Projection-based diagnostic test for generalized functional regression models

Guizhen Li[1], Mengying You[1], Ling Zhou[1], Hua Liang[2] and Huazhen Lin[1]

[1] *Southwestern University of Finance and Economics, Chengdu, China*

[2] *George Washington University, Washington, D.C., USA*

Supplementary Material

In this document, we present notation (in Section S.1), the proposed PD test for the functional error (in Section S.2), the proofs for the main theorems (in Section S.3), several corollaries and their proofs (in Section S.4), and the proofs of several lemmas (in Section S.5) that are used to prove the main results. We also present various auxiliary numerical results.

S.1 Notation

Let $q = Kd+1$. Denote $\boldsymbol{\varpi} = (\varpi_i)_{i=1}^n := \Big(\dot{\ell}(m(Y_i, \boldsymbol{\eta}_i))\Big)_{i=1}^n$, $\boldsymbol{V} = \mathrm{Diag}\{\varpi_1^2, \cdots, \varpi_n^2\}$, $\boldsymbol{D} = \{\boldsymbol{D}_i\}_{i=1}^n$ and $\tilde{\boldsymbol{D}} = \{(\ddot{\ell}(m(Y_i, \boldsymbol{\eta}_i)))^{1/2}\boldsymbol{D}_i\}_{i=1}^n$ are $n \times q$-dimensional matrices, $\boldsymbol{\Gamma} = \lim_{n\to\infty} \frac{\boldsymbol{D}^\top \boldsymbol{V}\boldsymbol{D}}{n} := \{\Gamma_{k,l}\}_{1\le k,l\le q}$, $\tilde{\boldsymbol{\Gamma}} = \lim_{n\to\infty}\left(\frac{\tilde{\boldsymbol{D}}^\top \tilde{\boldsymbol{D}}}{n}\right) = (\tilde{\Gamma}_{k,l})_{1\le k,l\le q}$, $\Xi = \tilde{\boldsymbol{\Gamma}}^{-1} = (\zeta_{j,k})_{1\le j,k\le q}$, $\sum_{k=1}^q \zeta_{j,k}^{(1/2)}\zeta_{k,l}^{(1/2)} = \zeta_{j,l}$.

For the local alternatives, let $Y_{i,n} = g_l\left(a, \langle X_{i,1}, \beta_1\rangle, \cdots, \langle X_{i,d}, \beta_d\rangle, \epsilon_i, n^{-1/2}\mathcal{F}(\boldsymbol{X}_i)\right)$, $Y_i = g_l\left(a, \langle X_{i1}, \beta_1\rangle, \cdots, \langle X_{id}, \beta_d\rangle, \epsilon_i, 0\right) := g(a, \langle X_{i,1}, \beta_1\rangle, \cdots, \langle X_{i,d}, \beta_d\rangle, \epsilon_i)$,

$g_l(\cdot)$ $\varepsilon(\boldsymbol{X}_i;\boldsymbol{\beta}) = m(Y_i, a, \langle X_{i,1},$

$\beta_1\rangle, \cdots, \langle X_{i,d}, \beta_d\rangle)$, and $\varepsilon(\boldsymbol{X}_i;\check{\boldsymbol{\beta}}) = m(Y_{i,n}, \check{a}, (\hat{\boldsymbol{\xi}}_{i1})^\top \check{\boldsymbol{b}}_1, \cdots, (\hat{\boldsymbol{\xi}}_{id})^\top \check{\boldsymbol{b}}_d)$. Denote $m_y(Y,\boldsymbol{\eta}) = \partial m(Y,\boldsymbol{\eta})/\partial Y$, $m_y(\cdot,\cdot)$ and $\dot{g}_l(\boldsymbol{\eta}_i, \epsilon_i, b) = \partial g_l(\boldsymbol{\eta}_i, \epsilon_i, b)/\partial b$. Let $\boldsymbol{\upsilon}_n = n^{-1}\Xi\sum_{i=1}^n \ddot{\ell}(m(Y_i,\boldsymbol{\eta}_i))m_y(Y_i,\boldsymbol{\eta}_i)\dot{g}_l$

$(\boldsymbol{\eta}_i, \epsilon_i, 0)\mathcal{F}(\boldsymbol{X}_i)\boldsymbol{D}_i$, and $\boldsymbol{\upsilon} = \mathbb{E}(\boldsymbol{\upsilon}_n)$ $\boldsymbol{\upsilon}$.

Let $\varsigma_{l,F}(\boldsymbol{\alpha}, u) = \mathbb{E}\left[\Delta(\boldsymbol{X}_j;\boldsymbol{\alpha},u)D_{j,l}\right]$, and

$$\sigma_F^2(\boldsymbol{\alpha}_{1,\bullet}, u_1, \boldsymbol{\alpha}_{2,\bullet}, u_2) = \sum_{k_1,k_2,l_1,l_2}^{q} \zeta_{l_1,k_1}\zeta_{l_2,k_2}\Gamma_{k_1,k_2}\varsigma_{l_1,F}(\boldsymbol{\alpha}_{1,\bullet}, u_1)\varsigma_{l_2,F}(\boldsymbol{\alpha}_{2,\bullet}, u_2),$$

$$\sigma_{c,F}(\boldsymbol{\alpha}_{1,\bullet}, u_1, v_1, \boldsymbol{\alpha}_{2,\bullet}, u_2) = \sum_{k,l=1}^{q} \zeta_{l,k}\varsigma_{l,F}(\boldsymbol{\alpha}_{2,\bullet}, u_2)\mathbb{E}\left[I(\epsilon_i \le v_1)\varpi_i D_{i,k}\Delta(\boldsymbol{X}_i;\boldsymbol{\alpha}_{1,\bullet}, u_1)\right].$$

S.2 PD test for the function-valued error

Under model (2.5), let $q_f = K^2 d + K$. Denote $\epsilon_i^f = \epsilon_i(\cdot)$, and $\boldsymbol{\varpi}_f = \left(\dot{\ell}(m(Y_i(t), \boldsymbol{\eta}_i^f(t)))\right)_{i=1}^n$ $\boldsymbol{\varpi}_f$, and $\boldsymbol{D}_f = \{\boldsymbol{D}_i^f\}_{i=1}^n$ is a $n \times q_f$-dimensional matrix. $\langle \boldsymbol{D}_f, \boldsymbol{\varpi}_f\rangle = \{\sum_{i=1}^n \langle D_{i,k}^f, \varpi_i^f\rangle\}_{k=1}^{q_f}$, $\boldsymbol{D}_f$ $\boldsymbol{\Gamma}_f := \{\Gamma_{k,l}^f\}_{1\le k,l\le q_f} := \{\lim_{n\to\infty} \frac{\sum_{i=1}^n \langle D_{i,k}^f, \varpi_i^f\rangle \langle D_{i,l}^f, \varpi_i^f\rangle}{n}\}_{1\le k,l\le q_f}$ is a $q_f \times q_f$-dimensional matrix.

Let $\tilde{\boldsymbol{\Gamma}}_n^f = \sum_{i=1}^n \int \ddot{\ell}(m(Y_i(t), \boldsymbol{\eta}_i^f(t)))\boldsymbol{D}_i^f(t)(\boldsymbol{D}_i^f(t))^\top dt$ be a $q_f \times q_f$-dimensional matrix, $\tilde{\boldsymbol{\Gamma}}_f\tilde{\boldsymbol{\Gamma}}_f = \lim_{n\to\infty}\left(\frac{\tilde{\boldsymbol{\Gamma}}_n^f}{n}\right) = (\tilde{\Gamma}_{k,l}^f)_{0\le k,l\le q_f}$, $\Xi_f = \tilde{\boldsymbol{\Gamma}}_f^{-1} = (\zeta_{j,k}^f)_{0\le j,k\le q_f}$, $\sum_{k=1}^{q_f}(\zeta_{j,k}^f)^{1/2}(\zeta_{k,l}^f)^{1/2} = \zeta_{j,l}^f$. $\zeta_{j,l}^f$ Let $\boldsymbol{\upsilon}_f = \mathbb{E}(\boldsymbol{\upsilon}_n^f)$ and

$$\boldsymbol{\upsilon}_n^f = n^{-1}\Xi_f\sum_{i=1}^n \int \ddot{\ell}(m(Y_i(t), \boldsymbol{\eta}_i^f(t)))m_y(Y_i(t), \boldsymbol{\eta}_i^f(t))\dot{g}_l(\boldsymbol{\eta}_i^f(t), \epsilon_i(t), 0)\mathcal{F}(\boldsymbol{X}_i)\boldsymbol{D}_i^f(t)dt.$$

For the local alternatives, let $Y_{i,n}(t) = g_l\left(a(t), \langle X_{i,1}, \beta_1\rangle, \cdots, \langle X_{i,d}, \beta_d\rangle, \epsilon_i(t), n^{-1/2}\mathcal{F}(\boldsymbol{X}_i)\right)$,

$$Y_i(t) = g_l\left(a(t), \langle X_{i,1}, \beta_1\rangle, \cdots, \langle X_{i,d}, \beta_d\rangle, \epsilon_i(t), 0\right) := g\left(a(t), \langle X_{i,1}, \beta_1\rangle, \cdots, \langle X_{i,d}, \beta_d\rangle, \epsilon_i(t)\right).$$

Let $\varepsilon^f(\boldsymbol{X}_i;\boldsymbol{\beta})(\cdot) = m\big(Y_i(\cdot), a(\cdot), \langle X_{i,1}, \beta_1(\cdot)\rangle, \cdots, \langle X_{i,d}, \beta_d(\cdot)\rangle\big)$, and

$$\varepsilon^f(\boldsymbol{X}_i;\check{\boldsymbol{\beta}})(\cdot) = m(Y_{i,n}(\cdot), (\check{\boldsymbol{a}}_f)^\top \hat{\boldsymbol{\phi}}(\cdot), (\hat{\boldsymbol{\xi}}_{i1} \otimes \hat{\boldsymbol{\phi}}(\cdot))^\top \check{\boldsymbol{b}}_1^f, \cdots, (\hat{\boldsymbol{\xi}}_{id} \otimes \hat{\boldsymbol{\phi}}(\cdot))^\top \check{\boldsymbol{b}}_d^f),$$

where $\otimes$ indicates the Kronecker product.

For a function-valued residual $\varepsilon^f(\boldsymbol{X}_i;\boldsymbol{\beta})(t)$, to construct a PD test, we consider to induce the projection of $\varepsilon^f(\boldsymbol{X}_i;\boldsymbol{\beta})(t)$ along the direction $\boldsymbol{\gamma} \in \mathcal{R}^p$, then similar procedure follows as that for a scalar residual. In particular, define

$$M_{n,\mathtt{F}}^f(\boldsymbol{\beta};\boldsymbol{\alpha},\boldsymbol{\gamma},u,v) = n^{-1/2}\sum_{i=1}^n \left[I\left(\langle \varepsilon^f(\boldsymbol{X}_i;\boldsymbol{\beta}), \boldsymbol{\gamma}\rangle_m \le v\right) - F_{n,\langle \epsilon^f, \boldsymbol{\gamma}\rangle_m}(v)\right] M_{n,\mathtt{F}}^f \\ \times I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}\rangle_m \le u\right).$$

To avoid any subjective selection of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$, which may cause the test to be inconsistent (Escanciano 2006), we consider integrating out all possible $\boldsymbol{\alpha}$'s and $\boldsymbol{\gamma}$'s. Particularly, we consider the following PD test statistic:

$$T_{n,\mathtt{F}}^f(\boldsymbol{\beta}) = \int_{\mathbb{S}^p}\int_{\mathbb{S}^{pd}}\iint_{\mathcal{R}^2}\left(M_{n,\mathtt{F}}^f(\boldsymbol{\beta};\boldsymbol{\alpha},\boldsymbol{\gamma},u,v)\right)^2 F_{n,\langle \boldsymbol{X},\boldsymbol{\alpha}\rangle_m}(du) F_{n,\langle \epsilon^f,\boldsymbol{\gamma}\rangle_m}(dv)\, d\boldsymbol{\alpha}\, d\boldsymbol{\gamma}. T_{n,\mathtt{F}}^f$$

To estimate $\boldsymbol{\beta}$, define the loss functions $\ell(\varepsilon^f(\boldsymbol{X};\boldsymbol{\beta}))$ as $\int \ell(\varepsilon^f(\boldsymbol{X};\boldsymbol{\beta})(t))dt$, where

$$\varepsilon^f(\boldsymbol{X};\boldsymbol{\beta})(t) := m\Bigg(Y(t), a(t), \langle X_1, \beta_1(,t)\rangle, \cdots, \langle X_d, \beta_d(,t)\rangle\Bigg),$$

is the residual between $Y(t)$ and the predictors $\boldsymbol{X}$ based on model (2.5). For model (2.5), assume $\beta_j(s,t) = \sum_{k,l=1} b_{j,kl}^f \phi_k(s)\phi_l(t)$, $j = 1, \cdots, d$ and

$a(t) = \sum_{j=1}^{K} a_j^f \phi_j(t)$. Denote $\boldsymbol{b}_j^f = \{b_{j,kl}^f\}_{k,l=1}^K$, for $j = 1, \cdots, d$, $\boldsymbol{a}_f = \{a_j^f\}_{j=1}^K$, and $\boldsymbol{\theta}_f = ((\boldsymbol{a}_f)^\top, (\boldsymbol{b}_1^f)^\top, \cdots, (\boldsymbol{b}_d^f)^\top)^\top$. Then, the estimating equation takes the form:

$$\boldsymbol{U}_f(\boldsymbol{\theta}_f) := \sum_{i=1}^{n} \int \dot{\ell}\left(m(Y_i(t), \hat{\boldsymbol{\eta}}_i^f(t))\right) \hat{\boldsymbol{D}}_i^f(t) dt = 0,$$

where $\hat{\boldsymbol{\eta}}_i^f(t) = \left((\boldsymbol{a}_f)^\top \hat{\boldsymbol{\phi}}(t), (\boldsymbol{b}_1^f)^\top \left(\hat{\boldsymbol{\xi}}_1^i \otimes \hat{\boldsymbol{\phi}}(t)\right), \cdots, (\boldsymbol{b}_d^f)^\top \left(\hat{\boldsymbol{\xi}}_d^i \otimes \hat{\boldsymbol{\phi}}(t)\right)\right)$, $\hat{\boldsymbol{\phi}}(t) = \{\hat{\phi}_j(t)\}_{j=1}^K$, $\hat{\boldsymbol{D}}_i^f(t) = \left(m_0(Y_i(t), \hat{\boldsymbol{\eta}}_i^f(t))\hat{\boldsymbol{\phi}}^\top(t), m_1(Y_i(t), \hat{\boldsymbol{\eta}}_i^f(t)) \left(\hat{\boldsymbol{\xi}}_1^i \otimes \hat{\boldsymbol{\phi}}(t)\right)^\top, \cdots m_d(Y_i(t), \hat{\boldsymbol{\eta}}_i^f(t)) \left(\hat{\boldsymbol{\xi}}_d^i \otimes \hat{\boldsymbol{\phi}}(t)\right)^\top\right)^\top$. $\hat{\boldsymbol{\theta}}_f$ is denoted as the solution to $\boldsymbol{U}_f(\boldsymbol{\theta}_f) = 0$. Thus, we obtain $\hat{\beta}_j(s,t) = \sum_{k,l=1}^K \hat{b}_{j,kl}^f \hat{\phi}_k(s)\hat{\phi}_l(t)$, $j = 1, \cdots, d$. Further, we assume $p \le K$. Since for $p \ge K$, the projection $\langle X_i, \boldsymbol{\gamma}\rangle_m = \sum_{k=1}^K \xi_{i,k}\gamma_k$, which is equal to the projection of X_i along the direction $\boldsymbol{\gamma} \in \mathcal{R}^K$.

Next we will show the details on calculating the test statistics. Let $B_{ijl}(\boldsymbol{\beta}) = \int_{\mathbb{S}^p} I\big(\langle \varepsilon^f(\boldsymbol{X}_i;\boldsymbol{\beta}), \boldsymbol{\gamma}\rangle_m \le \langle \varepsilon^f(\boldsymbol{X}_l;\boldsymbol{\beta}), \boldsymbol{\gamma}\rangle_m\big) I\big(\langle \varepsilon^f(\boldsymbol{X}_j;\boldsymbol{\beta}), \boldsymbol{\gamma}\rangle_m \le \langle \varepsilon^f(\boldsymbol{X}_l;\boldsymbol{\beta}), \boldsymbol{\gamma}\rangle_m\big) d\boldsymbol{\gamma}$. After some calculations, we can obtain that

$$T_{n,F}^f(\boldsymbol{\beta}) = n^{-3} \sum_{i,j,k,l} A_{ijk} B_{ijl}(\boldsymbol{\beta}) - 2n^{-4} \sum_{i,j,k,l} \sum_{s} A_{ijk} B_{isl}(\boldsymbol{\beta}) + n^{-5} \sum_{i,j,k,l} \sum_{s_1,s_2} A_{ijk} B_{s_1 s_2 l}(\boldsymbol{\beta}).$$

Note that $B_{ijl}(\boldsymbol{\beta})$ involves p-dimensional integrals, we solve the computational problem similarly to the way of obtaining A_{ijk}.

The following conditions are needed to establish asymptotic properties for model (2.5):

(C1f) $\{(Y_i, \{X_{i,j}\}_{j=1}^{d})\}_{i=1}^{n}$ is a sequence of i.i.d. function-valued random vectors with

$$0 < \inf_{0<t<1} \mathbb{E}|Y_i(t)| < \sup_{0<t<1} \mathbb{E}|Y_i(t)| < \infty.$$

(C2f) $\epsilon_i(\cdot)$ is an independent mean zero process and independent of $\{X_{i,j}(\cdot)\}_{j=1}^{d}$, and $X_{i,j}(\cdot) \in L^2[0,1]$, for $j = 1, \cdots, d$. The covariance function of each $X_{i,j}$ is positive definite with a spectral decomposition $\Sigma_j(s,t) = \sum_{k=1}^{\infty} \theta_{j,k}\phi_k(s)\phi_k(t)$, where $\theta_{j,1} > \theta_{j,2} > \cdots$ and $C^{-1}k^{-\alpha} \leq \theta_{j,k} \leq Ck^{-\alpha}$, $\theta_{j,k} - \theta_{j,k+1} \geq C^{-1}k^{-\alpha-1}$, for $\alpha > 1$, $j = 1, \cdots, d, k \geq 1$. Furthermore, the true coefficient function $\beta_j(s,t) = \sum_{k_1,k_2=1}^{\infty} b_{j,k_1k_2}\phi_{k_1}(t)\phi_{k_2}(s)$ with $|b_{j,k_1k_2}| \leq C_1 \min(k_1^{-\kappa}, k_2^{-\kappa})$, for $k_1, k_2 \geq 1$, $\kappa \geq \alpha + 3$.

(C3f) $\mathbb{E}(\langle \varpi_i^f, \boldsymbol{D}_i^f \rangle \mid \boldsymbol{X}_i) = \boldsymbol{0}$, $\sup_{1\leq k\leq q_f} \mathbb{E}(\langle \varpi_i^f, D_{i,k}^f \rangle^2) < c_2 < \infty$, and

$$0 < c_l \leq \inf_{0<t<1, 1\leq i\leq n} \ddot{\ell}(m(Y_i(t), \boldsymbol{\eta}_i^f(t))) \leq \sup_{0<t<1, 1\leq i\leq n} \ddot{\ell}(m(Y_i(t), \boldsymbol{\eta}_i^f(t))) \leq c_u < \infty.$$

(C4f) $m(\cdot)$ has continuous bounded first-order derivatives, and $\|\boldsymbol{D}_f\|_\infty \leq c < \infty$, $\tilde{\boldsymbol{\Gamma}}_f$ is positive definite, and has a bounded maximum eigenvalue.

(C5f) $Kn^{-\frac{1}{2\kappa+\alpha-1}}$ is bounded away from zero and infinity as $n \to \infty$. The following equations hold:

$$\sum_{k_1,k_2,k_3,k_4=1}^{q_f} \mathbb{E}\left(\langle D^f_{i,k_1}, \varpi^f_i\rangle\langle D^f_{i,k_2}, \varpi^f_i\rangle\langle D^f_{i,k_3}, \varpi^f_i\rangle\langle D^f_{i,k_4}, \varpi^f_i\rangle \zeta^f_{k_1,k_2}\zeta^f_{k_3,k_4}\right) = o(n/q_f^2),$$

$$\sum_{k_1,\cdots,k_8=1}^{q_f} \mathbb{E}\left(\langle D^f_{i,k_1}, \varpi^f_i\rangle\langle D^f_{i,k_3}, \varpi^f_i\rangle\langle D^f_{i,k_5}, \varpi^f_i\rangle\langle D^f_{i,k_7}, \varpi^f_i\rangle\right)$$
$$\times \mathbb{E}\left(\langle D^f_{i,k_2}, \varpi^f_i\rangle\langle D^f_{i,k_4}, \varpi^f_i\rangle\langle D^f_{i,k_6}, \varpi^f_i\rangle\langle D^f_{i,k_8}, \varpi^f_i\rangle\right) \zeta^f_{k_1,k_2}\zeta^f_{k_3,k_4}\zeta^f_{k_5,k_6}\zeta^f_{k_7,k_8} = o(n^2q_f^2).$$

Conditions (C1f)–(C4f) are general conditions easy to be satisfied in practice. Particularly, Condition (C1f) is the regularity condition on the upper bound of the expectation of the outcome. Condition (C2f) is about the independence between error and covariates, and smoothness of the covariance functions of covariates $\boldsymbol{X}$ and true coefficient function β_j. Noting that condition (C2f) is stronger than Condition (C2) for the scalar response model in terms of $\kappa \geq \alpha+3$ instead of $\kappa \geq \alpha+2$ required in Condition (C2). This is because the dimension of parameters involved in function-valued response model is larger than that in scaler response model. That is, $q_f \hat{=} K^2d + K$ in comparison to $q \hat{=} Kd + K$. Conditions (C3f) and (C4f) are conditions for the link function g and loss function ℓ, which are expressed through the conditions on $m(\cdot)$ and ℓ since function $m(\cdot)$ is determined by the link function g. For example, the continuous bounded first-order derivatives required for $m(\cdot)$ imply $g(\cdot)$ has continuous bounded first-order derivatives. In addition, the order of the second and fourth moment of estimation equations, which involves $m(\cdot)$ and $\ell(\cdot)$, required by condition (C5f) is also related to the link function g. These conditions are easily to be satisfied by commonly used canonical link functions $g(\cdot)$ which are monotonic with high-order smoothness. Condition (C5f) is to ensure that plug-in $\hat{\boldsymbol{\beta}}$ does not affect the convergence property and the order of local power. With the condition (C5f), the number of the parameters $\boldsymbol{\theta}_f$ satisfies $q_f := K^2 d + K \asymp n^{\frac{2}{2\kappa+\alpha-1}}$, it follows that $\int_0^1 \left(\hat{\beta}_j(s,t) - \beta_j(s,t)\right)^2 dsdt = O_p(n^{-\frac{2\kappa-\alpha-3}{\alpha+2\kappa-1}})$ and $\mathbb{E}\int_0^1 \left(\langle x, \hat{\beta}_j(t)\rangle - \langle x, \beta_j(t)\rangle\right)^2 dt = O(n^{-\frac{2\kappa+\alpha-3}{\alpha+2\kappa-1}})$ for $j = 1, \cdots, d$ and for any fixed function $x(t) = \sum_{k=1}^{\infty} x_k \phi_k(t)$ with $|x_k| \le Ck^{-\alpha/2}$.

Theorem S.1. *Under conditions (C1f) and (C2f), and the null hypothesis (3), if $p = o(n)$, for model (2.5) and any $m \in \mathcal{R}$,*

$$Prob(T^f_{n,\mathrm{F}}(\boldsymbol{\beta}_0) < m) - Prob(T^{0,f}_{\infty,F} < m) \to 0,$$

where $T^{0,f}_{\infty,F} := \int_{\mathbb{S}^{pd}} \int_{\mathbb{S}^p} \iint_{\mathcal{R}^2} \left(M^{0,f}_{\infty,F}(\boldsymbol{\alpha}, \boldsymbol{\gamma}, u, v)\right)^2 F_{\langle \epsilon^f, \boldsymbol{\gamma}\rangle_m}(dv) F_{\langle \boldsymbol{X}, \boldsymbol{\alpha}\rangle_m}(du) d\boldsymbol{\alpha} d\boldsymbol{\gamma}$, *and* $M^{0,f}_{\infty,F}(\cdot,\cdot,\cdot,\cdot)$ *is a Gaussian process with zero mean and covariance function*

$$K\left((\boldsymbol{\alpha}_{1,\bullet}, \boldsymbol{\gamma}_1, u_1, v_1), (\boldsymbol{\alpha}_{2,\bullet}, \boldsymbol{\gamma}_2, u_2, v_2)\right) = \Big\{\mathbb{E}\left[I\left(\langle \epsilon_i^f, \boldsymbol{\gamma}_1\rangle_m \le v_1\right) I\left(\langle \epsilon_i^f, \boldsymbol{\gamma}_2\rangle_m \le v_2\right)\right]$$
$$-F_{\langle \epsilon^f, \boldsymbol{\gamma}_1\rangle_m}(v_1) F_{\langle \epsilon^f, \boldsymbol{\gamma}_2\rangle_m}(v_2)\Big\}\, \mathbb{E}\left(\Delta(\boldsymbol{X}_i; \boldsymbol{\alpha}_{1,\bullet}, u_1)\Delta(\boldsymbol{X}_i; \boldsymbol{\alpha}_{2,\bullet}, u_2)\right).$$

Next we establish the asymptotic distribution of test statistics with estimated parameters. Let $\varsigma^f_{l,F}(\boldsymbol{\alpha},\boldsymbol{\gamma},u) = \mathbb{E}\left[\Delta(\boldsymbol{X}_i;\boldsymbol{\alpha},u)\langle D^f_{i,l},\boldsymbol{\gamma}\rangle_m\right]$, and

$$\sigma^{2f}_F(\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2) = \sum_{k_1,k_2,l_1,l_2}^{q_f} \zeta^f_{j_1,k_1}\zeta^f_{j_2,k_2}\Gamma^f_{k_1,k_2}\varsigma^f_{l_1,F}(\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1)\varsigma^f_{l_2,F}(\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2),$$

$$\sigma^f_{c,F}(\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,v_1,\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2) = \sum_{k,l=1}^{q_f}\zeta^f_{l,k}\varsigma^f_{l,F}(\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2)\mathbb{E}\Big[I(\langle\epsilon^f_i,\boldsymbol{\gamma}_1\rangle_m \le v_1) \\ \langle D^f_{i,k},\varpi^f_i\rangle\Delta(\boldsymbol{X}_i;\boldsymbol{\alpha}_{1,\bullet},u_1)\Big].$$

Theorem S.2. *For model (2.5), if conditions (C1f)–(C5f) hold, and under null hypothesis (3), for any* $m \in \mathcal{R}$,

$$Prob(T^f_{n,\mathbf{F}}(\hat{\boldsymbol{\beta}}) < m) - Prob(T^{1,f}_{\infty,F} < m) \to 0,$$

where $T^{1,f}_{\infty,F} := \int_{\mathbb{S}^p}\int_{\mathbb{S}^{pd}}\iint_{\mathcal{R}^2}\left(M^{1,f}_{\infty,F}(\boldsymbol{\alpha},\boldsymbol{\gamma},u,v)\right)^2 F_{\langle\epsilon^f,\boldsymbol{\gamma}\rangle_m}(dv)F_{\langle\boldsymbol{X},\boldsymbol{\alpha}\rangle_m}(du)d\boldsymbol{\alpha}d\boldsymbol{\gamma}$, *and* $M^{1,f}_{\infty,F} \equiv M^{0,f}_{\infty,F} + M^{e,f}_{\infty,F}$, $M^{e,f}_{\infty,F}(\cdot,\cdot,\cdot,\cdot)$ *is a Gaussian process with zero mean and covariance function* $K_1\left((\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,v_1),(\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2,v_2)\right) =$
$\sigma^{2f}_F(\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2)$
$f_{\langle\epsilon^f,\boldsymbol{\gamma}_1\rangle_m}(v_1)f_{\langle\epsilon^f,\boldsymbol{\gamma}_2\rangle_m}(v_2)$, *and cov* $\left(M^{0,f}_{\infty,F}(\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,v_1), M^{e,f}_{\infty,F}(\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2,v_2)\right) =$
$\sigma^f_{c,F}$
$(\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,v_1,\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2)f_{\langle\epsilon^f,\boldsymbol{\gamma}_2\rangle_m}(v_2)$.

Now we study the asymptotic distribution of $T^f_{n,\mathbf{F}}$ under a sequence of local alternatives converging to null at a parametric rate $n^{-1/2}$. We consider the local alternative

$$H^f_{A,n}: Y_{i,n}(t) \;=\; g_l\left(a(t),\langle X_{i,1},\beta_1\rangle,\cdots,\langle X_{i,d},\beta_d\rangle,\epsilon_i(t),n^{-1/2}\mathcal{F}(\boldsymbol{X}_i)\right), \quad \text{(S.1)}$$

where $g_l\left(a, \langle X_{i,1}, \beta_1\rangle, \cdots, \langle X_{i,d}, \beta_d\rangle, \epsilon_i(t), 0\right) := g\left(a, \langle X_{i,1}, \beta_1\rangle, \cdots, \langle X_{i,d}, \beta_d\rangle, \epsilon_i(t)\right)$.

Theorem S.3. *For model (2.5), under conditions (C1f)–(C5f) and (C6), and local alternative* (S.1)*, For any* $m \in \mathcal{R}$*,*

$$Prob(T^f_{n,\mathrm{F}}(\check{\boldsymbol{\beta}}) < m) - Prob(T^{a,f}_{\infty,F} < m) \to 0,$$

where $T^{a,f}_{\infty,F} := \int_{\mathbb{S}^{pd}} \int_{\mathbb{S}^p} \iint_{\mathcal{R}^2} \left(M^{a,f}_{\infty,F}(\boldsymbol{\alpha}, \boldsymbol{\gamma}, u, v)\right)^2 F_{\langle \epsilon^f, \boldsymbol{\gamma}\rangle_m}(dv) F_{\langle \boldsymbol{X}, \boldsymbol{\alpha}\rangle_m}(du) d\boldsymbol{\alpha} d\boldsymbol{\gamma}$*,*

and $\check{\boldsymbol{\beta}}$ *is the estimate obtained from model (2.5) using data* $\{Y_{i,n}(t), \boldsymbol{X}_i\}_{i=1}^n$*,*

$M^{a,f}_{\infty,F} \equiv M^{1,f}_{\infty,F} - D^{a,f}_F$*, and*

$$D^{a,f}_F(\boldsymbol{\alpha}, \boldsymbol{\gamma}, u, v) = \mathbb{E}\left\{\Delta(\boldsymbol{X}_i; \boldsymbol{\alpha}, u)\left(\left\langle m_y(Y_i(t), \boldsymbol{\eta}^f_i(t))\dot{g}_l(\boldsymbol{\eta}^f_i(t), \epsilon_i(t), 0)\mathcal{F}(\boldsymbol{X}_i), \boldsymbol{\gamma}\right\rangle_m \right.\right.$$
$$\left.\left. - \sum_{j=1}^q \langle D^f_{i,j}, \boldsymbol{\gamma}\rangle_m \upsilon^f_j\right)\right\} f_{\langle \epsilon^f, \boldsymbol{\gamma}\rangle_m}(v).$$

Consider the following local alternative hypothetical model:

$$H^f_{A,n} : Y_{i,n}(t) \;=\; g_l\left(a(t), \langle X_{i,1}, \beta_1\rangle, \cdots, \langle X_{i,d}, \beta_d\rangle, \epsilon_i(t), n^{\nu}\mathcal{F}(\boldsymbol{X}_i)\right). \quad \text{(S.2)}$$

Corollary S.1. *Under the conditions of Theorem S.3 and the alternative* (S.2) *with* $-1/2 < \nu \le 0$*, we have* $Prob(T^f_{n,\mathrm{F}} \to \infty) = 1$ *as* $n \to \infty$*.*

S.3 Proofs

Proof of Proposition 1. Denote $X_{i,j}^{(N)}(t) = \sum_{k=1}^{N} \langle X_{i,j}, \phi_k \rangle \phi_k(t)$. For any p and $\boldsymbol{\gamma} \in \mathcal{R}^p$, let

$$F_{\langle U,\boldsymbol{\gamma}\rangle_m | \boldsymbol{X}^{(N)}}(u) = \text{Prob}\left(\langle U_i, \boldsymbol{\gamma}\rangle_m \leq u \mid X_{i,1}^{(N)}, \cdots, X_{i,d}^{(N)}\right),$$
$$F_{\langle U,\boldsymbol{\gamma}\rangle_m | \langle \boldsymbol{X}^{(N)}, \boldsymbol{\alpha}\rangle_m}(u) = \text{Prob}\left(\langle U_i, \boldsymbol{\gamma}\rangle_m \leq u \mid \langle \boldsymbol{X}^{(N)}, \boldsymbol{\alpha}\rangle_m\right).$$

According to Lemma 2.1 (A) in Lavergne & Patilea (2008) and $U, X_j \in L^2[0,1]$, we obtain that, for any N,

$$F_{\langle U,\boldsymbol{\gamma}\rangle_m | \boldsymbol{X}^{(N)}}(u) \equiv \mathbb{E}\left[I(\langle U_i, \boldsymbol{\gamma}\rangle_m \leq u) \mid \boldsymbol{X}^{(N)}\right] = \mathbb{E}\left(I(\langle U_i, \boldsymbol{\gamma}\rangle_m \leq u)\right) \equiv F_{\langle U,\boldsymbol{\gamma}\rangle_m}(u),$$

holds if and only if

$$\begin{aligned} F_{\langle U,\boldsymbol{\gamma}\rangle_m | \langle \boldsymbol{X}^{(N)}, \boldsymbol{\alpha}\rangle_m}(u) &\equiv \mathbb{E}\left[I(\langle U_i, \boldsymbol{\gamma}\rangle_m \leq u) \mid \langle \boldsymbol{X}^{(N)}, \boldsymbol{\alpha}\rangle_m\right] \\ &= \mathbb{E}\left(I(\langle U_i, \boldsymbol{\gamma}\rangle_m \leq u)\right) \equiv F_{\langle U,\boldsymbol{\gamma}\rangle_m}(u), \quad \forall \boldsymbol{\alpha}_j \in \mathbb{S}^p, p \geq 1. \end{aligned}$$

On the other hand, as $N \to \infty$, $F_{\langle U,\boldsymbol{\gamma}\rangle_m | \boldsymbol{X}^{(N)}}(u) \to F_{\langle U,\boldsymbol{\gamma}\rangle_m | \boldsymbol{X}}(u)$. It then follows from Cramer-Wold theorem that $F_{\langle U,\boldsymbol{\gamma}\rangle_m | \boldsymbol{X}^{(N)}}(u) \to F_{\langle U,\boldsymbol{\gamma}\rangle_m | \boldsymbol{X}}$ holds if and only if $F_{\langle U,\boldsymbol{\gamma}\rangle_m | \langle \boldsymbol{X}^{(N)}, \boldsymbol{\alpha}\rangle_m}(u) \to F_{\langle U,\boldsymbol{\gamma}\rangle_m | \langle \boldsymbol{X}, \boldsymbol{\alpha}\rangle_m}(u), \forall \boldsymbol{\alpha}_j \in \mathbb{S}^p, p \geq 1$. Thus, Proposition 1 follows. □

Proofs of Theorems 1 and S.1. Note that under the null hypothesis (1.3), we have $\varepsilon(\boldsymbol{X}; \boldsymbol{\beta}_0) \equiv \epsilon$ for a scalar response, and $\varepsilon^f(\boldsymbol{X}; \boldsymbol{\beta}_0)(\cdot) \equiv \epsilon^f$ for a function-valued response.

We first prove Theorem 1. It follows from Theorem 1 of Escanciano (2006), and $\mathbb{E}[I(\epsilon_i \le v) - F_\epsilon(v)] = 0$, $\mathbb{E}\left[I(\epsilon_i \le v) - F_\epsilon(v)\right]^2 = F_\epsilon(v)(1 - F_\epsilon(v))$, that

$$M_{n,\mathtt{F}}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},u,v) = n^{-1/2}\sum_{i=1}^{n}\left[I(\epsilon_i \le v) - F_{n,\epsilon}(v)\right] I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}\rangle_m \le u\right)$$
$$= n^{-1/2}\sum_{i=1}^{n}\left[I(\epsilon_i \le v) - F_{\epsilon}(v)\right] I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}\rangle_m \le u\right)$$
$$-n^{-1/2}\left(F_{n,\epsilon}(v) - F_\epsilon(v)\right)\sum_{i=1}^{n} I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}\rangle_m \le u\right) := T_{n,1}(\boldsymbol{\alpha},u,v) - T_{n,2}(\boldsymbol{\alpha},u,v).$$

Note that $\mathbb{E}\left(T_{n,1}(\boldsymbol{\alpha},u,v)\right)^2 = F_\epsilon(v)(1-F_\epsilon(v))F_{\langle \boldsymbol{X},\boldsymbol{\alpha}\rangle_m}(u)$,

$$K_{T_{n,1}}\left((\boldsymbol{\alpha}_{1,\bullet},u_1,v_1),(\boldsymbol{\alpha}_{2,\bullet},u_2,v_2)\right) := \mathbb{E}\left(T_{n,1}(\boldsymbol{\alpha}_{1,\bullet},u_1,v_1)T_{n,1}(\boldsymbol{\alpha}_{2,\bullet},u_2,v_2)\right)$$
$$= \{\mathbb{E}\left[I(\epsilon_i \le v_1) I(\epsilon_i \le v_2)\right]$$
$$-F_\epsilon(v_1)F_\epsilon(v_2)\}\,\mathbb{E}[I(\langle \boldsymbol{X}_i,\boldsymbol{\alpha}_{1,\bullet}\rangle_m \le u_1)I(\langle \boldsymbol{X}_i,\boldsymbol{\alpha}_{2,\bullet}\rangle_m \le u_2)],$$

$$\mathbb{E}\left(T_{n,2}(\boldsymbol{\alpha},u,v)\right)^2 = F_\epsilon(v)(1-F_\epsilon(v))F^2_{\langle \boldsymbol{X},\boldsymbol{\alpha}\rangle_m}(u)\left(1+O(n^{-1})\right),$$

$$K_{T_{n,2}}((\boldsymbol{\alpha}_{1,\bullet},u_1,v_1),(\boldsymbol{\alpha}_{2,\bullet},u_2,v_2)) = \{\mathbb{E}\left[I(\epsilon_i \le v_1) I(\epsilon_i \le v_2)\right] - F_\epsilon(v_1)F_\epsilon(v_2)\}$$
$$\times F_{\langle \boldsymbol{X},\boldsymbol{\alpha}_{1,\bullet}\rangle_m}(u_1)F_{\langle \boldsymbol{X},\boldsymbol{\alpha}_{2,\bullet}\rangle_m}(u_2),$$

$$\mathbb{E}\left\{T_{n,1}(\boldsymbol{\alpha},u,v)T_{n,2}(\boldsymbol{\alpha},u,v)\right\} = F_\epsilon(v)(1-F_\epsilon(v))F^2_{\langle \boldsymbol{X},\boldsymbol{\alpha}\rangle_m}(u)\left(1+O(n^{-1})\right),$$

$$K_{T_{n,1}T_{n,2}}((\boldsymbol{\alpha}_1,u_1,v_1),(\boldsymbol{\alpha}_2,u_2,v_2)) = \{\mathbb{E}\left[I(\epsilon_i \le v_1) I(\epsilon_i \le v_2)\right] - F_\epsilon(v_1)F_\epsilon(v_2)\}$$
$$\times F_{\langle \boldsymbol{X},\boldsymbol{\alpha}_{1,\bullet}\rangle_m}(u_1)F_{\langle \boldsymbol{X},\boldsymbol{\alpha}_{2,\bullet}\rangle_m}(u_2).$$

Thus, $M_{n,\mathtt{F}}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},u,v)$ converges to $M^0_{\infty,F}(\boldsymbol{\alpha},u,v)$. Consequently, we have $M_{n,F}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},u,v) \xrightarrow{d} M^0_{\infty,F}(\boldsymbol{\alpha},u,v)$ for any fixed u, v, and $\boldsymbol{\alpha}_j \in \mathbb{S}^p := \{\boldsymbol{\alpha} \in$

$R^p : \|\boldsymbol{\alpha}\|_p = 1\}$, for $j = 1, \cdots, d$, where $M^0_{\infty,F}$ is a Gaussian process with zero mean and covariance function

$$K\left((\boldsymbol{\alpha}_{1,\bullet}, u_1, v_1), (\boldsymbol{\alpha}_{2,\bullet}, u_2, v_2)\right) = \left\{\mathbb{E}\left[I(\epsilon_i \leq v_1)I(\epsilon_i \leq v_2)\right] - F_\epsilon(v_1)F_\epsilon(v_2)\right\}$$
$$\times\mathbb{E}\left[\left(I(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}_{1,\bullet}\rangle_m \leq u_1) - F_{\langle \boldsymbol{X}, \boldsymbol{\alpha}_{1,\bullet}\rangle_m}(u_1)\right)\left(I(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}_{2,\bullet}\rangle_m \leq u_2) - F_{\langle \boldsymbol{X}, \boldsymbol{\alpha}_{2,\bullet}\rangle_m}(u_2)\right)\right].$$

Furthermore, based on the derivation, we have

$$\left|Prob(M_{n,F}(\boldsymbol{\beta}_0; \boldsymbol{\alpha}, u, v) < m) - Prob(M^0_{\infty,F}(\boldsymbol{\alpha}, u, v) < m)\right| = O(n^{-1/2}).$$

Since the dimension of $\boldsymbol{\alpha}$ is pd, which changes with n. From the central limit theorem for classes of functions changing with n (Van Der Vaart & Wellner 1996, Thm. 2.11.23), as long as

$$P^* F_n^2 = O(1),$$
$$P^* F_n^2\{F_n > \eta\sqrt{n}\} \to 0, \quad \text{for every} \;\; \eta > 0,$$
$$\sup_{\rho((\boldsymbol{\alpha}_{1,\bullet}, u_1, v_1), (\boldsymbol{\alpha}_{2,\bullet}, u_2, v_2)) < \delta_n} P\left(M_{n,F}(\boldsymbol{\beta}_0; \boldsymbol{\alpha}_{1,\bullet}, u_1, v_1) - M_{n,F}(\boldsymbol{\beta}_0; \boldsymbol{\alpha}_{2,\bullet}, u_2, v_2)\right) \to 0, \quad \text{for every} \;\; \delta_n \downarrow 0,$$
$$\int_0^{\delta_n} \sqrt{\log N_{[\,]}(\epsilon\|F_n\|_{P,2}, \mathcal{F}_n, L_2(P))}\, d\epsilon \to 0, \quad \text{for every} \;\; \delta_n \downarrow 0, \tag{S.3}$$

where F_n are envelope functions, then we have the sequence $\{M_{n,F}(\boldsymbol{\beta}_0; \boldsymbol{\alpha}, u, v)\}$ is asymptotically tight in $\ell^\infty(\mathbb{S}^{pd} \times R^2)$ and converges in distribution to a Gaussian process provided point-wised convergence on $\ell^\infty(\mathbb{S}^{pd} \times R^2)$. Since $\mathbb{S}^{pd}$ is a pd-unit ball, combining with the probability measure, the ℓ^∞ norm indicates that

$$
\begin{aligned}
& \left|Prob(T_{n,F}(\boldsymbol{\beta}_0) < m) - Prob(T^0_{\infty,F} < m)\right| \\
:= & \left|Prob\left(\int_{\mathbb{S}^{pd}}\int_{R^2} M_{n,F}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},u,v)F_{n,\epsilon}(dv)F_{n,\langle \boldsymbol{X},\boldsymbol{\alpha}\rangle_m}(du)d\boldsymbol{\alpha} < m\right)\right. \\
& \left. -Prob\left(\int_{\mathbb{S}^{pd}}\int_{R^2} M^0_{\infty,F}(\boldsymbol{\alpha},u,v)F_{\epsilon}(dv)F_{\langle \boldsymbol{X},\boldsymbol{\alpha}\rangle_m}(du)d\boldsymbol{\alpha} < m\right)\right| \\
\to & \ 0.
\end{aligned}
$$

Taking F_n as a constant function, and $\delta_n = n^{-1/2}$, then conditions (S.3) hold from

$$
\log N_{[\,]}(\epsilon\|F_n\|_{P,2}, \mathcal{F}_n, L_2(P)) = pd\log(\frac{1}{\epsilon}),
$$

and $pd = o(n)$. Therefore, as long as $pd = o(n)$, Theorem 1 follows.

Next, we prove Theorem S.1. Recall that $\varepsilon^f(\boldsymbol{X}_i;\boldsymbol{\beta}_0) = \epsilon_i^f$, then we have

$$
\begin{aligned}
M^f_{n,\mathtt{F}}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},\boldsymbol{\gamma},u,v) &= n^{-1/2}\sum_{i=1}^n \left[I\left(\langle \epsilon_i^f,\boldsymbol{\gamma}\rangle_m \le v\right) - F_{n,\langle \epsilon^f,\boldsymbol{\gamma}\rangle_m}(v)\right] I\left(\langle \boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m \le u\right) \\
&= n^{-1/2}\sum_{i=1}^n \left[I\left(\langle \epsilon_i^f,\boldsymbol{\gamma}\rangle_m \le v\right) - F_{\langle \epsilon^f,\boldsymbol{\gamma}\rangle_m}(v)\right] I\left(\langle \boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m \le u\right) \\
&\quad -n^{1/2}\left(F_{n,\langle \epsilon^f,\boldsymbol{\gamma}\rangle_m}(v) - F_{\langle \epsilon^f,\boldsymbol{\gamma}\rangle_m}(v)\right) F_{n,\langle \boldsymbol{X},\boldsymbol{\alpha}\rangle_m}(u) \\
&:= T^f_{n,1}(\boldsymbol{\alpha},\boldsymbol{\gamma},u,v) - T^f_{n,2}(\boldsymbol{\alpha},\boldsymbol{\gamma},u,v),
\end{aligned}
$$

and

$$\mathbb{E}\left(T^f_{n,1}(\boldsymbol{\alpha},\boldsymbol{\gamma},u,v)\right)^2 \to F_{\langle\epsilon^f,\boldsymbol{\gamma}\rangle_m}(v)\left(1-F_{\langle\epsilon^f,\boldsymbol{\gamma}\rangle_m}(v)\right)F_{\langle\boldsymbol{X},\boldsymbol{\alpha}\rangle_m}(u),$$

$$K_{T^f_{n,1}}((\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,v_1),(\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2,v_2)) \to \Big\{\mathbb{E}\left[I\left(\langle\epsilon^f_i,\boldsymbol{\gamma}_1\rangle_m \le v_1\right)I\left(\langle\epsilon^f_i,\boldsymbol{\gamma}_2\rangle_m \le v_2\right)\right]$$
$$-F_{\langle\epsilon^f,\boldsymbol{\gamma}_1\rangle_m}(v_1)F_{\langle\epsilon^f,\boldsymbol{\gamma}_2\rangle_m}(v_2)\Big\}\,\mathbb{E}\left[I\left(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}_{1,\bullet}\rangle_m \le u_1\right)I\left(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}_{2,\bullet}\rangle_m \le u_2\right)\right],$$

$$\mathbb{E}\left(T^f_{n,2}(\boldsymbol{\alpha},\boldsymbol{\gamma},u,v)\right)^2 \to F_{\langle\epsilon^f,\boldsymbol{\gamma}\rangle_m}(v)(1-F_{\langle\epsilon^f,\boldsymbol{\gamma}\rangle_m}(v))F^2_{\langle\boldsymbol{X},\boldsymbol{\alpha}\rangle_m}(u),$$

$$K_{T^f_{n,2}}((\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,v_1),(\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2,v_2)) \to \Big\{\mathbb{E}\left[I\left(\langle\epsilon^f_i,\boldsymbol{\gamma}_1\rangle_m \le v_1\right)I\left(\langle\epsilon^f_i,\boldsymbol{\gamma}_2\rangle_m \le v_2\right)\right]$$
$$-F_{\langle\epsilon^f,\boldsymbol{\gamma}_1\rangle_m}(v_1)F_{\langle\epsilon^f,\boldsymbol{\gamma}_2\rangle_m}(v_2)\Big\}\,F_{\langle\boldsymbol{X},\boldsymbol{\alpha}_{1,\bullet}\rangle_m}(u_1)F_{\langle\boldsymbol{X},\boldsymbol{\alpha}_{2,\bullet}\rangle_m}(u_2),$$

$$\mathbb{E}\left(T^f_{n,1}(\boldsymbol{\alpha},\boldsymbol{\gamma},u,v)T^f_{n,2}(\boldsymbol{\alpha},\boldsymbol{\gamma},u,v)\right) \to F_{\langle\epsilon^f,\boldsymbol{\gamma}\rangle_m}(v)(1-F_{\langle\epsilon^f,\boldsymbol{\gamma}\rangle_m}(v))F^2_{\langle\boldsymbol{X},\boldsymbol{\alpha}\rangle_m}(u),$$

$$K_{T^f_{n,1}T^f_{n,2}}((\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,v_1),(\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2,v_2)) \to \Big\{\mathbb{E}\left[I\left(\langle\epsilon^f_i,\boldsymbol{\gamma}_1\rangle_m \le v_1\right)I\left(\langle\epsilon^f_i,\boldsymbol{\gamma}_2\rangle_m \le v_2\right)\right]$$
$$-F_{\langle\epsilon^f,\boldsymbol{\gamma}_1\rangle_m}(v_1)F_{\langle\epsilon^f,\boldsymbol{\gamma}_2\rangle_m}(v_2)\Big\}\,F_{\langle\boldsymbol{X},\boldsymbol{\alpha}_{1,\bullet}\rangle_m}(u_1)F_{\langle\boldsymbol{X},\boldsymbol{\alpha}_{2,\bullet}\rangle_m}(u_2).$$

Then, Theorem S.1 follows from similar arguments to those in the proof of Theorem 1. In particular, we have $M^f_{n,F}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},\boldsymbol{\gamma},u,v) \xrightarrow{d} M^{0,f}_{\infty,F}(\boldsymbol{\alpha},\boldsymbol{\gamma},u,v)$ for any fixed u, v, and $\boldsymbol{\alpha}_j,\boldsymbol{\gamma} \in \mathbb{S}^p := \{\boldsymbol{\alpha} \in R^p : \|\boldsymbol{\alpha}\|_p = 1\}$, for $j = 1,\cdots,d$, where $M^{0,f}_{\infty,F}$ is a Gaussian process with zero mean and covariance function

$$K\left((\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,v_1),(\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2,v_2)\right)$$
$$=\Big\{\mathbb{E}\left[I(\langle\epsilon^f_i,\boldsymbol{\gamma}_1\rangle_m \le v_1)I(\langle\epsilon^f_i,\boldsymbol{\gamma}_2\rangle_m \le v_2)\right]-F_{\langle\epsilon^f,\boldsymbol{\gamma}_1\rangle_m}(v_1)F_{\langle\epsilon^f,\boldsymbol{\gamma}_2\rangle_m}(v_2)\Big\}$$
$$\times\mathbb{E}\left[\left(I(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}_{1,\bullet}\rangle_m \le u_1)-F_{\langle\boldsymbol{X},\boldsymbol{\alpha}_{1,\bullet}\rangle_m}(u_1)\right)\left(I(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}_{2,\bullet}\rangle_m \le u_2)-F_{\langle\boldsymbol{X},\boldsymbol{\alpha}_{2,\bullet}\rangle_m}(u_2)\right)\right].$$

Furthermore, based on the derivation, we have

$$\left|Prob(M^{f}_{n,F}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},\boldsymbol{\gamma},u,v)<m)-Prob(M^{0,f}_{\infty,F}(\boldsymbol{\alpha},\boldsymbol{\gamma},u,v)<m)\right| = O(n^{-1/2}).$$

Similarly to the proof of Theorem 1, as long as $p(d+1)=o(n)$, Theorem S.1 follows.

□

Proofs of Theorems 2 and S.2. We first prove Theorem 2. Note that

$$\begin{aligned}
&M_{n,\mathtt{F}}(\hat{\boldsymbol{\beta}};\boldsymbol{\alpha},u,v) = n^{-1/2}\sum_{i=1}^{n}\left[I\left(\varepsilon(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}})\le v\right)-\hat{F}_{n,\epsilon}(v)\right] I\left(\langle \boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m \le u\right)\\
&= M_{n,\mathtt{F}}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},u,v) + n^{-1/2}\sum_{i=1}^{n}\left[I\left(\varepsilon(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}})\le v\right)-I(\epsilon_i\le v)\right]\\
&\times\left[I\left(\langle \boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m \le u\right)-n^{-1}\sum_{j=1}^{n} I\left(\langle \boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m \le u\right)\right] := M_{n,\mathtt{F}}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},u,v)+T_n^e.
\end{aligned}$$

It follows from Lemma S.1 that $\varepsilon(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}})=\epsilon_i+h_i$, where $\mathbb{E}(h_i)=0$, and $h=\max_j h_j = O_p(q^{1/2}n^{-1/2})$. Then, we have

$$\begin{aligned}
&\mathbb{E}\left[I(v-h_i<\epsilon_i\le v)\right] = h_i f_\epsilon(v) - 0.5h_i^2\dot{f}_\epsilon(v)+o(h_i^2),\\
&\mathrm{Var}\left[I(v-h_i<\epsilon_i\le v)\right] = h_i f_\epsilon(v) - h_i^2\left(0.5\dot{f}_\epsilon(v)+f_\epsilon^2(v)\right)+o(h_i^2).
\end{aligned}$$

Substituting the above equalities into T_n^e, we have

$$
\begin{aligned}
T_n^e &= -n^{-1/2}\sum_{i=1}^n I(v-h_i<\epsilon_i\le v)\Delta_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u)\\
&= -n^{-1/2}\sum_{i=1}^n \Delta_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u)\left(h_i f_\epsilon(v)-0.5h_i^2\dot{f}_\epsilon(v)+o(h_i^2)\right)\\
&\qquad -n^{-1/2}\sum_{i=1}^n \Delta_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u)\left\{I(v-h_i<\epsilon_i\le v)-\mathbb{E}\left[I(v-h_i<\epsilon_i\le v)\right]\right\}\\
&= -n^{-1/2}\sum_{i=1}^n \Delta_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u)\left(h_i f_\epsilon(v)-0.5h_i^2\dot{f}_\epsilon(v)\right)+O_p(h^{1/2}),\\
(T_n^e)^2 &= n^{-1}\sum_{i,j}\Delta_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u)\Delta_{n,j}(\boldsymbol{X};\boldsymbol{\alpha},u)\left(h_i f_\epsilon(v)-0.5h_i^2\dot{f}_\epsilon(v)\right)\\
&\qquad\qquad \left(h_j f_\epsilon(v)-0.5h_j^2\dot{f}_\epsilon(v)\right)+O_p(h),\\
M_{n,\mathtt{F}}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},u,v)T_n^e &= -n^{-1}\sum_{i,j}\left[I\left(\epsilon_i\le v\right)-F_{n,\epsilon}(v)\right]I(v-h_j<\epsilon_j\le v)\\
&\qquad \times I\left(\langle \boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m\le u\right)\Delta_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u)\\
= -n^{-1}\sum_{i=1}^n &\left(h_i f_\epsilon(v)-0.5h_i^2\dot{f}_\epsilon(v)\right)I\left(\langle \boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m\le u\right)\Delta_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u)\\
&+n^{-2}\sum_{i,j}^n I(v-h_i<\epsilon_i\le v)I(\epsilon_j\le v)I\left(\langle \boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m\le u\right)\Delta_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u)\\
&-n^{-1}\sum_{i\ne j}^n I\left(\epsilon_i\le v\right)I(v-h_j<\epsilon_j\le v)I\left(\langle \boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m\le u\right)\Delta_{n,j}(\boldsymbol{X};\boldsymbol{\alpha},u)\\
&+n^{-2}\sum_{i\ne j}^n\sum_{k=1}^n I(v-h_j<\epsilon_j\le v)I(\epsilon_k\le v)I\left(\langle \boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m\le u\right)\Delta_{n,j}(\boldsymbol{X};\boldsymbol{\alpha},u),
\end{aligned}
$$

where $\Delta_{n,j}(\boldsymbol{X};\boldsymbol{\alpha},u)=I\left(\langle \boldsymbol{X}_j,\boldsymbol{\alpha}\rangle_m\le u\right)-F_{n,\langle \boldsymbol{X},\boldsymbol{\alpha}\rangle_m}(u)$. Note that

$$
n^{-1}\sum_{i,j}^n \Delta_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u)\Delta_{n,j}(\boldsymbol{X};\boldsymbol{\alpha},u)T_{i2}T_{j2}=o_p(1),\quad\text{and}\quad n^{-1}\sum_{i,j} I(\epsilon_i\le v)T_{j2}=o_p(1).
$$

After some calculations, we can obtain that

$$\mathbb{E}(T_n^e) = O(h^{1/2} + n^{1/2}h^2), \quad \mathbb{E}(T_n^e)^2 = \sigma_F^2(\boldsymbol{\alpha}, u, \boldsymbol{\alpha}, u) f_\epsilon^2(v) + O(h),$$

$$\mathbb{E}\left(M_{n,\mathtt{F}}(\boldsymbol{\beta}_0; \boldsymbol{\alpha}, u, v) T_n^e\right) = \sigma_{c,F}(\boldsymbol{\alpha}, u, v, \boldsymbol{\alpha}, u) f_\epsilon(v).$$

where $\sigma_{c,F}(\boldsymbol{\alpha}_{1,\bullet}, u_1, v_1, \boldsymbol{\alpha}_{2,\bullet}, u_2) = \sum_{k,l=1}^q \zeta_{l,k} \varsigma_{l,F}(\boldsymbol{\alpha}_2, u_2) \mathbb{E}\left[I(\epsilon_i \le v_1) \varpi_i D_{i,k} \Delta(\boldsymbol{X}_i; \boldsymbol{\alpha}_1, u_1)\right]$, $\sigma_F^2(\boldsymbol{\alpha}_{1,\bullet}, u_1, \boldsymbol{\alpha}_{2,\bullet}, u_2) = \sum_{k_1,k_2,l_1,l_2=1}^q \zeta_{l_1,k_1} \zeta_{l_2,k_2} \Gamma_{k_1,k_2} \varsigma_{l_1,F}(\boldsymbol{\alpha}_{1,\bullet}, u_1) \varsigma_{l_2,F}(\boldsymbol{\alpha}_{2,\bullet}, u_2)$, $\varsigma_{l,F}(\boldsymbol{\alpha}, u) = \mathbb{E}\left[\Delta(\boldsymbol{X}_j; \boldsymbol{\alpha}, u) D_{j,l}\right]$, $\Delta(\boldsymbol{X}_i; \boldsymbol{\alpha}, u) = I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha} \rangle_m \le u\right) - F_{\langle \boldsymbol{X}, \boldsymbol{\alpha} \rangle_m}(u)$. Consequently, we have $M_{n,F}(\hat{\boldsymbol{\beta}}; \boldsymbol{\alpha}, u, v) \xrightarrow{d} M_{\infty,F}^1(\boldsymbol{\alpha}, u, v)$ for any fixed u, v, and $\boldsymbol{\alpha}_j \in \mathbb{S}^p := \{\boldsymbol{\alpha} \in R^p : \|\boldsymbol{\alpha}\|_p = 1\}$, for $j = 1, \cdots, d$, where $M_{\infty,F}^1 \equiv M_{\infty,F}^0 + M_{\infty,F}^e$, $M_{\infty,F}^e$ is a Gaussian process with zero mean and covariance function

$$K_1\left((\boldsymbol{\alpha}_{1,\bullet}, u_1, v_1), (\boldsymbol{\alpha}_{2,\bullet}, u_2, v_2)\right) = \sigma_F^2(\boldsymbol{\alpha}_{1,\bullet}, u_1, \boldsymbol{\alpha}_{2,\bullet}, u_2) f_\epsilon(v_1) f_\epsilon(v_2),$$

$$\operatorname{cov}\left(M_{\infty,F}^0(\boldsymbol{\alpha}_{1,\bullet}, u_1, v_1), M_{\infty,F}^e(\boldsymbol{\alpha}_{2,\bullet}, u_2, v_2)\right) = \sigma_{c,F}(\boldsymbol{\alpha}_{1,\bullet}, u_1, v_1, \boldsymbol{\alpha}_{2,\bullet}, u_2) f_\epsilon(v_2).$$

Furthermore, based on the derivation, we have

$$\left|Prob(M_{n,F}(\hat{\boldsymbol{\beta}}; \boldsymbol{\alpha}, u, v) < m) - Prob(M_{\infty,F}^1(\boldsymbol{\alpha}, u, v) < m)\right| = O(n^{-1/2} + h^{1/2}).$$

Again, similar to the proof of Theorem 1, as long as $p^2 d^2 q = o(n)$, Theorem 2 follows.

Now we prove Theorem S.2. Note that

$$
\begin{aligned}
&M^f_{n,\mathtt{F}}(\hat{\boldsymbol{\beta}};\boldsymbol{\alpha},\boldsymbol{\gamma},u,v) = n^{-1/2}\sum_{i=1}^n \left[I\left(\langle \varepsilon^f(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}}),\boldsymbol{\gamma}\rangle_m \le v\right) - \hat{F}_{n,\langle \epsilon^f,\boldsymbol{\gamma}\rangle_m}(v)\right] I\left(\langle \boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m \le u\right)\\
&= M^f_{n,\mathtt{F}}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},\boldsymbol{\gamma},u,v) - n^{-1/2}\sum_{i=1}^n I\left(v - \langle h^f_i,\boldsymbol{\gamma}\rangle_m < \langle \epsilon^f_i,\boldsymbol{\gamma}\rangle_m \le v\right)\Delta_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u)\\
&:= M^f_{n,\mathtt{F}}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},\boldsymbol{\gamma},u,v) - T^{e,f}_n.
\end{aligned}
$$

Let $f_{\langle \epsilon^f,\boldsymbol{\gamma}\rangle_m}(v) = \mathbb{E}\left[I\left(\langle \epsilon^f_i,\boldsymbol{\gamma}\rangle_m = v\right)\right]$. Let $\langle h_f,\boldsymbol{\gamma}\rangle_m = \sup_j \langle h^f_j,\boldsymbol{\gamma}\rangle_m$. Similarly, using Lemma S.2, we obtain

$$
\begin{aligned}
&\mathbb{E}(T^{e,f}_n) = O(\langle h_f,\boldsymbol{\gamma}\rangle_m^{1/2}), \quad \mathbb{E}(T^{e,f}_n)^2 = \sigma^{2f}_F(\boldsymbol{\alpha},\boldsymbol{\gamma},u,\boldsymbol{\alpha},\boldsymbol{\gamma},u) f^2_{\langle \epsilon^f,\boldsymbol{\gamma}\rangle_m}(v) + (\langle h_f,\boldsymbol{\gamma}\rangle_m),\\
&\mathbb{E}\left(M^f_{n,\mathtt{F}}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},\boldsymbol{\gamma},u,v)T^{e,f}_n\right) = \sigma^f_{c,F}(\boldsymbol{\alpha},\boldsymbol{\gamma},u,v,\boldsymbol{\alpha},\boldsymbol{\gamma},u) f_{\langle \epsilon^f,\boldsymbol{\gamma}\rangle_m}(v),
\end{aligned}
$$

where

$$
\begin{aligned}
&\sigma^{2f}_F(\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2) = \sum_{k_1,k_2,l_1,l_2} \zeta^f_{j_1,k_1}\zeta^f_{j_2,k_2}\Gamma^f_{k_1,k_2}\varsigma^f_{l_1,F}(\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1)\varsigma^f_{l_2,F}(\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2),\\
&\sigma^f_{c,F}(\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,v_1,\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2) = \sum_{k,l=1}^{q_f} \zeta^f_{l,k}\varsigma^f_{l,F}(\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2)\times\\
&\qquad\qquad \mathbb{E}\left[I(\langle \epsilon^f_i,\boldsymbol{\gamma}_1\rangle_m \le v_1)\langle D^f_{i,k},\varpi^f_i\rangle \Delta(\boldsymbol{X}_i;\boldsymbol{\alpha}_{1,\bullet},u_1)\right],\\
&\varsigma^f_{l,F}(\boldsymbol{\alpha},\boldsymbol{\gamma},u) = \mathbb{E}\left[\Delta(\boldsymbol{X}_i;\boldsymbol{\alpha},u)\langle D^f_{i,l},\boldsymbol{\gamma}\rangle_m\right].
\end{aligned}
$$

Then, Theorem S.2 follows. In particular, we have $M^f_{n,F}(\hat{\boldsymbol{\beta}};\boldsymbol{\alpha},\boldsymbol{\gamma},u,v) \xrightarrow{d} M^{1,f}_{\infty,F}(\boldsymbol{\alpha},\boldsymbol{\gamma},u,v)$ for any fixed u, v, and $\boldsymbol{\alpha}_j,\boldsymbol{\gamma} \in \mathbb{S}^p := \{\boldsymbol{\alpha} \in R^p : \|\boldsymbol{\alpha}\|_p = 1\}$, $j = 1,\cdots,d$, where $M^{1,f}_{\infty,F} \equiv M^{0,f}_{\infty,F} + M^{e,f}_{\infty,F}$, $M^{e,f}_{\infty,F}$ is a Gaussian process with zero mean and covariance function

$$K_1\left((\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,v_1),(\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2,v_2)\right)=\sigma_F^{2f}(\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2)f_{\langle\epsilon^f,\boldsymbol{\gamma}_1\rangle_m}(v_1)f_{\langle\epsilon^f,\boldsymbol{\gamma}_2\rangle_m}(v_2),$$

$$\operatorname{cov}\left(M^{0,f}_{\infty,F}(\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,v_1),M^{e,f}_{\infty,F}(\boldsymbol{\alpha}_{2,\bullet},\boldsymbol{\gamma}_2,u_2,v_2)\right)=\sigma^f_{c,F}(\boldsymbol{\alpha}_{1,\bullet},\boldsymbol{\gamma}_1,u_1,v_1,\boldsymbol{\alpha}_{2,\cdot},\boldsymbol{\gamma}_2,u_2)f_{\langle\epsilon^f,\boldsymbol{\gamma}_2\rangle_m}(v_2).$$

Furthermore, based on the derivation, we have

$$\left|Prob(M^f_{n,F}(\hat{\boldsymbol{\beta}};\boldsymbol{\alpha},\boldsymbol{\gamma},u,v)<m)-Prob(M^{1,f}_{\infty,F}(\boldsymbol{\alpha},\boldsymbol{\gamma},u,v)<m)\right|=O(n^{-1/2}+\langle h_f,\boldsymbol{\gamma}\rangle_m^{1/2}).$$

Similarly to the proof of Theorem S.1, as long as $p^2(d+1)^2q_f=o(n)$, Theorem S.2 follows.

□

Proofs of Theorems 3 and S.3. We first prove Theorem 3. Note that

$$\check{F}_{n,\epsilon}(v)=n^{-1}\sum_{i=1}^n I\left(\varepsilon(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}})+\check{h}_i-h_i\le v\right).$$

Under the local alternative (3.6), and Lemma S.3, we have that

$$\begin{aligned}
&M_{n,\mathtt{F}}(\check{\boldsymbol{\beta}};\boldsymbol{\alpha},u,v)=n^{-1/2}\sum_{i=1}^n\left[I(\varepsilon(\boldsymbol{X}_i;\check{\boldsymbol{\beta}})\le v)-\check{F}_{n,\epsilon}(v)\right]I\left(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m\le u\right)\\
&=M_{n,\mathtt{F}}(\hat{\boldsymbol{\beta}};\boldsymbol{\alpha},u,v)-n^{-1/2}\sum_{i=1}^n\left[I\left(v-\check{h}_i<\epsilon_i\le v-h_i\right)\right]\Delta_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u)\\
&=M_{n,\mathtt{F}}(\hat{\boldsymbol{\beta}};\boldsymbol{\alpha},u,v)-n^{-1/2}\sum_{i=1}^n\Delta_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u)\left[(\check{h}_i-h_i)f_\epsilon(v)-0.5(\check{h}_i-h_i)^2\dot{f}_\epsilon(v)\right]\\
&=M_{n,\mathtt{F}}(\hat{\boldsymbol{\beta}};\boldsymbol{\alpha},u,v)-n^{-1}\sum_{i=1}^n\Delta_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u)f_\epsilon(v)\left(m_y(Y_i,\boldsymbol{\eta}_i)\dot{g}_l(\boldsymbol{\eta}_i,\epsilon_i,0)\mathcal{F}(\boldsymbol{X}_i)-\sum_{j=1}^q D_{i,j}\upsilon_j\right).
\end{aligned}$$

Then, Theorem 3 follows from similar arguments to those in the derivation of Theorem 2.

For Theorem S.3, it follows from Lemma **S.4** that

$$\check{F}_{n,\langle\epsilon^f,\boldsymbol{\gamma}\rangle_m}(v) = n^{-1}\sum_{i=1}^{n} I\left(\langle\varepsilon^f(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}}),\boldsymbol{\gamma}\rangle_m + \langle\check{h}_i^f - h_i^f,\boldsymbol{\gamma}\rangle_m \le v\right).$$

Similar to the proof of Theorem 2, we have

$$\begin{aligned}
&M_{n,\mathtt{F}}^f(\check{\boldsymbol{\beta}};\boldsymbol{\alpha},\boldsymbol{\gamma},u,v) = n^{-1/2}\sum_{i=1}^{n}\left[I(\langle\varepsilon^f(\boldsymbol{X}_i;\check{\boldsymbol{\beta}}),\boldsymbol{\gamma}\rangle_m \le v) - \check{F}_{n,\langle\epsilon^f,\boldsymbol{\gamma}\rangle_m}(v)\right] I\left(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m \le u\right) \\
&= M_{n,\mathtt{F}}^f(\hat{\boldsymbol{\beta}};\boldsymbol{\alpha},\boldsymbol{\gamma},u,v) - n^{-1/2}\sum_{i=1}^{n}\left[I\left(v - \langle\check{h}_i^f,\boldsymbol{\gamma}\rangle_m < \langle\epsilon_i^f,\boldsymbol{\gamma}\rangle_m \le v - \langle h_i^f,\boldsymbol{\gamma}\rangle_m\right)\right]\Delta_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u) \\
&= M_{n,\mathtt{F}}^f(\hat{\boldsymbol{\beta}};\boldsymbol{\alpha},\boldsymbol{\gamma},u,v) - n^{-1}\sum_{i=1}^{n}\Delta_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u) f_{\langle\epsilon^f,\boldsymbol{\gamma}\rangle_m}(v) \\
&\qquad\times\left(\left\langle m_y(Y_i(t),\boldsymbol{\eta}_i^f(t))\dot{g}_l(\boldsymbol{\eta}_i^f(t),\epsilon_i(t),0)\mathcal{F}(\boldsymbol{X}_i),\boldsymbol{\gamma}\right\rangle_m - \sum_{j=1}^{q}\langle D_{i,j}^f,\boldsymbol{\gamma}\rangle_m v_j^f\right).
\end{aligned}$$

Then, Theorem S.3 follows. □

Proof of Theorem 4. Define $\hat{\boldsymbol{\eta}}_i(\boldsymbol{\theta}) = \left(\hat{a},\hat{\boldsymbol{b}}_1^\top\hat{\boldsymbol{\xi}}_{i1},\cdots,\hat{\boldsymbol{b}}_d^\top\hat{\boldsymbol{\xi}}_{id}\right)$, and $\hat{\boldsymbol{D}}_i^*(\hat{\boldsymbol{\theta}}) =$

$(m_0(Y_{i,n},\hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}})),$

$m_1(Y_{i,n},\hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}}))(\hat{\boldsymbol{\xi}}_{i1})^\top,\cdots,m_d(Y_{i,n},\hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}}))(\hat{\boldsymbol{\xi}}_{id})^\top)^\top\varrho_i$. Note that

$$\begin{aligned}
M_{n,\mathtt{F}}^* &= n^{-1/2}\sum_{i=1}^{n}\left[I\left(\varepsilon^*(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}}^*) \le v\right) - F_{n,\epsilon}^*(v)\right] I\left(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m \le u\right) \\
&= M_{n,\mathtt{F}}(\varrho) + n^{-1/2}\sum_{i=1}^{n}\left[I(h_i^* + \epsilon_i\varrho_i \le v) - I(\epsilon_i\varrho_i \le v)\right]\triangle_{n,i}(\boldsymbol{X};\boldsymbol{\alpha},u),
\end{aligned}$$

where $h_i^* = (\hat{\boldsymbol{D}}_i^*(\hat{\boldsymbol{\theta}}))^\top(\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}}) + \left[m(Y_{i,n},\hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}})) - m(Y_i,\hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}})) + \boldsymbol{D}_i^\top(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\right]\varrho_i$,

and $M_{n,\mathtt{F}}(\varrho) = n^{-1/2}\sum_{i=1}^{n}\left[I(\epsilon_i\varrho_i \le v) - F_{n,\epsilon,\varrho}(v)\right] I\left(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m \le u\right)$, and $F_{n,\epsilon,\varrho}(v) =$

$n^{-1}\sum_{i=1}^{n} I(\epsilon_i\varrho_i \le v)$. It follows from Taylor expansion that

$$\begin{aligned}
0 &\equiv n^{-1/2}\sum_{i=1}^{n} \dot{\ell}\left(m\left(Y_{i,n}^{*}, \hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}}^{*})\right)\right) \hat{\boldsymbol{D}}_i^{*}(\hat{\boldsymbol{\theta}}^{*}) \\
&= n^{-1/2}\sum_{i=1}^{n} \dot{\ell}\left(m\left(Y_{i,n}^{*}, \hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}})\right)\right) \boldsymbol{D}_i^{*}(\hat{\boldsymbol{\theta}}) + n^{1/2}\tilde{\boldsymbol{\Gamma}}^{*}(\hat{\boldsymbol{\theta}}^{*} - \hat{\boldsymbol{\theta}}) + o_p(1) \\
&= n^{-1/2}\sum_{i=1}^{n} \dot{\ell}\left(m\left(Y_{i,n}, \boldsymbol{\eta}_i\right)\varrho_i\right) \boldsymbol{D}_i + n^{1/2}\tilde{\boldsymbol{\Gamma}}^{*}(\hat{\boldsymbol{\theta}}^{*} - \hat{\boldsymbol{\theta}}) \\
&\qquad + n^{-1/2}\sum_{i=1}^{n} \ddot{\ell}\left(m\left(Y_{i,n}, \boldsymbol{\eta}_i\right)\varrho_i\right) \boldsymbol{D}_i\boldsymbol{D}_i^{\top}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\varrho_i + o_p(1),
\end{aligned}$$

where $\tilde{\boldsymbol{\Gamma}}^{*} = n^{-1}\sum_{i=1}^{n} \ddot{\ell}\left(m\left(Y_{i,n}^{*}, \hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}})\right)\right) \boldsymbol{D}_i^{*}(\hat{\boldsymbol{\theta}})(\boldsymbol{D}_i^{*}(\hat{\boldsymbol{\theta}}))^{\top}$. Then from Lemma **S.3**, we have

$$\begin{aligned}
\hat{\boldsymbol{\theta}}^{*} - \hat{\boldsymbol{\theta}} = -n^{-1}(\tilde{\boldsymbol{\Gamma}}^{*})^{-1}\sum_{i=1}^{n} \Big[&\dot{\ell}(\epsilon_i\varrho_i) \\
&+ \left(\dot{\ell}\left(m\left(Y_{i,n}, \boldsymbol{\eta}_i\right)\varrho_i\right) - \dot{\ell}(m(Y_i, \boldsymbol{\eta}_i)\varrho_i)\right)\Big] \boldsymbol{D}_i(1 + o_p(1)). \quad \text{(S.4)}
\end{aligned}$$

Substituting expression (S.4) into h_i^{*} and using condition (C7), we obtain

$$h_i^{*} = -n^{-1}\sum_{s=1}^{n}\sum_{k,j=1}^{q} \zeta_{j,k} D_{s,k} D_{i,j}\dot{\ell}(\epsilon_s\varrho_s).$$

Then Theorem 4 follows using similar arguments to the proof of Theorem 2 and condition (C7). □

S.4 Corollaries

Corollary S.2. *Under conditions of Theorem 1 and null hypothesis (1.1), if $p = o(n)$,* *(a) For model (2.4), for any* $m \in \mathcal{R}$, *Prob*$(T_{n,\mathtt{M}}(\boldsymbol{\beta}_0) < m) -$ *Prob*$(T^0_{\infty,M} < m) \to 0$, *where* $T^0_{\infty,M} := \int_{\mathbb{S}^{pd}} \int_{\mathcal{R}} \left(M^0_{\infty,M}(\boldsymbol{\alpha}, u)\right)^2 F_{\langle \boldsymbol{X}, \boldsymbol{\alpha}\rangle_m}(du) d\boldsymbol{\alpha}$, *and* $M^0_{\infty,M}(\cdot,\cdot)$ *is a Gaussian process with zero mean and covariance function*

$$K_M\left(\boldsymbol{\alpha}_{1,\bullet}, u_1, \boldsymbol{\alpha}_{2,\bullet}, u_2\right) = \mathbb{E}(\epsilon_i^2)\mathbb{E}\left[I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}_{1,\bullet}\rangle_m \le u_1\right) I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}_{2,\bullet}\rangle_m \le u_2\right)\right].$$

(b) For model (2.5), for any $m \in \mathcal{R}$, *Prob*$(T^f_{n,\mathtt{M}}(\boldsymbol{\beta}_0) < m) -$ *Prob*$(T^{0,f}_{\infty,M} < m) \to 0$, *where* $T^{0,f}_{\infty,M} := \int_{\mathbb{S}^{pd}} \int_{\mathcal{R}} \int_0^1 \left(M^{0,f}_{\infty,M}(\boldsymbol{\alpha}, u; t)\right)^2 dt F_{\langle \boldsymbol{X}, \boldsymbol{\alpha}\rangle_m}(du) d\boldsymbol{\alpha}$, *where* $M^{0,f}_{\infty,M}(\cdot,\cdot;t)$ *is a Gaussian process with zero mean and covariance function*

$$K^f_M\left(\boldsymbol{\alpha}_{1,\bullet}, u_1, \boldsymbol{\alpha}_{2,\bullet}, u_2; t)\right) = \mathbb{E}\left[(\epsilon_i(t))^2\right]\mathbb{E}\left[I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}_{1,\bullet}\rangle_m \le u_1\right) I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}_{2,\bullet}\rangle_m \le u_2\right)\right].$$

Proof. For model (2.4), it follows from Theorem 1 in Escanciano (2006), and

$$M^0_{n,M}(\boldsymbol{\alpha}, u) = n^{-1/2} \sum_{i=1}^n \epsilon_i I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}\rangle_m \le u\right),$$

that $M_{n,\mathtt{M}}(\boldsymbol{\beta}_0; \boldsymbol{\alpha}, u)$ converges to $M^0_{\infty,M}(\boldsymbol{\alpha}, u)$. For model (2.5), similarly, we have $M^f_{n,\mathtt{M}}(\boldsymbol{\beta}_0; \boldsymbol{\alpha}, u, t)$ converges to $M^{0,f}_{\infty,M}(\boldsymbol{\alpha}, u; t)$ from

$$M^f_{n,\mathtt{M}}(\boldsymbol{\beta}_0; \boldsymbol{\alpha}, u, t) = n^{-1/2} \sum_{i=1}^n \epsilon_i(t) I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}\rangle_m \le u\right).$$

Then, Corollary S.2 holds from similar arguments as those in the proof of Theorems 1 and S.1.

□

Corollary S.3. *Under conditions (C1)–(C2),*

(a) for model (2.4), If further conditions (C3)–(C5) hold, and under null hypothesis (1.1), for any $m \in \mathcal{R}$,

$$Prob(T_{n,\mathtt{M}}(\hat{\boldsymbol{\beta}}) < m) - Prob(T^1_{\infty,M} < m) \to 0,$$

where $T^1_{\infty,M} := \int_{\mathbb{S}^{pd}} \int_{\mathcal{R}} \left(M^1_{\infty,M}(\boldsymbol{\alpha}, u)\right)^2 F_{\langle \boldsymbol{X}, \boldsymbol{\alpha}\rangle_m}(du) d\boldsymbol{\alpha}$, *and* $M^1_{\infty,M} \equiv M^0_{\infty,M} + M^e_{\infty,M}$, $M^e_{\infty,M}$ *is a Gaussian process with zero mean and covariance function* $K_{1,M}(\boldsymbol{\alpha}_{1,\bullet}, u_1, \boldsymbol{\alpha}_{2,\bullet}, u_2) = \sigma^2_M(\boldsymbol{\alpha}_{1,\bullet}, u_1, \boldsymbol{\alpha}_{2,\bullet}, u_2)$, *and*

$$cov\left(M^0_{\infty,M}(\boldsymbol{\alpha}_{1,\bullet}, u_1), M^e_{\infty,M}(\boldsymbol{\alpha}_{2,\bullet}, u_2)\right) = -\sigma_{c,M}(\boldsymbol{\alpha}_{1,\bullet}, u_1, \boldsymbol{\alpha}_{2,\bullet}, u_2).$$

(b) for model (2.5), if further conditions (C3f)–(C5f) hold, and under null hypothesis (1.1), for any $m \in \mathcal{R}$,

$$Prob(T^f_{n,\mathtt{M}}(\hat{\boldsymbol{\beta}}) < m) - Prob(T^{1,f}_{\infty,M} < m) \to 0,$$

where $T^{1,f}_{\infty,M} := \int_{\mathbb{S}^{pd}} \int_{\mathcal{R}} \int_0^1 \left(M^{1,f}_{\infty,M}(\boldsymbol{\alpha}, u; t)\right)^2 dt F_{\langle \boldsymbol{X}, \boldsymbol{\alpha}\rangle_m}(du) d\boldsymbol{\alpha}$, *and* $M^{1,f}_{\infty,M} \equiv M^{0,f}_{\infty,M} + M^{e,f}_{\infty,M}$, $M^{e,f}_{\infty,M}$ *is a Gaussian process with zero mean and covariance function* $K^f_{1,M}(\boldsymbol{\alpha}_{1,\bullet}, u_1, \boldsymbol{\alpha}_{2,\bullet}, u_2; t) = \sigma^{2f}_M(\boldsymbol{\alpha}_{1,\bullet}, u_1, \boldsymbol{\alpha}_{2,\bullet}, u_2; t)$, *and*

$$cov\left(M^{0,f}_{\infty,M}(\boldsymbol{\alpha}_{1,\bullet}, u_1; t), M^{e,f}_{\infty,M}(\boldsymbol{\alpha}_{2,\bullet}, u_2; t)\right) = -\sigma^f_{c,M}(\boldsymbol{\alpha}_{1,\bullet}, u_1, \boldsymbol{\alpha}_{2,\bullet}, u_2; t).$$

Proof. We first show part (a). Note that

$$\begin{aligned} M_{n,\mathtt{M}}(\hat{\boldsymbol{\beta}};\boldsymbol{\alpha},u) &= n^{-1/2}\sum_{i=1}^{n}\varepsilon(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}})I\left(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m\le u\right) \\ &= M_{n,\mathtt{M}}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},u)+n^{-1/2}\sum_{i=1}^{n}\left(\varepsilon(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}})-\epsilon_i\right)I\left(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m\le u\right) \\ &:= M_{n,\mathtt{M}}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},u)+T_n^{e,u}. \end{aligned}$$

It follows from Lemma **S.1** that $\varepsilon(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}})=\epsilon_i+h_i$. Then, we have $\mathbb{E}(T_n^{e,u})=o_p(1)$, $\mathbb{E}(T_n^{e,u})^2=\sigma_M^2(\boldsymbol{\alpha}_{1,\bullet},u_1,\boldsymbol{\alpha}_{2,\bullet},u_2)+o(1)$, and

$$\begin{aligned} \mathbb{E}\left(M_{n,\mathtt{M}}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},u)T_n^{e,u}\right) &= \mathbb{E}\left[n^{-1}\sum_{i,j}^{n}\epsilon_i h_j I\left(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m\le u\right)I\left(\langle\boldsymbol{X}_j,\boldsymbol{\alpha}\rangle_m\le u\right)\right] \\ &= -\sigma_{c,M}(\boldsymbol{\alpha},u,\boldsymbol{\alpha},u). \end{aligned}$$

where $\sigma_{c,M}(\boldsymbol{\alpha}_{1,\bullet},u_1,\boldsymbol{\alpha}_{2,\bullet},u_2)=\sum_{k,l=1}^{q}\zeta_{l,k}\varsigma_{l,M}(\boldsymbol{\alpha}_{2,\bullet},u_2)\mathbb{E}\left[\epsilon_i\varpi_i D_{i,k}I\left(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}_{1,\bullet}\rangle_m\le u_1\right)\right]$, $\sigma_M^2(\boldsymbol{\alpha}_{1,\bullet},u_1,\boldsymbol{\alpha}_{2,\bullet},u_2)=\sum_{k_1,k_2,l_1,l_2=1}^{q}\zeta_{l_1,k_1}\zeta_{l_2,k_2}\Gamma_{k_1,k_2}\varsigma_{l_1,M}(\boldsymbol{\alpha}_{1,\bullet},u_1)\varsigma_{l_2,M}(\boldsymbol{\alpha}_{2,\bullet},u_2)$, and $\varsigma_{l,M}(\boldsymbol{\alpha},u)=\mathbb{E}\left[I\left(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m\le u\right)D_{i,l}\right]$.

For part (b), similarly, we have

$$\begin{aligned} &M_{n,\mathtt{M}}^{f}(\hat{\boldsymbol{\beta}};\boldsymbol{\alpha},u,t)=n^{-1/2}\sum_{i=1}^{n}\varepsilon^f(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}})I\left(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m\le u\right) \\ &=M_{n,\mathtt{M}}^{f}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},u,t)+n^{-1/2}\sum_{i=1}^{n}h_i^f I\left(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m\le u\right):=M_{n,\mathtt{M}}^{f}(\boldsymbol{\beta}_0;\boldsymbol{\alpha},u,t)+T_n^{e,u,f}. \end{aligned}$$

It follows from Lemma **S.2** that $\varepsilon^f(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}})=\epsilon_i^f+h_i^f$. Then, we have

$\mathbb{E}(T_n^{e,u,f}) = o_p(1)$, $\mathbb{E}(T_n^{e,u,f})^2 = \sigma_M^{2f}(\boldsymbol{\alpha}, u, \boldsymbol{\alpha}, u; t) + o(1)$, and

$$\mathbb{E}\left(M_{n,\mathtt{M}}^f(\boldsymbol{\beta}_0; \boldsymbol{\alpha}, u, t) T_n^{e,u,f}\right) = \mathbb{E}\left[n^{-1}\sum_{i,j}^{n} \epsilon_i(t) h_j^f I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}\rangle_m \le u\right) I\left(\langle \boldsymbol{X}_j, \boldsymbol{\alpha}\rangle_m \le u\right)\right]$$

$$= -\sum_{k,l=1}^{q_f} \zeta_{l,k}^f \varsigma_{l,M}^f(\boldsymbol{\alpha}, u; t)\mathbb{E}\left[\epsilon_i(t)\langle D_{i,k}^f, \varpi_i^f\rangle I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}\rangle_m \le u\right)\right],$$

where

$$\sigma_{c,M}^f(\boldsymbol{\alpha}_{1,\bullet}, u_1, \boldsymbol{\alpha}_{2,\bullet}, u_2; t) = \sum_{k,l=1}^{q} \zeta_{l,k}^f \varsigma_{l,M}^f(\boldsymbol{\alpha}_{2,\bullet}, u_2; t)\mathbb{E}\left[\epsilon_i(t)\langle D_{i,k}^f, \varpi_i^f\rangle I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}_{1,\bullet}\rangle_m \le u_1\right)\right],$$

$$\sigma_M^{2f}(\boldsymbol{\alpha}_{1,\bullet}, u_1, \boldsymbol{\alpha}_{2,\bullet}, u_2; t) = \sum_{k_1,k_2,l_1,l_2} \zeta_{j_1,k_1}^f \zeta_{j_2,k_2}^f \Gamma_{k_1,k_2}^f \varsigma_{l_1,M}^f(\boldsymbol{\alpha}_{1,\bullet}, u_1; t)\varsigma_{l_2,M}^f(\boldsymbol{\alpha}_{2,\bullet}, u_2; t),$$

$$\varsigma_{l,M}^f(\boldsymbol{\alpha}, u; t) = \mathbb{E}\left[I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}\rangle_m \le u\right) D_{i,l}^f(t)\right].$$

□

Corollary S.4. *Under conditions of Theorem 3,*

(a) for model (2.4), under conditions (C1)–(C6), and local alternative (3.6), for any $m \in \mathcal{R}$,

$$Prob(T_{n,\mathtt{M}}(\check{\boldsymbol{\beta}}) < m) - Prob(T_{\infty,M}^a < m) \to 0,$$

where $T_{\infty,M}^a := \int_{\mathbb{S}^{pd}} \int_{\mathcal{R}} \left(M_{\infty,M}^a(\boldsymbol{\alpha}, u)\right)^2 F_{\langle \boldsymbol{X},\boldsymbol{\alpha}\rangle_m}(du)d\boldsymbol{\alpha}$, and $\check{\boldsymbol{\beta}}$ is the estimate obtained from model (2.4) using data $\{Y_{i,n}, \boldsymbol{X}_i\}_{i=1}^n$, $M_{\infty,M}^a(\boldsymbol{\alpha}, u) \equiv M_{\infty,M}^1(\boldsymbol{\alpha}, u) + D_M^a(\boldsymbol{\alpha}, u)$, and

$$D_M^a(\boldsymbol{\alpha}, u) = \mathbb{E}\left\{I\left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha}\rangle_m \le u\right)\left(m_y(Y_i, \boldsymbol{\eta}_i)\dot{g}_l(\boldsymbol{\eta}_i, \epsilon_i, 0)\mathcal{F}(\boldsymbol{X}_i) - \sum_{j=1}^{q} D_{i,j}\upsilon_j\right)\right\},$$

where $m_y, \dot{g}_l, \upsilon_j$ are defined in Section S.1.

(b) for model (2.5), under conditions (C1)–(C2), (C3f)–(C5f), (C6), and local alternative (S.1)*, for any* $m \in \mathcal{R}$*,*

$$Prob(T^{f}_{n,\mathtt{M}}(\check{\boldsymbol{\beta}}) < m) - Prob(T^{a,f}_{\infty,M} < m) \to 0,$$

where $T^{a,f}_{\infty,M} := \int_{\mathbb{S}^{pd}} \int_{\mathcal{R}} \int_0^1 \left(M^{a,f}_{\infty,M}(\boldsymbol{\alpha}, u; t) \right)^2 dt F_{\langle \boldsymbol{X}, \boldsymbol{\alpha} \rangle_m}(du) d\boldsymbol{\alpha}$*, and* $\check{\boldsymbol{\beta}}$ *is the estimate obtained from model (2.5) using data* $\{Y_{i,n}(t), \boldsymbol{X}_i\}_{i=1}^n$*,* $M^{a,f}_{\infty,M}(\boldsymbol{\alpha}, u; t) \equiv M^{1,f}_{\infty,M}(\boldsymbol{\alpha}, u; t) + D^{a,f}_M(\boldsymbol{\alpha}, u; t)$*, and*

$$D^{a,f}_M(\boldsymbol{\alpha}, u; t) = \mathbb{E} \left\{ I \left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha} \rangle_m \le u \right) \left(m_y(Y_i(t), \boldsymbol{\eta}^f_i(t)) \dot{g}_l(\boldsymbol{\eta}^f_i(t), \epsilon_i(t), 0) \mathcal{F}(\boldsymbol{X}_i) - \sum_{j=1}^q D^f_{i,j}(t) \upsilon^f_j \right) \right\},$$

where υ^f_j *is defined in Appendix A.1.*

Proof. For model (2.4), under the local alternative (3.6), and Lemma **S.3**, we have that

$$M_{n,\mathtt{M}}(\check{\boldsymbol{\beta}}; \boldsymbol{\alpha}, u) = n^{-1/2} \sum_{i=1}^n \varepsilon(\boldsymbol{X}_i; \check{\boldsymbol{\beta}}) I \left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha} \rangle_m \le u \right)$$
$$= M_{n,\mathtt{M}}(\hat{\boldsymbol{\beta}}; \boldsymbol{\alpha}, u) + n^{-1} \sum_{i=1}^n \left(m_y(Y_i, \boldsymbol{\eta}_i) \dot{g}_l(\boldsymbol{\eta}_i, \epsilon_i, 0) \mathcal{F}(\boldsymbol{X}_i) - \sum_{j=1}^q D_{i,j} \upsilon_j \right) I \left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha} \rangle_m \le u \right).$$

For model (2.5), similarly, we have

$$M^f_{n,\mathtt{M}}(\check{\boldsymbol{\beta}}; \boldsymbol{\alpha}, u) = n^{-1/2} \sum_{i=1}^n \epsilon^f(\boldsymbol{X}_i; \check{\boldsymbol{\beta}}) I \left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha} \rangle_m \le u \right)$$
$$= M^f_{n,\mathtt{M}}(\hat{\boldsymbol{\beta}}; \boldsymbol{\alpha}, u) + n^{-1} \sum_{i=1}^n \left(m_y(Y_i(t), \boldsymbol{\eta}^f_i(t)) \dot{g}_l(\boldsymbol{\eta}^f_i(t), \epsilon_i(t), 0) \mathcal{F}(\boldsymbol{X}_i) - \sum_{j=1}^q D^f_{i,j}(t) \upsilon^f_j \right) I \left(\langle \boldsymbol{X}_i, \boldsymbol{\alpha} \rangle_m \le u \right).$$

□

Corollary S.5. *Under the null hypothesis (1.1) or alternative hypothesis (3.7), if Conditions (C1)–(C7) are satisfied, the conditional distribution of* $T^*_{n,M}$ *converges in distribution to the null limiting distribution* $T_{n,\mathtt{M}}$ *giving* $\{Y_i, \boldsymbol{X}_i\}_{i=1}^n$.

Proof. We consider the alternative (3.7) with $\nu \leq 0$. For the null hypothesis, we assume $\mathcal{F}(\boldsymbol{X}_i) = 0$. The result can be proved similarly and the details are omitted. Define $\hat{\boldsymbol{\eta}}_i(\boldsymbol{\theta}) = \left(\hat{a}, \hat{\boldsymbol{b}}_1^\top \hat{\boldsymbol{\xi}}_{i1}, \cdots, \hat{\boldsymbol{b}}_d^\top \hat{\boldsymbol{\xi}}_{id}\right)$. Note that

$$
\begin{aligned}
M^*_{n,M} &= n^{-1/2} \sum_{i=1}^n \varepsilon(X_i; \hat{\boldsymbol{\beta}}^*) I(\langle \boldsymbol{X}_i, \boldsymbol{\alpha} \rangle_m \leq u) \\
&= n^{-1/2} \sum_{i=1}^n \epsilon_i \varrho_i I(\langle \boldsymbol{X}_i, \boldsymbol{\alpha} \rangle_m \leq u) \\
&\quad + n^{-1/2} \sum_{i=1}^n \left[m\left(Y^*_{i,n}, \hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}}^*)\right) - m\left(Y^*_{i,n}, \hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}})\right)\right] I(\langle \boldsymbol{X}_i, \boldsymbol{\alpha} \rangle_m \leq u) \\
&\quad + n^{-1/2} \sum_{i=1}^n \left[m\left(Y_{i,n}, \hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}})\right) - \epsilon_i \right] \varrho_i I(\langle \boldsymbol{X}_i, \boldsymbol{\alpha} \rangle_m \leq u) \equiv M_1^* + M_2^* + M_3^*.
\end{aligned}
$$

It follows from Lemma **S.3** and random variable sequence $\{\varrho_i, i = 1, \cdots, n\}$ are i.i.d. with mean zero, and ϱ_i is independent of $(Y_i, \boldsymbol{X}_i)$ that,

$$
\begin{aligned}
M_2^* &= n^{-1/2} \sum_{i=1}^n (\hat{\boldsymbol{D}}_i^*(\hat{\boldsymbol{\theta}}))^\top (\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}}) I(\langle \boldsymbol{X}_i, \boldsymbol{\alpha} \rangle_m \leq u)(1 + o_p(1)), \\
M_3^* &= n^{-1/2} \sum_{i=1}^n \left[m\left(Y_{i,n}, \hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}})\right) - m\left(Y_i, \hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}})\right) + \boldsymbol{D}_i^\top (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\right] V_i I(\langle \boldsymbol{X}_i, \boldsymbol{\alpha} \rangle_m \leq u)(1 + o_p(1)),
\end{aligned}
$$

where $\hat{\boldsymbol{D}}_i^*(\hat{\boldsymbol{\theta}}) = \left(m_0(Y_{i,n}, \hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}})), m_1(Y_{i,n}, \hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}}))(\hat{\boldsymbol{\xi}}_{i1})^\top, \cdots, m_d(Y_{i,n}, \hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}}))(\hat{\boldsymbol{\xi}}_{id})^\top\right)^\top \varrho_i$.

Next we show the asymptotic expression for $\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}}$. It follows from Taylor expansion that

$$
\begin{aligned}
0 &\equiv n^{-1/2}\sum_{i=1}^{n}\dot{\ell}\left(m\left(Y^*_{i,n},\hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}}^*)\right)\right)\hat{\boldsymbol{D}}^*_i(\hat{\boldsymbol{\theta}}^*)\\
&= n^{-1/2}\sum_{i=1}^{n}\dot{\ell}\left(m\left(Y^*_{i,n},\hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}})\right)\right)\boldsymbol{D}^*_i(\hat{\boldsymbol{\theta}}) + n^{1/2}\tilde{\boldsymbol{\Gamma}}^*(\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}}) + o_p(1)\\
&= n^{-1/2}\sum_{i=1}^{n}\dot{\ell}\left(m\left(Y_{i,n},\boldsymbol{\eta}_i\right)\varrho_i\right)\boldsymbol{D}_i + n^{1/2}\tilde{\boldsymbol{\Gamma}}^*(\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}})\\
&\qquad + n^{-1/2}\sum_{i=1}^{n}\ddot{\ell}\left(m\left(Y_{i,n},\boldsymbol{\eta}_i\right)\varrho_i\right)\boldsymbol{D}_i\boldsymbol{D}_i^{\top}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\varrho_i + o_p(1),
\end{aligned}
$$

where $\tilde{\boldsymbol{\Gamma}}^* = n^{-1}\sum_{i=1}^{n}\ddot{\ell}\left(m\left(Y^*_{i,n},\hat{\boldsymbol{\eta}}_i(\hat{\boldsymbol{\theta}})\right)\right)\boldsymbol{D}^*_i(\hat{\boldsymbol{\theta}})(\boldsymbol{D}^*_i(\hat{\boldsymbol{\theta}}))^{\top}$. Then from Lemma **S.3**, we have

$$
\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}} = -n^{-1}(\tilde{\boldsymbol{\Gamma}}^*)^{-1}\sum_{i=1}^{n}\left[\dot{\ell}(\epsilon_i\varrho_i) + \left(\dot{\ell}\left(m\left(Y_{i,n},\boldsymbol{\eta}_i\right)\varrho_i\right) - \dot{\ell}(m(Y_i,\boldsymbol{\eta}_i)\varrho_i)\right)\right]\boldsymbol{D}_i(1 + o_p(1)).
$$

Substituting above expression into M^*_2, under the condition that ε_i and $\varepsilon_i\varrho_i$ have the same distribution and ϱ_i has the variance 1 and mean 0, we have

$$
\begin{aligned}
M^*_{n,M} &= n^{-1/2}\sum_{i=1}^{n}\epsilon_i\varrho_i I(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m \le u)\\
&\qquad - n^{-1/2}\sum_{i=1}^{n}(n^{-1}\sum_{s=1}^{n}\sum_{k,j}^{q}\zeta_{j,k}D_{s,k}D_{i,j}\dot{\ell}(\epsilon_s\varrho_s))I(\langle\boldsymbol{X}_i,\boldsymbol{\alpha}\rangle_m \le u) + o_p(1).
\end{aligned}
$$

□

S.5 Preliminary Lemmas

Let $q = Kd + 1$. Denote $\boldsymbol{\varpi} = \left(\dot{\ell}(m(Y_i, \boldsymbol{\eta}_i))\right)_{i=1}^n$, $\boldsymbol{V} = \text{Diag}\{\varpi_1^2, \cdots, \varpi_n^2\}$, $\boldsymbol{D} = \{\boldsymbol{D}_i\}_{i=1}^n$ and $\tilde{\boldsymbol{D}} = \{(\ddot{\ell}(m(Y_i, \boldsymbol{\eta}_i)))^{1/2}\boldsymbol{D}_i\}_{i=1}^n$ are $n \times q$-dimensional matrices, $\boldsymbol{\Gamma} = \lim_{n\to\infty} \frac{\boldsymbol{D}^\top \boldsymbol{V}\boldsymbol{D}}{n} := \{\Gamma_{k,l}\}_{1\le k,l\le q}$, $\tilde{\boldsymbol{\Gamma}} = \lim_{n\to\infty}\left(\frac{\tilde{\boldsymbol{D}}^\top\tilde{\boldsymbol{D}}}{n}\right) = (\tilde{\Gamma}_{k,l})_{1\le k,l\le q}$, $\Xi = \tilde{\boldsymbol{\Gamma}}^{-1} = (\zeta_{j,k})_{1\le j,k\le q}$, $\sum_{k=1}^q \zeta_{j,k}^{(1/2)}\zeta_{k,l}^{(1/2)} = \zeta_{j,l}$. Denote $T_{i2} = \sum_{j=1}^d m_j(Y_i, \boldsymbol{\eta}_i)\left(\hat{\boldsymbol{\xi}}_{ij} - \boldsymbol{\xi}_{ij}\right)^\top \boldsymbol{b}_j$.

Lemma S.1. *Under conditions (C1)–(C5),* $\varepsilon(\boldsymbol{X}_i; \hat{\boldsymbol{\beta}}) = \epsilon_i + h_i$, *where*

$$h_i = -n^{-1}\sum_{s=1}^n \sum_{k,j=1}^q \zeta_{j,k} D_{s,k} D_{i,j}\varpi_s + T_{i2},$$

and $\mathbb{E}(h_i) = 0$, $h_i = O_p(q^{1/2}n^{-1/2})$.

Proof. Using a Taylor expansion, for a $\tilde{\boldsymbol{\theta}}$ between $\boldsymbol{\theta}$ and $\hat{\boldsymbol{\theta}}$, we have

$$\boldsymbol{U}(\boldsymbol{\theta}) = \boldsymbol{U}(\hat{\boldsymbol{\theta}}) - \boldsymbol{J}_{\tilde{\boldsymbol{\theta}}}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) = -\left[\tilde{\boldsymbol{D}}^\top\tilde{\boldsymbol{D}} + (\boldsymbol{J}_{\tilde{\boldsymbol{\theta}}} - \boldsymbol{J}_{\boldsymbol{\theta}}) + \left(\boldsymbol{J}_{\boldsymbol{\theta}} - \tilde{\boldsymbol{D}}^\top\tilde{\boldsymbol{D}}\right)\right]\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right),$$

where

$$\boldsymbol{J}_{\boldsymbol{\theta}} = \tilde{\boldsymbol{D}}^\top\tilde{\boldsymbol{D}} + \sum_{i=1}^n \varpi_i \frac{\partial \boldsymbol{D}_i}{\partial\boldsymbol{\theta}} + \sum_{i=1}^n \partial\left\{\dot{\ell}(m(Y_i, \hat{\boldsymbol{\eta}}_i))\hat{\boldsymbol{D}}_i - \varpi_i\boldsymbol{D}_i\right\}/\partial\boldsymbol{\theta}.$$

Denote $\mathcal{Z}_n = \left(\frac{\tilde{\boldsymbol{D}}^\top\tilde{\boldsymbol{D}}}{n}\right)^{-1}\frac{\boldsymbol{D}^\top\boldsymbol{\varpi}}{\sqrt{n}}$. Then it follows that

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = -\left[I_p + \left(\frac{\tilde{\boldsymbol{D}}^\top\tilde{\boldsymbol{D}}}{n}\right)^{-1}\left(\frac{\boldsymbol{J}_{\tilde{\boldsymbol{\theta}}} - \boldsymbol{J}_{\boldsymbol{\theta}}}{n}\right) + \left(\frac{\tilde{\boldsymbol{D}}^\top\tilde{\boldsymbol{D}}}{n}\right)^{-1}\left(\frac{\boldsymbol{J}_{\boldsymbol{\theta}} - \tilde{\boldsymbol{D}}^\top\tilde{\boldsymbol{D}}}{n}\right)\right]^{-1}$$
$$\times\left[\mathcal{Z}_n + \left(\frac{\tilde{\boldsymbol{D}}^\top\tilde{\boldsymbol{D}}}{n}\right)^{-1}\frac{\boldsymbol{U}(\boldsymbol{\theta}) - \boldsymbol{D}^\top\boldsymbol{\varpi}}{\sqrt{n}}\right].$$

Note that under conditions (C3)-(C4), $\mathbb{E}\left(\left\|\frac{\boldsymbol{J}_{\boldsymbol{\theta}}-\tilde{\boldsymbol{D}}^{\top}\tilde{\boldsymbol{D}}}{n}\right\|_2^2\right)=O(q^2/n)$. Combining with condition (C3)–(C5) leads to

$$\begin{aligned}
&\left\|\left(\frac{\tilde{\boldsymbol{D}}^{\top}\tilde{\boldsymbol{D}}}{n}\right)^{-1}\left(\frac{\boldsymbol{J}_{\tilde{\boldsymbol{\theta}}}-\boldsymbol{J}_{\boldsymbol{\theta}}}{n}\right)\left(\frac{\tilde{\boldsymbol{D}}^{\top}\tilde{\boldsymbol{D}}}{n}\right)^{-1}\frac{\boldsymbol{D}^{\top}\boldsymbol{\varpi}}{\sqrt{n}}\right\|_2=o_p(1),\\
&\left\|\left(\frac{\tilde{\boldsymbol{D}}^{\top}\tilde{\boldsymbol{D}}}{n}\right)^{-1}\left(\frac{\boldsymbol{J}_{\boldsymbol{\theta}}-\tilde{\boldsymbol{D}}^{\top}\tilde{\boldsymbol{D}}}{n}\right)\left(\frac{\tilde{\boldsymbol{D}}^{\top}\tilde{\boldsymbol{D}}}{n}\right)^{-1}\frac{\boldsymbol{D}^{\top}\boldsymbol{\varpi}}{\sqrt{n}}\right\|_2=o_p(1),\\
&\left\|\left(\frac{\tilde{\boldsymbol{D}}^{\top}\tilde{\boldsymbol{D}}}{n}\right)^{-1}\frac{\boldsymbol{U}(\boldsymbol{\theta})-\boldsymbol{D}^{\top}\boldsymbol{\varpi}}{\sqrt{n}}\right\|_2\\
&=\left\|n^{-1/2}\sum_{i=1}^{n}\left(\frac{\tilde{\boldsymbol{D}}^{\top}\tilde{\boldsymbol{D}}}{n}\right)^{-1}\left\{\dot{\ell}(m(Y_i,\hat{\boldsymbol{\eta}}_i))\hat{\boldsymbol{D}}_i-\varpi_i\boldsymbol{D}_i\right\}\right\|_2=O_p(qn^{-1/2}).
\end{aligned}$$

Thus, $\left\|\sqrt{n}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)+\mathcal{Z}_n\right\|_2=o_p(1)$, which implies the asymptotically term to be $\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\sim-\mathcal{Z}_n$. Denote $\mathcal{X}_n=\frac{\Xi^{1/2}\boldsymbol{D}^{\top}\boldsymbol{\varpi}}{\sqrt{n}}$, that is, $\mathcal{X}_n=\left(\sum_{i=1}^{n}\sum_{k=1}^{q}\zeta_{j,k}^{(1/2)}D_{i,k}\varpi_i/\sqrt{n}\right)_{j=1}^{q}$. Thus, it follows that

$$\begin{aligned}
\mathcal{X}_n^T\mathcal{X}_n &= \frac{1}{n}\sum_{j=1}^{q}\sum_{i_1,i_2}^{n}\sum_{k_1,k_2}^{q}\varpi_{i_1}\varpi_{i_2}D_{i_1,k_1}D_{i_2,k_2}\zeta_{j,k_1}^{(1/2)}\zeta_{j,k_2}^{(1/2)}\\
&= \frac{1}{n}\sum_{i=1}^{n}\varpi_i^2\sum_{k_1,k_2}^{q}D_{i,k_1}D_{i,k_2}\zeta_{k_1,k_2}+\frac{1}{n}\sum_{i_1\neq i_2}^{n}\sum_{k_1,k_2}^{q}\varpi_{i_1}\varpi_{i_2}D_{i_1,k_1}D_{i_2,k_2}\zeta_{k_1,k_2}:=A_n+B_n.
\end{aligned}$$

Using Condition (C5) that q and $n^{1/(2\kappa+\alpha-1)}$ are at the same order and Condition (C2) that $\kappa\geq\alpha+2$ and $\alpha>1$, it follows that $q^4/n\leq n^{\frac{4}{2\kappa+\alpha-1}-1}\leq n^{\frac{4}{3\kappa+3}-1}<n^{\frac{4}{6}-1}=o(1)$. Combining the regularity Conditions (C3)–(C4) and using similar arguments as those in Lemmas 7.3-7.5 in Muller & Stadtmuller (2005), we have $A_n=\mathbb{E}\left[\mathrm{Tr}\left(\Xi\boldsymbol{\Gamma}\right)\right]+O_p(\sqrt{q})=O_p(q)$, $B_n=O_p(\sqrt{q})$. Therefore, we obtain that $n(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})^{\top}\tilde{\boldsymbol{\Gamma}}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})=O_p(q)$.

For $\varepsilon(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}})$, we have that

$$\begin{aligned}
\varepsilon(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}}) &= \epsilon_i + \varepsilon(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}}) - \epsilon_i \\
&= \epsilon_i + m\left(Y_i, \hat{a}, \hat{\boldsymbol{b}}_1^\top \hat{\boldsymbol{\xi}}_{i1}, \cdots, \hat{\boldsymbol{b}}_d^\top \hat{\boldsymbol{\xi}}_{id}\right) - m\left(Y_i, a_0, \langle X_{i,1}, \beta_{0,1}\rangle, \cdots, \langle X_{i,d}, \beta_{0,d}\rangle\right) \\
&= \epsilon_i + \boldsymbol{D}_i^\top \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) + \sum_{j=1}^{d} m_j(Y_i, \boldsymbol{\eta}_i) \left\{(\hat{\boldsymbol{\xi}}_{ij} - \boldsymbol{\xi}_{ij})^\top \boldsymbol{b}_j - \sum_{k=K+1}^{\infty} \xi_{ij,k} b_{j,k}\right\} \\
&:= \varepsilon_i + T_{i1} + T_{i2} + T_{i3},
\end{aligned}$$

where $T_{i1} = -n^{-1}\sum_{s=1}^{n}\sum_{k,j=1}^{q}\zeta_{j,k}D_{s,k}D_{i,j}\varpi_s$, and $T_{i3} = \sum_{j=1}^{d} m_j(Y_i, \boldsymbol{\eta}_i)\sum_{k=K+1}^{\infty}\xi_{ij,k}b_{j,k}$. Note that $\mathbb{E}(T_{i2}) = \mathbb{E}(T_{i3}) = 0$. Under condition (C2) that $\alpha > 1$ and $\kappa \geq \alpha + 2$ and condition (C5) that q and $n^{\frac{1}{2\kappa+\alpha-1}}$ are at the same order as $n \to \infty$, it follows from Cai & Hall (2006) that $T_{i3}^2 = O(n^{-\frac{2\kappa+\alpha-2}{\alpha+2\kappa-1}})$ (the discussion on page 2169) and $T_{i2}^2 = O_p(n^{-1})$ (inequality (5.20)). On the other hand, we have $\mathbb{E}(T_{i1}) = 0$, and

$$\begin{aligned}
\mathbb{E}\left(T_{i1}^2\right) &= \mathbb{E}\left(n^{-2}\sum_{s=1}^{n}\varpi_s^2 \sum_{k_1,k_2,j_1,j_2=1}^{q} \zeta_{j_1,k_1}\zeta_{j_2,k_2}D_{s,k_1}D_{s,k_2}D_{i,j_1}D_{i,j_2}\right) \\
&= n^{-1}\mathbb{E}\left(\sum_{k_1,k_2,j_1,j_2=1}^{q} \zeta_{j_1,k_1}\Gamma_{k_1,k_2}\zeta_{k_2,j_2}D_{i,j_1}D_{i,j_2}\right) = O(qn^{-1}).
\end{aligned}$$

Thus, $\varepsilon(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}}) = \epsilon_i + h_i = \epsilon_i + O_p(q^{1/2}n^{-1/2})$, where $h_i = -n^{-1}\sum_{s=1}^{n}\sum_{k,j=1}^{q}\zeta_{j,k}D_{s,k}D_{i,j}\varpi_s + T_{i2}$, and $\mathbb{E}(h_i) = 0$. □

Recall that $q_f = K^2 d + K$, $\boldsymbol{\varpi}_f = \left(\dot{\ell}(m(Y_i(\cdot), \boldsymbol{\eta}_i^f(\cdot)))\right)_{i=1}^{n}$, and $\boldsymbol{D}_f =$

$\{\boldsymbol{D}_i^f\}_{i=1}^n$ is a $n \times q_f$-dimensional matrix.

$$\langle \boldsymbol{D}_f, \boldsymbol{\varpi}_f \rangle = \{\sum_{i=1}^n \langle D_{i,k}^f, \varpi_i^f \rangle\}_{k=1}^{q_f},$$

$$\boldsymbol{\Gamma}_f := \{\Gamma_{k,l}^f\}_{1\le k,l\le q_f} := \{\lim_{n\to\infty} \frac{\sum_{i=1}^n \langle D_{i,k}^f, \varpi_i^f\rangle \langle D_{i,l}^f, \varpi_i^f\rangle}{n}\}_{1\le k,l\le q_f}.$$

$\tilde{\boldsymbol{\Gamma}}_n^f = \sum_{i=1}^n \int \ddot{\ell}(m(Y_i(t), \boldsymbol{\eta}_i^f(t)))\boldsymbol{D}_i^f(t)(\boldsymbol{D}_i^f(t))^\top dt$ be a $q_f \times q_f$-dimensional matrix, $\tilde{\boldsymbol{\Gamma}}_f = \lim_{n\to\infty}\left(\frac{\tilde{\boldsymbol{\Gamma}}_n^f}{n}\right) = (\tilde{\Gamma}_{k,l}^f)_{0\le k,l\le q_f}$, $\Xi_f = \tilde{\boldsymbol{\Gamma}}_f^{-1} = (\zeta_{j,k}^f)_{0\le j,k\le q_f}$, $\sum_{k=1}^{q_f}(\zeta_{j,k}^f)^{1/2}(\zeta_{k,l}^f)^{1/2} = \zeta_{j,l}^f$. Denote $T_{i2}^f = m_0(Y_i(\cdot), \boldsymbol{\eta}_i^f(\cdot))\left(\hat{\boldsymbol{\phi}}(\cdot) - \boldsymbol{\phi}(\cdot)\right)^\top \boldsymbol{a}_f + \sum_{j=1}^d m_j(Y_i(\cdot), \boldsymbol{\eta}_i^f(\cdot))\left(\hat{\boldsymbol{\xi}}_{ij}\otimes\hat{\boldsymbol{\phi}}(\cdot) - \boldsymbol{\xi}_{ij}\otimes\boldsymbol{\phi}(\cdot)\right)^\top \boldsymbol{b}_j^f$.

Lemma S.2. *Under conditions (C1)–(C2), (C3f)–(C5f), for any* $\boldsymbol{\gamma} \in \mathcal{R}^p$, $\langle \varepsilon^f(\boldsymbol{X}_i; \hat{\boldsymbol{\beta}}), \boldsymbol{\gamma}\rangle_m = \langle \epsilon_i^f, \boldsymbol{\gamma}\rangle_m + \langle h_i^f, \boldsymbol{\gamma}\rangle_m$, *where*

$$\langle h_i^f, \boldsymbol{\gamma}\rangle_m = -n^{-1}\sum_{s=1}^n \sum_{k,j=1}^{q_f} \zeta_{j,k}^f \langle D_{s,k}^f, \varpi_s^f\rangle \langle D_{i,j}^f, \boldsymbol{\gamma}\rangle_m + \langle T_{i2}^f, \boldsymbol{\gamma}\rangle_m,$$

and $\mathbb{E}(\langle h_i^f, \boldsymbol{\gamma}\rangle_m) = 0$, $\langle h_i^f, \boldsymbol{\gamma}\rangle_m = O_p(q_f^{1/2} n^{-1/2})$.

Proof. Using a Taylor expansion, for a $\tilde{\boldsymbol{\theta}}_f$ between $\boldsymbol{\theta}_f$ and $\hat{\boldsymbol{\theta}}_f$, we have

$$\boldsymbol{U}_f(\boldsymbol{\theta}_f) = \boldsymbol{U}_f(\hat{\boldsymbol{\theta}}_f) - \boldsymbol{J}_{\tilde{\boldsymbol{\theta}}_f}\left(\hat{\boldsymbol{\theta}}_f - \boldsymbol{\theta}_f\right) = -\left[\tilde{\boldsymbol{\Gamma}}_n^f + \left(\boldsymbol{J}_{\tilde{\boldsymbol{\theta}}_f} - \boldsymbol{J}_{\boldsymbol{\theta}_f}\right) + \left(\boldsymbol{J}_{\boldsymbol{\theta}_f} - \tilde{\boldsymbol{\Gamma}}_n^f\right)\right]\left(\hat{\boldsymbol{\theta}}_f - \boldsymbol{\theta}_f\right),$$

where $\boldsymbol{J}_{\boldsymbol{\theta}_f} = \tilde{\boldsymbol{\Gamma}}_n^f + \sum_{i=1}^n \int \varpi_i(t)\frac{\partial \boldsymbol{D}_i^f(t)}{\partial \boldsymbol{\theta}_f}dt + \sum_{i=1}^n \int \partial\left[\dot{\ell}\left(m(Y_i(t), \hat{\boldsymbol{\eta}}_i(t))\right)\hat{\boldsymbol{D}}_i^f(t) - \varpi_i(t)\boldsymbol{D}_i^f(t)\right]/\partial\boldsymbol{\theta}_f dt$. Denote $\mathcal{Z}_n^f = \left(\frac{\tilde{\boldsymbol{\Gamma}}_n^f}{n}\right)^{-1}\frac{\langle \boldsymbol{D}_f, \boldsymbol{\varpi}_f\rangle}{\sqrt{n}}$. Note that

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_f - \boldsymbol{\theta}_f) = -\left[I_{q_f} + \left(\frac{\tilde{\boldsymbol{\Gamma}}_n^f}{n}\right)^{-1}\left(\frac{\boldsymbol{J}_{\tilde{\boldsymbol{\theta}}_f} - \boldsymbol{J}_{\boldsymbol{\theta}_f}}{n}\right) + \left(\frac{\tilde{\boldsymbol{\Gamma}}_n^f}{n}\right)^{-1}\left(\frac{\boldsymbol{J}_{\boldsymbol{\theta}_f} - \tilde{\boldsymbol{\Gamma}}_n^f}{n}\right)\right]^{-1}$$
$$\times\left\{\mathcal{Z}_n^f + \left(\frac{\tilde{\boldsymbol{\Gamma}}_n^f}{n}\right)^{-1}\frac{\boldsymbol{U}_f(\boldsymbol{\theta}_f) - \langle \boldsymbol{D}_f, \boldsymbol{\varpi}_f\rangle}{\sqrt{n}}\right\}.$$

Under conditions (C3f)–(C5f), we have $\left\|\sqrt{n}\left(\hat{\boldsymbol{\theta}}_f - \boldsymbol{\theta}_f\right) + \mathcal{Z}_n^f\right\|_2 = o_p(1)$. Thus, the asymptotically term is seen to be $\sqrt{n}(\hat{\boldsymbol{\theta}}_f - \boldsymbol{\theta}_f) \sim -\mathcal{Z}_n^f$. Denote $\mathcal{X}_n^f = \frac{\Xi_f^{1/2}\langle \boldsymbol{D}_f, \boldsymbol{\varpi}_f\rangle}{\sqrt{n}}$, that is, $\mathcal{X}_n^f = \left(\sum_{i=1}^n \sum_{k=1}^{q_f} (\zeta_{j,k}^f)^{1/2} \langle D_{i,k}^f, \varpi_i^f\rangle / \sqrt{n}\right)_{j=1}^{q_f}$. Thus, it follows that

$$\begin{aligned}
(\mathcal{X}_n^f)^\top \mathcal{X}_n^f &= \frac{1}{n}\sum_{j=1}^{q_f}\sum_{i_1,i_2}^{n}\sum_{k_1,k_2}^{q_f} \langle D_{i_1,k_1}^f, \varpi_{i_1}^f\rangle \langle D_{i_2,k_2}^f, \varpi_{i_2}^f\rangle (\zeta_{j,k_1}^f)^{1/2} (\zeta_{j,k_2}^f)^{1/2} \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{k_1,k_2}^{q_f} \langle D_{i,k_1}^f, \varpi_i^f\rangle \langle D_{i,k_2}^f, \varpi_i^f\rangle \zeta_{k_1,k_2}^f + \frac{1}{n}\sum_{i_1\neq i_2}^{n}\sum_{k_1,k_2}^{q_f} \langle D_{i_1,k_1}^f, \varpi_{i_1}^f\rangle \langle D_{i_2,k_2}^f, \varpi_{i_2}^f\rangle \zeta_{k_1,k_2}^f \\
&:= A_n^f + B_n^f.
\end{aligned}$$

Using condition (C5f) that K and $n^{1/(2\kappa+\alpha-1)}$ are at the same order and condition (C2) that $\kappa \geq \alpha+3$ and $\alpha > 1$, it follows that $q_f^4/n \leq n^{\frac{8}{2\kappa+\alpha-1}-1} \leq n^{\frac{8}{3\alpha+5}-1} = o(1)$. Then, using similar arguments as those in Lemmas 7.3-7.5 in Muller & Stadtmuller (2005), under condition (C5f), we have $A_n^f = \mathbb{E}\left[\mathrm{Tr}\left(\Xi_f \boldsymbol{\Gamma}_f\right)\right] + O_p(\sqrt{q_f}) = O_p(q_f)$, and $B_n^f = O_p(\sqrt{q_f})$. Therefore, $n(\hat{\boldsymbol{\theta}}_f - \boldsymbol{\theta}_f)^\top \tilde{\boldsymbol{\Gamma}}_f (\hat{\boldsymbol{\theta}}_f - \boldsymbol{\theta}_f) = O_p(q_f)$.

For $\langle \varepsilon^f(\boldsymbol{X}_i; \hat{\boldsymbol{\beta}}), \boldsymbol{\gamma}\rangle_m$, we have that

$$
\begin{aligned}
&\langle \varepsilon^f(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}}),\boldsymbol{\gamma}\rangle_m = \langle \epsilon_i^f,\boldsymbol{\gamma}\rangle_m + \langle \varepsilon^f(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}}),\boldsymbol{\gamma}\rangle_m - \langle \epsilon_i^f,\boldsymbol{\gamma}\rangle_m \\
&= \langle \epsilon_i^f,\boldsymbol{\gamma}\rangle_m + \left\langle m\left(Y_i(\cdot),\hat{\boldsymbol{\eta}}_i^f(\cdot)\right) - m\Big(Y_i(\cdot), a_0(\cdot), \langle X_{i,1},\beta_{0,1}(\cdot)\rangle,\cdots,\langle X_{i,d},\beta_{0,d}(\cdot)\rangle\Big),\boldsymbol{\gamma}\right\rangle_m \\
&= \langle \epsilon_i^f,\boldsymbol{\gamma}\rangle_m + \left\langle (\boldsymbol{D}_i^f)^\top\left(\hat{\boldsymbol{\theta}}_f - \boldsymbol{\theta}_f\right),\boldsymbol{\gamma}\right\rangle_m \\
&\quad + \left\langle m_0(Y_i(\cdot),\boldsymbol{\eta}_i^f(\cdot))\left\{\left(\hat{\boldsymbol{\phi}}(\cdot) - \boldsymbol{\phi}(\cdot)\right)^\top \boldsymbol{a}_f - \sum_{k=K+1}^{\infty}\phi_k(\cdot)a_k^f(\cdot)\right\},\boldsymbol{\gamma}\right\rangle_m \\
&\quad + \sum_{j=1}^{d}\left\langle m_j(Y_i(\cdot),\boldsymbol{\eta}_i^f(\cdot))\left\{\left(\hat{\boldsymbol{\xi}}_{ij}\otimes\hat{\boldsymbol{\phi}}(\cdot) - \boldsymbol{\xi}_{ij}\otimes\boldsymbol{\phi}(\cdot)\right)^\top \boldsymbol{b}_j^f - \sum_{k=K+1}^{\infty}\sum_{l=1}^{K} b_{j,lk}^f\xi_{ij,k}\phi_l(\cdot)\right.\right. \\
&\quad \left.\left. - \sum_{l=K+1}^{\infty}\sum_{k=1}^{K} b_{j,lk}^f\xi_{ij,k}\phi_l(\cdot) - \sum_{k=K+1}^{\infty}\sum_{l=K+1}^{\infty} b_{j,lk}^f\xi_{ij,k}\phi_l(\cdot)\right\},\boldsymbol{\gamma}\right\rangle_m \\
&:= \langle \epsilon_i^f,\boldsymbol{\gamma}\rangle_m + \langle T_{i1}^f,\boldsymbol{\gamma}\rangle_m + \langle T_{i2}^f,\boldsymbol{\gamma}\rangle_m + \langle T_{i3}^f,\boldsymbol{\gamma}\rangle_m,
\end{aligned}
$$

where $\langle T_{i1}^f,\boldsymbol{\gamma}\rangle_m = -n^{-1}\sum_{s=1}^n\sum_{k,j=1}^{q_f}\zeta_{j,k}^f\langle D_{s,k}^f,\varpi_s^f\rangle\langle D_{i,j}^f,\boldsymbol{\gamma}\rangle_m$, and $T_{i3}^f =$ $m_0(Y_i(\cdot),\boldsymbol{\eta}_i^f(\cdot))$ $\times\sum_{k=K+1}^{\infty}\phi_k(\cdot)a_k^f(\cdot)+\sum_{j=1}^d m_j(Y_i(\cdot),\boldsymbol{\eta}_i^f(\cdot))\Big(\sum_{k=K+1}^{\infty}\sum_{l=1}^K b_{j,lk}^f\xi_{ij,k}\phi_l(\cdot) + \sum_{l=K+1}^{\infty}\sum_{k=1}^K b_{j,lk}^f\xi_{ij,k}\phi_l(\cdot)$ $+\sum_{k=K+1}^{\infty}\sum_{l=K+1}^{\infty} b_{j,lk}^f\xi_{ij,k}\phi_l(\cdot)\Big)$. Following similarly arguments as the proof of Lemma **S.1**, we obtain that $\mathbb{E}\Big(\langle T_{i2}^f,\boldsymbol{\gamma}\rangle_m\Big) = \mathbb{E}\Big(\langle T_{i3}^f,\boldsymbol{\gamma}\rangle_m\Big) = 0$, $\left\langle T_{i3}^f,\boldsymbol{\gamma}\right\rangle_m^2 = O_p(q_f n^{-1})$, and $\left\langle T_{i2}^f,\boldsymbol{\gamma}\right\rangle_m^2 = O_p(q_f n^{-1})$. On the other hand, we have $\mathbb{E}\left(\left\langle T_{i1}^f,\boldsymbol{\gamma}\right\rangle_m\right) = 0$, and

$$
\begin{aligned}
\mathbb{E}\left((\langle T_{i1}^f,\boldsymbol{\gamma}\rangle_m)^2\right) &= n^{-2}\sum_{s=1}^{n}\sum_{k_1,k_2,j_1,j_2=1}^{q_f}\mathbb{E}\left(\zeta_{j_1,k_1}^f\zeta_{j_2,k_2}^f\langle D_{s,k_1}^f,\varpi_s^f\rangle\langle D_{s,k_2}^f,\varpi_s^f\rangle\langle D_{i,j_1}^f,\boldsymbol{\gamma}\rangle_m\langle D_{i,j_2}^f,\boldsymbol{\gamma}\rangle_m\right) \\
&= n^{-1}\mathbb{E}\left(\sum_{k_1,k_2,j_1,j_2=1}^{q_f}\zeta_{j_1,k_1}^f\Gamma_{k_1,k_2}^f\zeta_{j_2,k_2}^f\langle D_{i,j_1}^f,\boldsymbol{\gamma}\rangle_m\langle D_{i,j_2}^f,\boldsymbol{\gamma}\rangle_m\right) = O(q_f n^{-1}).
\end{aligned}
$$

Thus, we have $\langle \varepsilon^f(\boldsymbol{X}_i;\hat{\boldsymbol{\beta}}),\boldsymbol{\gamma}\rangle = \langle \epsilon_i^f,\boldsymbol{\gamma}\rangle_m + \langle h_i^f,\boldsymbol{\gamma}\rangle_m = \langle \epsilon_i^f,\boldsymbol{\gamma}\rangle_m + O_p(q_f^{1/2} n^{-1/2})$, where $\langle h_i^f,\boldsymbol{\gamma}\rangle_m = -n^{-1}\sum_{s=1}^n \sum_{k,j=1}^{q_f} \zeta_{j,k}^f \langle D_{s,k}^f, \varpi_s^f\rangle \langle D_{i,j}^f,\boldsymbol{\gamma}\rangle_m + \langle T_{i2}^f,\boldsymbol{\gamma}\rangle_m$, and $\mathbb{E}(\langle h_i^f,\boldsymbol{\gamma}\rangle_m) = 0$. □

Recall that $Y_{i,n} = g_l\left(a, \langle X_{i,1},\beta_1\rangle, \cdots, \langle X_{i,d},\beta_d\rangle, \epsilon_i, n^{-1/2}\mathcal{F}(\boldsymbol{X}_i)\right)$, $Y_i = g_l\left(a, \langle X_{i,1},\beta_1\rangle, \cdots, \langle X_{i,d},\beta_d\rangle, \epsilon_i, 0\right) := g\left(a, \langle X_{i,1},\beta_1\rangle, \cdots, \langle X_{i,d},\beta_d\rangle\right)$, $\varepsilon(\boldsymbol{X}_i;\boldsymbol{\beta}) = m(Y_i, a, \langle X_{i,1},\beta_1\rangle, \cdots, \langle X_{i,d},\beta_d\rangle)$, and $\varepsilon(\boldsymbol{X}_i;\check{\boldsymbol{\beta}}) = m(Y_{i,n}, \check{a}, (\hat{\boldsymbol{\xi}}_{i1})^\top \check{\boldsymbol{b}}_1, \cdots, (\hat{\boldsymbol{\xi}}_{id})^\top \check{\boldsymbol{b}}_d)$.

Lemma S.3. *Under conditions (C1)–(C6),* $\varepsilon(\boldsymbol{X}_i;\check{\boldsymbol{\beta}}) = \epsilon_i + \check{h}_i$, *where*

$$\check{h}_i = -n^{-1}\sum_{s=1}^n \sum_{k,j=1}^{q} \zeta_{j,k} D_{s,k} D_{i,j} \varpi_s + T_{i2} + n^{-1/2}\left(m_y(Y_i,\boldsymbol{\eta}_i)\dot{g}_l(\boldsymbol{\eta}_i,\epsilon_i,0)\mathcal{F}(\boldsymbol{X}_i) - \sum_{j=1}^q D_{i,j}\upsilon_j\right).$$

Proof. Note that $\check{\boldsymbol{\beta}}$ is obtained by solving the following estimating equation,

$$\begin{aligned}
\boldsymbol{U}(\boldsymbol{\theta}) &= \sum_{i=1}^n \dot{\ell}\left(m(Y_{ni},\hat{\boldsymbol{\eta}}_i)\right)\hat{\boldsymbol{D}}_i \\
&= \sum_{i=1}^n \dot{\ell}\left(m(Y_{ni},\boldsymbol{\eta}_i)\right)\boldsymbol{D}_i + \left[\sum_{i=1}^n \dot{\ell}\left(m(Y_{ni},\hat{\boldsymbol{\eta}}_i)\right)\hat{\boldsymbol{D}}_i - \sum_{i=1}^n \dot{\ell}\left(m(Y_{ni},\boldsymbol{\eta}_i)\right)\boldsymbol{D}_i\right] \\
&= \sum_{i=1}^n \dot{\ell}(m(Y_i,\boldsymbol{\eta}_i))\boldsymbol{D}_i + \sum_{i=1}^n \ddot{\ell}(m(Y_i,\boldsymbol{\eta}_i))m_y(Y_i,\boldsymbol{\eta}_i)\dot{g}_l(\boldsymbol{\eta}_i,\epsilon_i,0)n^{-1/2}\mathcal{F}(\boldsymbol{X}_i)\boldsymbol{D}_i \\
&\quad + \left[\sum_{i=1}^n \dot{\ell}\left(m(Y_{ni},\hat{\boldsymbol{\eta}}_i)\right)\hat{\boldsymbol{D}}_i - \sum_{i=1}^n \dot{\ell}\left(m(Y_{ni},\boldsymbol{\eta}_i)\right)\boldsymbol{D}_i\right],
\end{aligned}$$

where $m_y(Y,\boldsymbol{\eta}) = \partial m(Y,\boldsymbol{\eta})/\partial Y$, and $\dot{g}_l(\boldsymbol{\eta}_i,\epsilon_i,b) = \partial g_l(\boldsymbol{\eta}_i,\epsilon_i,b)/\partial b$,

Similar to Lemma **S.1**, we have that

$$\sqrt{n}(\check{\boldsymbol{\theta}} - \boldsymbol{\theta}) \approx -\mathcal{Z}_n - \boldsymbol{\upsilon}_n,$$

where $\boldsymbol{\upsilon}_n = n^{-1}\Xi\sum_{i=1}^{n}\ddot{\ell}(m(Y_i,\boldsymbol{\eta}_i))m_y(Y_i,\boldsymbol{\eta}_i)\dot{g}_l(\boldsymbol{\eta}_i,\epsilon_i,0)\mathcal{F}(\boldsymbol{X}_i)\boldsymbol{D}_i$.

For $\varepsilon(\boldsymbol{X}_i;\check{\boldsymbol{\beta}})$, it is easy to obtain that

$$\begin{aligned}
&\varepsilon(\boldsymbol{X}_i;\check{\boldsymbol{\beta}}) = \varepsilon_i + \varepsilon(\boldsymbol{X}_i;\check{\boldsymbol{\beta}}) - \varepsilon_i \\
&= \varepsilon_i + m(Y_{i,n},\check{a},(\hat{\boldsymbol{\xi}}_{i1})^\top\check{\boldsymbol{b}}_1,\cdots,(\hat{\boldsymbol{\xi}}_{id})^\top\check{\boldsymbol{b}}_d) - m(Y_i,a,\langle X_{i,1},\beta_1\rangle,\cdots,\langle X_{i,d},\beta_d\rangle) \\
&= \varepsilon_i + m_y(Y_i,\boldsymbol{\eta}_i)\dot{g}_l(\boldsymbol{\eta}_i,\epsilon_i,0)n^{-1/2}\mathcal{F}(\boldsymbol{X}_i) + \boldsymbol{D}_i^\top(\check{\boldsymbol{\theta}}-\boldsymbol{\theta}) + T_{i2} + T_{i3}.
\end{aligned}$$

Thus, it follows from Lemma **S.1** that

$$\check{h}_i = -n^{-1}\sum_{s=1}^{n}\sum_{k,j=1}^{q}\zeta_{j,k}D_{s,k}D_{i,j}\varpi_s + T_{i2} + n^{-1/2}\left(m_y(Y_i,\boldsymbol{\eta}_i)\dot{g}_l(\boldsymbol{\eta}_i,\epsilon_i,0)\mathcal{F}(\boldsymbol{X}_i) - \sum_{j=1}^{q}D_{i,j}\upsilon_j\right).$$

□

Recall that $Y_{i,n}(t) = g_l\left(a(t),\langle X_{i,1},\beta_1\rangle,\cdots,\langle X_{i,d},\beta_d\rangle,\epsilon_i(t),n^{-1/2}\mathcal{F}(\boldsymbol{X}_i)\right)$, $Y_i(t) = g_l\left(a(t),\langle X_{i,1},\beta_1\rangle,\cdots,\langle X_{i,d},\beta_d\rangle,\epsilon_i(t),0\right) := g\left(a(t),\langle X_{i,1},\beta_1\rangle,\cdots,\langle X_{i,d},\beta_d\rangle\right)$, $\varepsilon_i^f = m(Y_i(\cdot),a(\cdot),\langle X_{i,1},\beta_1(\cdot)\rangle,\cdots,\langle X_{i,d},\beta_d(\cdot)\rangle)$, and

$$\varepsilon^f(\boldsymbol{X}_i;\check{\boldsymbol{\beta}}) = m(Y_{i,n}(\cdot),(\check{\boldsymbol{a}}_f)^\top\hat{\boldsymbol{\phi}}(\cdot),(\hat{\boldsymbol{\xi}}_{i1}\otimes\hat{\boldsymbol{\phi}}(\cdot))^\top\check{\boldsymbol{b}}_1^f,\cdots,(\hat{\boldsymbol{\xi}}_{id}\otimes\hat{\boldsymbol{\phi}}(\cdot))^\top\check{\boldsymbol{b}}_d^f).$$

Lemma S.4. *Under conditions (C1)–(C2), (C3f)–(C5f), and (C6),* $\langle\varepsilon^f(\boldsymbol{X}_i;\check{\boldsymbol{\beta}}),\boldsymbol{\gamma}\rangle_m = \langle\epsilon_i^f,\boldsymbol{\gamma}\rangle_m + \langle\check{h}_i^f,\boldsymbol{\gamma}\rangle_m$, *where*

$$\begin{aligned}
\langle\check{h}_i^f,\boldsymbol{\gamma}\rangle_m &= -n^{-1}\sum_{s=1}^{n}\sum_{k,j=1}^{q_f}\zeta_{j,k}^f\langle D_{s,k}^f,\varpi_s^f\rangle\langle D_{i,j}^f,\boldsymbol{\gamma}\rangle_m + \langle T_{i2}^f,\boldsymbol{\gamma}\rangle_m \\
&\quad + n^{-1/2}\left(\langle m_y(Y_i(\cdot),\boldsymbol{\eta}_i^f(\cdot))\dot{g}_l(\boldsymbol{\eta}_i^f(\cdot),\epsilon_i(\cdot),0)\mathcal{F}(\boldsymbol{X}_i),\boldsymbol{\gamma}\rangle_m - \sum_{j=1}^{q}\langle D_{i,j}^f,\boldsymbol{\gamma}\rangle_m\upsilon_j^f\right).
\end{aligned}$$

Proof. Note that $\check{\boldsymbol{\beta}}$ is obtained by solving the following estimating equation,

$$
\begin{aligned}
\boldsymbol{U}_f(\boldsymbol{\theta}_f) &= \sum_{i=1}^{n} \int \dot{\ell}\left(m(Y_{i,n}(t), \hat{\boldsymbol{\eta}}_i^f(t))\right) \hat{\boldsymbol{D}}_i^f(t)dt \\
&= \sum_{i=1}^{n} \int \dot{\ell}\left(m(Y_{i,n}(t), \boldsymbol{\eta}_i^f(t))\right) \boldsymbol{D}_i^f(t)dt \\
&\quad + \left[\sum_{i=1}^{n} \int \dot{\ell}\left(m(Y_{i,n}(t), \hat{\boldsymbol{\eta}}_i^f(t))\right) \hat{\boldsymbol{D}}_i^f(t)dt - \sum_{i=1}^{n} \int \dot{\ell}\left(m(Y_{i,n}(t), \boldsymbol{\eta}_i^f(t))\right) \boldsymbol{D}_i^f(t)dt\right] \\
&= \sum_{i=1}^{n} \int \dot{\ell}\left(m(Y_i(t), \boldsymbol{\eta}_i^f(t))\right) \boldsymbol{D}_i^f(t)dt \\
&\quad + \sum_{i=1}^{n} \int \ddot{\ell}(m(Y_i(t), \boldsymbol{\eta}_i^f(t))) m_y(Y_i(t), \boldsymbol{\eta}_i^f(t)) \dot{g}_l(\boldsymbol{\eta}_i^f(t), \epsilon_i(t), 0) n^{-1/2} \mathcal{F}(\boldsymbol{X}_i) \boldsymbol{D}_i^f(t)dt \\
&\quad + \left[\sum_{i=1}^{n} \int \dot{\ell}\left(m(Y_{i,n}(t), \hat{\boldsymbol{\eta}}_i^f(t))\right) \hat{\boldsymbol{D}}_i^f(t)dt - \sum_{i=1}^{n} \int \dot{\ell}\left(m(Y_{i,n}(t), \boldsymbol{\eta}_i^f(t))\right) \boldsymbol{D}_i^f(t)dt\right].
\end{aligned}
$$

Similar to Lemma **S.2**, we have that

$$
\sqrt{n}(\check{\boldsymbol{\theta}}_f - \boldsymbol{\theta}_f) \approx -\mathcal{Z}_n^f - \boldsymbol{\upsilon}_n^f,
$$

where $\boldsymbol{\upsilon}_n^f = n^{-1}\Xi_f \sum_{i=1}^n \int \ddot{\ell}(m(Y_i(t), \boldsymbol{\eta}_i^f(t))) m_y(Y_i(t), \boldsymbol{\eta}_i^f(t)) \dot{g}_l(\boldsymbol{\eta}_i^f(t), \epsilon_i(t), 0) \mathcal{F}(\boldsymbol{X}_i) \boldsymbol{D}_i^f(t)dt$.

For $\langle \varepsilon^f(\boldsymbol{X}_i; \check{\boldsymbol{\beta}}), \boldsymbol{\gamma}\rangle_m$, it is easy to obtain that

$$
\begin{aligned}
&\langle \varepsilon^f(\boldsymbol{X}_i; \check{\boldsymbol{\beta}}), \boldsymbol{\gamma}\rangle_m = \langle \epsilon_i^f, \boldsymbol{\gamma}\rangle_m + \langle \varepsilon_i^f(\boldsymbol{X}_i, \check{\boldsymbol{\beta}}), \boldsymbol{\gamma}\rangle_m - \langle \epsilon_i^f, \boldsymbol{\gamma}\rangle_m \\
&= \langle \epsilon_i^f, \boldsymbol{\gamma}\rangle_m + \langle m(Y_{i,n}(\cdot), (\check{\boldsymbol{a}}_f)^\top \hat{\boldsymbol{\phi}}(\cdot), (\hat{\boldsymbol{\xi}}_{i1} \otimes \hat{\boldsymbol{\phi}}(\cdot))^\top \check{\boldsymbol{b}}_1^f, \cdots, (\hat{\boldsymbol{\xi}}_{id} \otimes \hat{\boldsymbol{\phi}}(\cdot))^\top \check{\boldsymbol{b}}_d^f) \\
&\quad - m(Y_i(\cdot), a_0(\cdot), \langle X_{i,1}, \beta_{0,1}(\cdot)\rangle, \cdots, \langle X_{i,d}, \beta_{0,d}(\cdot)\rangle), \boldsymbol{\gamma}\rangle_m \\
&= \langle \epsilon_i^f, \boldsymbol{\gamma}\rangle_m + n^{-1/2}\langle m_y(Y_i(\cdot), \boldsymbol{\eta}_i^f(\cdot)) \dot{g}_l(\boldsymbol{\eta}_i^f(\cdot), \epsilon_i(\cdot), 0) \mathcal{F}(\boldsymbol{X}_i), \boldsymbol{\gamma}\rangle_m \\
&\quad + \langle (\boldsymbol{D}_i^f)^\top (\check{\boldsymbol{\theta}}_f - \boldsymbol{\theta}_f), \boldsymbol{\gamma}\rangle_m + \langle T_{i2}^f, \boldsymbol{\gamma}\rangle_m + \langle T_{i3}^f, \boldsymbol{\gamma}\rangle_m.
\end{aligned}
$$

Thus, it follows from Lemma **S.2** that

$$\langle \check{h}_i, \boldsymbol{\gamma} \rangle_m = -n^{-1} \sum_{s=1}^{n} \sum_{k,j=1}^{q_f} \zeta_{j,k}^f \langle D_{s,k}^f, \varpi_s^f \rangle \langle D_{i,j}^f, \boldsymbol{\gamma} \rangle_m + \langle T_{i2}^f, \boldsymbol{\gamma} \rangle_m$$
$$+ n^{-1/2} \left(\langle m_y(Y_i(\cdot), \boldsymbol{\eta}_i^f(\cdot)) \dot{g}_l(\boldsymbol{\eta}_i^f(\cdot), \epsilon_i(\cdot), 0) \mathcal{F}(\boldsymbol{X}_i), \boldsymbol{\gamma} \rangle_m - \sum_{j=1}^{q} \langle D_{i,j}^f, \boldsymbol{\gamma} \rangle_m \upsilon_j^f \right).$$

□

S.6 Additional results from numerical study

In this section, we provide more simulation results. The following part includes the non-Gaussian noise cases, Brownian Bridge and Pareto noise with sample size $n = 40$. We also list the computation time for different test methods for Examples 1.1 of the main paper and 1.2 of the Supplementary Material.

Example 1.2 (FLMsR) We consider a model where the response, Y_i, is a scalar and the predictor, $X_i(t)$, is a univariate functional predictor. It is given by a functional linear model as follows:

$$Y_i = \int_0^1 c_1 \cdot \beta(t) X_i(t) dt + \int_0^1 c_2 \cdot \beta(t) X_i^2(t) dt + \left(\int_0^1 \{X_i(t)\}^{c_3} dt \right) \epsilon_i, \quad 1 \le i \le n, \tag{S.5}$$

where $\{X_i(t)\}_{i=1}^n$ are generated independently from Brownian bridges, $\{\epsilon_i\}_{i=1}^n$ follows $N(0, 0.1^2)$, $\beta(t) = \exp(t^2)/2$ and $c_1 = 0.25$. We consider $c_2 = 0$ and $c_3 = 0$ for the null hypothesis, and $c_2 = n^{-1/2}, n^{-2/5}$ or $c_3 = 2$ for the alter-

Table 1: Simulation results for **Example 1.2** based on the proposed test and FMDD under models (S.5) respectively. The rows of β_0 and $\hat{\beta}$show the results based on using the true $\beta_0(t)$ and the estimated value $\hat{\beta}(t)$, respectively.

test	β	Level=10%		Level=5%		Level=10%		Level=5%	
		$n=40$	$n=100$	$n=40$	$n=100$	$n=40$	$n=100$	$n=40$	$n=100$
		$c_2=0, c_3=0$				$c_2=n^{-2/5}, c_3=0$			
$T_{n,F}$	β_0	0.100	0.100	0.038	0.052	0.754	0.848	0.646	0.770
	$\hat{\beta}$	0.074	0.086	0.042	0.049	1.000	1.000	1.000	1.000
$T_{n,M}$	β_0	0.102	0.090	0.050	0.056	0.962	0.994	0.934	0.990
	$\hat{\beta}$	0.136	0.120	0.078	0.058	1.000	1.000	1.000	1.000
FMDD	β_0	0.098	0.078	0.052	0.052	0.470	0.562	0.354	0.380
		$c_2=n^{-1/2}, c_3=0$				$c_2=0, c_3=2$			
$T_{n,F}$	β_0	0.350	0.344	0.260	0.256	0.298	0.490	0.171	0.292
	$\hat{\beta}$	0.958	1.000	0.910	1.000	0.932	1.000	0.892	1.000
$T_{n,M}$	β_0	0.846	0.856	0.822	0.834	0.002	0.004	0.002	0.002
	$\hat{\beta}$	1.000	1.000	1.000	1.000	0.030	0.042	0.020	0.026
FMDD	β_0	0.296	0.266	0.212	0.154	0.004	0.007	0	0.004

natives.In this setting, the estimation $\hat{\beta}(t)$ is obtained by using the function fregre.basis in the fda.usc package.

Table 1 shows the empirical sizes and power of our proposed test compared with FMDD based on 500 repetitions. Similar conclusions with those for Table 1 in the main text can be obtained.

Example 1.3 (FLMfR) Consider the same linear regression settings as in the main paper. The response, Y_i, is a functional response and the predictor, $X_i(t)$, is a univariate functional predictor. It is given by a functional linear model as follows:

$$Y_i(t) = \int_0^1 c_1 \cdot \beta(s,t) X_i(s) ds + \int_0^1 c_2 \cdot \beta(s,t) X_i^2(s) ds + \{X_i(t)\}^{c_3} \epsilon_i(t), \quad 1 \leq i \leq n, \tag{S.6}$$

where $\{X_i(t)\}_{i=1}^n$ are generated independently from Brownian bridges, $\beta(s,t) = \exp(s^2 + t^2)/2$, and $c_1 = 0.25$. $\{\epsilon_i(t)\}_{i=1}^n$ follows Pareto distribution with location 0.01 and shape 2.5, and Brownian bridge, respectively. Setting homogeneous scenario ($c_3 = 0$), we consider $c_2 = 0$ for the null hypothesis, and $c_2 = 0.05, n^{-1/2}, n^{-2/5}$, and 1 for the alternatives, respectively.

The left block of Table 2 shows the results for the Pareto distribution while the right block shows the results for the Brownian bridge. It presents our proposed test's empirical sizes and power compared with fdapss and FMDD, when sample size $n = 40$. We use 500 bootstrap samples for each original sample to compute the critical value. Five hundred samples are generated to approximate the percentages of rejection. The results are based on the true value of coefficient function $\beta_0(s,t)$, we demonstrate our proposed test for comparison with the test fdapss proposed by Patilea et al. (2016) and the test FMDD proposed by Lee et al. (2020). The number of basic components K is chosen for each simulated sample so that the percentage of explained variance is larger than 95%, and $p = K$ is taken. Since the Pareto distribution is a skewed positive distribution, we first centralize the random measurement error to satisfy the mean zero assumption for the conditional mean test.

From Table 2, we can see that the empirical size of $T^f_{n,\mathtt{M}}$ is zero for the Pareto distribution. The reason is that when we use wild bootstrap to calculate the distribution, the skewed Pareto distribution may be concentrated around zero. The wild bootstrap can not generate the empirical distribution evenly. When the null hypothesis does not hold, $T^f_{n,\mathtt{M}}$ is more powerful than $T^f_{n,\mathtt{F}}$. When the alternative part becomes more significant as c_2 increases, $T^f_{n,\mathtt{F}}$ performs slightly better than FMDD in test power.

Further, to consider the impact of the heterogeneous variance, we generate data with $c_2 = 0, c_3 = 0$ for the null and $c_2 = 0, c_3 = 2$ for the alternative. The bottom right block of Table 2 presents the results of the test statistics $T^f_{n,\mathtt{F}}$ and $T^f_{n,\mathtt{M}}$ under heterogeneity compared with FMDD and fdapss. The results show that under model (S.6), only the distribution-based test $T^f_{n,\mathtt{F}}$ can detect the heterogeneity, whereas both $T^f_{n,\mathtt{M}}$, FMDD and fdapss fail to do so.

Compared to the fdapss test and FMDD test, we show the CPU time of computations for our methods in Table 3. It is conducted on a personal computer (Processor: Inter Core i7-9700, CPU 3.00GHz, RAM: 16GB). Our approach is realized by R. When the sample size is small, our $T^f_{n,\mathtt{M}}$ test is comparable to the FMDD test. And our $T^f_{n,\mathtt{M}}$ test is slightly slower than fdapss as the fdapss test uses FORTRAN routines for inner loop iterations. Although when sample size increases, fdapss and FMDD tests outperform our $T^f_{n,\mathtt{F}}$ test method significantly in terms of computing speed. And our $T^f_{n,\mathtt{F}}$ test method needs to apply parallel methods as the sample size rises to complete the 500 simulations. Our $T^f_{n,\mathtt{F}}$ test can detect heterogeneity, whereas both $T^f_{n,\mathtt{M}}$ and FMDD fail to do so.

Table 2: Simulation results for **Example 1.3** based on the proposed test, FMDD, and fdapss and $\epsilon \sim$ Pareto distribution with location 0.01 and shape 2.5 and Brownian Bridge, respectively. The rows of β_0 show the results based on using the true $\beta_0(s,t)$.

test	β	Level=10%	Level=5%	Level=10%	Level=5%	Level=10%	Level=5%	Level=10%	Level=5%
		$n=40$	$n=40$	$n=40$	$n=40$	$n=40$	$n=40$	$n=40$	$n=40$
		Pareto				Brownian bridge			
		$c_2=0, c_3=0$		$c_2=n^{-2/5}, c_3=0$		$c_2=0, c_3=0$		$c_2=n^{-2/5}, c_3=0$	
$T^f_{n,F}$	β_0	0.098	0.056	0.970	0.798	0.092	0.044	0.080	0.039
$T^f_{n,M}$	β_0	0.000	0.000	1.000	1.000	0.078	0.039	0.167	0.095
FMDD	β_0	0.158	0.080	0.880	0.554	0.122	0.084	0.134	0.080
fdapss	β_0	0.032	0.010	0.818	0.642	0.107	0.048	0.110	0.060
		$c_2=0.05, c_3=0$		$c_2=1, c_3=0$		$c_2=0.05, c_3=0$		$c_2=1, c_3=0$	
$T^f_{n,F}$	β_0	0.990	0.918	0.988	0.866	0.084	0.040	0.156	0.076
$T^f_{n,M}$	β_0	1.000	1.000	1.000	1.000	0.088	0.045	0.974	0.964
FMDD	β_0	0.866	0.556	0.880	0.554	0.128	0.084	0.296	0.150
fdapss	β_0	0.710	0.528	0.808	0.648	0.109	0.054	0.126	0.066
		$c_2=n^{-1/2}, c_3=0$		$c_2=0, c_3=2$		$c_2=n^{-1/2}, c_3=0$		$c_2=0, c_3=2$	
$T^f_{n,F}$	β_0	0.992	0.904	0.252	0.146	0.078	0.042	0.264	0.154
$T^f_{n,M}$	β_0	1.000	1.000	0.004	0.004	0.115	0.065	0.000	0.000
FMDD	β_0	0.880	0.556	0.006	0.006	0.128	0.078	0.002	0.002
fdapss	β_0	0.808	0.638	0.028	0.016	0.108	0.056	0.087	0.044

Table 3: Computation time for **Examples 1.1 and 1.2** based on the proposed test, FMDD, and fdapss under models FLMfR and FLMsR, respectively. The rows of β_0 and $\hat{\beta}$ show the results using the true $\beta_0(s,t)$, $\beta_0(t)$, the estimated value $\hat{\beta}(s,t)$, and $\hat{\beta}(t)$(seconds for one replicate), respectively.

test	β	Example 1.1 $n=40$	Example 1.1 $n=100$	Example 1.2 $n=40$	Example 1.2 $n=100$
		$c_2=0, c_3=0$			
$T^f_{n,F}$	β_0	480.94	16129.95	267.98	15300.31
$T^f_{n,M}$	β_0	4.91	22.22	0.83	7.31
FMDD	β_0	5.23	18.42	1.27	7.93
fdapss	β_0	0.61	2.47	-	-
		$c_2=\{0.05, n^{-1/2}, n^{-2/5}, n^0\}, c_3=0$			
$T^f_{n,F}$	β_0	495.17	16144.86	264.24	14806.5
$T^f_{n,M}$	β_0	4.83	21.06	0.79	7.01
FMDD	β_0	5.84	35.94	2.48	15.86
fdapss	β_0	0.63	2.58	-	-
		$c_2=0, c_3=2$			
$T^f_{n,F}$	β_0	493.07	16236.44	263.78	14988.11
$T^f_{n,M}$	β_0	4.77	20.80	0.82	7.01
FMDD	β_0	5.81	37.22	2.53	15.87
fdapss	β_0	0.61	2.55	-	-
		$c_2=0, c_3=0$			
$T^f_{n,F}$	$\hat{\beta}$	485.12	16349.13	281.68	15519.18
$T^f_{n,M}$	$\hat{\beta}$	20.69	36.99	3.80	11.06
		$c_2=\{0.05, n^{-1/2}, n^{-2/5}, n^0\}, c_3=0$			
$T^f_{n,F}$	$\hat{\beta}$	495.28	16497.24	280.17	15583.53
$T^f_{n,M}$	$\hat{\beta}$	20.66	37.41	4.73	11.08
		$c_2=0, c_3=2$			
$T^f_{n,F}$	$\hat{\beta}$	489.90	15853.81	280.97	15142.93
$T^f_{n,M}$	$\hat{\beta}$	20.20	37.22	4.81	11.83

References

Cai, T. T. & Hall, P. (2006), 'Prediction in functional linear regression', *Ann. Stat.* **34**(5), 2159–2179.

Escanciano, J. C. (2006), 'A consistent diagnostic test for regression models using projections', *Econ. Theory.* **22**(6), 1030–1051.

Lavergne, P. & Patilea, V. (2008), 'Breaking the curse of dimensionality in nonparametric testing', *J. Econom.* **143**(1), 103–122.

Lee, C., Zhang, X. & Shao, X. (2020), 'Testing conditional mean independence for functional data', *Biometrika* **107**(2), 331–346.

Muller, H.-G. & Stadtmuller, U. (2005), 'Generalized functional linear models', *Ann. Stat.* **33**, 774–805.

Patilea, V., Sánchez-Sellero, C. & Saumard, M. (2016), 'Testing the predictor effect on a functional response', *J. Am. Stat. Assoc.* **111**(516), 1684–1695.

Van Der Vaart, A. W. & Wellner, J. A. (1996), *Weak Convergence and Empirical Processes*, Springer-Verlag.