\xi|^r/x^r,$$

where c is a positive constant depending only on r.

Lemma S.5 (Lemma S.14 of Wang et al. (2020)). *Let $W_i \sim N(0, \sigma_i^2), \sigma_i > 0, i = 1, \ldots, n$, for $a > 2$*

$$\mathbb{P}\left(\max_{1\le i\le n}|W_i/\sigma_i| > a\sqrt{\log n}\right) < \sqrt{\pi/2}\,n^{1-a^2/2}.$$

The Lévy concentration function of real random variable ξ and $\varepsilon > 0$ according to Chernozhukov et al. (2015) is

$$\mathcal{L}(\xi, \varepsilon) = \sup_{z\in\mathbb{R}}\mathbb{P}\{|\xi - z| \le \varepsilon\}.$$

The next lemma concerns also distributions of Gaussian maxima. For $p \in \mathbb{N}_+$, let covariance matrices $\Sigma_X = (\sigma_{jk,X})_{j,k=1}^p, \Sigma_Y = (\sigma_{jk,Y})_{j,k=1}^p$ be standardized so that $\sigma_{jj,X} = \sigma_{jj,Y} = 1, 1 \leq j \leq p$, and let random vectors

$$(X_1, \ldots, X_p)^\mathsf{T} \sim N(0_p, \Sigma_X), (Y_1, \ldots, Y_p)^\mathsf{T} \sim N(0_p, \Sigma_Y).$$

Denote

$$\Delta = \max_{1 \leq j,k \leq p} |\sigma_{jk,X} - \sigma_{jk,Y}|, \quad a_{2p} = \mathbb{E}\left[\max_{1 \leq j \leq p} \left(|Y_j| \, \sigma_{jj,Y}^{-1/2}\right)\right].$$

Then according to Chernozhukov et al. (2015),

$$a_{2p} = \mathbb{E}\left[\max_{1 \leq j \leq p} \left(Y_j \sigma_{jj,Y}^{-1/2}, -Y_j \sigma_{jj,Y}^{-1/2}\right)\right] \leq \sqrt{2 \log 2p}.$$

Lemma S.6. *For any $\varepsilon > 0$,*

$$\mathcal{L}(\max_{1 \leq j \leq p} |X_j|, \varepsilon) \leq 4\varepsilon (a_{2p} + 1). \tag{S3.1}$$

Furthermore, there is a universal constant $C > 0$ such that

$$\begin{aligned} &\sup_{x \in (0,+\infty)} \left| \mathbb{P}\left(\max_{1 \leq j \leq p} |X_j| \leq x\right) - \mathbb{P}\left(\max_{1 \leq j \leq p} |Y_j| \leq x\right) \right| \\ \leq \;& C \Delta^{1/3} \log^{1/3}(2p) \left\{\max\left(1, a_{2p}^2, -\log \Delta\right)\right\}^{1/3} \end{aligned}$$

$$\leq \quad C\Delta^{1/3} \max\left(1, \log^{2/3}(2p/\Delta)\right). \tag{S3.2}$$

PROOF. The proof follows from arguments in the proofs of Theorems 1 and 3 and Comment 5 in Chernozhukov et al. (2015). □

The next lemma provides spline approximation of $\mathcal{C}^{q,\mu}[0,1]^D$.

Lemma S.7 (Theorem 12.8 of Schumaker (1981))**.** *There is an absolute constant $C_{q,\mu} > 0$ such that for every $\phi \in \mathcal{C}^{q,\mu}[0,1]^D$ for some $\mu \in (0,1]$, there exists a function $g \in \mathscr{H}^{q,D}$ for which$\|g - \phi\|_\infty \leq C_{q,\mu}\|\phi\|_{q,\mu} h_s^{q+\mu}$.*

The next result follows from p. 96 of De Boor (2001)

Lemma S.8 (p96 of De Boor (2001))**.** *For the spline function $B_{J_d,p}(x_d), x_d \in [0,1]$,*

$$\max_{1-p \leq J_d \leq N_{s_d}} \sum_{j_d=1}^{N_d} B_{J_d,p}\left(\frac{j_d}{N_d}\right) = \mathcal{O}\left(NN_s^{-1}\right),$$

$$\max_{1 \leq j_d \leq N_d} \sum_{J_d=1-p}^{N_{s_d}} B_{J_d,p}\left(\frac{j_d}{N_d}\right) \leq p + 1 = \mathcal{O}(1).$$

For $p > 1$, there exists a constant $C_p > 0$,only depend on p such that

$$\max_{\substack{1-p \leq J_d, J_d' \leq N_{s_d},\\ 1 \leq j_d \leq N_d}} \sup_{x_d \in [(i-1)/N, i/N]} \left| B_{J_d,p}\left(\frac{j_d}{N_d}\right) B_{J_d',p}\left(\frac{j_d}{N_d}\right) - B_{J_d,p}(x_d) B_{J_d',p}(x_d) \right| \leq C_p N^{-1} N_s.$$

which also implies that $\|B_{J_d,p}\|_{0,1} = \mathcal{O}(N_s), \forall J = 1, \ldots, N_s$.

Lemma S.9. *The constraints (3.5), (3.6),(3.7),(3.8) are consistent and guarantee that there exists* γ *satisfied (3.9). Under Assumptions (A3) and (A6), as* $N \to \infty$*,*

$$
\begin{aligned}
N_s^{-p^*} (n \log n)^{2/\omega_0} &= \mathcal{O}(1), && \text{(S3.3)} \\
N N_s^{-1} N^{-\nu} &= \mathcal{O}(1), && \text{(S3.4)} \\
N^{-D/2} N_s^{D/2} \log^{1/2} N &= \mathcal{O}\left(n^{-1/2}\right), && \text{(S3.5)} \\
N^{D\beta_2 - 1} N_s &= \mathcal{O}\left(n^{-1/2}\right). && \text{(S3.6)}
\end{aligned}
$$

PROOF. First, note that Hölder continuity indices μ and ν are merely specified appropriate range in (3.5). Then, since $\theta < 2p^*, \theta < 2p^*/(1+p^*)$ required in (3.6), it implies that $p^* > \theta/2$, $1 - \theta/(2D) - \theta/(2p^* D)$, and thus there exists β_2 that satisfies (3.7). Next, since (3.7) compels $\beta_2 < (1 - \theta/2 - \theta/(2p^*))/D$, hence,

$$
2p^*(1 - D\beta_2 - \theta/2) - \theta > 0 \qquad \text{(S3.7)}
$$

and ω_0 exists. Finally, according to (3.7), (3.8) and (S3.7), one obtains $1 - \nu < 1 - D\beta_2 - \theta/2$ and $\theta(1 + 4/\omega_0)/2p^* < 1 - D\beta_2 - \theta/2$, thus

$$
\max\left\{\frac{\theta}{2p^*} + \frac{2\theta}{p^*\omega_0}, 1 - \nu\right\} < 1 - D\beta_2 - \frac{\theta}{2}.
$$

thus γ exists.

Noticing that Assumptions (A3) and (A6) ensure that

$$N_s^{-p^*}(n\log n)^{2/\omega_0}n^{1/2} = N^{-p^*\gamma+2\theta/\omega_0+\theta}(\log^\tau N)^{-p^*}\left(\log n\right)^{2/\omega_0}.$$

Following from (3.9, (S3.3) is proved.

Since $N_s N^{-\gamma} + N_s^{-1}N^{\gamma} = \mathcal{O}\left(\log^\tau N\right)$, one computes,

$$\begin{aligned} NN_s^{-1}N^{-\nu} &= \mathcal{O}(N^{1-\gamma-\nu}\log^{-\tau p^*} N) = \mathcal{O}\left(N^{-1-\gamma-\nu}\log^{-\tau} N\right) \\ N^{-D/2}N_s^{D/2}\log^{1/2} Nn^{1/2} &= \mathcal{O}(N^{-D/2+D\gamma/2+\theta/2}\log^{(\tau D+1)/2} N) \\ &= \mathcal{O}\left(N^{-D/2+D\gamma/2+\theta/2}\log^{(\tau D+1)/2} N\right) \\ N^{D\beta_2-1}N_s n^{1/2} &= \mathcal{O}(N^{D\beta_2-1+\gamma+\theta/2}\log^\tau N) = \mathcal{O}\left(N^{D\beta_2-1+\gamma+\theta/2}\log^\tau N\right) \end{aligned}$$

which are $o\left(1\right)$ based on (3.9) and (S3.4) - (S3.6). The proof is completed. □

Lemma S.10. *Assumption (A5') implies Assumptions (A5)*

PROOF. Under Assumption (A5), applying Lemma S.4, for $k = 1, \ldots, \kappa_n$, one obtains $\left\{\tilde{\xi_{ik}}\right\}_{i=1}^n$ equal in distribution to $\left\{\xi_{ik}\right\}_{i=1}^n$ and standard Gaussian variables

$\{Z_{ik,\xi}\}_{i=1}^{n}$ on a probability space $\left(\tilde{\Omega}_{k,\xi}, \tilde{\mathcal{A}}_{k,\xi}, \tilde{\mathbb{P}}_{k,\xi}\right)$, such that

$$\tilde{\mathbb{P}}_{k,\zeta}\left(\max_{1\le t\le n}\left|\sum_{i=1}^{t}\tilde{\xi_{ik}} - \sum_{i=1}^{t} Z_{ik,\xi}\right| > n^{\beta_1}\right) \le c_{\varpi_1} n \mathbb{E}|\xi_{1k}|^{\varpi_1} n^{-\beta_1\varpi_1}.$$

Similarly, for $\beta_2 \in (0, \varpi_3/D)$, such that $\varpi_3\beta_2 > D+1+\theta$, one obtains $\left\{\tilde{\varepsilon}_{i,j_1(j)\ldots j_D(j)}\right\}_{j=1}^{T_D}$ equal in distribution to $\left\{\varepsilon_{i,j_1(j)\ldots j_D(j)}\right\}_{j=1}^{T_D}$, and standard Gaussian variables $\left\{Z_{i,j_1(j)\ldots j_D(j),\varepsilon}\right\}_{j=1}^{T_D}$ on a probability space $\left(\tilde{\Omega}_{i,\varepsilon}, \tilde{\mathcal{A}}_{i,\varepsilon}, \tilde{\mathbb{P}}_{i,\varepsilon}\right)$ such that

$$\mathbb{P}_{i,\varepsilon}\left\{\max_{1\le t\le T_D}\left|\sum_{j=1}^{t}\tilde{\varepsilon}_{i,j_1(j)\ldots j_D(j)} - \sum_{j=1}^{t} Z_{i,j_1(j)\ldots j_D(j),\varepsilon}\right| > N^{D\beta_2}\right\}$$

$$\le c_{\varpi_2} N^D \mathbb{E}|\varepsilon_{1,1\ldots 1}|^{\varpi_2} N^{-D\beta_2\varpi_2}.$$

Since Assumption (A5') stipulates the independence of $\{\xi_{ik}\}_{i=1,k=1}^{n,\infty}$ and $\left\{\varepsilon_{i,j_1(j)\ldots j_D(j)}\right\}_{i=1,j=1}^{n,T_D}$, the independence is automatically preserved for $\left\{\tilde{\xi}_{ik}\right\}_{i=1,k=1}^{n,\kappa_n}$ and $\left\{\tilde{\varepsilon}_{i,j_1(j)\ldots j_D(j)}\right\}_{i=1,j=1}^{n,T_D}$ if their new probability space $\left(\tilde{\Omega}_{k,\xi}, \tilde{\mathcal{A}}_{k,\xi}, \tilde{\mathbb{P}}_{k,\xi}\right)$, $k \ge 1$ and $\left(\tilde{\Omega}_{i,\varepsilon}, \tilde{\mathcal{A}}_{i,\varepsilon}, \tilde{\mathbb{P}}_{i,\varepsilon}\right)$, $i \ge 1$ are all independently embedded into a product probability space

$$\left(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{\mathbb{P}}\right) \equiv \left(\left(\bigotimes_{k=1}^{\infty}\tilde{\Omega}_{k,\xi}\right)\bigotimes\left(\bigotimes_{i=1}^{\infty}\tilde{\Omega}_{i,\varepsilon}\right), \left(\bigotimes_{k=1}^{\infty}\tilde{\mathcal{A}}_{k,\xi}\right)\bigotimes\left(\bigotimes_{i=1}^{\infty}\tilde{\mathcal{A}}_{i,\varepsilon}\right), \tilde{\mathbb{P}}\right),$$

according to Ionescu-Tulcea Theorem (Theorem 14.32 in Klenke (2014)). This independent embedding also ensures that all Gaussian random vectors $\{Z_{ik,\xi}\}_{i=1,k=1}^{n,\kappa_n}$,

$\{Z_{ij,\varepsilon}\}_{i=1,j=1}^{n,T_D}$ remain independent in the new product probability space, as required in the Assumption (A5).

In what follows, with some abuse of notations, we will not distinguish ξ_{ik}, ε_{ij} on the original probability space from $\tilde{\xi}_{ik}$, $\tilde{\varepsilon}_{ij}$ on the above product probability space, nor the original probability measure $\mathbb{P}$ from $\tilde{\mathbb{P}}$ on the product space.

Since $\sup_{k\geq 1} \mathbb{E}\,|\xi_{1,k}|^{\varpi_1} < \infty$ by Assumption (A5'), there exists a common $c_{\varpi_1} > 0$, such that

$$\mathbb{P}\left(\max_{1\leq k\leq \kappa_n}\max_{1\leq t\leq n}\left|\sum_{i=1}^{t}\xi_{ik}-\sum_{i=1}^{t}Z_{ik,\xi}\right|>n^{\beta_1}\right)\leq c_{\varpi_1}\sup_{t,k}\mathbb{E}|\xi_{1k}|^{\varpi_1}\kappa_n n^{1-\beta_1\varpi_1}.$$

Noticing that $\varpi_1 > (4+2\omega)$, so there exists some $\beta \in (0, 1/2)$, $\varpi_1 > (2+\omega)/\beta_1$ and $\gamma_1 = \varpi_1\beta_1 - 1 - \omega > 1$, the first assertion is proved.

Also, one has that

$$\mathbb{P}\left\{\max_{1\leq i\leq n, 1\leq t\leq T_D}\left|\sum_{j=1}^{t}\varepsilon_{i,j_1(j)\ldots j_D(j)}-\sum_{j=1}^{t}Z_{i,j_1(j)\ldots j_D(j),\varepsilon}\right|>N^{D\beta_2}\right\}$$

$$\leq c_{\varpi_2}\mathbb{E}|\varepsilon_{1,1\ldots 1}|^{\varpi_2}N^{D(1-\beta_2\varpi_2)+\theta},$$

and based on Assumption (A5'), there exists β_2 in (3.7) satisfying $D\varpi_2\beta_2 > D+1+\theta$ ensuring $D\,(1-\beta_2\varpi_2)+\theta < -1$. The proof is completed. $\square$

Denote $\widetilde{\boldsymbol{Z}}_{ip,\varepsilon}(\cdot) = \mathbf{B}_p^\top(\cdot)\left(\mathbf{X}^\top\mathbf{X}\right)^{-1}\mathbf{X}^\top\boldsymbol{Z}_i$, where as in equation (3.10), $\boldsymbol{Z}_i = \left(\sigma\left(\boldsymbol{x}_{j_1,\dots,j_D}\right) Z_{i,j_1\dots j_D}\right)_{j_d=1\dots N}$

Lemma S.11. *Under Assumptions (A2) -(A6), as $N\to\infty$,*

$$\max_{1\le i\le n}\left\|\widetilde{\boldsymbol{Z}}_{ip,\varepsilon} - \widetilde{e}_{ip}\right\|_\infty = \mathcal{O}_{a.s}\left(N^{D\beta_2-1}N_s\right).$$

PROOF. This follows from

$$\begin{aligned}
&\sum_{j=1}^{T_D} B^{[D]}_{J_1\dots J_D,p}\left(\boldsymbol{x}_{j_1(j)\dots j_D(j)}\right)\sigma_i\left(\boldsymbol{x}_{j_1(j)\dots j_D(j)}\right)\left(Z_{i,j_1(j)\dots j_D(j),\varepsilon} - \varepsilon_{i,j_1(j)\dots j_D(j)}\right)\\
=&\sum_{j=1}^{T_D-1}\{\Big(B^{[D]}_{J_1\dots J_D,p}\left(\boldsymbol{x}_{j_1(j)\dots j_D(j)}\right)\sigma_i\left(\boldsymbol{x}_{j_1(j)\dots j_D(j)}\right)\\
&-B^{[D]}_{J_1\dots J_D,p}\left(\boldsymbol{x}_{j_1(j+1)\dots j_D(j+1)}\right)\sigma_i\left(\boldsymbol{x}_{j_1(j+1)\dots j_D(j+1)}\right)\Big)\sum_{k=1}^{j}\left(Z_{i,j_1(k)\dots j_D(k),\varepsilon} - \varepsilon_{i,j_1(k)\dots j_D(k)}\right)\}\\
&+B^{[D]}_{J_1\dots J_D,p}\left(\boldsymbol{x}_{j_1(T_D)\dots j_D(T_D)}\right)\sigma_i\left(\boldsymbol{x}_{j_1(T_D)\dots j_D(T_D)}\right)\sum_{k=1}^{T_D}\left(\left(Z_{ij_1(k)\dots j_D(k),\varepsilon} - \varepsilon_{ij_1(k)\dots j_D(k)}\right)\right).
\end{aligned}$$

Notice that the difference between these two vectors, $B_{J_1\dots J_D,p}\left(\boldsymbol{x}_{j_1(j)\dots j_D(j)}\right)$ and $B_{J_1\dots J_D,p}\left(\boldsymbol{x}_{j_1(j+1)\dots j_D(j+1)}\right)$, is only one component.

Applying Lemma S.8, one has

$$\left|B_{J_1\dots J_D,p}\left(\boldsymbol{x}_{j_1(j)\dots j_D(j)}\right) - B_{J_1\dots J_D,p}\left(\boldsymbol{x}_{j_1(j+1)\dots j_D(j+1)}\right)\right| \le CN_sN^{-1}.$$

Under Assumption (A3), the Hölder continuity ensures that

$$\left|\sigma_i\left(\boldsymbol{x}_{j_1(j)\ldots j_D(j)}\right)-\sigma_i\left(\boldsymbol{x}_{j_1(j+1)\ldots j_D(j+1)}\right)\right| \leq CN^{-\nu}\max_{1\leq i\leq n}\|\sigma_i\|_{0,\nu} \leq CN_sN^{-1},$$

where these bounds are uniformly over $1 \leq J_1 \ldots J_D \leq N_s$. Then we can get

$$\begin{aligned} & \left\|T_D^{-1}\mathbf{X}^\top\left(\boldsymbol{Z}_i-\boldsymbol{e}_i\right)\right\|_\infty \\ \leq & \left\{\max_{1\leq i\leq n}\max_{1\leq j\leq T_D}\left|T_D^{-1}\sum_{t=1}^{j}\left(\varepsilon_{i,j_1(j)\ldots j_D(j)}-Z_{i,j_1(j)\ldots j_D(j),\varepsilon}\right)\right|\right\} \\ & \times\{CN^DN_s^{-D}N_sN^{-1}\}+T_D^{-1}\sum_{k=1}^{T_D}\left(\left(Z_{ij_1(k)\ldots j_D(k),\varepsilon}-\varepsilon_{ij_1(k)\ldots j_D(k)}\right)\right) \\ = & \mathcal{O}_{a.s.}\left(N^{D\beta_2-1}N_s^{1-D}+N^{D\beta_2-D}\right). \end{aligned}$$

Then Lemma S.3 implies that

$$\begin{aligned} \max_{1\leq i\leq n}\left\|\widetilde{\boldsymbol{Z}}_{ip,\varepsilon}-\widetilde{e}_{ip}\right\|_\infty & \leq C\left\|\widehat{V}_{p,D}\right\|_\infty\left\|N^{-D}\mathbf{X}^\top\left(\boldsymbol{Z}_i-\boldsymbol{e}_i\right)\right\|_\infty \\ & = \mathcal{O}_{a.s.}\left(N^{D\beta_2-1}N_s\right). \end{aligned}$$

The proof is completed. □

Lemma S.12. *Under Assumptions (A1)-(A3), (A5)-(A6), as* $N \to \infty$

$$\max_{1\le i\le n} \|\widetilde{e}_{ip}\|_\infty = \mathcal{O}_{a.s.}\left(N^{D\beta_2-1}N_s + N^{-D/2}N_s^{D/2}\log^{1/2}N\right).$$

PROOF. We just need to control the decayed rate of $\widetilde{\boldsymbol{Z}}_{ip,\varepsilon}(\boldsymbol{x})$. Note that the random vector $T_D^{-1}\widehat{\boldsymbol{V}}_{p,D}^{-1}\mathbf{X}^\top\boldsymbol{Z}_i$ is $K_n = \prod_{d=1}^{D}(N_{s_d}+p)$ (see Section 4) dimensional normal with covariance matrix $T_D^{-2}\widehat{\boldsymbol{V}}_{p,D}^{-1}\mathbf{X}^\top\mathrm{Var}(\boldsymbol{Z}_i)\,\mathbf{X}^\top\widehat{\boldsymbol{V}}_{p,D}^{-1}$ is bounded by

$$\left\|T_D^{-2}\widehat{\boldsymbol{V}}_{p,D}^{-1}\widehat{\boldsymbol{V}}_{p,D}\widehat{\boldsymbol{V}}_{p,D}^{-1}\right\|_\infty \le CT_D^{-1}\left\|\widehat{\boldsymbol{V}}_{p,2}^{-1}\right\|_\infty = \mathcal{O}\left(N^{-D}N_s^D\right).$$

bounding the tail probability of entries of $T_D^{-1}\widehat{\boldsymbol{V}}_{p,2}^{-1}\mathbf{X}^\top\boldsymbol{Z}_i$ by Lemma S.5 and applying the Borel-Cantelli lemma leads to

$$\left\|T_D^{-1}\widehat{\boldsymbol{V}}_{p,D}^{-1}\mathbf{X}^\top\boldsymbol{Z}_i\right\|_\infty = \mathcal{O}_{a.s.}\left(N^{-D/2}N_s^{D/2}\log^{1/2}n\right).$$

which implies

$$\left\|\widetilde{\boldsymbol{Z}}_{ip,\varepsilon}\right\|_\infty = \mathcal{O}_{a.s.}\left(N^{-D/2}N_s^{D/2}\log^{1/2}N\right).$$

hence, the triangle inequality implies the initial assertion. $\square$

For any function $\phi \in C[0,1]^D$ denote the vector $\boldsymbol{\phi} = \left(\phi\left(\boldsymbol{x}_{j_1\ldots j_D}\right)\right)^\top$as the same order as (3.10) and function $\widetilde{\phi}(\cdot) = \mathbf{B}_p^\top(\cdot)\left(\mathbf{X}^\top\mathbf{X}\right)^{-1}\mathbf{X}^\top\boldsymbol{\phi}$.

Lemma S.13. *There exists* $c_{\phi,p} \in (0,\infty)$ *such that when* n *is large enough,* $\left\|\widetilde{\phi}\right\|_\infty \leq c_{\phi,p}\|\phi\|_\infty$ *for any* $\phi \in C[0,1]^D$*. Furthermore, if* $\phi \in \mathcal{C}^{q,\mu}[0,1]^D$ *for some* $\mu \in (0,1]$*, then there exists* $\widetilde{C}_{q,\mu}$

$$\left\|\widetilde{\phi} - \phi\right\|_\infty \leq \widetilde{C}_{p-1,\mu}\|\phi\|_{q,\mu} h_s^{\mu+q}.$$

PROOF. Note that for any $\boldsymbol{x} \in [0,1]^D$ at most $(p+1)^D$ the vector component of $\mathbf{B}_p(\cdot)$ are in $(0,1)$, others being 0 based on Lemma S.8, so

$$\begin{aligned}\left\|\widetilde{\phi}\right\|_\infty &\leq (p+1)^D \left\|\left(\mathbf{X}^\top\mathbf{X}\right)^{-1}\mathbf{X}^\top\boldsymbol{\phi}\right\|_\infty \leq (p+1)^D T_D^{-1}\left\|\widehat{\boldsymbol{V}}_{p,D}^{-1}\mathbf{X}^\top\boldsymbol{\phi}\right\| \\ &\leq M_p^D T_D^{-1}(p+1)^d N_s^D\|\phi\|_\infty \left\|\mathbf{X}^\top\mathbf{1}_{T_D}\right\|,\end{aligned}$$

in which $\mathbf{1}_{T_D} = (1,\ldots,1)^\top$ is a T_D-dimensional vector of 1's. Clearly, Lemma S.8 ensures that

$$\left\|\mathbf{X}^\top\mathbf{1}_{T_D}\right\|_\infty = \max_{1-p\leq J_j\leq N_s, 1\leq d\leq D}\sum_{j_1=1}^{N_1}\cdots\sum_{j_D=1}^{N_D} B_{J_1\ldots J_D}^{[D]}\left(\boldsymbol{x}_{j_1\ldots j_D}\right) \leq CN^D N_s^{-D},$$

which implies $\left\|\widetilde{\phi}\right\|_\infty \leq c_{\phi,p}\|\phi\|_\infty$.

Now if $\phi \in \mathcal{C}^{q,\mu}[0,1]^D$ for some $\mu \in (0,1]$, let $g \in \mathscr{H}^{p-1,D}$ be such that $\|g-\phi\|_\infty \leq C_{q,\mu}\|\phi\|_{q,\nu} h_s^{q+\mu}$ according to Lemma S.7, the $\widetilde{g} \equiv g$ as $g \in \mathscr{H}^{p-1,D}$, hence,

$$\left\|\widetilde{\phi} - \phi\right\|_\infty = \left\|\widetilde{\phi} - \widetilde{g}\right\|_\infty + \|\phi - g\|_\infty$$

$$\leq (c_{\phi,p}+1)\|\phi-g\|_\infty \leq \widetilde{C}_{q,\mu}\|\phi\|_{q,\mu}h_s^{\mu+q}.$$

The proof is completed. □

Lemma S.14. *Under Assumptions (A1),(A3),(A6) as $N \to \infty$*

$$\sqrt{n}\left\|\widetilde{m}_p - m\right\|_\infty = o(1).$$

PROOF. According to Lemma S.13 and the Assumptions (A1), (A3),(A6), $\sqrt{n}h_s^{p^*} \to 0$. □

Lemma S.15. *Under Assumptions (A2) - (A6), as $N \to \infty$*

$$\sqrt{n}\left\|\widetilde{e}_p\right\|_\infty = o_p(1).$$

PROOF. Let $\widetilde{\boldsymbol{Z}}_{p,\varepsilon}(\cdot) = \mathbf{B}_p^\top(\cdot)\left(\mathbf{X}^\top\right)^{-1}\mathbf{X}^\top\boldsymbol{Z}, \boldsymbol{Z} = \sum_{i=1}^n \boldsymbol{Z}_i/n$. Applying Lemma S.11,

$$\begin{aligned}\left\|\widetilde{\boldsymbol{Z}}_{p,\varepsilon} - \widetilde{e}_p\right\|_\infty &\leq n^{-1}\sum_{i=1}^n \max_{1\leq i\leq n}\left\|\widetilde{\boldsymbol{Z}}_{ip,\varepsilon} - \widetilde{e}_{ip}\right\| \\ &= \mathcal{O}_{a.s.}\left(N^{D\beta_2-1}N_s\right).\end{aligned}$$

Next, we would derive the uniform convergence rate of $\widetilde{\boldsymbol{Z}}_{p,\varepsilon}(\boldsymbol{x})$ analogously, because the variance of $\left(\mathbf{X}^\top\mathbf{X}\right)^{-1}\mathbf{X}^\top\boldsymbol{Z}$ are bounded by

$$\left\|N^{-2D}\widehat{\boldsymbol{V}}_{p,D}^{-1}\widehat{\boldsymbol{V}}_{p,D}\widehat{\boldsymbol{V}}_{p,D}^{-1}\right\|_\infty \le Cn^{-1}N^{-D}\left\|\widehat{\boldsymbol{V}}_{p,2}^{-D}\right\|_\infty = \mathcal{O}\left(n^{-1}N^{-D}N_s^D\right).$$

Lemma S.5 ensures that the uniform convergence rate of $\left(\mathbf{X}^\top\mathbf{X}\right)^{-1}\mathbf{X}^\top\boldsymbol{Z}$ is $\mathcal{O}_{a.s.}\left(n^{-1/2}N^{-D}N_s^D\log^{1/2}n\right)$, Finally, the proof is completed by triangle inequality and (S3.5) - (S3.6)

$$\begin{aligned}\sqrt{n}\left\|\widetilde{e}\right\|_\infty &\le \sqrt{n}\left\|\widetilde{\boldsymbol{Z}}_{p,\varepsilon}\right\|_\infty + \sqrt{n}\left\|\widetilde{\boldsymbol{Z}}_{p,\varepsilon} - \widetilde{e}_p\right\|_\infty \\ &= \mathcal{O}_{a.s.}\left(N^{D\beta_2-1}N_s n^{1/2} + N^{-D}N_s^D\log^{1/2}n\right) = o_p(1).\end{aligned}$$

□

Lemma S.16. *Under Assumptions (A1)-(A6), for i.i.d.* $N(0,1)$ *variables* $\overline{Z}_{\cdot,k,\xi} = n^{-1}\sum_{i=1}^n Z_{i,k,\xi}$, $Z_{i,k,\xi}$*'s as in Assumptions (A5), as* $n\to\infty$,

$$\left\|\sum_{k=1}^\infty \sqrt{n}\overline{Z}_{\cdot,k,\xi}\phi_k - \sqrt{n}\{\overline{m} - m\}\right\|_\infty = o_p(1).$$

For any $\alpha\in(0,1)$ *and i.i.d.* $N(0,1)$ *variables* $Z_k, k\in\mathbb{N}_+$, *as* $n\to\infty$,

$$\mathbb{P}\left\{\sup_{\boldsymbol{x}\in[0,1]^D}\left|\sqrt{n}\{\overline{m}(\boldsymbol{x}) - m(\boldsymbol{x})\}G(\boldsymbol{x},\boldsymbol{x})^{-1/2}\right| \le Q_{1-\alpha}\right\}$$

$$\to \mathbb{P}\left\{\sup_{\boldsymbol{x}\in[0,1]^D}\left|\sum_{k=1}^{\infty} Z_k\phi_k(\boldsymbol{x})\, G(\boldsymbol{x},\boldsymbol{x})^{-1/2}\right| \le Q_{1-\alpha}\right\} = 1-\alpha.$$

PROOF. We denote $\widetilde{\zeta}_k(\cdot) = \overline{Z}_{\cdot k,\xi}\phi_k(\cdot), 1 \le k < \infty$ and define

$$\widetilde{\Xi}(\cdot) = \sqrt{n}\left[\sum_{k=1}^{\infty}\phi_l^2(\cdot)\right]^{-1/2}\sum_{k=1}^{\infty}\widetilde{\zeta}_k(\cdot) = \sqrt{n}G(\cdot,\cdot)^{-1/2}\sum_{l=1}^{\infty}\widetilde{\zeta}_k(\cdot).$$

It is clear that $\widetilde{\Xi}(\cdot)$ is a Gaussian field with the same distribution as $\Xi(\cdot)$, i.e., $\mathbb{E}\widetilde{\Xi}(\cdot) = 0, \mathbb{E}\widetilde{\Xi}^2(\cdot) = 1$ and

$$\mathbb{E}\widetilde{\Xi}(\boldsymbol{x})\,\widetilde{\Xi}(\boldsymbol{x}') = G(\boldsymbol{x},\boldsymbol{x}')\{G(\boldsymbol{x},\boldsymbol{x})\,G(\boldsymbol{x}',\boldsymbol{x}')\}^{-1/2}, \boldsymbol{x}, \boldsymbol{x}'^{D}.$$

Notice that for $\overline{\xi}_{\cdot k} = n^{-1}\sum_{i=1}^{n}\xi_{ik}$

$$\mathbb{E}\left|\overline{\xi}_{\cdot k}\right| \le \left(\mathbb{E}\left|\overline{\xi}_{\cdot k}\right|^2\right) = \mathcal{O}\left(n^{-1/2}\right).$$

In addition, Assumption (A5) entails that

$$\mathbb{P}\left(\max_{1\le k\le\kappa_n}\left|\overline{\xi}_{\cdot k} - \overline{Z}_{\cdot,k,\xi}\right| > C_1 n^{\beta_1-1}\right) \le C_2 n^{-\gamma_1},$$

which implies

$$\max_{1\le k\le\kappa_n}\left|\overline{\xi}_{\cdot k} - \overline{Z}_{\cdot,k,\xi}\right| = \mathcal{O}_{a.s.}\left(n^{\beta_1-1}\right).$$

Similarly, under Assumption (A4) it is clearly that

$$\mathbb{E}\left|\overline{\xi}_{\cdot k}-\overline{Z}_{\cdot,k,\xi}\right|\leq\mathbb{E}\left|\overline{\xi}_{\cdot,k}\right|+\mathbb{E}\left|\overline{Z}_{\cdot,k,\xi}\right|=\mathcal{O}(n^{-1/2})$$

and

$$\begin{aligned}
& \sqrt{n}\sup_{\boldsymbol{x}\in[0,1]^D}G\left(\boldsymbol{x},\boldsymbol{x}\right)^{-1/2}\left|\sum_{k=1}^{\kappa_n}\left(\overline{Z}_{\cdot k,\xi}-\overline{\xi}_{\cdot k}\right)\phi_k\left(\boldsymbol{x}\right)\right| \\
\leq\; & C\max_{1\leq k\leq\kappa_n}\left|\overline{\xi}_{\cdot,k}-\overline{Z}_{\cdot,k,\xi}\right|\sum_{k=1}^{\kappa_n}\|\phi_k\|_\infty=\mathcal{O}_{a.s.}(1) \\
& \mathbb{E}\sqrt{n}\sup_{\boldsymbol{x}\in[0,1]^D}G\left(\boldsymbol{x},\boldsymbol{x}\right)^{-1/2}\left|\sum_{k=\kappa_n+1}^{\infty}\left(\overline{Z}_{\cdot k,\xi}-\overline{\xi}_{\cdot k}\right)\phi_k\left(\boldsymbol{x}\right)\right| \\
\leq\; & Cn^{\beta_1-1/2}\sum_{k=\kappa_n+1}^{\infty}\|\phi_k\|_\infty=\mathcal{O}(1),
\end{aligned}$$

hence

$$\sqrt{n}\sup_{\boldsymbol{x}\in[0,1]^D}G\left(\boldsymbol{x},\boldsymbol{x}\right)^{-1/2}\left|\sum_{k=1}^{\infty}\left(\overline{Z}_{\cdot k,\xi}-\overline{\xi}_{\cdot k}\right)\phi_k\left(\boldsymbol{x}\right)\right|=\mathcal{O}_p(1).$$

Note that

$$G\left(\cdot,\cdot\right)^{1/2}\widetilde{\Xi}\left(\cdot\right)-\sqrt{n}\{\overline{m}\left(\cdot\right)-m\left(\cdot\right)\}=\sqrt{n}\sum_{k=1}^{\infty}\left(\overline{Z}_{\cdot k,\xi}-\overline{\xi}_{\cdot k}\right)\phi_k\left(\cdot\right),$$

hence

$$\sup_{\boldsymbol{x}\in[0,1]^D}\left|\widetilde{\Xi}\left(\boldsymbol{x}\right)-\sqrt{n}G\left(\boldsymbol{x},\boldsymbol{x}\right)^{-1/2}\left\{\overline{m}\left(\boldsymbol{x}\right)-m\left(\boldsymbol{x}\right)\right\}\right|=\mathcal{O}_p(1).$$

The proof is complete. □

Proof of Theorem 1.

For any $k \in \mathbb{N}_+$, let $\widetilde{\phi}_k(\cdot) = \mathbf{B}_p(\cdot)\left(\mathbf{X}^\top\mathbf{X}\right)^{-1}\mathbf{X}\boldsymbol{\phi}_k$. According to equation (3.13),

$$\widehat{\eta}_i(\cdot) - \eta_i(\cdot) = \widetilde{R}_{ip}(\cdot) - R_i(\cdot) + \widetilde{m}_p(\cdot) - m(\cdot) + \widetilde{e}_i(\cdot). \tag{S3.8}$$

By Lemma S.13, there exist universal constant $C_{q,\mu}$ such that

$$\begin{aligned} \|\widetilde{m}_p - m\|_\infty &\leq C_{q,\mu}\|m\|_{q,\mu}N_s^{-p^*}, \\ \left\|\widetilde{\phi}_k - \phi_k\right\|_\infty &\leq C_{q,\mu}\|\phi_k\|_{q,\mu}N_s^{-p^*}, k = 1\ldots,\kappa_n, \end{aligned}$$

which implies that

$$\left\|\widetilde{R}_{ip} - R_i\right\|_\infty = \sum_{k=1}^{\infty}|\xi_{ik}|\left\|\widetilde{\phi}_k - \phi_k\right\|_\infty \leq C_{q,\mu}W_iN_s^{-p^*},$$

where $W_i = \sum_{k=1}^{\infty}|\xi_{ik}|\|\phi_k\|_{q,\mu}, i = 1,\ldots,n$ are i.i.d nonnegative random variables with ω_0-th finite absolute moment under Assumptions (A4), (A5). Hence

$$\mathbb{P}\left\{\max_{1\leq i\leq n} W_i > (n\log n)^{2/\omega_0}\right\} \leq n\frac{\mathbb{E}W_i^{\omega_0}}{(n\log n)^2} = \mathbb{E}W_i^{\omega_0}n^{-1}(\log n)^{-2},$$

thus, the Borel-Cantelli lemma ensures that $\max\limits_{1\leq i\leq n} W_i = \mathcal{O}_{a.s}\left\{(n\log n)^{2/\varpi_0}\right\}$. Combined with equations (S3.8), (S3.3), (S3.6) and Lemma S.12

$$\begin{aligned}
& \max_{1\le i\le n} \|\widehat{\eta}_i - \eta_i\|_\infty \\
\le & \max_{1\le i\le n} \left\|\widetilde{R}_{ip} - R_i\right\|_\infty + \max_{1\le i\le n} \|\widetilde{m}_p - m\|_\infty + \max_{1\le i\le n} \|\widetilde{e}_i\|_\infty \\
= & \mathcal{O}_{a.s.}\left(N_s^{-p^*} + N_s^{-p^*}\left(n\log n\right)^{2/\omega_0} + N^{D\beta_2-1}N_s + N^{-D/2}N_s^{D/2}\log^{\frac{1}{2}} N\right) \\
= & \mathcal{O}_{a.s.}\left(N_s^{-p^*}\left(n\log n\right)^{2/\varpi_0} + N^{D\beta_2-1}N_s\right).
\end{aligned}$$

□

Proof of Theorem 2.

Notice the decomposition

$$|\overline{m}(\cdot) - \widehat{m}_p(\cdot)| \le \max_{1\le i\le n} |\widetilde{\eta}_{ip}(\cdot) - \eta(\cdot)| + \max_{1\le i\le n} |\widetilde{e}_{ip}(\cdot)|.$$

The maximum deviation $|\overline{m}(\cdot) - \widehat{m}_p(\cdot)|$ is controlled by the uniform approximate power of trajectories and the convergence rate of error term. Therefore, applying Lemmas S.14, S.15, and S.16, the proof is completed. □

Proof of Theorem 3.

Lemmas S.14, S.15, and S.16 together imply Theorem 3. □

Proof of Theorem 4.

For the collection of i.i.d. random fields $\{\eta_i(\boldsymbol{x}), \boldsymbol{x} \in [0,1]^D\}_{i=1}^n$, we define the "infeasible"estimator of covariance function, for $\boldsymbol{x}, \boldsymbol{x}' \in [0,1]^D$.

$$\widehat{G}(\boldsymbol{x}, \boldsymbol{x}') = \frac{1}{n}\sum_{i=1}^{n}\left(\eta_i(\boldsymbol{x}) - \overline{m}(\boldsymbol{x})\right)\left(\eta_i(\boldsymbol{x}') - \overline{m}(\boldsymbol{x}')\right), \boldsymbol{x}, \boldsymbol{x}' \in [0,1]^D, \tag{S3.9}$$

where $\overline{m}(\boldsymbol{x})$ is the "infeasible"estimator of mean function defined in (2.2).

Applying Theorem 1, and Lemmas S.15, S.16, S.14 and Assumption (A6) one obtains that

$$\begin{aligned}
& \sup_{\boldsymbol{x},\boldsymbol{x}' \in [0,1]^D} \left|\widehat{G}(\boldsymbol{x}, \boldsymbol{x}') - \widehat{G}_p(\boldsymbol{x}, \boldsymbol{x}')\right| \\
= & \sup_{\boldsymbol{x},\boldsymbol{x}' \in [0,1]^D} \left|\left(\widehat{m}_p(\boldsymbol{x})\widehat{m}_p(\boldsymbol{x}') - \overline{m}(\boldsymbol{x})\overline{m}(\boldsymbol{x}')\right) + \frac{1}{n}\sum_{i=1}^{n}\left(\eta(\boldsymbol{x})\eta(\boldsymbol{x}') - \widehat{\eta}(\boldsymbol{x})\widehat{\eta}(\boldsymbol{x}')\right)\right| \\
\leq & \sup_{\boldsymbol{x},\boldsymbol{x}' \in [0,1]^D} \left|\widehat{m}_p(\boldsymbol{x})\widehat{m}_p(\boldsymbol{x}') - \overline{m}(\boldsymbol{x})\overline{m}(\boldsymbol{x}')\right| + \frac{1}{n}\sum_{i=1}^{n}\sup_{\boldsymbol{x},\boldsymbol{x}' \in [0,1]^D}\left|\eta(\boldsymbol{x})\eta(\boldsymbol{x}') - \widehat{\eta}(\boldsymbol{x})\widehat{\eta}(\boldsymbol{x}')\right| \\
\leq & \left\|\left(|\overline{m}| + |\widehat{m}_p|\right)|\widehat{m}_p - \overline{m}|\right\|_\infty + \max_{1\leq i\leq n}\left\|\left(|\eta_i| + |\widehat{\eta}_i|\right)|\widehat{\eta}_i - \eta_i|\right\|_\infty
\end{aligned}$$

Similarly, one recalls,

$$\begin{aligned}
\max_{1\leq i\leq n}\|\widehat{\eta}_i - \eta_i\|_\infty &= \mathcal{O}_p\left(N_s^{-q}(n\log n)^{2/\omega_0} + N^{D\beta_2 - 1}N_s\right) \\
\max_{1\leq i\leq n}\|\eta_i - m\|_\infty &= \mathcal{O}_p\left((n\log n)^{2/\omega_0}\right) \\
\|\overline{m} - m\|_\infty &= \mathcal{O}_p\left(n^{-1/2}\right).
\end{aligned}$$

Hence, by equations (S3.3) and (S3.6)

$$\begin{aligned}
& \|(|\overline{m}| + |\widehat{m}_p|)\,|\widehat{m}_p - \overline{m}|\|_\infty + \max_{1\le i\le n} \|(|\eta_i| + |\widehat{\eta}_i|)\,|\widehat{\eta}_i - \eta_i|\|_\infty \\
= \;& \mathcal{O}_p\left(n^{-1/2} + N_s^{-q}\left(n\log n\right)^{4/\omega_0} + \left(n\log n\right)^{2/\omega_0} N^{D\beta_2-1} N_s\right) \\
= \;& \scriptstyle{\mathcal{O}}_p\left(n^{-\varrho}\right)
\end{aligned}$$

for some $\varrho > 0$. Similar to the proof of Proposition 1 in Cao et al. (2016), one can show that

$$\left\|\widehat{G} - G\right\|_\infty = \scriptstyle{\mathcal{O}}_p(n^{-\varrho}).$$

The proof is complete. □

Proof of Theorem 5.

Recalling the definition of $\Xi(\cdot)$ and $\widehat{\Xi}_{K_n}(\cdot)$, and defining a transitional process $\Xi_{K_n}(\cdot)$,

$$\begin{aligned}
\Xi(\cdot) &= \sum_{k=1}^{\infty} Z_k \phi_k(\cdot) / G^{1/2}(\cdot,\cdot), \\
\widehat{\Xi}_{K_n}(\cdot) &= \sum_{k=1}^{\kappa_n} Z_k \widehat{\phi}_k(\cdot) / \widehat{G}_p^{1/2}(\cdot,\cdot), \\
\Xi_{K_n}(\cdot) &= \sum_{k=1}^{K_n} Z_k \phi_k(\cdot) / G_{K_n}^{1/2}(\cdot,\cdot),
\end{aligned}$$

where Z_k are i.i.d. $N(0,1)$ variables, generated independently from $\mathbf{Y}_i, 1 \leq i \leq n$, partial sum approximate covariance function $G_{K_n}(\boldsymbol{x}, \boldsymbol{x}') \equiv \sum_{k=1}^{K_n} \phi_k(\boldsymbol{x}) \phi_k(\boldsymbol{x}')$.

Noticed that $\Xi(\cdot)$ is the Gaussian process satisfies the assumptions in Theorem 1 in Yang (2025a) and Theorem 1 in Yang (2025b). Following Theorem 3, Theorem 1 in Yang (2025a) and Theorem 1 in Yang (2025b) and Pólya's Theorem, one has that as $n \to \infty$,

$$\sup_{z \in \mathbb{R}} \left| \mathbb{P}\left(\sup_{\boldsymbol{x} \in [0,1]^D} \sqrt{n} \left| \frac{\overline{m}(\boldsymbol{x}) - m(\boldsymbol{x})}{G(\boldsymbol{x}, \boldsymbol{x})^{1/2}} \right| \leq z \right) - \mathbb{P}\left(\|\Xi\|_\infty \leq z \right) \right| = \scriptstyle\mathcal{O}(1). \tag{S3.10}$$

Since $K_n = \prod_{d=1}^{D} (N_{s_d} + p) \to \infty$ as $N \to \infty$, Assumption (A6) ensures that as $N \to \infty$, $\|G_{K_n} - G\|_\infty = \mathcal{O}\left(n^{-\vartheta}\right)$ and $\inf_{\boldsymbol{x} \in [0,1]^D} G_{K_n}(\boldsymbol{x}, \boldsymbol{x}) \geq c_G/2$ for large enough N. Note that according to Theorem 4, as $N \to \infty$,

$$\left\| \widehat{G}_p - G \right\|_\infty = \scriptstyle\mathcal{O}_p\left(n^{-\varrho}\right).$$

for some $\varrho > 0$. Hence, as $N \to \infty$,

$$\left\| \widehat{G}_p - G_{K_n} \right\|_\infty = \scriptstyle\mathcal{O}_p\left(n^{-\varrho} + n^{-\vartheta}\right),$$

and for large enough N, $\inf_{\boldsymbol{x} \in [0,1]^D} \widehat{G}_p(\boldsymbol{x}, \boldsymbol{x}) \geq c_G/2$ with probability as close to 1 as specified. Thus as $N \to \infty$,

$$\sup_{\boldsymbol{x},\boldsymbol{x}'\in[0,1]^D}\left|\frac{\widehat{G}_p(\boldsymbol{x},\boldsymbol{x}')}{\widehat{G}_p(\boldsymbol{x},\boldsymbol{x})^{1/2}\,\widehat{G}_p(\boldsymbol{x}',\boldsymbol{x}')^{1/2}}-\frac{G_{K_n}(\boldsymbol{x},\boldsymbol{x}')}{G_{K_n}(\boldsymbol{x},\boldsymbol{x})^{1/2}\,G_{K_n}(\boldsymbol{x}',\boldsymbol{x}')^{1/2}}\right|$$
$$= \mathcal{O}_p\left(n^{-\varrho}+n^{-\vartheta}\right). \qquad \text{(S3.11)}$$

Therefore, combined with Theorem 2, Theorem 1 in Yang (2025a) and Theorem 1 in Yang (2025b), one obtains that

$$\sup_{z\in\mathbb{R}}\left|\mathbb{P}\left(\sup_{\boldsymbol{x}\in[0,1]^D}\sqrt{n}\left|\frac{\widehat{m}_p(\boldsymbol{x})-m(\boldsymbol{x})}{\widehat{G}_p^{1/2}(\boldsymbol{x},\boldsymbol{x})}\right|\le z\right)-\mathbb{P}\left(\sup_{\boldsymbol{x}\in[0,1]^D}\sqrt{n}\left|\frac{\overline{m}(\boldsymbol{x})-m(\boldsymbol{x})}{G(\boldsymbol{x},\boldsymbol{x})^{1/2}}\right|\le z\right)\right|=\mathcal{O}(1).$$

For large enough N

$$\mathbb{E}\left|\|\Xi\|_\infty-\|\Xi_{K_n}\|_\infty\right|\le\mathbb{E}\left\|\Xi-\Xi_{K_n}\right\|_\infty\le c_G^{-1}\sum_{k=K_n+1}^{\infty}\mathbb{E}|Z_k|\,\|\phi_k\|_\infty$$
$$+c_G^{-1}\left(c_G/2\right)^{-1}\left\|G_{K_n}-G\right\|_\infty\sum_{k=1}^{K_n}\mathbb{E}|Z_k|\,\|\phi_k\|_\infty=\mathcal{O}\left(n^{-\vartheta}\right).$$

By Assumption (A4) on $\|\phi_k\|_{0,\mu}$, for large enough N

$$\left|\|\Xi_{K_n}\|_\infty-\max_{\substack{1\le l_d\le n_d\\1\le d\le D}}\left|\Xi_{K_n}(\boldsymbol{t}_{l_1,\ldots,l_D})\right|\right|$$
$$\le\max_{\substack{1\le l_d\le n_d\\1\le d\le D}}\left|\sup_{\|\boldsymbol{t}\|_\infty\le\left(\min_{1\le d\le D}n_d\right)^{-1}}\left|\Xi_{K_n}(\boldsymbol{t}_{l_1,\ldots,l_D})-\Xi_{K_n}(\boldsymbol{t}_{l_1,\ldots,l_D}+\boldsymbol{t})\right|\right|$$

$$\leq C\left(\min_{1\leq d\leq D} n_d\right)^{-\mu}\sum_{k=1}^{K_n}|Z_k|\left(\|\phi_k\|_\infty+\|\phi_k\|_{0,\mu}\right)=\mathcal{O}_p\left(n^{-\mu}\right).$$

Combining the above, one obtains that

$$\left|\|\Xi\|_\infty-\max_{\substack{1\leq l_d\leq n_d\\1\leq d\leq D}}\left|\Xi_{K_n}(\boldsymbol{t}_{l_1,\ldots,l_D})\right|\right|=\mathcal{O}_p\left(n^{-\mu}+n^{-\vartheta}\right). \tag{S3.12}$$

To examine the distribution of $\max_{\substack{1\leq l_d\leq n_d\\1\leq d\leq D}}|\Xi_{K_n}(\boldsymbol{t}_{l_1,\ldots,l_D})|$, recall that the $(1-\alpha)$-th quantile $\widehat{Q}_{1-\alpha}$ of $\max_{\substack{1\leq l_d\leq n_d\\1\leq d\leq D}}\left|\widehat{\Xi}_{K_n}(\boldsymbol{t}_{l_1,\ldots,l_D})\right|$ exists and is unique by Theorem 1 in Yang (2025a) and Theorem 1 in Yang (2025b). Conditional on $\mathbf{Y}_i, 1\leq i\leq n$, the centered $\tilde{T}_D$-dimensional Gaussian vectors $\left\{\widehat{\Xi}_{K_n}(\boldsymbol{t}_{l_1,\ldots,l_D})\right\}_{\substack{1\leq l_d\leq n_d,\\1\leq d\leq D}}$ and $\left\{\Xi_{K_n}\left(\boldsymbol{t}_{l_1,\ldots,l_D}\right)\right\}_{1\leq l_d\leq n_d,1\leq d\leq D}$ have covariance matrices,

$$\begin{aligned}\Sigma_1&=\left\{\frac{\widehat{G}_p\left(\boldsymbol{t}_{l_1,\ldots,l_D},\boldsymbol{t}_{l'_1,\ldots,l'_D}\right)}{\widehat{G}_p\left(\boldsymbol{t}_{l_1,\ldots,l_D},\boldsymbol{t}_{l_1,\ldots,l_D}\right)^{1/2}\widehat{G}_p\left(\boldsymbol{t}_{l'_1,\ldots,l'_D},\boldsymbol{t}_{l'_1,\ldots,l'_D}\right)^{1/2}}\right\},\\ \Sigma_2&=\left\{\frac{G_{K_n}\left(\boldsymbol{t}_{l_1,\ldots,l_D},\boldsymbol{t}_{l'_1,\ldots,l'_D}\right)}{G_{K_n}\left(\boldsymbol{t}_{l_1,\ldots,l_D},\boldsymbol{t}_{l_1,\ldots,l_D}\right)^{1/2}G_{K_n}\left(\boldsymbol{t}_{l'_1,\ldots,l'_D},\boldsymbol{t}_{l'_1,\ldots,l'_D}\right)^{1/2}}\right\}.\end{aligned}$$

Applying (S3.2) in Lemma S.6, one has that

$$\begin{aligned}&\sup_{z\in\mathbb{R}}\left|\mathbb{P}\left(\max_{\substack{1\leq l_d\leq n_d\\1\leq d\leq D}}\left|\Xi_{K_n}(\boldsymbol{t}_{l_1,\ldots,l_D})\right|\leq z\right)\right.\\&\left.-\mathbb{P}\left(\max_{\substack{1\leq l_d\leq n_d\\1\leq d\leq D}}\left|\widehat{\Xi}_{K_n}(\boldsymbol{t}_{l_1,\ldots,l_D})\right|\leq z\,|\{\mathbf{Y}_i\}_{i=1}^n\right)\right|\\ \leq\;&\mathcal{C}\Upsilon^{1/3}\max\left\{1,\log^{2/3}\left(\widetilde{T}_D\Upsilon^{-1}\right)\right\},\end{aligned} \tag{S3.13}$$

in which

$$\Upsilon = \|\mathrm{Vec}(\Sigma_2 - \Sigma_1)\|_\infty = \mathcal{O}_p\left(n^{-\varrho} + n^{-\vartheta}\right), \tag{S3.14}$$

according to (S3.11).

$$\sup_{z\in\mathbb{R}}\left|\mathbb{P}\left(\|\Xi\|_\infty \le z\right) - \mathbb{P}\left(\max_{\substack{1\le l_d\le n_d\\1\le d\le D}}\left|\widehat{\Xi}_{K_n}(\boldsymbol{t}_{l_1,\ldots,l_D})\right| \le z\left|\{\mathbf{Y}_i\}_{i=1}^n\right.\right)\right| = \mathcal{O}_p(1),$$

thus

$$\mathbb{P}\left(\|\Xi\|_\infty \le \widehat{Q}_{1-\alpha}\right) - \mathbb{P}\left(\max_{\substack{1\le l_d\le n_d\\1\le d\le D}}\left|\widehat{\Xi}_{K_n}(\boldsymbol{t}_{l_1,\ldots,l_D})\right| \le \widehat{Q}_{1-\alpha}\left|\{\mathbf{Y}_i\}_{i=1}^n\right.\right) = \mathcal{O}_p(1),$$

or

$$\mathbb{P}\left(\|\Xi\|_\infty \le \widehat{Q}_{1-\alpha}\right) - (1-\alpha) = \mathcal{O}_p(1),$$

which implies that

$$\left|\widehat{Q}_{1-\alpha} - Q_{1-\alpha}\right| \to_p 0,$$

$$\sup_{\boldsymbol{x}\in[0,1]^D}\left|\widehat{G}_p\left(\boldsymbol{x},\boldsymbol{x}\right)^{1/2}\widehat{Q}_{1-\alpha} - G\left(\boldsymbol{x},\boldsymbol{x}\right)^{1/2}Q_{1-\alpha}\right| \to_p 0.$$

Hence, (S3.10), (S3.12), (S3.13) and (S3.14) imply that

$$\sup_{z\in\mathbb{R}}\left|\mathbb{P}\left(\sup_{\boldsymbol{x}\in[0,1]^D}\sqrt{n}\left|\frac{\widehat{m}_p\left(\boldsymbol{x}\right) - m\left(\boldsymbol{x}\right)}{\widehat{G}_p^{1/2}\left(\boldsymbol{x},\boldsymbol{x}\right)}\right| \le z\right) - \mathbb{P}\left(\max_{\substack{1\le l_d\le n_d\\1\le d\le D}}\left|\widehat{\Xi}_{K_n}(\boldsymbol{t}_{l_1,\ldots,l_D})\right| \le z\left|\{\mathbf{Y}_i\}_{i=1}^n\right.\right)\right| = \mathcal{O}_p(1),$$

The proof is completed. □

References

Cao, G., Wang, L., Li, Y., and Yang, L. (2016). Oracle-efficient confidence envelopes for covariance functions in dense functional data. *Statist. Sinica* **26** 359–383.

Cao, G., Yang, L., and Todem, D. (2012). Simultaneous inference for the mean function based on dense functional data. *J. Nonparametr. Stat.* **24** 359–377.

Chernozhukov, V., Chetverikov, D., and Kato, K. (2015). Comparison and anti-concentration bounds for maxima of Gaussian random vectors. *Probab. Theory Relat. Fields* **162** 47–70.

De Boor, C. (2001). *A Practical Guide to Splines*. Springer, New York.

Götze, F., and Zaitsev, A. Y. (2010). Rates of approximation in the multidimensional invariance principle for sums of i.i.d. random vectors with finite moments. *J. Math. Sci.* **167** 495–500.

Klenke, A. (2014). *Probability Theory*, 2nd edn. Springer, London.

Schumaker, L. (1981). *Spline Functions: Basic Theory*. Wiley.

Wang, J., Cao, G., Wang, L., and Yang, L. (2020). Simultaneous confidence band for stationary covariance function of dense functional data. *J. Multivariate Anal.* **176** 104584.

Yang, L. (2025a). Strict monotonicity of stochastic process extreme distributions. *Statist. Probab. Lett.* **217** 110292.

Yang, L. (2025b). Continuity of Gaussian extreme distributions. *Statist. Probab. Lett.* **215** 110274.

Department of Statistics and Data Science, Tsinghua University, Beijing 100084, China.

E-mail: hqr20@mails.tsinghua.edu.cn

Department of Statistics and Data Science, Tsinghua University, Beijing 100084, China.

E-mail: yanglijian@tsinghua.edu.cn