Supplement to "A Theoretical Framework for Bayesian Nonparametric Regression"

A Proof of Theorem 3.1

Proof of Theorem 3.1. First define the following quantity:

$$\epsilon = \frac{2}{3}\left(t - \frac{\alpha}{2\alpha+1}\right), \quad \delta = \frac{2\alpha}{2\alpha+1} - 1 + 2\epsilon, \quad \zeta = \frac{\alpha}{2\alpha+1} + \frac{\epsilon}{2}.$$

It follows from simple algebra that $2t > \delta + 1 > 2\zeta > -2\alpha\delta$ and $2\zeta > 1 - \zeta/\alpha$. Set $m_n = \lceil n^{1/(2\alpha+1)}(\log n)^\delta \rceil$, $\epsilon_n = n^{-\alpha/(2\alpha+1)}(\log n)^t$, $\underline{\epsilon}_n = n^{-\alpha/(2\alpha+1)}(\log n)^\zeta$ and denote $f_m(x) = \sum_{k=1}^m \beta_k \psi_k(x)$ given that $N = m$, *i.e.*, $\beta_k = 0$ for al $k > m$.

We first verify condition (2.5) with $\omega = 1$ and $k_n = \lceil \underline{\epsilon}_n^{-1/\alpha} \rceil$. Clearly, $k_n \underline{\epsilon}_n^2 = O(1)$. For sufficiently large n, write

$$\begin{aligned}
&\left\{ f_{k_n}(x) = \sum_{k=1}^{k_n} \beta_k \psi_k(x) : \|f_0 - f_{k_n}\|_2 \le \underline{\epsilon}_n, \sum_{k=k_n+1}^{\infty} |\beta_k - \beta_{0k}| \le \omega \right\} \\
&= \left\{ f_{k_n}(x) = \sum_{k=1}^{k_n} \beta_k \psi_k(x) : \sum_{k=1}^{k_n} (\beta_k - \beta_{0k})^2 + \sum_{k=k_n+1}^{\infty} \beta_{0k}^2 \le \underline{\epsilon}_n^2 \right\} \\
&\supset \left\{ f_{k_n}(x) = \sum_{k=1}^{k_n} \beta_k \psi_k(x) : \sum_{k=1}^{k_n} (\beta_k - \beta_{0k})^2 \le \frac{\underline{\epsilon}_n^2}{2} \right\},
\end{aligned}$$

since $f_0 \in \mathcal{H}_\alpha(Q)$ and for sufficiently large n,

$$\sum_{k=k_n+1}^{\infty} \beta_{0k}^2 \le \frac{1}{\lceil \underline{\epsilon}_n^{-1/\alpha} \rceil^{2\alpha}} \sum_{k=k_n+1}^{\infty} \beta_{0k}^2 k^{2\alpha} \le \frac{1}{2} \lceil \underline{\epsilon}_n^{-1/\alpha} \rceil^{-2\alpha} \le \frac{1}{2} \underline{\epsilon}_n^2,$$

We proceed to bound

$$\left\{ f_{k_n}(x) = \sum_{k=1}^{k_n} \beta_k \psi_k(x) : \sum_{k=1}^{k_n} (\beta_k - \beta_{0k})^2 \le \frac{\underline{\epsilon}_n^2}{2} \right\} \supset \bigcap_{k=1}^{k_n} \left\{ \beta_k : |\beta_k - \beta_{0k}| \le \frac{\underline{\epsilon}_n}{k_n} \right\}.$$

Observe that

$$\inf_{k \in \mathbb{N}_+} \inf_{|\beta_k - \beta_{0k}| < \underline{\epsilon}_n / k_n} g(\beta_k) \ge \inf_{k \in \mathbb{N}_+} g(|\beta_{0k}| + 1) = g\left(\sup_{k \ge 1} |\beta_{0k}| + 1 \right) > 0.$$

It follows that for sufficiently large n,

$$\begin{aligned}
\Pi(\|f_{k_n} - f_0\|_2 \le \underline{\epsilon}_n \mid N = k_n) &\ge \prod_{k=1}^{k_n} \Pi\left(\beta_k : |\beta_k - \beta_{0k}| \le \frac{\underline{\epsilon}_n}{k_n} \,\middle|\, N = k_n \right) = \prod_{k=1}^{k_n} \int_{\beta_{0k} - \underline{\epsilon}_n/k_n}^{\beta_{0k} + \underline{\epsilon}_n/k_n} g(\beta_k) \mathrm{d}\beta_k \\
&\ge \prod_{k=1}^{k_n} \left(\frac{2\underline{\epsilon}_n}{k_n} \right) \left[\inf_{k \in \mathbb{N}_+} \inf_{|\beta_k - \beta_{0k}| < \underline{\epsilon}_n / k_n} g(\beta_k) \right] \ge \exp\left[-2k_n \log k_n - k_n \left(\log \frac{1}{\underline{\epsilon}_n} \right) \right].
\end{aligned}$$

Therefore

$$
\begin{aligned}
\Pi\left(\|f-f_0\|_2 \le \underline{\epsilon}_n, \sum_{k=k_n+1}^{\infty}|\beta_k-\beta_{0k}|\le\omega\right) &\ge \Pi\left(\|f_{k_n}-f_0\|_2\le\underline{\epsilon}_n \mid N=k_n\right)\pi_N(k_n)\\
&\ge \exp\left[-Dn^{1/(2\alpha+1)}(\log n)^{1-\zeta/\alpha}\right]\\
&\ge \exp\left[-Dn^{1/(2\alpha+1)}(\log n)^{2\zeta}\right]=\exp(-Dn\underline{\epsilon}_n^2)
\end{aligned}
$$

for some constant $D>0$.

Now set $\delta=Q$ and construct the sieve $\mathcal{F}_{m_n}(\delta)$ as follows:

$$
\mathcal{F}_{m_n}(Q)=\left\{f(x)=\sum_{k=1}^{\infty}\beta_k\psi_k(x):\sum_{k=m_n+1}^{\infty}k^{2\alpha}(\beta_k-\beta_{0k})^2\le Q^2\right\}.
$$

We claim that $\mathcal{F}_{m_n}(Q)$ satisfies (2.2). In fact, by assumption $\sup_k\|\psi_k\|_\infty<\infty$, and it folows from Cauchy-Schwarz inequality that

$$
\begin{aligned}
\|f-f_0\|_\infty &\lesssim \sum_{k=1}^{\infty}|\beta_k-\beta_{0k}|\le\sum_{k=1}^{m_n}|\beta_k-\beta_{0k}|+\sum_{k=m+1}^{\infty}|\beta_k-\beta_{0k}|\\
&\le\sqrt{m_n}\left(\sum_{k=1}^{\infty}(\beta_k-\beta_{0k})^2\right)^{1/2}+\left(\sum_{k=m_n+1}^{\infty}\frac{1}{k^{2\alpha}}\right)^{1/2}\left(\sum_{k=m_n+1}^{\infty}k^{2\alpha}(\beta_k-\beta_{0k})^2\right)^{1/2}\\
&\le\sqrt{m_n}\|f-f_0\|_2+\left(\sum_{k=1}^{\infty}\frac{1}{k^{2\alpha}}\right)^{1/2}Q.
\end{aligned}
$$

This shows that $\|f-f_0\|_\infty^2\lesssim m\|f-f_0\|_2^2+\delta^2$ with $\delta=(\sum_{k=1}^{\infty}k^{-2\alpha})^{1/2}Q<\infty$. Next we verify condition (2.3). Note $N_{nj}\le\mathcal{N}(\xi j\epsilon_n,\mathcal{F}_{m_n}(\delta)\cap\{\|f-f_0\|_2\le(j+1)\epsilon_n\},\|\cdot\|_2)$. Write

$$
\begin{aligned}
\mathcal{F}_{m_n}(\delta)\cap\{\|f-f_0\|_2\le(j+1)\epsilon_n\}\subset&\left\{f(x)=\sum_{k=1}^{\infty}\beta_k\psi_k(x):\sum_{k=1}^{m_n}(\beta_k-\beta_{0k})^2\le(j+1)^2\epsilon_n^2\right\}\\
&\cap\left\{f(x)=\sum_{k=1}^{\infty}\beta_k\psi_k(x):\sum_{k=m_n+1}^{\infty}(\beta_k-\beta_{0k})^2k^{2\alpha}\le Q^2\right\}.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&\mathcal{N}(\xi j\epsilon_n,\mathcal{F}_{m_n}(\delta)\cap\{\|f-f_0\|_2\le(j+1)\epsilon_n\},\|\cdot\|_2)\\
&\quad\le\mathcal{N}\left(\frac{\xi j\epsilon_n}{2},\left\{(\beta_1,\cdots,\beta_{m_n}):\sum_{k=1}^{m_n}(\beta_k-\beta_{0k})^2\le(j+1)^2\epsilon_n^2\right\},\|\cdot\|_2\right)\\
&\qquad\times\mathcal{N}\left(\frac{\xi j\epsilon_n}{2},\left\{(\beta_{m_n+1},\cdots,):\sum_{k=m_n+1}^{\infty}(\beta_k-\beta_{0k})^2k^{2\alpha}\le Q^2\right\},\|\cdot\|_2\right).
\end{aligned}
$$

We now bound the two covering number separately. For the first covering number, computation of covering number in Euclidean space due to lemma 4.1 in Pollard (1990) yields

$$\mathcal{N}\left(\frac{\xi j\epsilon_n}{2}, \left\{(\beta_1,\cdots,\beta_{m_n}) : \sum_{k=1}^{m_n}(\beta_k-\beta_{0k})^2 \le (j+1)^2\epsilon_n^2\right\}, \|\cdot\|_2\right) \le \left(\frac{6(j+1)\epsilon_n}{\xi j\epsilon_n/2}\right)^{m_n} \le \exp\left(m_n \log\frac{24}{\xi}\right).$$

For the second covering number, we see that

$$\begin{aligned}
&\mathcal{N}\left(\frac{\xi j\epsilon_n}{2}, \left\{(\beta_{m_n+1},\cdots,) : \sum_{k=m_n+1}^{\infty}(\beta_k-\beta_{0k})^2k^{2\alpha} \le Q^2\right\}, \|\cdot\|_2\right)\\
&\quad\le \mathcal{N}\left(\frac{\xi j\epsilon_n}{2}, \left\{(\beta_1,\beta_2,\cdots,) : \sum_{k=1}^{\infty}(\beta_k-\beta_{0k})^2k^{2\alpha} \le Q^2\right\}, \|\cdot\|_2\right) \le \exp\left\{\log[4(2\mathrm{e})^{2\alpha}]\left(\frac{6Q}{\xi j\epsilon_n}\right)^{1/\alpha}\right\},
\end{aligned}$$

where the last inequality is due to the covering number of Sobolev ball (see lemma 6.4 in Belitser and Ghosal (2003)). We conclude that $N_{nj} \lesssim \exp\left[D_1 n^{1/(2\alpha+1)}(\log n)^{\delta}\right]$ for some constant $D_1 > 0$, since $\epsilon_n^{-1/\alpha} \asymp n^{1/(2\alpha+1)}(\log n)^{-t/\alpha} \le n^{1/(2\alpha+1)}(\log n)^{\delta} \asymp m_n$. Hence

$$\begin{aligned}
\sum_{j=M}^{\infty} N_{nj}\exp(-Dnj^2\epsilon_n^2) &\le \exp\left[D_1 n^{1/(2\alpha+1)}(\log n)^{\delta}\right]\sum_{j=M}^{\infty}\int_{j-1}^{j}\exp(-Dn\epsilon_n^2x^2)\mathrm{d}x\\
&\le \exp\left[D_1 n^{1/(2\alpha+1)}(\log n)^{\delta}\right]\int_{M-1}^{\infty}\exp(-Dn\epsilon_n^2x^2)\mathrm{d}x\\
&\lesssim \exp\left[D_1 n^{1/(2\alpha+1)}(\log n)^{\delta}\right]\exp\left[-\frac{1}{2}D(M-1)^2n^{1/(2\alpha+1)}(\log n)^{2t}\right] \to 0.
\end{aligned}$$

for sufficiently large n, and hence, condition (2.3) holds.

We are now left to show that $\mathcal{F}_{m_n}(Q)$ satisfies (2.4) with the same constant D. Write

$$\begin{aligned}
\mathcal{F}_{m_n}^c(Q) &= \left\{f(x)=\sum_{k=1}^{\infty}\beta_k\psi_k(x) : \sum_{k=m_n+1}^{\infty}(\beta_k-\beta_{0k})^2k^{2\alpha} > Q^2\right\}\\
&\subset \left\{f(x)=\sum_{k=1}^{\infty}\beta_k\psi_k(x) : \sum_{k=m_n+1}^{\infty}\beta_k^2k^{2\alpha} > \frac{Q^2}{4}\right\}
\end{aligned}$$

since by definition $f_0 \in \mathcal{H}_\alpha(Q)$ and $\sum_{k=m_n+1}^{\infty}\beta_{0k}^2k^{2\alpha} < Q^2/4$ for sufficiently large n. Next write

$$\begin{aligned}
\mathcal{F}_{m_n}^c(Q) &\subset \left\{f(x)=\sum_{k=1}^{\infty}\beta_k\psi_k(x) : \sum_{k=m_n+1}^{\infty}\beta_k^2k^{2\alpha} > \frac{Q^2}{4}\right\}\\
&\subset \bigcup_{m=1}^{\infty}\left\{f(x)=\sum_{k=1}^{\infty}\beta_k\psi_k(x) : \sum_{k=m_n+1}^{\infty}\beta_k^2k^{2\alpha} > \frac{Q^2}{4}, N=m\right\}\\
&\subset \bigcup_{m=1}^{\infty}\left\{f(x)=\sum_{k=1}^{m}\beta_k\psi_k(x) : \sum_{k=m_n+1}^{\infty}\beta_k^2k^{2\alpha} > \frac{Q^2}{4}, N=m\right\}\\
&= \bigcup_{m=m_n+1}^{\infty}\left\{f(x)=\sum_{k=1}^{m}\beta_k\psi_k(x) : \sum_{k=m_n+1}^{m}\beta_k^2k^{2\alpha} > \frac{Q^2}{4}, N=m\right\} \subset \bigcup_{m=m_n+1}\{N=m\}.
\end{aligned}$$

It follows that

$$\Pi(\mathcal{F}^c_{m_n}(Q)) \le \sum_{m=m_n+1}^{\infty} \pi_N(m) \le \exp(-b_1 m_n \log m_n) \le \exp\left[-D_2 n^{1/(2\alpha+1)} (\log n)^{\delta+1}\right]$$
$$\le \exp\left[-\left(2D + \frac{1}{\sigma^2}\right) n^{1/(2\alpha+1)} (\log n)^{2\zeta}\right] = \exp\left[-\left(2D + \frac{1}{\sigma^2}\right) n\underline{\epsilon}_n^2\right]$$

for some constant $D_2 > 0$ when n is sufficiently large. Hence condition (2.4) holds with the same constant D. □

B Proof of Theorem 3.2

Define $K_\alpha = \lceil (8Q^2)^{1/(2\alpha)} n^{1/(2\alpha+1)} \rceil$. Let L_n to be the smallest integer such that $\mathrm{e}^{L_n} > K_\alpha$, and define $k_n = \lceil \mathrm{e}^{L_n} \rceil$.

Lemma B.1 *For k_n defined above, $\epsilon_n = n^{-1/(2\alpha+1)}$, and $f_0 \in \mathcal{H}_\alpha(Q)$,*

$$\Pi\left(\sum_{k=k_n+1}^{\infty} (\beta_k - \beta_{0k})^2 \le \frac{\epsilon_n^2}{2}, \sum_{k=k_n+1}^{\infty} |\beta_k - \beta_{0k}| \le Q\right) \ge \frac{1}{2}$$

holds for sufficiently large n.

Proof. First write by the union bound

$$\Pi\left(\sum_{k=k_n+1}^{\infty} (\beta_k - \beta_{0k})^2 \le \frac{\epsilon_n^2}{2}, \sum_{k=k_n+1}^{\infty} |\beta_k - \beta_{0k}| \le Q\right)$$
$$\ge \Pi\left(\sum_{k=k_n+1}^{\infty} (\beta_k - \beta_{0k})^2 \le \frac{\epsilon_n^2}{2}\right) + \Pi\left(\sum_{k=k_n+1}^{\infty} |\beta_k - \beta_{0k}| \le Q\right) - 1$$
$$= \Pi\left(\sum_{k=k_n+1}^{\infty} (\beta_k - \beta_{0k})^2 \le \frac{\epsilon_n^2}{2}\right) - \Pi\left(\sum_{k=k_n+1}^{\infty} |\beta_k - \beta_{0k}| > Q\right).$$

By lemma 5.4 in Gao and Zhou (2016a), we know that the fist term on the right-hand side of the proceeding display is $1 - o(1)$. Hence it suffices to show that the second term on the right-hand side is $o(1)$. Write

$$\Pi\left(\sum_{k=k_n+1}^{\infty} |\beta_k - \beta_{0k}| > Q\right) \le \frac{1}{Q} \sum_{k=k_n+1}^{\infty} \mathbb{E}_\Pi |\beta_k - \beta_{0k}| \le \frac{1}{Q} \sum_{k=k_n+1}^{\infty} \left[2\mathbb{E}_\Pi\left(\beta_k^2\right) + 2\beta_{0k}^2\right]^{1/2}$$
$$\le \frac{\sqrt{2}}{Q} \sum_{k=k_n+1}^{\infty} \sqrt{\mathbb{E}_\Pi(\beta_k^2)} + \frac{\sqrt{2}}{Q} \sum_{k=k_n+1}^{\infty} |\beta_{0k}|$$
$$\le \frac{\sqrt{2}}{Q} \sum_{k=k_n+1}^{\infty} \sqrt{\mathbb{E}_\Pi(\beta_k^2)} + \frac{\sqrt{2}}{Q} \left(\sum_{k=k_n+1}^{\infty} \beta_{0k}^2 k^{2\alpha}\right)^{1/2} \left(\sum_{k=k_n+1}^{\infty} \frac{1}{k^{2\alpha}}\right)^{1/2}$$
$$\le \frac{\sqrt{2}}{Q} \sum_{k=k_n+1}^{\infty} \sqrt{\mathbb{E}_\Pi(\beta_k^2)} + o(1).$$

We are now left with showing $\sum_{k>k_n} \sqrt{\mathbb{E}_\Pi(\beta_k^2)} = o(1)$:

$$\sum_{k=k_n+1}^{\infty} \sqrt{\mathbb{E}_\Pi(\beta_k^2)} \leq \sum_{\ell=L_n-1}^{\infty} \sum_{k=k_\ell}^{k_{\ell+1}-1} \sqrt{\mathbb{E}_\Pi(A_\ell)} \leq \sum_{\ell=L_n-1}^{\infty} 2\mathrm{e}^{\ell+1-c_2\ell^2/2} = o(1),$$

where the last inequality is due to (3.3). □

Lemma B.2 *For the block prior Π defined in section 3.2 with $f_0 \in \mathcal{H}_\alpha(Q)$ for some $\alpha > 1/2$ and $Q > 0$, there exists some constant $D > 0$ such that $\Pi(B_n(k_n, \epsilon_n, Q)) \geq \exp(-Dn\epsilon_n^2)$, where $\epsilon_n = n^{-\alpha/(2\alpha+1)}$, $k_n = \lceil \mathrm{e}^{L_n} \rceil$, L_n is the smallest integer such that $\mathrm{e}^{L_n} > K_\alpha$, and $K_\alpha = \lceil (8Q^2)^{1/(2\alpha)} n^{1/(2\alpha+1)} \rceil$.*

Proof. By lemma B.1 we have

$$\begin{aligned}
\Pi(B_n(k_n, \epsilon_n, Q)) &\geq \Pi\left(\sum_{k=1}^{\infty}(\beta_k - \beta_{0k})^2 \leq \epsilon_n^2, \sum_{k=k_n+1}^{\infty} |\beta_k - \beta_{0k}| \leq Q\right) \\
&\geq \Pi\left(\sum_{k>k_n}(\beta_k - \beta_{0k})^2 \leq \frac{\epsilon_n^2}{2}, \sum_{k=k_n+1}^{\infty} |\beta_k - \beta_{0k}| \leq Q\right) \Pi\left(\sum_{k=1}^{k_n}(\beta_k - \beta_{0k})^2 \leq \frac{\epsilon_n^2}{2}\right) \\
&\geq \frac{1}{2}\Pi\left(\sum_{k=1}^{k_n}(\beta_k - \beta_{0k})^2 \leq \frac{\epsilon_n^2}{2}\right),
\end{aligned}$$

where the second inequality is due to the fact that $(\beta_k)_{k>k_n}$ and $(\beta_k)_{k=1}^{k_n}$ are independent under the prior distribution Π. Exploiting the proof of (6) in theorem 2.1 in Gao and Zhou (2016a), we see that

$$\Pi\left(\sum_{k=1}^{k_n}(\beta_k - \beta_{0k})^2 \leq \frac{\epsilon_n^2}{2}\right) \geq \exp(-D'n\epsilon_n^2)$$

for some constant $D' > 0$, and thus the proof is completed. □

Proof of theorem 3.2. Set $\epsilon_n = \underline{\epsilon}_n = n^{-\alpha/(2\alpha+1)}$, $\delta = Q$, and $m_n = \lceil (n\kappa^{-1})^{1/(2\alpha+1)} \rceil$, where κ is a constant determined later. Let $(k_n)_{n=1}^{\infty}$ be defined as in lemma B.2 and $\omega = Q$. Clearly, $m_n\epsilon_n^2 \to 0$, $k_n\underline{\epsilon}_n^2 = O(1)$, and $\delta = O(1)$, since $\alpha > 1/2$. By assumption $f_0 \in \mathcal{H}_\alpha(Q)$, and hence yields the following series expansion

$$f_0(x) = \sum_{k=1}^{\infty} \beta_{0k}\psi_k(x), \quad \text{where} \quad \sum_{k=1}^{\infty} k^{2\alpha}\beta_{0k}^2 \leq Q^2.$$

For condition (2.5), by lemma B.2 we see that it holds for some constant $D > 0$ with $k_n\epsilon_n^2 = O(1)$ and $\omega = Q = O(1)$. For condition (2.3) and (2.4), We use the following sieve $\mathcal{F}_{m_n}(Q)$

$$\mathcal{F}_{m_n}(Q) = \left\{ f(x) = \sum_{k=1}^{\infty} \beta_k\psi_k(x) : \sum_{k=m_n+1}^{\infty} k^{2\alpha}(\beta_k - \beta_{0k})^2 \leq Q^2 \right\}.$$

We next verify condition (2.3) with the same constant $D > 0$. Following the proof of theorem 3.1, we have

$$N_{nj} \leq \mathcal{N}\left(\frac{\xi j \epsilon_n}{2}, \left\{(\beta_1, \cdots, \beta_{m_n}) : \sum_{k=1}^{m_n} (\beta_k - \beta_{0k})^2 \leq (j+1)^2 \epsilon_n^2 \right\}, \|\cdot\|_2\right)$$
$$\times \mathcal{N}\left(\frac{\xi j \epsilon_n}{2}, \left\{(\beta_{m_n+1}, \cdots,) : \sum_{k=m_n+1}^{\infty} (\beta_k - \beta_{0k})^2 k^{2\alpha} \leq Q^2 \right\}, \|\cdot\|_2\right).$$

For the first covering number, lemma 4.1 in Pollard (1990) yields

$$\mathcal{N}\left(\frac{\xi j \epsilon_n}{2}, \left\{(\beta_1, \cdots, \beta_{m_n}) : \sum_{k=1}^{m_n} (\beta_k - \beta_{0k})^2 \leq (j+1)^2 \epsilon_n^2 \right\}, \|\cdot\|_2\right) \leq \exp\left(m_n \log \frac{24}{\xi}\right).$$

For the second covering number, we obtain by lemma 6.4 in Belitser and Ghosal (2003) that

$$\mathcal{N}\left(\frac{\xi j \epsilon_n}{2}, \left\{(\beta_{m_n+1}, \cdots,) : \sum_{k=m_n+1}^{\infty} (\beta_k - \beta_{0k})^2 k^{2\alpha} \leq Q^2 \right\}, \|\cdot\|_2\right) \leq \exp\left\{\log[4(2\mathrm{e})^{2\alpha}] \left(\frac{6Q}{\xi j \epsilon_n}\right)^{1/\alpha}\right\},$$

We conclude that $N_{nj} \leq \exp(D_1 n \epsilon_n^2)$ for some constant $D_1 > 0$ by applying the fact that $\epsilon_n^{-1/\alpha} \asymp m_n \asymp n\epsilon_n^2 = n^{1/(2\alpha+1)}$, and hence,

$$\sum_{j=M}^{\infty} N_{nj} \exp(-Dnj^2\epsilon_n^2) \leq \exp(D_1 n\epsilon_n^2) \sum_{j=M}^{\infty} \int_{j-1}^{j} \exp(-Dn\epsilon_n^2 x^2)\mathrm{d}x \leq \exp(D_1 n\epsilon_n^2) \int_{M-1}^{\infty} \exp(-Dn\epsilon_n^2 x^2)\mathrm{d}x$$
$$\lesssim \exp(D_1 n\epsilon_n^2) \exp\left[-\frac{1}{2}D(M-1)^2 n\epsilon_n^2\right] \to 0.$$

as long as M is sufficiently large. Hence condition (2.3) holds.

We are now left to verify condition (2.4) with the same constant D. Set

$$\kappa = \min\left\{\left[\frac{8\mathrm{e}^2}{c_3}\left(2D + \frac{1}{\sigma^2}\right)\right]^{2\alpha+1}, \left[\frac{32\mathrm{e}^2}{Q^2}\left(2D + \frac{1}{\sigma^2}\right)\right]^{-(2\alpha+1)}\right\},$$

denote $\Pi^A(\cdot) = \Pi(\cdot \mid A)$ for any sequence $(A_\ell)_{\ell=1}^\infty$, and define the set $\mathcal{A}_n = \{A_\ell \leq \mathrm{e}^{-\ell^2}$ for all $\ell \geq \lfloor \log(m_n/2) \rfloor - 1\}$. It follows that

$$\Pi(\mathcal{F}_{m_n}^c(Q)) = \Pi(\mathcal{F}_{m_n}^c(Q) \mid A \in \mathcal{A}_n)\Pi(\mathcal{A}_n) + \Pi(\mathcal{F}_{m_n}^c(Q) \mid A \in \mathcal{A}_n^c)\Pi(\mathcal{A}_n^c) \leq \sup_{A \in \mathcal{A}_n} \Pi^A(\mathcal{F}_{m_n}^c(Q)) + \Pi(\mathcal{A}_n^c).$$

Using a similar argument as that in page 340 in Gao and Zhou (2016a), for sufficiently large n we have

$$\Pi(\mathcal{A}_n^c) \leq \sum_{\ell \geq \lfloor \log m_n/2 \rfloor - 1} \Pi\left(A_\ell > \mathrm{e}^{-\ell^2}\right) \leq \sum_{\ell \geq \lfloor \log m_n/2 \rfloor - 1} \exp\left(-c_3 \mathrm{e}^{\ell}\right) \leq \exp\left[-\frac{1}{2}c_3 \exp(\lfloor \log m_n/2 \rfloor - 1)\right]$$
$$\leq \exp\left[-\frac{1}{2}c_3 \exp(\log m_n - \log 2 - 2)\right] \leq \exp\left[-\frac{c_3}{8\mathrm{e}^2}\kappa^{-1/(2\alpha+1)} n^{1/(2\alpha+1)}\right] = \exp\left[-\left(2D + \frac{1}{\sigma^2}\right) n\epsilon_n^2\right],$$

where we use the fact that $\lfloor \log m_n \rfloor \geq \log m_n - 1$ and $m_n \geq (n\kappa^{-1})^{1/(2\alpha+1)}/2$ for sufficiently large n. For any $A \in \mathcal{A}_n$ and sufficiently large n, the following holds:

$$\Pi^A(\mathcal{F}^c_{m_n}(Q)) \leq \Pi\left(2\sum_{k=m_n+1}^{\infty}\beta_k^2 k^{2\alpha} + 2\sum_{k=m_n+1}^{\infty}\beta_{0k}^2 k^{2\alpha} > Q^2\right) \leq \Pi\left(\sum_{k=m_n+1}^{\infty}\beta_k^2 k^{2\alpha} > \frac{Q^2}{4}\right),$$

since for sufficiently large n, $\sum_{k=m_n+1}^{\infty}\beta_{0k}^2 k^{2\alpha} < Q^2/4$. Write

$$\begin{aligned}
\left\{\sum_{k=m_n+1}^{\infty}\beta_k^2 k^{2\alpha} > \frac{Q^2}{4}\right\} &\subset \left\{\sum_{\ell:k_{\ell+1}\geq m_n}\sum_{k=k_\ell}^{k_{\ell+1}-1}\beta_k^2 k^{2\alpha} > \frac{Q^2}{4}\right\} \subset \left\{\sum_{\ell=\lfloor\log(m_n/2)\rfloor-1}^{\infty} k_{\ell+1}^{2\alpha}\sum_{k=k_\ell}^{k_{\ell+1}-1}\beta_k^2 > \frac{Q^2}{4}\right\} \\
&\subset \left\{\sum_{\ell=\lfloor\log(m_n/2)\rfloor-1}^{\infty} k_{\ell+1}^{2\alpha}\|\boldsymbol{\beta}_\ell\|^2 > \sum_{\ell=\lfloor\log(m_n/2)\rfloor-1}^{\infty}\frac{Q^2}{\ell^2}\right\} \\
&\subset \bigcup_{\ell=\lfloor\log(m_n/2)\rfloor-1}^{\infty}\left\{\ell^2 k_{\ell+1}^{2\alpha}\|\boldsymbol{\beta}_\ell\|^2 > Q^2\right\} = \bigcup_{\ell=\lfloor\log(m_n/2)\rfloor-1}^{\infty}\left\{\ell^2 k_{\ell+1}^{2\alpha}A_\ell\chi_\ell^2(n_\ell) > Q^2\right\},
\end{aligned}$$

since for sufficiently large n, $\sum_{\ell=\lfloor\log(m_n/2)\rfloor-1}^{\infty}\ell^{-2} < 1/4$, and $\chi_\ell^2(n_\ell)$ are independent $\chi^2(n_\ell)$-random variables. We proceed to compute

$$\begin{aligned}
\Pi^A(\mathcal{F}^c_{m_n}(Q)) &\leq \sum_{\ell=\lfloor\log(m_n/2)\rfloor-1}^{\infty}\Pi\left(2^{2\alpha}\mathrm{e}^{2\alpha(\ell+1)}\ell^2\mathrm{e}^{-\ell^2}\chi_\ell^2(n_\ell) > Q^2\right) \\
&= \sum_{\ell=\lfloor\log(m_n/2)\rfloor-1}^{\infty}\Pi\left(\exp(2\alpha\log 2 + 2\alpha(\ell+1) + 2\log\ell - \ell^2)\chi_\ell^2(n_\ell) > Q^2\right) \\
&\leq \sum_{\ell=\lfloor\log(m_n/2)\rfloor-1}^{\infty}\Pi\left(\exp\left(-\frac{1}{2}\ell^2\right)\chi_\ell^2(n_\ell) > Q^2\right)
\end{aligned}$$

for sufficiently large n. By the Chernoff bound for χ^2-random variables, we obtain for sufficiently large n

$$\begin{aligned}
\Pi^A(\mathcal{F}^c_{m_n}(Q)) &\leq \sum_{\ell=\lfloor\log(m_n/2)\rfloor-1}^{\infty}\exp\left[-\frac{n_\ell}{4}\log n_\ell - \frac{Q^2\mathrm{e}^{\ell^2/2}}{2} + \frac{n_\ell}{2}\log(Q^2\mathrm{e}^{\ell^2/2})\right] \\
&\leq \sum_{\ell=\lfloor\log(m_n/2)\rfloor-1}^{\infty}\exp\left[-\frac{\mathrm{e}^\ell}{4}\log\mathrm{e}^\ell - \frac{Q^2\mathrm{e}^{\ell^2/2}}{2} + \mathrm{e}^\ell\log(Q^2\mathrm{e}^{\ell^2/2})\right] \\
&\leq \sum_{\ell=\lfloor\log(m_n/2)\rfloor-1}^{\infty}\exp\left(-\frac{Q^2}{4}\mathrm{e}^{\ell^2/2}\right) \leq \exp\left[-\frac{Q^2}{16}\exp\left(\lfloor\log(m_n/2)\rfloor - 1\right)\right] \\
&\leq \exp\left[-\left(2D + \frac{1}{\sigma^2}\right)n\epsilon_n^2\right],
\end{aligned}$$

where in the second inequality we use the fact that $\mathrm{e}^\ell \leq n_\ell \leq 2\mathrm{e}^\ell$ for sufficiently large n when $\ell \geq$

$\lfloor \log(m_n/2) \rfloor - 1$. Therefore we conclude that

$$\Pi(\mathcal{F}^c_{m_n}(Q)) \le \sup_{A \in \mathcal{A}_n} \Pi^A(\mathcal{F}^c_{m_n}(Q)) + \Pi(\mathcal{A}_n^c) \le 2 \exp\left[-\left(2D + \frac{1}{\sigma^2}\right) n\epsilon_n^2 \right],$$

and condition (2.4) holds with the same constant D. $\square$

C Proof of Theorem 3.3

Define

$$j_n = \left\lceil \frac{1}{2\alpha} \log_2\left(\frac{4^\alpha}{4^\alpha - 1} 8Q^2 \right) + \frac{1}{2\alpha+1} \log_2(n) \right\rceil.$$

Clearly one has $2^{j_n}\epsilon_n^2 \to 0$ as $n \to \infty$ since $\alpha > 1/2$. For any $\alpha > 1/2$ and $Q > 0$, the $(2, 2, \alpha)$-Besov ball $\mathfrak{B}^\alpha_{2,2}(Q)$ is defined as follows ((Gao and Zhou, 2016a)):

$$\mathfrak{B}^\alpha_{2,2}(Q) = \left\{ f(x) = \sum_{j=0}^\infty \sum_{k \in I_j} \beta_{jk} \psi_{jk}(x) : \sum_{j=0}^\infty 2^{2\alpha j} \sum_{k \in I_j} \beta_{jk}^2 \le Q^2 \right\},$$

Lemma C.1 *For j_n defined above, $\epsilon_n = n^{-1/(2\alpha+1)}$, and $f_0 \in \mathcal{B}^\alpha_{2,2}(Q)$ with $\alpha > 1/2$,*

$$\Pi\left(\sum_{j=j_n}^\infty \|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|^2 \le \frac{\epsilon_n^2}{2}, \sum_{j=j_n}^\infty 2^{j/2} \|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|_\infty \le Q \right) \ge \frac{1}{2}$$

holds for sufficiently large n.

Proof. The proof is similar to Lemma B.1 and is included here for the sake of completeness. First write by the union bound

$$\begin{aligned}
&\Pi\left(\sum_{j=j_n}^\infty \|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|^2 \le \frac{\epsilon_n^2}{2}, \sum_{j=j_n}^\infty 2^{j/2} \|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|_\infty \le Q \right) \\
&\ge \Pi\left(\sum_{j=j_n}^\infty \|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|^2 \le \frac{\epsilon_n^2}{2} \right) + \Pi\left(\sum_{j=j_n}^\infty 2^{j/2} \|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|_\infty \le Q \right) - 1 \\
&= \Pi\left(\sum_{j=j_n}^\infty \|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|^2 \le \frac{\epsilon_n^2}{2} \right) - \Pi\left(\sum_{j=j_n}^\infty 2^{j/2} \|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|_\infty > Q \right).
\end{aligned}$$

By lemma G.4 in Gao and Zhou (2016b), we know that the fist term on the right-hand side of the proceeding display is $1 - o(1)$. Hence it suffices to show that the second term on the right-hand side is $o(1)$. Write

$$\Pi\left(\sum_{j=j_n}^\infty 2^{j/2} \|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|_\infty > Q \right) \le \frac{1}{Q} \sum_{j=j_n}^\infty 2^{j/2} \mathbb{E}_\Pi \|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|_2 \le \frac{1}{Q} \sum_{j=j_n}^\infty 2^{j/2} \left(2\mathbb{E}_\Pi \|\boldsymbol{\beta}_j\|_2^2 + 2\|\boldsymbol{\beta}_{0j}\|_2^2 \right)^{1/2}$$

$$\le \frac{\sqrt{2}}{Q}\sum_{j=j_n}^{\infty} 2^{j/2}\sqrt{\mathbb{E}_\Pi \|\boldsymbol{\beta}_j\|_2^2} + \frac{\sqrt{2}}{Q}\sum_{j=j_n}^{\infty} 2^{j/2}\|\boldsymbol{\beta}_{0j}\|$$
$$\le \frac{\sqrt{2}}{Q}\sum_{j=j_n}^{\infty} 2^{j/2}\sqrt{\mathbb{E}_\Pi \|\boldsymbol{\beta}_j\|_2^2} + \frac{\sqrt{2}}{Q}\left(\sum_{j=j_n}^{\infty} 2^{(1-2\alpha)j}\right)^{1/2}\left(\sum_{j=j_n}^{\infty} 2^{2\alpha j}\|\boldsymbol{\beta}_{0j}\|^2\right)^{1/2}$$
$$= \frac{\sqrt{2}}{Q}\sum_{j=j_n}^{\infty} 2^{j/2}\sqrt{\mathbb{E}_\Pi \|\boldsymbol{\beta}_j\|_2^2} + o(1),$$

where the last inequality is due to the fact that $f_0 \in \mathcal{B}_{2,2}^{\alpha}(Q)$ with $\alpha > 1/2$. We are now left with showing $\sum_{j\ge j_n} 2^{j/2}\sqrt{\mathbb{E}_\Pi\|\boldsymbol{\beta}_j\|^2} = o(1)$:

$$\sum_{j=j_n}^{\infty} 2^{j/2}\sqrt{\mathbb{E}_\Pi\|\boldsymbol{\beta}_j\|^2} \le \sum_{j=j_n}^{\infty} 2^{j}\sqrt{\mathbb{E}_\Pi(A_j)} \lesssim \sum_{j=j_n}^{\infty} 2^{j/2 - c_2 j^2/2} = o(1),$$

where the last inequality is due to (3.3). □

The following Lemma is the wavelet counterpart of Lemma B.2. The proof is included for completeness.

Lemma C.2 *For the block prior Π for the wavelet series defined in section 3.2 with $f_0 \in \mathcal{H}_\alpha(Q)$ for some $\alpha > 1/2$ and $Q > 0$, there exists some constant $D > 0$ such that $\Pi\left(B_n(2^{j_n}, \epsilon_n, Q)\right) \ge \exp(-Dn\epsilon_n^2)$, where $\epsilon_n = n^{-\alpha/(2\alpha+1)}$ and*

$$j_n = \left\lceil \frac{1}{2\alpha}\log_2\left(\frac{4^\alpha}{4^\alpha - 1}8Q^2\right) + \frac{1}{2\alpha+1}\log_2(n)\right\rceil.$$

Proof. Observe that for any $f(x) = \sum_{j=0}^{\infty}\sum_{k\in I_j}\beta_{0jk}\psi_{jk}(x)$ such that

$$\sum_{j=0}^{\infty}\|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|_2^2 \le \epsilon_n^2, \quad \sum_{j=j_n}^{\infty} 2^{j/2}\|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|_\infty \le Q,$$

we have,

$$\|f - f_0\|_\infty \le \sum_{j=0}^{j_n-1} 2^{j/2}\max_{k\in I_j}|\beta_{jk} - \beta_{0jk}| + \omega \le \sum_{j=0}^{j_n-1} 2^{j/2}\left[\sum_{k\in I_j}(\beta_{jk} - \beta_{0jk})^2\right]^{1/2} + \omega$$
$$\le \left(\sum_{j=0}^{j_n-1} 2^j\right)^{1/2}\left[\sum_{j=0}^{j_n-1}\sum_{k\in I_j}(\beta_{jk} - \beta_{0jk})^2\right]^{1/2} + \omega \le 2^{j_n/2}\epsilon_n + \omega = O(1).$$

Then by lemma C.1 we obtain

$$\Pi(B_n(2^{j_n}, \epsilon_n, Q)) \ge \Pi\left(\sum_{j=0}^{\infty}\|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|_2^2 \le \epsilon_n^2, \sum_{j=j_n}^{\infty} 2^{j/2}\|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|_\infty \le Q\right)$$
$$\ge \Pi\left(\sum_{j=j_n}^{\infty}\|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|_2^2 \le \frac{\epsilon_n^2}{2}, \sum_{j=j_n}^{\infty} 2^{j/2}\|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|_\infty \le Q\right)\Pi\left(\sum_{j=0}^{j_n-1}\|\boldsymbol{\beta}_j - \boldsymbol{\beta}_{0j}\|_2^2 \le \frac{\epsilon_n^2}{2}\right)$$

$$\geq \frac{1}{2}\Pi\left(\sum_{j=0}^{j_n-1}\|\boldsymbol{\beta}_j-\boldsymbol{\beta}_{0j}\|_2^2\leq\frac{\epsilon_n^2}{2}\right).$$

Exploiting the proof of (6) in theorem 2.1 in Gao and Zhou (2016a) and together with lemma G.2 and lemma G.3 in Gao and Zhou (2016b), we see that

$$\Pi\left(\sum_{j=0}^{j_n-1}\|\boldsymbol{\beta}_j-\boldsymbol{\beta}_{0j}\|_2^2\leq\frac{\epsilon_n^2}{2}\right)\geq\exp(-D'n\epsilon_n^2)$$

for some constant $D'>0$, and thus the proof is completed. □

Proof of Theorem 3.3. The proof of Theorem 3.3 is very similar to that of Theorem 3.2 and is included here for completeness. We use basically the same setup as that in the proof of theorem 3.2. Set $\epsilon_n=\underline{\epsilon}_n=n^{-\alpha/(2\alpha+1)}$, $\delta=\omega=Q$, j_n defined as in lemma C.2, $J_n=\lceil\log_2(n\kappa^{-1})/(2\alpha+1)\rceil$, $m_n=2^{J_n}$, and $k_n=2^{j_n}$, where κ is a constant determined later. Clearly, $J_n\leq\log_2(n\kappa^{-1})/(2\alpha+1)+1\leq\log_2[2(n\kappa^{-1})^{1/(2\alpha+1)}]$, and hence $m_n\epsilon_n^2=2^{J_n}\epsilon_n^2\leq 2(n\kappa^{-1})^{1/(2\alpha+1)}n^{-2\alpha/(2\alpha+1)}\to 0$, $k_n\underline{\epsilon}_n^2=2^j_n\underline{\epsilon}_n^2=O(1)$, and $\delta=\omega=O(1)$, since $\alpha>1/2$. By assumption f_0 is α-Sobolev. Using the relation between Besov space and Sobolev space, we see that there exists some $Q>0$ such that $f\in\mathfrak{B}_{2,2}^{\alpha}(Q)$, and hence yields the following series expansion

$$f_0(x)=\sum_{j=0}^{\infty}\sum_{k\in I_j}\beta_{0jk}\psi_{jk}(x),\quad\text{where}\quad\sum_{j=0}^{\infty}2^{2\alpha j}\sum_{k\in I_j}\beta_{0jk}^2\leq Q^2.$$

Denote $\lambda=2^j+k$ for each (j,k)-pair, and write $\beta_{2^j+k}=\beta_{jk}$, $\beta_{0,2^j+k}=\beta_{0jk}$, $\psi_\lambda(x)=\psi_{jk}(x)$. Since $I_j=\{0,1,\cdots,2^j-1\}$, $(j,k)\mapsto\lambda=2^j+k$ is one-to-one and hence the two index notations are equivalent. Thus we shall use the two indexes interchangeably. For condition (2.5), by lemma C.2 we see that it holds for some constant $D>0$ with $k_n\underline{\epsilon}_n^2=2^{j_n}\underline{\epsilon}_n^2=O(1)$ and $\omega=Q=O(1)$. For condition (2.3) and (2.4), We use a slightly different sieve $\mathcal{F}_{J_n}(Q)$ than that in the proof of theorem 3.2 as follows:

$$\mathcal{F}_{J_n}(Q)=\left\{f(x)=\sum_{j=0}^{\infty}\sum_{k\in I_j}\beta_{jk}\psi_{jk}(x):\sum_{j=J_n}^{\infty}2^{2\alpha j}\sum_{k\in I_j}(\beta_k-\beta_{0k})^2\leq Q^2\right\}.$$

We first argue that $\mathcal{F}_{J_n}(Q)$ satisfies (2.2) with $m=2^{J_n}$ and $\delta=Q$. In fact, for sufficiently large n, $\sum_{j=J_n}^{\infty}2^{(1-2\alpha)j}\leq 1$, and hence

$$\begin{aligned}\sum_{j=J_n}^{\infty}2^{j/2}\max_{k\in I_j}|\beta_{jk}-\beta_{0jk}|&\leq\sum_{j=J_n}^{\infty}2^{j/2-\alpha j}\left[2^{2\alpha j}\sum_{k\in I_j}(\beta_{jk}-\beta_{0jk})^2\right]^{1/2}\\&\leq\left[\sum_{j=J_n}^{\infty}2^{(1-2\alpha)j}\right]^{1/2}\left[\sum_{j=J_n}^{\infty}2^{2\alpha j}\sum_{k\in I_j}(\beta_{jk}-\beta_{0jk})^2\right]^{1/2}\leq Q.\end{aligned}$$

It follows that

$$
\begin{aligned}
\|f-f_0\|_\infty &\le \sum_{j=0}^{\infty} 2^{j/2}\max_{k\in I_j}|\beta_{jk}-\beta_{0jk}| \le \sum_{j=0}^{J_n-1} 2^{j/2}\max_{k\in I_j}|\beta_{jk}-\beta_{0jk}| + Q\\
&\le \sum_{j=0}^{J_n-1} 2^{j/2}\left[\sum_{k\in I_j}(\beta_{jk}-\beta_{0jk})^2\right]^{1/2} + Q\\
&\le \left(\sum_{j=0}^{J_n-1} 2^{j}\right)^{1/2}\left[\sum_{j=0}^{J_n-1}\sum_{k\in I_j}(\beta_{jk}-\beta_{0jk})^2\right]^{1/2} + Q \le 2^{J_n/2}\|f-f_0\|_2+\delta,
\end{aligned}
$$

implying that $\|f-f_0\|_\infty^2 \lesssim 2^{J_n}\|f-f_0\|_2^2 + Q^2$. Namely, $\mathcal{F}_{J_n}(Q)$ satisfies the property (2.2) with $m = 2^{J_n}$ and $\delta = Q$.

We next verify condition (2.3). Similar to the proof of theorem 4.2, we have

$$
\begin{aligned}
N_{nj} \le{}& \mathcal{N}\left(\frac{\xi j\epsilon_n}{2}, \left\{(\beta_1,\cdots,\beta_{2^{J_n}-1}) : \sum_{\lambda=1}^{2^{J_n}-1}(\beta_\lambda-\beta_{0\lambda})^2 \le (j+1)^2\epsilon_n^2\right\}, \|\cdot\|_2\right)\\
&\times \mathcal{N}\left(\frac{\xi j\epsilon_n}{2}, \left\{(\beta_{2^{J_n}},\cdots,) : \sum_{j=J_n}^{\infty} 2^{2\alpha j}\sum_{k\in I_j}(\beta_{jk}-\beta_{0jk})^2 \le Q^2\right\}, \|\cdot\|_2\right).
\end{aligned}
$$

We now bound the two covering number separately. For the first covering number, lemma 4.1 in Pollard (1990) yields

$$
\mathcal{N}\left(\frac{\xi j\epsilon_n}{2}, \left\{(\beta_1,\cdots,\beta_{2^{J_n}-1}) : \sum_{\lambda=1}^{2^{J_n}-1}(\beta_\lambda-\beta_{0\lambda})^2 \le (j+1)^2\epsilon_n^2\right\}, \|\cdot\|_2\right) \le \exp\left(2^{J_n}\log\frac{24}{\xi}\right).
$$

For the second covering number, we first observe that

$$
\begin{aligned}
\sum_{\lambda=2^{J_n}}^{\infty}\lambda^{2\alpha}(\beta_\lambda-\beta_{0\lambda})^2 &= \sum_{j=J_n}^{\infty}\sum_{k\in I_j}(2^j+k)^{2\alpha}(\beta_{jk}-\beta_{0jk})^2 \le \sum_{j=J_n}^{\infty}\sum_{k\in I_j}2^{2\alpha(j+1)}(\beta_{jk}-\beta_{0jk})^2\\
&= 2^{2\alpha}\sum_{j=J_n}^{\infty}2^{2\alpha j}\sum_{k\in I_j}(\beta_{jk}-\beta_{0jk})^2 \le 2^{2\alpha}Q^2,
\end{aligned}
$$

and then apply lemma 6.4 in Belitser and Ghosal (2003) to obtain

$$
\begin{aligned}
&\mathcal{N}\left(\frac{\xi j\epsilon_n}{2}, \left\{(\beta_{2^{J_n}},\cdots) : \sum_{j=J_n}^{\infty}2^{2\alpha j}\sum_{k\in I_j}(\beta_k-\beta_{0k})^2 \le Q^2\right\}, \|\cdot\|_2\right)\\
&\quad\le \mathcal{N}\left(\frac{\xi j\epsilon_n}{2}, \left\{(\beta_{2^{J_n}},\cdots) : \sum_{\lambda=2^{J_n}}^{\infty}\lambda^{2\alpha}(\beta_\lambda-\beta_{0\lambda})^2 \le (2^\alpha Q)^2\right\}, \|\cdot\|_2\right) \le \exp\left\{\log[4(2\mathrm{e})^{2\alpha}]\left(\frac{2^\alpha 6Q}{\xi j\epsilon_n}\right)^{1/\alpha}\right\}.
\end{aligned}
$$

We conclude that $N_{nj} \le \exp(D_1 n\epsilon_n^2)$ for some constant $D_1 > 0$ by applying the fact that $\epsilon_n^{-1/\alpha} \asymp 2^{J_n} \asymp$

$n\epsilon_n^2 = n^{1/(2\alpha+1)}$, and hence,

$$\sum_{j=M}^{\infty} N_{nj}\exp(-Dnj^2\epsilon_n^2) \le \exp(D_1 n\epsilon_n^2)\sum_{j=M}^{\infty}\int_{j-1}^{j}\exp(-Dn\epsilon_n^2 x^2)\mathrm{d}x \le \exp(D_1 n\epsilon_n^2)\int_{M-1}^{\infty}\exp(-Dn\epsilon_n^2 x^2)\mathrm{d}x$$

$$\lesssim \exp(D_1 n\epsilon_n^2)\exp\left[-\frac{1}{2}D(M-1)^2 n\epsilon_n^2\right] \to 0$$

as long as M is sufficiently large. Hence condition (2.3) holds.

We are now left to verify condition (2.4) with the same constant D. Set

$$\kappa = \min\left\{\left[\frac{4}{\log 2}\left(2D+\frac{1}{\sigma^2}\right)\right]^{-(2\alpha+1)}, \left[\frac{32}{Q^2}\left(2D+\frac{1}{\sigma^2}\right)\right]^{-(2\alpha+1)}\right\},$$

denote $\Pi^A(\cdot) = \Pi(\cdot \mid A)$ for any sequence $(A_j)_{j=0}^{\infty}$, and define the set

$$\mathcal{A}_n = \left\{A_j \le \exp(-j^2\log 2) \text{ for all } j \ge J_n\right\}.$$

It follows that

$$\Pi(\mathcal{F}_{J_n}^c(Q)) = \Pi(\mathcal{F}_{J_n}^c(Q) \mid A \in \mathcal{A}_n)\Pi(\mathcal{A}_n) + \Pi(\mathcal{F}_{J_n}^c(Q) \mid A \in \mathcal{A}_n^c)\Pi(\mathcal{A}_n^c) \le \sup_{A\in\mathcal{A}_n}\Pi^A(\mathcal{F}_{J_n}^c(Q)) + \Pi(\mathcal{A}_n^c).$$

Using a similar argument as that in page 340 in Gao and Zhou (2016a), for sufficiently large n we have

$$\Pi(\mathcal{A}_n^c) \le \sum_{j\ge J_n}\Pi\left(A_j > \exp(-j^2\log 2)\right) \le \sum_{j\ge J_n}\exp\left(-2^j\log 2\right) \le \exp\left(-\frac{1}{2}2^{J_n}\log 2\right)$$

$$\le \exp\left[-\frac{\log 2}{4}\kappa^{-1/(2\alpha+1)}n^{1/(2\alpha+1)}\right] \le \exp\left[-\left(2D+\frac{1}{\sigma^2}\right)n\epsilon_n^2\right],$$

where we use the fact that $J_n \ge \log_2[(n\kappa^{-1})^{1/(2\alpha+1)}] - 1$ and $2^{J_n} \ge (n\kappa^{-1})^{1/(2\alpha+1)}/2$ for sufficiently large n. For any $A \in \mathcal{A}_n$ and sufficiently large n, the following holds:

$$\Pi^A(\mathcal{F}_{J_n}^c(Q)) \le \Pi\left(2\sum_{j=J_n}^{\infty}2^{2\alpha j}\sum_{k\in I_j}\beta_{jk}^2 + 2\sum_{j=J_n}^{\infty}2^{2\alpha j}\sum_{k\in I_j}\beta_{0jk}^2 > Q^2\right) \le \Pi\left(\sum_{j=J_n}^{\infty}2^{2\alpha j}\sum_{k\in I_j}\beta_{jk}^2 > \frac{Q^2}{4}\right),$$

since for sufficiently large n, $\sum_{j=J_n}^{\infty}2^{2\alpha j}\sum_{k\in I_j}\beta_{0jk}^2 < Q^2/4$. Write

$$\left\{\sum_{j=J_n}^{\infty}2^{2\alpha j}\sum_{k\in I_j}\beta_{jk}^2 > \frac{Q^2}{4}\right\} = \left\{\sum_{j=J_n}^{\infty}2^{2\alpha j}\|\boldsymbol{\beta}_j\|^2 > \frac{Q^2}{4}\right\} \subset \left\{\sum_{j=J_n}^{\infty}2^{2\alpha j}\|\boldsymbol{\beta}_j\|^2 > \sum_{j=J_n}^{\infty}\frac{Q^2}{j^2}\right\}$$

$$\subset \bigcup_{j=J_n}^{\infty}\left\{j^2 2^{2\alpha j}\|\boldsymbol{\beta}_j\|^2 > Q^2\right\} = \bigcup_{j=J_n}^{\infty}\left\{j^2 2^{2\alpha j}A_j\chi_j^2(2^j) > Q^2\right\},$$

since for sufficiently large n, $\sum_{j=J_n}^{\infty}j^{-2} < 1/4$, and $\chi_j^2(2^j)$ are independent $\chi^2(2^j)$-random variables. We proceed to compute

$$\begin{aligned}\Pi^A(\mathcal{F}_{J_n}^c(Q)) &\leq \sum_{j=J_n}^{\infty}\Pi\left(j^2 2^{2\alpha j}A_j\chi_j^2(2^j) > Q^2\right) \leq \sum_{j=J_n}^{\infty}\Pi\left(\exp(2\log j + 2\alpha j\log 2 - j^2\log 2)\chi_j^2(2^j) > Q^2\right)\\ &\leq \sum_{j=J_n}^{\infty}\Pi\left(\exp\left(-\frac{j^2}{2}\right)\chi_j^2(2^j) > Q^2\right)\end{aligned}$$

for sufficiently large n. By the Chernoff bound for χ^2-random variables, we obtain for sufficiently large n

$$\begin{aligned}\Pi^A(\mathcal{F}_{J_n}^c(Q)) &\leq \sum_{j=J_n}^{\infty}\exp\left[-\frac{2^j}{4}j\log 2 - \frac{Q^2 \mathrm{e}^{j^2/2}}{2} + \frac{2^j}{2}\log\left(Q^2\mathrm{e}^{j^2/2}\right)\right] \leq \sum_{j=J_n}^{\infty}\exp\left(-\frac{Q^2}{4}\mathrm{e}^{j^2/2}\right)\\ &\leq \exp\left(-\frac{Q^2}{16}2^{J_n}\right) \leq \exp\left[-\frac{Q^2 n^{1/(2\alpha+1)}\kappa^{-1/(2\alpha+1)}}{32}\right] \leq \exp\left[-\left(2D + \frac{1}{\sigma^2}\right)n\epsilon_n^2\right],\end{aligned}$$

where we use the fact that $2^{J_n} \geq (n\kappa^{-1})^{1/(2\alpha+1)}/2$ for sufficiently large n when $j \geq J_n$. Therefore we conclude that

$$\Pi(\mathcal{F}_{J_n}^c(Q)) \leq \sup_{A\in\mathcal{A}_n}\Pi^A(\mathcal{F}_{J_n}^c(Q)) + \Pi(\mathcal{A}_n^c) \leq 2\exp\left[-\left(2D + \frac{1}{\sigma^2}\right)n\epsilon_n^2\right],$$

and condition (2.4) holds with the same constant D. □

D Proof of Theorem 3.4

Proof of Theorem 3.4. Take $\epsilon_n = \underline{\epsilon}_n = n^{-\alpha/(2\alpha+1)}(\log n)^{1/2}$ and $m = m_n = n^{1/(2\alpha+1)}$. Clearly, $m_n^{-\alpha} \lesssim \epsilon_n$. It follows from Lemma 3.2 in De Jonge and Van Zanten (2012) that there exists some constant $D > 0$, such that

$$\Pi(\|f - f_0\|_2 < \epsilon_n) \geq \exp\left[-Dm_n\left(\log\frac{1}{\epsilon_n}\right)\right] \geq \exp\left[-Dn^{1/(2\alpha+1)}\log n\right] = \exp(-Dn\epsilon_n^2).$$

Now set the sieve to be $\mathcal{F}_n = \{f(x) = \sum_{k=1}^{m_n}\beta_k B_k(x) : \beta_1, \dots, \beta_{m_n} \in \mathbb{R}\}$. It suffices to show that condition (2.3) holds with the same constant D, as condition (2.4) automatically holds since $\mathcal{F}_n^c = \varnothing$.

To begin with, we first show that $\mathcal{F}_n$ satisfies condition (2.2) with $m = m_n$ and some constant $\delta > 0$. Let $f(x) = \sum_{k=1}^{m_n}\beta_k B_k(x) \in \mathcal{F}_n$. Since $f_0 \in \mathfrak{C}_\alpha(Q)$, then by Lemma 3.1 there exists $\beta_{01}, \dots, \beta_{0m_n} \in \mathbb{R}$ such that

$$\left\|\sum_{k=1}^{m_n}\beta_{0k}B_k(x) - f_0\right\|_\infty \leq C(f_0)m_n^{-\alpha}$$

for some constant $C(f_0)$ only depending on f_0. Furthermore, for the spline functions $f(x) = \sum_{k=1}^{m_n}\beta_k B_k(x)$ and $g_0(x) = \sum_{k=1}^{m_n}\beta_{0k}B_k(x)$, Lemma 3.1 shows that

$$\max_{1\leq k\leq m_n}|\beta_k - \beta_{0k}| \asymp \|f - g_0\|_\infty, \quad \left[\sum_{k=1}^{m_n}(\beta_k - \beta_{0k})^2\right]^{1/2} \asymp \sqrt{m_n}\|f - g_0\|_2. \tag{1}$$

Therefore,

$$
\begin{aligned}
\|f-f_0\|_\infty &\le \|f-g_0\|_\infty + \|g_0-f_0\|_\infty \lesssim \max_{1\le k\le m_n} |\beta_k-\beta_{0k}| + C(f_0)m_n^{-\alpha} \\
&\le \left[\sum_{k=1}^{m_n}(\beta_k-\beta_{0k})^2\right]^{1/2} + C(f_0)m_n^{-\alpha} \lesssim \sqrt{m_n}\|f-g_0\|_2 + m_n^{-\alpha} \\
&\le \sqrt{m_n}(\|f-f_0\|_2 + \|f_0-g_0\|_\infty) + m_n^{-\alpha} \\
&\lesssim \sqrt{m_n}\|f-f_0\|_2 + m_n^{1/2-\alpha} + m_n^{-\alpha} \le \sqrt{m_n}\|f-f_0\|_2 + 1
\end{aligned}
$$

for sufficiently large n, as we are assuming that $\alpha > 1/2$. This shows that $\mathcal{F}_{m_n}$ satisfies condition (2.2).

We next estimate the covering number $N_{nj} = \mathcal{N}(\xi j\epsilon_n, \mathcal{S}_{nj}(\epsilon_n), \|\cdot\|_2)$. By construction and (1),

$$
\begin{aligned}
\mathcal{S}_{nj}(\epsilon_n) &\subset \left\{ f(x) = \sum_{k=1}^{m_n}\beta_k B_k(x) : \|f-f_0\|_2 \le (j+1)\epsilon_n \right\} \\
&\subset \left\{ f(x) = \sum_{k=1}^{m_n}\beta_k B_k(x) : \|f-g_0\|_2 - \|g_0-f_0\|_\infty \le (j+1)\epsilon_n \right\} \\
&= \left\{ f(x) = \sum_{k=1}^{m_n}\beta_k B_k(x) : \|f-g_0\|_2 \le (j+1)\epsilon_n + C(f_0)m_n^{-\alpha} \right\} \\
&\subset \left\{ f(x) = \sum_{k=1}^{m_n}\beta_k B_k(x) : \|f-g_0\|_2 \le 2j\epsilon_n \right\} \\
&\subset \left\{ f(x) = \sum_{k=1}^{m_n}\beta_k B_k(x) : \left[\sum_{k=1}^{m_n}(\beta_k-\beta_{0k})^2\right]^{1/2} \le C_2 j\sqrt{m_n}\epsilon_n \right\}
\end{aligned}
$$

for some constant $C_2 > 0$. Hence, applying the covering number bound for Euclidean ball and the fact that

$$
\left\|\sum_{k=1}^{\infty}\beta_k B_k(x)\right\|_2 \le \frac{1}{C_1\sqrt{m_n}}\left(\sum_{k=1}^{\infty}\beta_k^2\right)^{1/2}
$$

for some constant $C_1 > 0$, we obtain

$$
\begin{aligned}
\mathcal{N}_{jn} &\le \mathcal{N}\left(\xi j\epsilon_n, \left\{\sum_{k=1}^{m_n}\beta_k B_k(x) : \left[\sum_{k=1}^{m_n}(\beta_k-\beta_{0k})^2\right]^{1/2} \le C_2 j\sqrt{m_n}\epsilon_n\right\}, \|\cdot\|_2\right) \\
&\le \mathcal{N}\left(\xi j\epsilon_n, \left\{(\beta_1,\ldots,\beta_{m_n}) : \left[\sum_{k=1}^{m_n}(\beta_k-\beta_{0k})^2\right]^{1/2} \le C_2 j\sqrt{m_n}\epsilon_n\right\}, \frac{\|\cdot\|_2}{C_1\sqrt{m_n}}\right) \\
&\le \left(\frac{6C_2 j\sqrt{m_n}\epsilon_n}{C_1\xi j\sqrt{m_n}\epsilon_n}\right)^{m_n} = \exp\left(m_n \log\frac{6C_2}{C_1\xi}\right).
\end{aligned}
$$

Hence, there exists some constant $C_3 > 0$ such that

$$\begin{aligned}\sum_{j=M}^{\infty} N_{nj}\exp(-Dnj^2\epsilon_n^2) &\le \exp(C_3 m_n)\sum_{j=M}^{\infty}\int_{j-1}^{j}\exp(-Dn\epsilon_n^2 x^2)\mathrm{d}x\\ &= \exp(C_3 m_n)\int_{M-1}^{\infty}\exp(-Dn\epsilon_n^2 x^2)\mathrm{d}x\\ &\le \exp\left[C_3 m_n - \frac{1}{2}D(M-1)^2 n\epsilon_n^2\right]\int_{M-1}^{\infty}\exp\left(-\frac{1}{2}Dn\epsilon_n^2 x^2\right)\mathrm{d}x\\ &= \exp\left[C_3 m_n - \frac{1}{2}D(M-1)^2 n\epsilon_n^2\right]\sqrt{\frac{2\pi}{Dn\epsilon_n^2}} \to 0.\end{aligned}$$

This shows that condition (2.3) also holds. The proof is thus completed. □

E Proof of Theorem 4.1

The proof of Theorem 3.4 is immediate by combining the proof of Theorem 2.1, Lemma F.1, Lemma E.1, Lemma E.2, and Lemma E.3. The proofs of these lemmas are similar to their counterparts in the manuscript, and are presented here for completeness. The following lemma is the key ingredient in bridging the gap between the empirical L_2-distance and the integrated L_2-distance.

Lemma E.1 *Suppose the design points $(x_i)_{i=1}^n$ are fixed and satisfy (4.1). Let $\mathcal{F}_{m_n}(\delta)$ be defined as (4.2) with $m_n \to \infty$, $m_n/n \to 0$, and δ is some constant. Then for any $f_1 \in \mathcal{F}_{m_n}(\delta)$ with $\sqrt{n}\|f_1 - f_0\|_2 > 1$, there exists a test function $\phi_n : \mathcal{Y}^n \to [0,1]$ such that*

$$\mathbb{E}_0\phi_n \le \exp\left(-Cn\|f_1 - f_0\|_2^2\right), \quad \sup_{\{f\in\mathcal{F}_{m_n}(\delta):\|f-f_1\|_2^2\le\xi^2\|f_0-f_1\|_2^2\}}\mathbb{E}_f(1-\phi_n) \le \exp\left(-Cn\|f_1 - f_0\|_2^2\right)$$

for some constant $C > 0$ and ξ in $(0,1)$.

Proof. Take $\xi = 1/8$. Since $\|f_0 - f_1\| > 1/\sqrt{n}$. We obtain from the assumption that

$$\left|\mathbb{P}_n(f_0 - f_1)^2 - \|f_0 - f_1\|_2^2\right| \le \eta\left(\frac{m_n}{n}\|f_0 - f_1\|_2^2 + \frac{\delta}{\sqrt{n}}\|f_1 - f_0\|_2^2\right) \le \frac{1}{4}\|f_1 - f_0\|_2^2$$

when n is sufficiently large, implying that $(3/4)\|f_1 - f_0\|_2^2 \le \mathbb{P}_n(f_0 - f_1)^2 \le (5/4)\|f_1 - f_0\|_2^2$. On the other hand,

$$\begin{aligned}\mathbb{P}_n(f - f_1)^2 &\le \left(1 + \frac{\eta m_n}{n}\right)\|f - f_1\|_2^2 + \frac{\eta\delta}{n}\|f - f_1\|_2 \le \frac{5}{256}\|f_0 - f_1\|_2^2 + \frac{\eta\xi\delta}{n}\|f_0 - f_1\|_2\\ &\le \frac{20/3}{256}\mathbb{P}_n(f_0 - f_1)^2 + \frac{\eta\xi\delta}{n}\|f_0 - f_1\|_2 \le \frac{1}{32}\mathbb{P}_n(f_0 - f_1)^2 + \frac{\eta\xi\delta}{\sqrt{n}}\|f_0 - f_1\|_2^2 \le \frac{1}{16}\mathbb{P}_n(f_1 - f_0)^2.\end{aligned}$$

Define the test function to be $\phi_n = \mathbb{1}\{T_n > 0\}$, where

$$T_n = \sum_{i=1}^n y_i(f_1(x_i) - f_0(x_i)) - \frac{1}{2}n\mathbb{P}_n(f_1^2 - f_0^2).$$

We first consider the type I error probability. Under $\mathbb{P}_0$, we have $y_i = f_0(x_i) + e_i$, where e_i's are i.i.d. Gaussian errors with $\mathbb{E}e_i = 0$ and $\mathrm{Var}(e_i) = \sigma^2$. Namely, there exists a constant $C_1 > 0$ such that $\mathbb{P}_0(e_i > t) \leq \exp(-4C_1 t^2)$ for all $t > 0$. Then for a sequence $(a_i)_{i=1}^n \in \mathbb{R}^n$, Chernoff bound yields

$$\mathbb{P}_0\left(\sum_{i=1}^n a_i e_i \geq t\right) \leq \exp\left(-\frac{4C_1 t^2}{\sum_{i=1}^n a_i^2}\right).$$

Now we set $a_i = f_1(\mathbf{x}_i) - f_0(\mathbf{x}_i)$ and $t = n\mathbb{P}_n(f_1 - f_0)^2/2$. Then under $\mathbb{P}_0$, we have

$$\mathbb{E}_0(\phi_n) = \mathbb{P}_0\left(T_n > 0\right) \leq \exp\left(-C_1 n \mathbb{P}_n(f_0 - f_1)^2\right) \leq \exp\left(-\frac{C_1}{16} n\|f_0 - f_1\|_2^2\right).$$

We next consider the type II error probability. Under $\mathbb{P}_f$, we have $y_i = f(x_i) + e_i$ with e_i's being i.i.d. mean-zero sub-Gaussian. Since $\mathbb{P}_n(f - f_1)^2 \leq \mathbb{P}_n(f_1 - f_0)^2/16$, we obtain

$$\begin{aligned}
T_n &= \sum_{i=1}^n e_i\left[f_1(x_i) - f_0(x_i)\right] + n\mathbb{P}_n(f - f_1)(f_1 - f_0) + \frac{1}{2}n\mathbb{P}_n(f_1 - f_0)^2 \\
&\geq \sum_{i=1}^n e_i\left[f_1(x_i) - f_0(x_i)\right] + \frac{1}{2}n\mathbb{P}_n(f_1 - f_0)^2 - n\sqrt{\mathbb{P}_n(f - f_1)^2\mathbb{P}_n(f_1 - f_0)^2} \\
&\geq \sum_{i=1}^n e_i\left[f_1(x_i) - f_0(x_i)\right] + \frac{1}{4}n\mathbb{P}_n(f_1 - f_0)^2.
\end{aligned}$$

Hence we use the sub-Gaussian tail bound to obtain

$$\mathbb{P}_f\left(T_n < 0\right) \leq \mathbb{P}\left(\sum_{i=1}^n e_i[f_1(x_i) - f_0(x_i)] \leq -\frac{1}{4}n\mathbb{P}_n(f_1 - f_0)^2\right) \leq \exp\left(-\frac{C_1}{4}n\|f_1 - f_0\|_2^2\right).$$

Hence we obtain the following exponential bound for type I and type II error probabilities:

$$\mathbb{E}_0\phi_n \leq \exp(-Cn\|f_1 - f_0\|_2^2), \quad \mathbb{E}_f(1 - \phi_n) \leq \exp(-Cn\|f_1 - f_0\|_2^2)$$

for some constant $C > 0$ for any $f \in \{f \in \mathcal{F}_m(\delta) : \|f - f_1\|_2^2 \leq \|f_1 - f_0\|_2^2/64\}$. The proof is completed by taking the supremum over $\{f \in \mathcal{F}_m(\delta) : \|f - f_1\|_2^2 \leq \|f_1 - f_0\|_2^2/64\}$. $\square$

Lemma E.2 *Suppose the design points $(x_i)_{i=1}^n$ are fixed and satisfy (4.1). Let the sieve $\mathcal{F}_m(\delta)$ satisfy (4.2). Let $(\epsilon_n)_{n=1}^\infty$ be a sequence with $n\epsilon_n^2 \to \infty$. Then there exists a sequence of test functions $(\phi_n)_{n=1}^\infty$ such that*

$$\mathbb{E}_0\phi_n \leq \sum_{j=M}^\infty N_{nj}\exp\left(-Cnj^2\epsilon_n^2\right), \qquad \sup_{\{f\in\mathcal{F}_m(\delta):\|f-f_0\|_2 > M\epsilon_n\}} \mathbb{E}_f(1 - \phi_n) \leq \exp(-CM^2 n\epsilon_n^2),$$

where $N_{nj} = \mathcal{N}(\xi j\epsilon_n, \mathcal{S}_{nj}(\epsilon_n), \|\cdot\|_2)$, $\mathcal{S}_{nj}(\epsilon_n) = \{f \in \mathcal{F}_m(\delta) : j\epsilon_n < \|f - f_0\|_2 \leq (j+1)\epsilon_n\}$, M can be sufficiently large, and C is some positive constant.

Proof. The proof is exactly the same as that of Lemma 2.2 and is omitted here. $\square$

Lemma E.3 *Suppose the design points $(x_i)_{i=1}^n$ are fixed. If $n\epsilon_n^2 \to \infty$ then for any constant $C > 0$,*

$$\mathbb{P}_0\left(\int \prod_{i=1}^n \frac{\phi_\sigma(y_i - f(\mathbf{x}_i))}{\phi_\sigma(y_i - f_0(\mathbf{x}_i))}\Pi(\mathrm{d}f) \le \Pi\left(\|f - f_0\|_\infty < \epsilon_n\right)\exp\left[-\left(C + \frac{1}{\sigma^2}\right)\epsilon_n^2\right]\right) \to 0.$$

Proof. Denote the re-normalized restriction of Π on $B_n = \{\|f - f_0\|_\infty < \epsilon_n^2\}$ to be $\Pi(\cdot \mid B_n)$. Then by Jensen's inequality

$$\begin{aligned}
\mathcal{H}_n^c := & \left\{\int \prod_{i=1}^n \frac{\phi_\sigma(y_i - f(\mathbf{x}_i))}{\phi_\sigma(y_i - f_0(\mathbf{x}_i))}\Pi(\mathrm{d}f) \le \Pi(B_n)\exp\left[-\left(C + \frac{1}{\sigma^2}\right)n\epsilon_n^2\right]\right\} \\
\subset & \left\{\int \prod_{i=1}^n \frac{\phi_\sigma(y_i - f(\mathbf{x}_i))}{\phi_\sigma(y_i - f_0(\mathbf{x}_i))}\Pi(\mathrm{d}f \mid B_n) \le \exp\left[-\left(C + \frac{1}{\sigma^2}\right)n\epsilon_n^2\right]\right\} \\
\subset & \left\{\frac{1}{\sigma^2}\sum_{i=1}^n e_i\left[f_0(\mathbf{x}_i) - \int f(\mathbf{x}_i)\Pi(\mathrm{d}f \mid B_n)\right] \ge \left(C + \frac{1}{2\sigma^2}\right)n\epsilon_n^2\right\},
\end{aligned}$$

where we have used the fact that on the event B_n, $\|f - f_0\|_\infty \le \epsilon_n$, which implies,

$$\sum_{i=1}^n \int (f(\mathbf{x}_i) - f_0(\mathbf{x}_i))^2\Pi(\mathrm{d}f \mid B_n) \le \int n\|f - f_0\|_\infty^2\Pi(\mathrm{d}f \mid B_n) < n\epsilon_n^2.$$

Now we use the tail bound for sub-Gaussian random variables to obtain

$$\begin{aligned}
\mathbb{P}_0(\mathcal{H}_n^c) & \le \exp\left\{-\left(C + \frac{1}{2\sigma^2}\right)^2\sigma^4 n\epsilon_n^4\left[\mathbb{P}_n\left(f_0 - \int f\Pi(\mathrm{d}f \mid B_n)\right)^2\right]^{-1}\right\} \\
& \le \exp\left\{-\left(C + \frac{1}{2\sigma^2}\right)^2\sigma^4 n\epsilon_n^4\left[\mathbb{P}_n\int (f - f_0)^2\Pi(\mathrm{d}f \mid B_n)\right]^{-1}\right\} \\
& \le \exp\left\{-\left(C + \frac{1}{2\sigma^2}\right)^2\sigma^4 n\epsilon_n^4\left[\int \|f - f_0\|_\infty^2\Pi(\mathrm{d}f \mid B_n)\right]^{-1}\right\} \le \exp\left[-\left(C + \frac{1}{2\sigma^2}\right)^2\sigma^4 n\epsilon_n^2\right] \to 0.
\end{aligned}$$

$\square$

F Proof of Theorem 4.2

Before proceeding to the proof of Theorem 4.2, we need to introduce several fundamental concepts and properties of the squared-exponential Gaussian process in order to apply Theorem 4.1. The eigen-system of the squared-exponential covariance function has been studied in the literature (see, for example, Yang and Pati (2017)). Under the aforementioned Fourier basis, the covariance function K yields the following eigen-expansion $K(x, x') = \sum_{k=1}^\infty \lambda_k\psi_k(x)\psi_k(x')$, where the eigenvalues $(\lambda_k)_{k=1}^\infty$ decay at $\lambda_k \asymp \exp(-k^2/4)$. It follows by the Karhunen-Loève theorem that f yields a series expansion $f(x) = \sum_{k=1}^\infty \beta_k\psi_k(x)$, where $\beta_k \sim \mathrm{N}(0, \lambda_k)$.

Lemma F.1 *Suppose the design points $(x_i)_{i=1}^n$ are fixed and satisfy (4.1). For any $m \in \mathbb{N}_+$ and $\delta > 0$, let*

$\mathcal{F}_m(\delta)$ be a function class such that

$$\mathcal{F}_m(\delta) \subset \left\{ f(x) = \sum_{k=1}^{\infty} \beta_k \psi_k(x) : \sum_{k=1}^{\infty} |\beta_k - \beta_{0k}| k^{\alpha} \leq \delta \right\}, \tag{2}$$

and $f, f_1 \in \mathcal{F}_m(\delta)$. Then

$$\left|\mathbb{P}_n(f - f_1)^2 - \|f - f_1\|_2^2\right| \leq \frac{Cm}{n}\|f - f_1\|_2^2 + \frac{C\delta}{n}\|f - f_1\|_2,$$
$$\left|\mathbb{P}_n(f_0 - f_1)^2 - \|f_0 - f_1\|_2^2\right| \leq \frac{Cm}{n}\|f_0 - f_1\|_2^2 + \frac{C\delta}{n}\|f_0 - f_1\|_2$$

hold for some universal constant $C > 0$ independent of f_0, f_1, and f.

Proof. Observe that for any $f = \sum_k \beta_k \psi_k \in \mathcal{F}_m(\delta)$, the term-by-term dffierentiation operation is permitted, since

$$\sup_{x \in [0,1]} \left| \sum_{k=m_0+1}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x} \beta_k \psi_k(x) \right| \lesssim \sum_{k=m_0+1}^{\infty} k|\beta_k| \|\psi_k\|_\infty \lesssim \sum_{k=m_0+1}^{\infty} k^{\alpha}|\beta_{0k}| + \sum_{k=m_0+1}^{\infty} |\beta_k - \beta_{0k}| k^{\alpha}.$$

As $m_0 \to \infty$, the first term on the right-hand side of the preceeding display converges to 0 by the definition of $\mathfrak{C}_\alpha(Q)$, and the second term also converges to 0 by the definition of $\mathcal{F}_m(Q)$. Hence the series $\sum_k \mathrm{d}[\beta_k \psi_k(x)]/\mathrm{d}x$ converges uniformly over $x \in [0, 1]$.

Now suppose $f(x) = \sum_{k=1}^{\infty} \beta_k \psi_k(x)$ permits term-by-term differentiation. We proceed to compute

$$\frac{\mathrm{d}}{\mathrm{d}x} f(x) = \sum_{k=1}^{\infty} \left(\beta_{2k+1} \frac{\mathrm{d}}{\mathrm{d}x} \psi_{2k+1}(x) + \beta_{2k} \frac{\mathrm{d}}{\mathrm{d}x} \psi_{2k}(x) \right) = \sum_{k=1}^{\infty} \left(-\pi k \beta_{2k+1} \psi_{2k}(x) + \pi k \beta_{2k} \psi_{2k}(x) \right),$$

and hence,

$$\int_0^1 \left| f(x) \frac{\mathrm{d}}{\mathrm{d}x} f(x) \right| \mathrm{d}x \leq \left(\sum_{k=1}^{\infty} \beta_k^2 \right)^{1/2} \left(\pi^2 \sum_{k=1}^{\infty} k^2 \beta_{2k-1}^2 + k^2 \beta_{2k}^2 \right)^{1/2} \lesssim \|f\|_2 \left(\sum_{k=1}^{\infty} k^2 \beta_k^2 \right)^{1/2}.$$

Since the design points satisfy (4.1), lemma 7.1 in Yoo et al. (2017) yields

$$|\mathbb{P}_n f^2 - \|f\|_2^2| = \left| \frac{1}{n} \sum_{i=1}^{n} f^2(x_i) - \int_0^1 f^2(x) \mathrm{d}x \right| \lesssim \frac{1}{n} \int_0^1 2 \left| f(x) \frac{\mathrm{d}}{\mathrm{d}x} f(x) \right| \mathrm{d}x \lesssim \frac{1}{n} \|f\|_2 \left(\sum_{k=1}^{\infty} k^2 \beta_k^2 \right)^{1/2}.$$

Observing that f, f_0, and f_1 are all term-by-term differentiable, we obtain

$$\begin{aligned}
\left|\mathbb{P}_n(f - f_1)^2 - \|f - f_1\|_2^2\right| &\lesssim \frac{1}{n}\|f - f_1\|_2 \left[m^2 \sum_{k=1}^{m} (\beta_k - \beta_{1k})^2 + \sum_{k=m+1}^{\infty} k^2 (\beta_k - \beta_{1k})^2 \right]^{1/2} \\
&\leq \frac{1}{n}\|f - f_1\|_2 \left[m^2 \|f - f_1\|_2^2 + \sum_{k=m+1}^{\infty} k^2 (\beta_k - \beta_{1k})^2 \right]^{1/2}
\end{aligned}$$

$$\leq \frac{1}{n}\|f-f_1\|_2\left\{m\|f-f_1\|_2+\left[\sum_{k=m+1}^{\infty}k^2(\beta_k-\beta_{1k})^2\right]^{1/2}\right\},$$

and similarly,

$$\begin{aligned}\left|\mathbb{P}_n(f_0-f_1)^2-\|f_0-f_1\|_2^2\right| &\lesssim \frac{1}{n}\|f_0-f_1\|_2\left[m^2\sum_{k=1}^{m}(\beta_{0k}-\beta_{1k})^2+\sum_{k=m+1}^{\infty}k^2(\beta_{0k}-\beta_{1k})^2\right]^{1/2}\\ &\leq \frac{1}{n}\|f_0-f_1\|_2\left[m^2\|f_0-f_1\|_2^2+\sum_{k=m+1}^{\infty}k^2(\beta_{0k}-\beta_{1k})^2\right]^{1/2}\\ &\leq \frac{1}{n}\|f_0-f_1\|_2\left\{m\|f_0-f_1\|_2+\left[\sum_{k=m+1}^{\infty}k^2(\beta_{0k}-\beta_{1k})^2\right]^{1/2}\right\}.\end{aligned}$$

By the definition of $\mathcal{F}_m(\delta)$ and the fact that $\alpha > 1$, we have

$$\begin{aligned}\sum_{k=m+1}^{\infty}k^2(\beta_k-\beta_{1k})^2 &= \sum_{k=m+1}^{\infty}k^2(\beta_k-\beta_{0k}+\beta_{0k}-\beta_{1k})^2\\ &\leq 2\sum_{k=m+1}^{\infty}k^2(\beta_k-\beta_{0k})^2+2\sum_{k=m+1}^{\infty}k^2(\beta_{1k}-\beta_{0k})^2\\ &\leq 2\left(\sum_{k=m+1}^{\infty}k|\beta_k-\beta_{0k}|\right)^2+2\left(\sum_{k=m+1}^{\infty}k|\beta_{1k}-\beta_{0k}|\right)^2\leq 4\delta^2,\end{aligned}$$

and hence,

$$\begin{aligned}|\mathbb{P}_n(f-f_1)^2-\|f-f_1\|_2^2| &\leq \frac{Cm}{n}\|f-f_1\|_2^2+\frac{C\delta}{n}\|f-f_1\|_2,\\ |\mathbb{P}_n(f_0-f_1)^2-\|f_0-f_1\|_2^2| &\leq \frac{Cm}{n}\|f_0-f_1\|_2^2+\frac{C\delta}{n}\|f_0-f_1\|_2\end{aligned}$$

for some universal constant $C > 0$. □

The following lemma controling the covering number is also useful for proving Theorem 4.2.

Lemma F.2 *For all $\epsilon > 0$, it holds that*

$$\log\mathcal{N}\left(\epsilon,\left\{(\beta_1,\beta_2,\cdots)\in l^2:\sum_{k=1}^{\infty}\beta_k^2\mathrm{e}^{k^2/c}\leq Q^2\right\},\|\cdot\|_2\right)\lesssim\left(\log\frac{1}{\epsilon}\right)^{3/2}.$$

Proof. For a metric space (X, d), denote the packing number $\mathcal{D}(\epsilon, X, d)$ to be the maximum number of points in X that are at least ϵ away from each other. It is proved that $\mathcal{D}(\epsilon, X, d) \geq \mathcal{N}(\epsilon, X, d)$ Ghosal and Van Der Vaart (2001), and therefore it suffices to work with the packing number. The proof here is similar to that of lemma 6.4 in Belitser and Ghosal (2003).

Let

$$\Theta(Q) = \left\{ (\beta_1, \beta_2, \cdots) \in l^2 : \sum_{k=1}^{\infty} \beta_k^2 \exp\left(\frac{k^2}{c}\right) \leq Q^2 \right\}.$$

Suppose $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_m \in \Theta(Q)$ are such that $\|\boldsymbol{\beta}_i - \boldsymbol{\beta}_j\|_2 = \epsilon$ whenever $i \neq j$. It suffices to consider ϵ to be small enough, since for large values of ϵ, $\mathcal{D}(\epsilon, \Theta(Q), \|\cdot\|_2) = 1$. Fixed an integer N and denote

$$\Theta_N(Q) = \left\{ (\beta_1, \cdots, \beta_N, 0, \cdots) \in \mathbb{R}^N : \sum_{k=1}^{N} \beta_k^2 \exp\left(\frac{k^2}{c}\right) \leq Q^2 \right\}.$$

For any $\boldsymbol{\beta} = (\beta_1, \beta_2, \cdots) \in \Theta(Q)$, denote $\overline{\boldsymbol{\beta}} = (\beta_1, \cdots, \beta_N, 0, \cdots) \in \Theta_N(Q)$. Now set $N = \lfloor \sqrt{c \log(8Q^2/\epsilon^2)} \rfloor$. Clearly,

$$\|\boldsymbol{\beta} - \overline{\boldsymbol{\beta}}\|_2^2 = \sum_{k=N+1}^{\infty} \beta_k^2 \leq \exp\left[-\frac{(N+1)^2}{c}\right] \sum_{k=N+1}^{\infty} \beta_k^2 \exp\left(\frac{k^2}{c}\right) \leq \frac{\epsilon^2}{8}.$$

It follows that

$$\begin{aligned}
\epsilon^2 = \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_j\|_2^2 &= \|\overline{\boldsymbol{\beta}}_i - \overline{\boldsymbol{\beta}}_j\|_2^2 + \|(\boldsymbol{\beta}_i - \overline{\boldsymbol{\beta}}_i) - (\overline{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j)\|_2^2 \\
&\leq \|\overline{\boldsymbol{\beta}}_i - \overline{\boldsymbol{\beta}}_j\|_2^2 + 2\|\boldsymbol{\beta}_i - \overline{\boldsymbol{\beta}}_i\|_2^2 + 2\|\overline{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j\|_2^2 \leq \|\overline{\boldsymbol{\beta}}_i - \overline{\boldsymbol{\beta}}_j\|_2^2 + \frac{\epsilon^2}{2},
\end{aligned}$$

implying that $\|\overline{\boldsymbol{\beta}}_i - \overline{\boldsymbol{\beta}}_j\|_2 \geq \epsilon/\sqrt{2}$. For any $\boldsymbol{\beta}_i$ and $\mathbf{t} = (t_1, \cdots, t_N, 0, \cdots) \in B(\overline{\boldsymbol{\beta}_i}, \epsilon/(2\sqrt{2})) \subset \mathbb{R}^N$, one has

$$\sum_{k=1}^{N} t_k^2 \exp\left(\frac{k^2}{c}\right) \leq 2\sum_{k=1}^{N} \beta_k^2 \exp\left(\frac{k^2}{c}\right) + 2\sum_{k=1}^{N} (t_k - \beta_k)^2 \exp\left(\frac{k^2}{c}\right) \leq 2Q^2 + 2\exp\left(\frac{N^2}{c}\right)\frac{\epsilon^2}{8} \leq 4Q^2,$$

and thus

$$\bigcup_{j=1}^{m} B\left(\overline{\boldsymbol{\beta}}_i, \frac{\epsilon}{2\sqrt{2}}\right) \subset \Theta_N(2Q).$$

Since $B(\overline{\boldsymbol{\beta}}_i, \epsilon/(2\sqrt{2}))$'s overlap on each other only on a set of volume 0, then by denoting V_N the volume of the unit ball in $\mathbb{R}^N$ we obtain

$$m\left(\frac{\epsilon^2}{8}\right)^{N/2} V_N \leq (2Q)^N V_N \prod_{k=1}^{N} \exp\left(-\frac{k^2}{c}\right),$$

implying that

$$m \leq \exp\left[N \log\left(4\sqrt{2}Q\right) - \frac{1}{6c}N^3 + N\left(\log\frac{1}{\epsilon}\right)\right].$$

Since the maximum number of m is the packing number, the proof is completed by noticing that $N \asymp [\log(1/\epsilon)]^{1/2}$. □

Lemma F.3 *Suppose $f \sim \Pi = \mathrm{GP}(0, K)$ where K is the squared-exponential covariance function, and*

$f_0 \in \mathcal{A}_4(Q)$. *Then for sufficiently small* $\epsilon > 0$ *it holds that*

$$-\log \Pi(\|f - f_0\|_\infty < \epsilon) \lesssim \left(\log \frac{1}{\epsilon}\right)^2.$$

Proof. Denote $\mathbb{H}$ to be the reproducing kernel Hilbert space (RKHS) associated with K. Define the concentration function

$$\phi_{f_0}(\epsilon) = \frac{1}{2} \inf_{f \in \mathbb{H}: \|f - f_0\|_\infty < \epsilon} \|f\|_{\mathbb{H}} - \log \Pi(\|f\|_\infty < \epsilon).$$

By lemma 5.3 in van der Vaart and van Zanten (2008), it holds that $-\log \Pi(\|f - f_0\|_\infty < \epsilon) \leq \phi_{f_0}(\epsilon/2)$. By theorem 4.1 in van der Vaart and van Zanten (2008), $\mathbb{H}$ is the set of functions $f(x) = \sum_{k=1}^\infty \beta_k \psi_k(x)$ such that $\sum_{k=1}^\infty \beta_k^2/\lambda_k < \infty$. Since $\lambda_k \asymp \mathrm{e}^{-k^2/4}$, there exists some constants $\underline{\lambda}, \overline{\lambda} > 0$ such that $\underline{\lambda}\mathrm{e}^{-k^2/4} \leq \lambda_k \leq \overline{\lambda}\mathrm{e}^{-k^2/4}$. Using the fact that $f_0 \in \mathcal{A}_4(Q)$, we obtain

$$\sum_{k=1}^\infty \frac{\beta_{0k}^2}{\lambda_k} \leq \sum_{k=1}^\infty \frac{\beta_{0k}^2}{\underline{\lambda}} \exp\left(\frac{k^2}{4}\right) \leq \frac{1}{\underline{\lambda}} Q^2 < \infty.$$

Therefore $f_0 \in \mathbb{H}$, and the first term in $\phi_{f_0}(\epsilon)$ is upper bounded by $\|f_0\|_{\mathbb{H}}/2 = O(1)$. Furthermore, by lemma 4.6 in van der Vaart and van Zanten (2009), the second term in $\phi_{f_0}(\epsilon)$ is upper bounded by a constant multiple of $[\log(1/\epsilon)]^2$. The proof is thus completed. $\square$

Proof of theorem 4.2. Set $\epsilon_n = \underline{\epsilon}_n = n^{-1/2}(\log n)$, $\delta = Q$, and $m_n = \lceil(\log n)^2\rceil$. Clearly, $m_n \epsilon_n^2 \to 0$ and $\delta = O(1)$. By assumption $f_0 \in \mathcal{A}_4(Q)$, and hence yields the following series expansion

$$f_0(x) = \sum_{k=1}^\infty \beta_{0k}\psi_k(x), \quad \text{where} \quad \sum_{k=1}^\infty \beta_{0k}^2 \exp\left(\frac{k^2}{4}\right) \leq Q^2.$$

Still let constants $\underline{\lambda}, \overline{\lambda}$ be such that $\underline{\lambda}\mathrm{e}^{-k^2/8} \leq \lambda_k \leq \overline{\lambda}\mathrm{e}^{-k^2/8}$. Define the sieve $\mathcal{F}_{m_n}(Q)$ to be

$$\mathcal{F}_{m_n}(Q) = \left\{ f(x) = \sum_{k=1}^\infty \beta_k \psi_k(x) : \sum_{k=m_n+1}^\infty (\beta_k - \beta_{0k})^2 \exp\left(\frac{k^2}{8}\right) \leq Q^2 \right\}.$$

Clearly, $\mathcal{F}_{m_n}(Q)$ satisfies the property (2), and hence satisfies the property (4.2) by Lemma F.1. In fact, for any $f = \sum_k \beta_k \psi_k \in \mathcal{F}_{m_n}(Q)$, we directly compute by Cauchy-Schwartz inequality

$$\sum_{k > m_n} |\beta_k - \beta_{0k}| \leq \left[\sum_{k > m_n} (\beta_k - \beta_{0k})^2 \mathrm{e}^{k^2/8}\right]^{1/2} \left[\sum_{k > m_n}^\infty \frac{1}{\mathrm{e}^{k^2/8}}\right]^{1/2} \leq Q.$$

In light of Theorem 4.1, it suffices to verify the conditions (2.3), (2.4), and $\Pi(\|f - f_0\|_\infty < \underline{\epsilon}_n) \geq \mathrm{e}^{-Dn\underline{\epsilon}_n^2}$ for some constant $D > 0$.

By Lemma F.3, it holds for some constant $D' > 0$ and all sufficiently small $\epsilon > 0$ that

$$\Pi(\|f - f_0\|_\infty < \epsilon) \geq \exp\left[-D'\left(\log\frac{1}{\epsilon}\right)^2\right].$$

Using the fact that $n\underline{\epsilon}_n^2 \asymp [\log(1/\underline{\epsilon}_n)]^2 \asymp (\log n)^2$, it follows that there exists some constant $D > 0$, such that $\Pi(\|f - f_0\|_\infty < \underline{\epsilon}_n) \geq \exp(-Dn\underline{\epsilon}_n^2)$.

We next verify condition (2.3) with the same constant $D > 0$. Observe that

$$N_{nj} \leq \mathcal{N}(\xi j\epsilon_n, \mathcal{F}_{m_n}(\delta) \cap \{\|f - f_0\|_2 \leq (j+1)\epsilon_n\}, \|\cdot\|_2),$$

it suffices to bound the right-hand side of the preceeding display. Write

$$\begin{aligned}\mathcal{F}_{m_n}(\delta) \cap \{\|f - f_0\|_2 \leq (j+1)\epsilon_n\} \subset &\left\{f(x) = \sum_{k=1}^\infty \beta_k\psi_k(x) : \sum_{k=1}^{m_n}(\beta_k - \beta_{0k})^2 \leq (j+1)^2\epsilon_n^2\right\}\\ &\cap\left\{f(x) = \sum_{k=1}^\infty \beta_k\psi_k(x) : \sum_{k=m_n+1}^\infty (\beta_k - \beta_{0k})^2\exp\left(\frac{k^2}{8}\right) \leq Q^2\right\}.\end{aligned}$$

It follows that

$$\begin{aligned}&\mathcal{N}(\xi j\epsilon_n, \mathcal{F}_{m_n}(\delta) \cap \{\|f - f_0\|_2 \leq (j+1)\epsilon_n\}, \|\cdot\|_2)\\ &\quad\leq \mathcal{N}\left(\frac{\xi j\epsilon_n}{2}, \left\{(\beta_1, \cdots, \beta_{m_n}) : \sum_{k=1}^{m_n}(\beta_k - \beta_{0k})^2 \leq (j+1)^2\epsilon_n^2\right\}, \|\cdot\|_2\right)\\ &\qquad\times\mathcal{N}\left(\frac{\xi j\epsilon_n}{2}, \left\{(\beta_{m_n+1}, \cdots,) : \sum_{k>m_n}(\beta_k - \beta_{0k})^2 \mathrm{e}^{k^2/8} \leq Q^2\right\}, \|\cdot\|_2\right).\end{aligned}$$

We now bound the two covering numbers separately. For the first factor, computation of covering number in Euclidean space due to lemma 4.1 in Pollard (1990) yields

$$\mathcal{N}\left(\frac{\xi j\epsilon_n}{2}, \left\{(\beta_1, \cdots, \beta_{m_n}) : \sum_{k=1}^{m_n}(\beta_k - \beta_{0k})^2 \leq (j+1)^2\epsilon_n^2\right\}, \|\cdot\|_2\right) \leq \left(\frac{6(j+1)\epsilon_n}{\xi j\epsilon_n/2}\right)^{m_n} \leq \exp\left(m_n\log\frac{24}{\xi}\right).$$

For the second covering number, lemma F.2 yields

$$\mathcal{N}\left(\frac{\xi j\epsilon_n}{2}, \left\{(\beta_{m_n+1}, \cdots,) : \sum_{k=m_n+1}^\infty(\beta_k - \beta_{0k})^2\exp\left(\frac{k^2}{8}\right) \leq Q^2\right\}, \|\cdot\|_2\right) \leq \exp\left[D_1\left(\log\frac{1}{\xi j\epsilon_n}\right)^{3/2}\right]$$

for some constant $D_1 > 0$. We conclude that $N_{nj} \leq \exp(D_1 n\epsilon_n^2)$ for some constant $D_1 > 0$ by applying the fact that $[\log(1/\epsilon_n)]^{3/2} + m_n \asymp n\epsilon_n^2 \asymp (\log n)^2$, and hence,

$$\sum_{j=M}^\infty N_{nj}\exp(-Dnj^2\epsilon_n^2) \leq \exp(D_1 n\epsilon_n^2)\sum_{j=M}^\infty\int_{j-1}^j \exp(-Dn\epsilon_n^2x^2)\mathrm{d}x \leq \exp(D_1 n\epsilon_n^2)\int_{M-1}^\infty \exp(-Dn\epsilon_n^2x^2)\mathrm{d}x$$

$$\lesssim \exp(D_1 n\epsilon_n^2)\exp\left[-\frac{1}{2}D(M-1)^2 n\epsilon_n^2\right] \to 0$$

as long as M is sufficiently large. Hence condition (2.3) holds.

Finally we verify condition (2.4) with the same constant D. By definition

$$\Pi(\mathcal{F}_{m_n}^c(Q)) \le \Pi\left(2\sum_{k=m_n+1}^{\infty}\beta_k^2\exp\left(\frac{k^2}{8}\right) + 2\sum_{k=m_n+1}^{\infty}\beta_{0k}^2\exp\left(\frac{k^2}{8}\right) > Q^2\right).$$

For sufficiently large n, we have

$$\sum_{k=m_n+1}^{\infty}\beta_{0k}^2\exp\left(\frac{k^2}{8}\right) \le \sum_{k=m_n+1}^{\infty}\beta_{0k}^2\exp\left(\frac{k^2}{4}\right) < Q^2/4.$$

This is because by definition $f_0 \in \mathcal{H}_\alpha(Q)$ and $\sum_{k=1}^{\infty}\beta_{0k}^2\exp(k^2/4) \le Q^2$. Hence the preceeding display reduces to

$$\Pi(\mathcal{F}_{m_n}^c(Q)) \le \Pi\left(\sum_{k=m_n+1}^{\infty}\beta_k^2\exp\left(\frac{k^2}{8}\right) > \frac{Q^2}{4}\right),$$

and it suffices to bound the right-hand side. By the Markov's inequality, it holds for sufficiently large n that

$$\begin{aligned}
\Pi\left(\sum_{k=m_n+1}^{\infty}\beta_k^2\exp\left(\frac{k^2}{8}\right) > \frac{Q^2}{4}\right) &\le \frac{4}{Q^2}\sum_{k=m_n+1}^{\infty}\exp\left(\frac{k^2}{8}\right)\mathbb{E}(\beta_k^2) \le \frac{4}{Q^2}\sum_{k=m_n+1}^{\infty}\lambda_k\exp\left(\frac{k^2}{8}\right)\\
&\le \frac{4}{Q^2}\sum_{k=m_n+1}^{\infty}\overline{\lambda}\exp\left(-\frac{k^2}{8}\right) \le \frac{4\overline{\lambda}}{Q^2}\sum_{k=m_n+1}^{\infty}\int_{k-1}^{k}\exp\left(-\frac{x^2}{8}\right)\mathrm{d}x\\
&\le \frac{4\overline{\lambda}}{Q^2}\int_{m_n}^{\infty}\exp\left(-\frac{x^2}{8}\right)\mathrm{d}x \le \exp\left(-\frac{m_n^2}{16}\right).
\end{aligned}$$

Since $m_n^2 \asymp (\log n)^4 \ge (2D + 1/\sigma^2)(\log n)^2 = (2D+1/\sigma^2)n\epsilon_n^2$ when n is sufficiently large, it follows that (2.4) is satisfied with the same constant D. $\square$

G Proof of Theorem 4.3

The proof of Theorem 4.3 follows exactly the same lines of that of Theorem 2.1 with the assist of Lemmas 2.1, G.2, and G.3 below that are variations of Lemmas 2.1, G.2, and 2.3, respectively.

Lemma G.1 *Let $\mathcal{G}_m^A(\delta)$ satisfies (4.4). Then for any $f_1 \in \mathcal{G}_m^A(\delta)$ with $\sqrt{n}\|f_1 - f_0\|_2 > 1$, there exists a test function $\phi_n : (\mathcal{X}\times\mathcal{Y})^n \to [0,1]$ such that*

$$\mathbb{E}_0\phi_n \le \exp\left(-Cn\|f_1 - f_0\|_2^2\right),$$

$$\sup_{\{f\in\mathcal{G}_m^A(\delta):\|f-f_1\|_2^2\le\xi^2\|f_0-f_1\|_2^2\}}\mathbb{E}_f(1-\phi_n) \le \exp\left(-Cn\|f_1-f_0\|_2^2\right) + 2\exp\left(-\frac{Cn\|f_1-f_0\|_2^2}{A^2 m\|f_1-f_0\|_2^2+\delta^2}\right)$$

for some constant $C > 0$ and $\xi \in (0,1)$.

Proof. Recall the assumption that any $f \in \mathcal{G}_m(\delta, \mathbf{z})$ with $\|\mathbf{z}\|_1 \le A$ satisfies

$$\|f - f_0\|_\infty^2 \lesssim A^2 m \|f - f_0\|_2^2 + \delta^2. \tag{3}$$

The rest of the proof is similar to that of Lemma 2.1 and we only sketch the proof. Let us take $\xi = 1/(4\sqrt{2})$. Define the test function to be $\phi_n = \mathbb{1}\{T_n > 0\}$, where

$$T_n = \sum_{i=1}^n y_i(f_1(\mathbf{x}_i) - f_0(\mathbf{x}_i)) - \frac{1}{2} n \mathbb{P}_n(f_1^2 - f_0^2) - \frac{\sqrt{n}}{8\sqrt{2}} \|f_1 - f_0\|_2 \sqrt{n \mathbb{P}_n (f_1 - f_0)^2}.$$

We first consider the type I error probability. Following the proof of Lemma 2.1, it is immediate that

$$\mathbb{E}_0 \phi_n \le \exp\left(-\frac{C_1}{32} n \|f_1 - f_0\|_2^2\right).$$

We next consider the type II error probability. For any f with $\|f - f_1\|_2 \le \|f_0 - f_1\|_2/(4\sqrt{2}) \le \|f_0 - f_1\|_2/4$, following the proof of Lemma G.1, we have

$$\begin{aligned}
\mathbb{E}_f(1 - \phi_n) \le \exp\left(-\frac{C_1}{32} n \|f_1 - f_0\|_2^2\right) + \mathbb{P}\left(\mathbb{G}_n(f_1 - f_0)^2 < -\frac{\sqrt{n}}{2} \|f_1 - f_0\|_2^2\right) \\
+ \mathbb{P}\left(\mathbb{P}_n(f - f_1)^2 > \frac{1}{16} \mathbb{P}_n(f_1 - f_0)^2\right).
\end{aligned}$$

Using Bernstein's inequality, we obtain the tail probability of the empirical process $\mathbb{G}_n(f_1 - f_0)^2$

$$\mathbb{P}\left(\mathbb{G}_n(f_1 - f_0)^2 < -\frac{\sqrt{n}}{2} \|f_1 - f_0\|_2^2\right) \le \exp\left(-\frac{C' n \|f_1 - f_0\|_2^2}{A^2 m \|f_1 - f_0\|_2^2 + \delta^2}\right),$$

for some constant $C' > 0$, where we use the relation (3). On the other hand, when $\mathbb{P}_n(f - f_1)^2 > \mathbb{P}_n(f_1 - f_0)^2/16$, we again use Bernstein's inequality and the fact that $f \in \{f \in \mathcal{G}_m^A(\delta) : \|f - f_1\|_2^2 \le 2^{-5} \|f_0 - f_1\|_2^2\}$ to compute

$$\mathbb{P}\left(\mathbb{P}_n(f - f_1)^2 > \frac{1}{16} \mathbb{P}_n(f_1 - f_0)^2\right) \le \exp\left(-\frac{1}{4} \frac{n \|f_1 - f_0\|_2^4 / 1024}{\|g\|_2^2 + \|f_1 - f_0\|_2^2 \|g\|_\infty / 32}\right),$$

where $g = (f - f_1)^2 - (f_1 - f_0)^2/16$. We further compute

$$\begin{aligned}
\|g\|_2^2 &\le \left(\|(f - f_1)^2\|_2 + \frac{1}{16} \|(f_1 - f_0)^2\|_2\right)^2 \le \left(\|f - f_1\|_\infty \|f - f_1\|_2 + \frac{1}{16} \|f_1 - f_0\|_\infty \|f_1 - f_0\|_2\right)^2 \\
&\lesssim \|f - f_1\|_\infty^2 \|f - f_1\|_2^2 + \|f_1 - f_0\|_\infty^2 \|f_1 - f_0\|_2^2 \lesssim (A^2 m \|f_1 - f_0\|_2^2 + \delta^2) \|f_0 - f_1\|_2^2,
\end{aligned}$$

where we use (3), the fact that $\|f - f_1\|_2 \lesssim \|f_0 - f_1\|_2$, and that

$$\|f - f_1\|_\infty^2 \le 2\|f - f_0\|_\infty^2 + 2\|f_0 - f_1\|_\infty^2 \lesssim A^2 m \|f_1 - f_0\|_2^2 + \delta^2.$$

Similarly, we obtain on the other hand,

$$\|g\|_\infty = \|f - f_1\|_\infty^2 + \frac{1}{16}\|f_1 - f_0\|_\infty^2 \lesssim A^2 m\|f_0 - f_1\|_2^2 + \delta^2.$$

Therefore, we end up with

$$\mathbb{P}\left(\mathbb{P}_n(f - f_1)^2 > \frac{1}{16}\mathbb{P}_n(f_1 - f_0)^2\right) \le \exp\left(-\frac{\tilde{C}_2 n\|f_1 - f_0\|_2^2}{A^2 m\|f_1 - f_0\|_2^2 + \delta^2}\right),$$

where $\tilde{C}_2 > 0$ is some constant. Assembling all the pieces obtained above, we obtain the following exponential bound for type I and type II error probabilities:

$$\begin{aligned}
\mathbb{E}_0\phi_n &\le \exp(-Cn\|f_1 - f_0\|_2^2), \\
\mathbb{E}(1 - \phi_n) &\le \exp(-Cn\|f_1 - f_0\|_2^2) + 2\exp\left(-\frac{Cn\|f_1 - f_0\|_2^2}{A^2 m\|f_1 - f_0\|_2^2 + \delta^2}\right)
\end{aligned}$$

for some constant $C > 0$ whenever $\|f - f_1\|_2^2 \le \|f_1 - f_0\|_2^2/32$. Taking the supremum of the type II error over $f \in \{f \in \mathcal{G}_m(\delta) : \|f - f_1\|_2^2 \le \|f_1 - f_0\|_2^2/32\}$ completes the proof. $\square$

Exploiting the proof of Lemma 2.2, we obtain the following lemma for the sparse additive models immediately by combining Lemma G.1.

Lemma G.2 *Let $m \in \mathbb{N}_+$ be an positive integer, and $\delta, A > 0$ be positive. Suppose that $\mathcal{G}_m^A(\delta)$ satisfies (4.4). Let $(\epsilon_n)_{n=1}^\infty$ be a sequence with $n\epsilon_n^2 \to \infty$. Then there exists a sequence of test functions $(\phi_n)_{n=1}^\infty$ such that*

$$\mathbb{E}_0\phi_n \le \sum_{j=M}^\infty N_{nj}^A \exp\left(-Cnj^2\epsilon_n^2\right),$$

$$\sup_{\{f \in \mathcal{G}_m^A(\delta) : \|f - f_0\|_2 > M\epsilon_n\}} \mathbb{E}_f(1 - \phi_n) \le \exp(-CM^2 n\epsilon_n^2) + 2\exp\left(-\frac{CM^2 n\epsilon_n^2}{A^2 m M^2\epsilon_n^2 + \delta^2}\right),$$

where $N_{nj}^A = \mathcal{N}(\xi j\epsilon_n, \mathcal{S}_{nj}^A(\epsilon_n), \|\cdot\|_2)$ is the covering number of

$$\mathcal{S}_{nj}^A(\epsilon_n) = \left\{f \in \mathcal{G}_m^A(\delta) : j\epsilon_n < \|f - f_0\|_2 \le (j+1)\epsilon_n\right\},$$

and C is some positive constant.

The following lemma for the sparse additive models is immediate by exploiting the proof of Lemma 2.3 and observing the fact that $\|f - f_0\|_\infty = O(1)$ for any $f \in \widetilde{B}(k_n, \epsilon_n, \omega)$ given that $k_n\epsilon_n^2 = O(1)$.

Lemma G.3 *Let $\widetilde{B}(m, \epsilon, \omega)$ be defined as Theorem 4.3. Suppose sequences $(\epsilon_n)_{n=1}^\infty$ and $(k_n)_{n=1}^\infty$ satisfy $\epsilon_n \to 0$, $n\epsilon_n^2 \to \infty$, $k_n\epsilon_n^2 = O(1)$, and ω is some constant. Then for any constant $C > 0$,*

$$\mathbb{P}_0\left(\int \exp(\Lambda_n)\Pi(\mathrm{d}f) \le \Pi\left(\widetilde{B}(k_n, \epsilon_n, \omega)\right)\exp\left[-\left(C + \frac{1}{\sigma^2}\right)n\epsilon_n^2\right]\right) \to 0.$$

H Proof of Theorem 4.4

Lemma H.1 *Let m be an positive integer, $\delta, A > 0$, and $\xi \in (0,1)$ is some absolute constant. Assume that $f_{0j} \in \mathfrak{C}_\alpha(Q)$ for some $\alpha > 1/2$ and some $Q > 0$, $j \in \{j_1, \ldots, j_q\}$. Take $\mathcal{G}_m^A(qQ) = \bigcup_{\mathbf{z}:\|\mathbf{z}\|_1 \le Aq} \mathcal{G}_m(qQ, \mathbf{z})$ for some positive integer $A \lesssim q$, where*

$$\mathcal{G}_m(qQ, \mathbf{z}) = \left\{ f(\mathbf{x}) = \mu + \sum_{j=1}^p \sum_{k=1}^\infty z_j \beta_{jk} \psi_k(x_j), \beta_{j1} = -\sum_{k=2}^\infty \beta_{jk} \int_0^1 \psi_k(x_j) \mathrm{d}x_j, N = m \right\}. \tag{4}$$

Then $\mathcal{G}_m(qQ, \mathbf{z})$ satisfies (4.4), and

$$\log \mathcal{N}(\xi j \epsilon_n, \mathcal{S}_{nj}^A(\epsilon), \|\cdot\|_2) \le \left(2 \log \frac{12}{\xi}\right) Aqm + \log \binom{p}{Aq},$$

where $\mathcal{S}_{nj}^A(\epsilon) = \left\{ f \in \mathcal{G}_m^A(qQ) : j\epsilon_n < \|f - f_0\|_2 \le (j+1)\epsilon_n \right\}$.

Proof. We first show that $\mathcal{G}_m(qQ, \mathbf{z})$ satisfies (4.4) with $\delta = qQ$. In fact, by the Cauchy-Schwartz inequality,

$$\begin{aligned}
\|f - f_0\|_\infty &\lesssim |\mu - \mu_0| + \sum_{j=1}^p \sum_{k=1}^\infty |z_j \beta_{jk} - \beta_{0jk}| = |\mu - \mu_0| + \sum_{j \in \{j: z_j = 1\} \cup \{j_1, \cdots, j_q\}} \sum_{k=1}^m |z_j \beta_{jk} - \beta_{0jk}| \\
&\le |\mu - \mu_0| + \sum_{j \in \{j: z_j = 1\} \cup \{j_1, \cdots, j_q\}} \sqrt{m} \|f_j - f_{0j}\|_2,
\end{aligned}$$

and hence,

$$\begin{aligned}
\|f - f_0\|_\infty^2 &\lesssim (\mu - \mu_0)^2 + A^2 m \sum_{j=1}^p \|z_j f_j - f_{0j}\|_2^2 \le A^2 m (\mu - \mu_0)^2 + A^2 m \left\| \sum_{j=1}^p (z_j f_j - f_{0j}) \right\|_2^2 \\
&= A^2 m \left\| \left(\mu + \sum_{j=1}^p z_j f_j \right) - \left(\mu_0 + \sum_{j=1}^p f_{0j} \right) \right\|_2^2 \le A^2 m \|f - f_0\|_2^2 + 1.
\end{aligned}$$

This shows that $\mathcal{G}_m(qQ, \mathbf{z})$ satisfies (4.4). We next prove the covering number bounds. Observe that

$$\begin{aligned}
N_{nj}^A &\le \mathcal{N}(\xi j \epsilon_n, \mathcal{G}_m^A(qQ) \cap \{\|f - f_0\|_2 \le (j+1)\epsilon_n\}, \|\cdot\|_2) \\
&\le \sum_{\mathbf{z} \in \{0,1\}^p : \|\mathbf{z}\|_1 \le Aq} \mathcal{N}(\xi j \epsilon_n, \mathcal{G}_m(qQ, \mathbf{z}) \cap \{\|f - f_0\|_2 \le (j+1)\epsilon_n\}, \|\cdot\|_2),
\end{aligned}$$

and that

$$\|f - f_0\|_2^2 = (\mu - \mu_0) + \sum_{j=1}^p \sum_{k=1}^\infty (z_j \beta_{jk} - \beta_{0jk})^2,$$

due to the fact that $\int_0^1 f_j(x_j) \mathrm{d}x_j = \int_0^1 f_{0j}(x_j) \mathrm{d}x_j = 0$, $j = 1, \cdots, p$. Write

$$\mathcal{G}_m(qQ, \mathbf{z}) \cap \{\|f - f_0\|_2 \le (j+1)\epsilon_n\}$$

$$\subset \left\{ f(\mathbf{x}) = \mu + \sum_{j=1}^{p}\sum_{k=1}^{\infty} z_j\beta_{jk}\psi_k(x_j) : (\mu-\mu_0)^2 + \sum_{j:z_j=1}\sum_{k=1}^{m}(\beta_{jk}-\beta_{0jk})^2 \le (j+1)^2\epsilon_n^2 \right\},$$

It follows that

$$\begin{aligned}
&\mathcal{N}(\xi j\epsilon_n, \mathcal{G}_m(qQ,\mathbf{z}) \cap \{\|f-f_0\|_2 \le (j+1)\epsilon_n\}, \|\cdot\|_2) \\
&\quad\le \mathcal{N}\left(\xi j\epsilon_n, \left\{(\mu, \beta_{jk}: z_j = 1, k = 1,\cdots,m) : (\mu-\mu_0)^2 + \sum_{j:z_j=1}\sum_{k=1}^{m}(\beta_{jk}-\beta_{0jk})^2 \le (j+1)^2\epsilon_n^2\right\}, \|\cdot\|_2\right) \\
&\quad\le \left(\frac{6(j+1)\epsilon_n}{\xi j\epsilon_n}\right)^{Aqm+1} \le \exp\left[(Aqm+1)\log\frac{12}{\xi}\right].
\end{aligned}$$

Therefore,

$$\mathcal{N}(\xi j\epsilon_n, S_{nj}^A(\epsilon_n), \|\cdot\|_2) \le \binom{p}{Aq}\exp\left\{\left(2\log\frac{12}{\xi}\right)Aqm\right\}.$$

The proof is completed by taking the logarithm of the preceding display. □

Proof of Theorem 4.4. The proof is based on the proof of Theorem 4.3, along with several modifications. We begin by defining the following quantity:

$$\epsilon = \frac{2}{3}\left(t - \frac{\alpha}{2\alpha+1}\right), \quad \delta = \frac{2\alpha}{2\alpha+1} - 1 + 2\epsilon, \quad \zeta = \frac{\alpha}{2\alpha+1} + \frac{\epsilon}{2}.$$

It follows from simple algebra that $2t > \delta + 1 > 2\zeta > -2\alpha\delta$ and $2\zeta > 1 - \zeta/\alpha$. Without loss of generality we may also assume that ϵ is small so that $\zeta < 1$, since contraction for smaller t implies contraction for larger t. Set $m_n = \lceil n^{1/(4\alpha+2)}(\log n)^\delta\rceil$, $A_n = \lceil n^{1/(4\alpha+2)}\log n\rceil$, $\epsilon_n = n^{-\alpha/(2\alpha+1)}(\log n)^t$, $\underline{\epsilon}_n = n^{-(\alpha+1/4)/(2\alpha+1)}(\log n)^\zeta$ and denote

$$f_m(\mathbf{x}) = \mu + \sum_{j=1}^{p}\sum_{k=1}^{m}\xi_j\beta_{jk}\psi_k(x_j)$$

given that $N = m$, *i.e.*, $\beta_{jk} = 0$ for al $k > m$, $j = 1,\cdots,p$.

We first verify condition (4.7) with $\omega = 1$ and $k_n = \lceil n^{1/(4\alpha+2)}(\log n)^{-\zeta/\alpha}\rceil$. Clearly, $k_n\underline{\epsilon}_n^2 = O(1)$. Observe that $\int_0^1 \psi_k(x_j)\mathrm{d}x_j \asymp k^{-1}$, it follows from the Cauchy-Schwarz inequality that

$$\begin{aligned}
\|f-f_0\|_2^2 &= (\mu-\mu_0)^2 + \sum_{j=1}^{p}\left[\sum_{k=2}^{\infty}(z_j\beta_{jk}-\beta_{0jk})\int_0^1\psi_k(x_j)\mathrm{d}x_j\right]^2 + \sum_{j=1}^{p}\sum_{k=2}^{\infty}(\beta_{jk}-\beta_{0jk})^2 \\
&\lesssim (\mu-\mu_0)^2 + \sum_{j=1}^{p}\left[\sum_{k=2}^{\infty}(\beta_{jk}-\beta_{0jk})^2\right]\left[\sum_{k=2}^{\infty}\frac{1}{k^2}\right] + \sum_{j=1}^{p}\sum_{k=2}^{\infty}(\beta_{jk}-\beta_{0jk})^2 \\
&\lesssim (\mu-\mu_0)^2 + \sum_{j=1}^{p}\sum_{k=2}^{\infty}(z_j\beta_{jk}-\beta_{0jk})^2.
\end{aligned}$$

Namely, $\|f-f_0\|_2^2 \le C_\psi^{-2}[(\mu-\mu_0)^2 + \sum_{j=1}^{p}\sum_{k=2}^{\infty}(z_j\beta_{jk}-\beta_{0jk})^2]$ for some constant $C_\psi > 0$. For sufficiently large n, write

$$
\begin{aligned}
& B(k_n, \epsilon_n, 1) \\
& \supset \left\{ f(\mathbf{x}) = \mu + \sum_{j=1}^{p}\sum_{k=1}^{\infty} \xi_j \beta_{jk}\psi_k(x_j) : \|f - f_0\|_2 < \epsilon_n, \sum_{j=1}^{p}\sum_{k=k_n+1}^{\infty} |z_j\beta_{jk} - \beta_{0jk}| \leq 1 \right\} \cap \{\|\mathbf{z}\|_1 \leq 2q\} \\
& \supset \left\{ f_{k_n}(\mathbf{x}) : (\mu - \mu_0)^2 + \sum_{j:z_j=1}\sum_{k=2}^{k_n} (\beta_{jk} - \beta_{0jk})^2 \leq C_\psi \epsilon_n^2, N = k_n, z_{j_r} = 1, r = 1, \cdots, q, \|\mathbf{z}\|_1 = q \right\} \\
& \supset \left\{ |\mu - \mu_0| < \frac{C_\psi \epsilon_n}{2} \right\} \cap \bigcap_{j:z_j=1}\bigcap_{k=2}^{k_n} \left\{ \beta_{jk} : |\beta_{jk} - \beta_{0jk}| \leq \frac{C_\psi \epsilon_n}{k_n} \right\} \cap \bigcap_{j\in\{j_1,\cdots,j_q\}} \{z_j = 1\} \cap \bigcap_{j\notin\{j_1,\cdots,j_q\}} \{z_j = 0\}.
\end{aligned}
$$

Hence for sufficiently large n,

$$
\begin{aligned}
& \Pi\left(\widetilde{B}(k_n, \epsilon_n, \delta) \,\middle|\, N = k_n\right) \\
& \geq \Pi\left(|\mu - \mu_0| < \frac{C_\psi \epsilon_n}{2}\right) \left[\prod_{j:z_j=1}\prod_{k=2}^{k_n} \Pi\left(\beta_{jk} : |\beta_{jk} - \beta_{0jk}| \leq \frac{C_\psi \epsilon_n}{k_n} \,\middle|\, N = k_n\right)\right] \left(\frac{1}{p}\right)^q \left(1 - \frac{1}{p}\right)^{p-q} \\
& \geq \left[\frac{C_\psi \epsilon_n}{2} \min_{|\mu-\mu_0|\leq 1} \pi(\mu)\right] \left[\prod_{j:z_j=1}\prod_{k=2}^{k_n} \Pi\left(\beta_{jk} : |\beta_{jk} - \beta_{0jk}| \leq \frac{C_\psi \epsilon_n}{k_n} \,\middle|\, N = k_n\right)\right] \exp[-q\log p - \log(2\mathrm{e})] \\
& \geq \left[\frac{C_\psi \epsilon_n}{2} \min_{|\mu-\mu_0|\leq 1} \pi(\mu)\right] \left[\prod_{j:z_j=1}\prod_{k=2}^{k_n} \int_{\beta_{0k}-C_\psi\epsilon_n/k_n}^{\beta_{0k}+C_\psi\epsilon_n/k_n} g(\beta_k)\mathrm{d}\beta_k\right] \exp[-q\log p - \log(2\mathrm{e})] \\
& \geq \exp\left[-4qk_n\left(\log\frac{1}{\epsilon_n}\right) - 2k_n\log k_n - q\log p\right].
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\Pi\left(\widetilde{B}(k_n, \epsilon_n, \delta)\right) & \geq \Pi\left(\widetilde{B}(k_n, \epsilon_n, \delta) \,\middle|\, N = k_n\right) \pi_N(k_n) \\
& \geq \exp\left[-4q\lceil n^{1/(4\alpha+2)}(\log n)^{-\zeta/\alpha}\rceil \left(\log\frac{1}{\epsilon_n}\right) - q\log p - 2k_n\log k_n\right] \\
& \quad \times \exp\left[-b_0\lceil n^{1/(4\alpha+2)}(\log n)^{-\zeta/\alpha}\rceil \log\lceil n^{1/(4\alpha+2)}(\log n)^{-\zeta/\alpha}\rceil\right] \\
& \geq \exp\left\{-D\left[n^{1/(4\alpha+2)}(\log n)^{1-\zeta/\alpha}\right]\right\} \geq \exp(-Dn\epsilon_n^2)
\end{aligned}
$$

for some constant $D > 0$.

Now set $\delta = qQ$ and construct the sieve $\mathcal{G}_{m_n}^{A_n}(qQ) = \bigcup_{\|\mathbf{z}\|_1 \leq A_n q} \mathcal{G}_{m_n}(qQ, \mathbf{z})$, where $\mathcal{G}_{m_n}(qQ, \mathbf{z})$ is given as in (4). Clearly, for any $f(\mathbf{x}) = \sum_{j=1}^{p}\sum_{k=1}^{m_n} z_j\beta_{jk}\psi_k(x_j) \in \mathcal{G}_{m_n}(qQ, \mathbf{z})$,

$$
\sum_{j=1}^{p}\sum_{k=m_n+1}^{\infty} |z_j\beta_{jk} - \beta_{0jk}| = \sum_{j=1}^{p}\sum_{k=m_n+1}^{\infty} |\beta_{0jk}| \leq \sum_{j=1}^{p}\sum_{k=m_n+1}^{\infty} |\beta_{0jk}|k^\alpha \leq qQ.
$$

Therefore $\mathcal{G}_{m_n}^{A_n}(qQ)$ satisfies (4.4). Furthermore, $A_n m_n \epsilon_n^2 = n^{(1-2\alpha)/(2\alpha+1)}(\log n)^{2t+\delta+1} \to 0$. Next we verify condition (4.5). Invoking Lemma H.1, we see that

$$N_{nj}^{A_n} \le \exp\left\{D_1\left[n^{1/(2\alpha+1)}(\log n)^{\delta+1} + A_n \log p\right]\right\} \le \exp\left[D_1 n^{1/(2\alpha+1)}(\log n)^{2t}\right] = \exp(D_1 n\epsilon_n^2)$$

for some constant $D_1 > 0$, and hence

$$\begin{aligned}\sum_{j=M}^{\infty} N_{nj}^{A}\exp(-Dnj^2\epsilon_n^2) &\le \exp(D_1 n\epsilon_n^2)\sum_{j=M}^{\infty}\int_{j-1}^{j}\exp(-Dn\epsilon_n^2 x^2)\mathrm{d}x \le \exp\left(D_1 n\epsilon_n^2\right)\int_{M-1}^{\infty}\exp(-Dn\epsilon_n^2 x^2)\mathrm{d}x \\ &\lesssim \exp(D_1 n\epsilon_n^2)\exp\left[-\frac{1}{2}D(M-1)^2 n\epsilon_n^2\right] \to 0\end{aligned}$$

for sufficiently large n by taking $M \ge 1 + \sqrt{4D_1/D}$, and hence, condition (4.5) holds.

We are now left to show that $\mathcal{G}_{m_n}^{A_n}(qQ)$ satisfies (4.6). Following the proof of Theorem 3.1, write

$$\mathcal{G}_{m_n}^{A_n}(qQ)^c \subset \{\mathbf{z} : \|\mathbf{z}\|_1 \ge A_n q\} \cup \bigcup_{m=m_n+1} \{N = m\}.$$

A version of the Chernoff's inequality for binomial distribution is of the form

$$\mathbb{P}(X > ap) \le \left[\left(\frac{1}{ap}\right)^a \exp\left(a\right)\right]^p \quad \text{if } X \sim \text{Binomial}\left(p, \frac{1}{p}\right) \text{ and } a \ge 1/p.$$

Therefore,

$$\Pi\left(\mathbf{z} : \|\mathbf{z}\|_1 \ge A_n q\right) \le \left[\left(\frac{1}{A_n q}\right)^{A_n q/p}\exp\left(\frac{A_n q}{p}\right)\right]^p = \exp\left[A_n q - A_n q\log(A_n q)\right].$$

It follows that

$$\begin{aligned}\Pi(\mathcal{G}_{m_n}^{A_n}(qQ)^c) &\le \sum_{m=m_n+1}^{\infty}\pi_N(m) + \Pi(\mathbf{z} : \|\mathbf{z}\|_1 \ge A_n q) \\ &\lesssim \exp(-b_1 m_n\log m_n) + \exp[-A_n q(\log A_n - 1) - A_n q\log q] \\ &\lesssim \exp\left\{-D_2\min\left[n^{1/(4\alpha+2)}(\log n)^{\delta+1}, n^{1/(4\alpha+2)}(\log n)^2\right]\right\} \\ &\le \exp\left[-\left(2D + \frac{1}{\sigma^2}\right)n^{1/(4\alpha+2)}(\log n)^{2\zeta}\right] = \exp\left[-\left(2D + \frac{1}{\sigma^2}\right)n\underline{\epsilon}_n^2\right]\end{aligned}$$

for some constant $D_2 > 0$ when n is sufficiently large. Hence condition (2.4) holds with the same constant D. ∎

References

Belitser, E. and Ghosal, S. (2003). Adaptive Bayesian inference on the mean of an infinite-dimensional normal distribution. *The Annals of Statistics*, 31(2):536–559.

De Jonge, R. and Van Zanten, J. (2012). Adaptive estimation of multivariate functions using conditionally Gaussian tensor-product spline priors. *Electronic Journal of Statistics*, 6:1984–2001.

Gao, C. and Zhou, H. H. (2016a). Rate exact Bayesian adaptation with modified block priors. *The Annals of Statistics*, 44(1):318–345.

Gao, C. and Zhou, H. H. (2016b). Supplement to "rate exact bayesian adaptation with modified block priors".

Ghosal, S. and Van Der Vaart, A. W. (2001). Entropies and rates of convergence for maximum likelihood and Bayes estimation for mixtures of normal densities. *The Annals of Statistics*, 29(5):1233–1263.

Pollard, D. (1990). Empirical processes: theory and applications. In *NSF-CBMS regional conference series in probability and statistics*, pages i–86. JSTOR.

van der Vaart, A. W. and van Zanten, J. H. (2008). Reproducing kernel Hilbert spaces of Gaussian priors. In *Pushing the limits of contemporary statistics: contributions in honor of Jayanta K. Ghosh*, pages 200–222. Institute of Mathematical Statistics.

van der Vaart, A. W. and van Zanten, J. H. (2009). Adaptive Bayesian estimation using a Gaussian random field with inverse Gamma bandwidth. *The Annals of Statistics*, pages 2655–2675.

Yang, Y. and Pati, D. (2017). Bayesian model selection consistency and oracle inequality with intractable marginal likelihood. *arXiv preprint arXiv:1701.00311*.

Yoo, W. W., Rousseau, J., and Rivoirard, V. (2017). Adaptive supremum norm posterior contraction: Spike-and-slab priors and anisotropic besov spaces. *arXiv preprint arXiv:1708.01909*.