Supplementary Material: Convergence Rates for Kernel Regression in Infinite Dimensional Spaces

Joydeep Chowdhury · Probal Chaudhuri

1 Small ball probabilities of non-Gaussian processes

In Propositions 1, 2 and 3 below, we consider two random elements $\mathbf{T}$ and $\mathbf{G}$, and define $\phi_{\mathbf{T}}(\mathbf{t}, h) = \mathbb{P}[\|\mathbf{T} - \mathbf{t}\| \leq h]$ and $\phi_{\mathbf{G}}(\mathbf{g}, h) = \mathbb{P}[\|\mathbf{G} - \mathbf{g}\| \leq h]$, where $\mathbf{t}$ and $\mathbf{g}$ are some fixed elements and $h > 0$.

Proposition 1 *Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be separable Banach spaces, and $f(\cdot) : \mathcal{B}_2 \longrightarrow \mathcal{B}_1$ be a function such that for any $\mathbf{u} \in \mathcal{B}_2$, there exist constants $r, s > 0$, which may depend on $\mathbf{u}$, such that for any $\mathbf{v} \in \mathcal{B}_2$ sufficiently close to $\mathbf{u}$, we have $r\|\mathbf{v} - \mathbf{u}\| \leq \|f(\mathbf{v}) - f(\mathbf{u})\| \leq s\|\mathbf{v} - \mathbf{u}\|$. If $\mathbf{T}$ and $\mathbf{G}$ are random elements with $\mathbf{T} = f(\mathbf{G})$, and the small ball probability of $\mathbf{G}$ satisfies the bounds described in (9) in the main paper, then similar bounds also hold for $\mathbf{T}$.*

Proof Under the assumptions of the proposition, $f(\cdot)$ is a one-to-one function. Let $\mathbf{t}$ be an element in the range of $f(\cdot)$. Then, $\mathbf{t} = f(\mathbf{g})$ for some $\mathbf{g}$. Consequently, for some positive constants r and s, which may depend on $\mathbf{g}$, we have for all sufficiently small h,

$$\begin{aligned} &\mathbb{P}[s\|\mathbf{G} - \mathbf{g}\| \leq h] \leq \mathbb{P}[\|f(\mathbf{G}) - f(\mathbf{g})\| \leq h] \leq \mathbb{P}[r\|\mathbf{G} - \mathbf{g}\| \leq h] \\ &\iff \phi_{\mathbf{G}}\left(\mathbf{g}, \frac{h}{s}\right) \leq \phi_{\mathbf{T}}(\mathbf{x}, h) \leq \phi_{\mathbf{G}}\left(\mathbf{g}, \frac{h}{r}\right). \end{aligned} \tag{1}$$

The proof follows by applying the bounds in (9) in (1). □

Joydeep Chowdhury (corresponding author)
Statistics and Mathematics Unit, Indian Statistical Institute,
203, B.T. Road, Kolkata, India. PIN: 700108.
E-mail: joydeepchowdhury01@gmail.com

Probal Chaudhuri
Statistics and Mathematics Unit, Indian Statistical Institute,
203, B.T. Road, Kolkata, India. PIN: 700108.
E-mail: probal@isical.ac.in

Let $\mathbf{G}$ be a Gaussian process whose small ball probability $\phi_{\mathbf{G}}(\mathbf{g}, h)$ satisfies the bounds in (9) for sufficiently small h, so that

$$C_1 h^{t_1} \exp\left[-C_2(1/h)^{t_2}(\log(1/h))^{t_3}\right] \leq \phi_{\mathbf{G}}(\mathbf{g}, h) \leq C_3 h^{t_4} \exp\left[-C_4(1/h)^{t_2}(\log(1/h))^{t_3}\right]$$

as $h \longrightarrow 0^+$. Here, $C_1, C_2, C_3, C_4 > 0$ and $t_1, t_2, t_3, t_4 \geq 0$ are appropriate constants, all of which, except C_1, are independent of $\mathbf{g}$. C_1 may or may not depend on $\mathbf{g}$, but if it depends on $\mathbf{g}$ then $C_1 = C_1' \exp[-(1/2)\|\mathbf{g}\|^2]$ for some positive constant C_1'. Also, either $t_2 > 0$, or $t_3 > 1$ with $C_2 = C_4$.

In Proposition 2 and Proposition 3 below, we shall derive the bounds on the small ball probabilities of some non-Gaussian processes. There, we shall assume $C_1 = C_1' \exp[-(1/2)\|\mathbf{g}\|^2]$ for some positive constant C_1'. Since $C_1' \geq C_1' \exp[-(1/2)\|\mathbf{g}\|^2]$ for all $\mathbf{g}$, establishing the lower bound of the small ball probability, when $C_1 = C_1' \exp[-(1/2)\|\mathbf{g}\|^2]$, also gives an appropriate lower bound when C_1 does not depend on $\mathbf{g}$.

Proposition 2 *Let $\mathbf{T} = \mathbf{G}/\mathbf{U}$, where $\mathbf{G}$ is a Gaussian process whose small ball probability satisfies the bounds in* (9) *in the main paper, and $\mathbf{U}$ is a bounded positive random variable independent of $\mathbf{G}$. Then, the small ball probability of $\mathbf{T}$ also satisfies the bounds in* (9)*.*

Proof Note that

$$\phi_{\mathbf{T}}(\mathbf{t}, h) = \mathbb{P}\left[\|\mathbf{G} - \mathbf{t}\mathbf{U}\| \leq h\mathbf{U}\right] = \mathbb{E}\left[\phi_{\mathbf{G}}\left(\mathbf{t}\mathbf{U},\ h\mathbf{U}\right)\right]. \tag{2}$$

Let $0 \leq \mathbf{U} \leq u_0$ for some $u_0 > 0$. Recall from (10) in the main paper that $m(h) = C_2(1/h)^{t_2}(\log(1/h))^{t_3}$ for $0 < h < 1$. Since $m(hu_0) \leq m(h\mathbf{U})$ for all $h > 0$, we have

$$\begin{aligned}
\phi_{\mathbf{G}}\left(\mathbf{t}\mathbf{U},\ h\mathbf{U}\right) &\leq C_3(h\mathbf{U})^{t_4} \exp\left[-(C_4/C_2)m(h\mathbf{U})\right] \\
&\leq C_3(hu_0)^{t_4} \exp\left[-(C_4/C_2)m(hu_0)\right] \\
&= C_3 u_0^{t_4} h^{t_4} \exp\left[-C_4\left(\frac{1}{u_0}\right)^{t_2}\left(1 - \frac{\log u_0}{\log \frac{1}{h}}\right)^{t_3}\left(\frac{1}{h}\right)^{t_2}\left(\log\frac{1}{h}\right)^{t_3}\right] \\
&\leq C_3 u_0^{t_4} h^{t_4} \exp\left[-\frac{C_4}{2}\left(\frac{1}{u_0}\right)^{t_2}\left(\frac{1}{h}\right)^{t_2}\left(\log\frac{1}{h}\right)^{t_3}\right]
\end{aligned}$$

for all sufficiently small h. Hence, for all sufficiently small h,

$$\mathbb{E}\left[\phi_{\mathbf{G}}\left(\mathbf{t}\mathbf{U},\ h\mathbf{U}\right)\right] \leq C_3 u_0^{t_4} h^{t_4} \exp\left[-\frac{C_4}{2}\left(\frac{1}{u_0}\right)^{t_2}\left(\frac{1}{h}\right)^{t_2}\left(\log\frac{1}{h}\right)^{t_3}\right]. \tag{3}$$

Now, if $\mathbf{U}$ is a degenerate positive random variable, i.e., $\mathbb{P}[\mathbf{U} = u_0] = 1$, then the lower bound of $\phi_{\mathbf{G}}\left(\mathbf{t}\mathbf{U},\ h\mathbf{U}\right)$ trivially satisfies (9). So, we assume that $\mathbf{U}$ is non-degenerate, and $\mathbb{P}[l_0 \leq \mathbf{U} < u_0] > 0$ for some $l_0 > 0$. We consider the case where the constant C_1 depends on the center of the small ball probability of $\mathbf{G}$. The case when C_1 does not depend on the center of the small ball probability of $\mathbf{G}$ can be covered similarly. So, we have

$$\begin{aligned}
&\mathbb{E}\left[\phi_{\mathbf{G}}\left(\mathbf{t}\mathbf{U},\ h\mathbf{U}\right)\right] \\
&\geq \mathbb{E}\left[C_1' \exp[-(1/2)\|\mathbf{t}\mathbf{U}\|^2](h\mathbf{U})^{t_1} \exp\left[-m(h\mathbf{U})\right]\right] \\
&\geq \mathbb{E}\left[C_1' \exp[-(1/2)\mathbf{U}^2\|\mathbf{t}\|^2](h\mathbf{U})^{t_1} \exp\left[-m(h\mathbf{U})\right]\mathbb{I}(\mathbf{U}\geq l_0)\right] \\
&\geq \mathbb{E}\left[C_1' \exp[-(1/2)u_0^2\|\mathbf{t}\|^2](h\mathbf{U})^{t_1} \exp\left[-m(h\mathbf{U})\right]\mathbb{I}(\mathbf{U}\geq l_0)\right] \\
&\geq C_1' \exp[-(1/2)u_0^2\|\mathbf{t}\|^2] l_0^{t_1} h^{t_1} \exp\left[-m(hl_0)\right] \\
&= C_1' \exp[-(1/2)u_0^2\|\mathbf{t}\|^2] l_0^{t_1} h^{t_1} \exp\left[-\left(\frac{1}{l_0}\right)^{t_2}\left(1-\frac{\log l_0}{\log\frac{1}{h}}\right)^{t_3}\left(\frac{1}{h}\right)^{t_2}\left(\log\frac{1}{h}\right)^{t_3}\right] \\
&\geq C_1' \exp[-(1/2)u_0^2\|\mathbf{t}\|^2] l_0^{t_1} h^{t_1} \exp\left[-2\left(\frac{1}{l_0}\right)^{t_2}\left(\frac{1}{h}\right)^{t_2}\left(\log\frac{1}{h}\right)^{t_3}\right] \qquad (4)
\end{aligned}$$

for all sufficiently small h. The proof is completed combining (2), (3) and (4). □

Note that if $\mathbf{T}$ is an infinite dimensional t-process with degree k, it can be expressed as $\mathbf{T} = \mathbf{G}/\sqrt{\boldsymbol{\chi}/k}$, where $\mathbf{G}$ is an infinite dimensional Gaussian process, $\boldsymbol{\chi}$ follows a χ^2 distribution with degree of freedom k, and $\boldsymbol{\chi}$ is independent of $\mathbf{G}$. In the proposition below, we establish the bounds for the small ball probability of an infinite dimensional t-process $\mathbf{T}$.

Proposition 3 *Let $\mathbf{T}$ be an infinite dimensional t-process in some normed vector space with corresponding Gaussian process $\mathbf{G}$, and the small ball probability of $\mathbf{G}$ satisfies the bounds in (9) in the main paper with $t_2 > 0$. Then, the small ball probability of $\mathbf{T}$ also satisfies the bounds in (9).*

Proof We have

$$\begin{aligned}
\phi_{\mathbf{T}}(\mathbf{t},h) &= \mathbb{P}\left[\left\|\mathbf{G}-\mathbf{t}\sqrt{\boldsymbol{\chi}/k}\right\| \leq h\sqrt{\boldsymbol{\chi}/k}\right] \\
&= \mathbb{E}\left[\mathbb{P}\left[\left\|\mathbf{G}-\mathbf{t}\sqrt{\boldsymbol{\chi}/k}\right\| \leq h\sqrt{\boldsymbol{\chi}/k}\ \middle|\ \boldsymbol{\chi}\right]\right] \\
&= \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}\int_0^\infty \phi_{\mathbf{G}}\left(\mathbf{t}\sqrt{\frac{u}{k}},\ h\sqrt{\frac{u}{k}}\right)e^{-\frac{u}{2}}u^{\frac{k}{2}-1}du. \qquad (5)
\end{aligned}$$

Define $m_1(h) = (1/h)^{t_2}(\log(1/h))^{t_3}$ for $0 < h < 1$. Since $t_2 > 0$, $m_1(h) \longrightarrow \infty$ as $h \longrightarrow 0^+$. Let

$$t_5 = 1 + \frac{t_2}{2}. \qquad (6)$$

Define

$$U(h) = (m_1(h))^{\frac{1}{t_5}}. \qquad (7)$$

Clearly, $U(h) \longrightarrow \infty$ as $h \longrightarrow 0^+$. Also,

$$h\sqrt{U(h)} = h\left[\left(\frac{1}{h}\right)^{t_2}\left(\log\frac{1}{h}\right)^{t_3}\right]^{\frac{1}{2t_5}} = h^{\frac{1}{t_5}}\left(\log\frac{1}{h}\right)^{\frac{t_3}{2t_5}} \longrightarrow 0 \text{ as } h \longrightarrow 0^+. \tag{8}$$

So, from (9) in the main paper and (6), (7) and (8), we have for all sufficiently small h and for any $u \leq U(h)$,

$$\begin{aligned}
&\phi_{\mathbf{G}}\left(\mathbf{t}\sqrt{\frac{u}{k}},\, h\sqrt{\frac{u}{k}}\right)\\
&\leq C_3\left(h\sqrt{\frac{u}{k}}\right)^{t_4}\exp\left[-C_4 m_1\left(h\sqrt{\frac{u}{k}}\right)\right]\\
&= \frac{C_3}{k^{\frac{t_4}{2}}}u^{\frac{t_4}{2}}h^{t_4}\exp\left[-C_4 k^{\frac{t_2}{2}}u^{-\frac{t_2}{2}}\left(\frac{1}{h}\right)^{t_2}\left(\log\frac{1}{h}\right)^{t_3}\left(1+\frac{\log\sqrt{k}}{\log\frac{1}{h}}-\frac{\log\sqrt{u}}{\log\frac{1}{h}}\right)^{t_3}\right]\\
&\leq \frac{C_3}{k^{\frac{t_4}{2}}}u^{\frac{t_4}{2}}h^{t_4}\exp\left[-C_4 k^{\frac{t_2}{2}}\left(\frac{1}{t_2+2}\right)^{t_3}(m_1(h))^{\frac{1}{t_5}}\right],
\end{aligned} \tag{9}$$

since for all sufficiently small h and any $u \leq U(h)$,

$$1+\frac{\log\sqrt{k}}{\log\frac{1}{h}}-\frac{\log\sqrt{u}}{\log\frac{1}{h}} > \frac{1}{t_2+2}.$$

Hence, from (5) and (9), we have for all sufficiently small h,

$$\begin{aligned}
&\phi_{\mathbf{T}}(\mathbf{t},h)\\
&= \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}\int_0^{U(h)}\phi_{\mathbf{G}}\left(\mathbf{t}\sqrt{\frac{u}{k}},\, h\sqrt{\frac{u}{k}}\right)e^{-\frac{u}{2}}u^{\frac{k}{2}-1}du\\
&\quad+\frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}\int_{U(h)}^{\infty}\phi_{\mathbf{G}}\left(\mathbf{t}\sqrt{\frac{u}{k}},\, h\sqrt{\frac{u}{k}}\right)e^{-\frac{u}{2}}u^{\frac{k}{2}-1}du\\
&< \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}\frac{C_3}{k^{\frac{t_4}{2}}}\int_0^{U(h)}h^{t_4}\exp\left[-C_4 k^{\frac{t_2}{2}}\left(\frac{1}{t_2+2}\right)^{t_3}(m_1(h))^{\frac{1}{t_5}}\right]e^{-\frac{u}{2}}u^{\frac{t_4+k}{2}-1}du\\
&\quad+\frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}\int_{U(h)}^{\infty}\exp\left[-\frac{1}{4}U(h)\right]e^{-\frac{u}{4}}u^{\frac{k}{2}-1}du\\
&< \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}\frac{C_3}{k^{\frac{t_4}{2}}}\left[\int_0^{\infty}e^{-\frac{u}{2}}u^{\frac{t_4+k}{2}-1}du\right]h^{t_4}\exp\left[-C_4 k^{\frac{t_2}{2}}\left(\frac{1}{t_2+2}\right)^{t_3}(m_1(h))^{\frac{1}{t_5}}\right]\\
&\quad+\frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}\left[\int_0^{\infty}e^{-\frac{u}{4}}u^{\frac{k}{2}-1}du\right]\exp\left[-\frac{1}{4}(m_1(h))^{\frac{1}{t_5}}\right]\\
&= \left(\frac{\Gamma\left(\frac{t_4+k}{2}\right)}{\Gamma\left(\frac{k}{2}\right)}\left(\frac{2}{k}\right)^{\frac{t_4}{2}}C_3\right)h^{t_4}\exp\left[-C_4 k^{\frac{t_2}{2}}\left(\frac{1}{t_2+2}\right)^{t_3}(m_1(h))^{\frac{1}{t_5}}\right]+2^{\frac{k}{2}}\exp\left[-\frac{1}{4}(m_1(h))^{\frac{1}{t_5}}\right]
\end{aligned}$$

$$\leq \left(\frac{\Gamma\left(\frac{t_4+k}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} \left(\frac{2}{k}\right)^{\frac{t_4}{2}} C_3 + 2^{\frac{k}{2}} \right) \exp\left[-\min\left\{ C_4 k^{\frac{t_2}{2}} \left(\frac{1}{t_2+2}\right)^{t_3}, \frac{1}{4} \right\} \left(\frac{1}{h}\right)^{\frac{t_2}{t_5}} \left(\log\frac{1}{h}\right)^{\frac{t_3}{t_5}} \right]. \tag{10}$$

We now proceed to find a lower bound for $\phi_{\mathbf{T}}(\mathbf{t},h)$. From (9) in the main paper, (6), (7) and (8), we get that for all sufficiently small h and for any $U(h) \leq u \leq 2U(h)$,

$$\begin{aligned}
&\phi_{\mathbf{G}}\left(\mathbf{t}\sqrt{\frac{u}{k}},\, h\sqrt{\frac{u}{k}}\right) \\
&\geq C_1' \exp\left[-\frac{1}{2}\left\|\mathbf{t}\sqrt{\frac{u}{k}}\right\|^2\right] \left(h\sqrt{\frac{u}{k}}\right)^{t_1} \exp\left[-C_2 m_1\left(h\sqrt{\frac{u}{k}}\right)\right] \\
&= \frac{C_1'}{k^{\frac{t_1}{2}}} u^{\frac{t_1}{2}} h^{t_1} \exp\left[-u\frac{\|\mathbf{t}\|^2}{2k} - C_2 k^{\frac{t_2}{2}} u^{-\frac{t_2}{2}} \left(\frac{1}{h}\right)^{t_2} \left(\log\frac{1}{h}\right)^{t_3} \left(1 + \frac{\log\sqrt{k}}{\log\frac{1}{h}} - \frac{\log\sqrt{u}}{\log\frac{1}{h}}\right)^{t_3}\right] \\
&\geq \frac{C_1'}{k^{\frac{t_1}{2}}} u^{\frac{t_1}{2}} h^{t_1} \exp\left[-\frac{\|\mathbf{t}\|^2}{k}(m_1(h))^{\frac{1}{t_5}} - C_2 k^{\frac{t_2}{2}} \left(\frac{2}{t_5}\right)^{t_3} (m_1(h))^{1-\frac{t_2}{2t_5}}\right] \\
&= \frac{C_1'}{k^{\frac{t_1}{2}}} u^{\frac{t_1}{2}} h^{t_1} \exp\left[-\left(\frac{\|\mathbf{t}\|^2}{k} + C_2 k^{\frac{t_2}{2}} \left(\frac{2}{t_5}\right)^{t_3}\right) (m_1(h))^{\frac{1}{t_5}}\right],
\end{aligned} \tag{11}$$

since for all sufficiently small h and any $U(h) \leq u$,

$$1 + \frac{\log\sqrt{k}}{\log\frac{1}{h}} - \frac{\log\sqrt{u}}{\log\frac{1}{h}} < \frac{2}{t_5}.$$

From (5) and (11), we have for all sufficiently small h,

$$\begin{aligned}
&\phi_{\mathbf{T}}(\mathbf{t},h) \\
&\geq \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} \int_{U(h)}^{2U(h)} \phi_{\mathbf{G}}\left(\mathbf{t}\sqrt{\frac{u}{k}},\, h\sqrt{\frac{u}{k}}\right) e^{-\frac{u}{2}} u^{\frac{k}{2}-1} du \\
&\geq \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} \frac{C_1'}{k^{\frac{t_1}{2}}} \int_{U(h)}^{2U(h)} h^{t_1} \exp\left[-\left(\frac{\|\mathbf{t}\|^2}{k} + C_2 k^{\frac{t_2}{2}}\left(\frac{2}{t_5}\right)^{t_3}\right)(m_1(h))^{\frac{1}{t_5}}\right] e^{-\frac{u}{2}} u^{\frac{t_1+k}{2}-1} du \\
&= \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} \frac{C_1'}{k^{\frac{t_1}{2}}} \left[\int_{U(h)}^{2U(h)} e^{-\frac{u}{2}} u^{\frac{t_1+k}{2}-1} du\right] h^{t_1} \exp\left[-\left(\frac{\|\mathbf{t}\|^2}{k} + C_2 k^{\frac{t_2}{2}}\left(\frac{2}{t_5}\right)^{t_3}\right)(m_1(h))^{\frac{1}{t_5}}\right] \\
&\geq \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} \frac{C_1'}{k^{\frac{t_1}{2}}} \left[\int_{U(h)}^{2U(h)} e^{-\frac{U(h)}{2}} (U(h))^{\frac{t_1+k}{2}-1} du\right] h^{t_1} \exp\left[-\left(\frac{\|\mathbf{t}\|^2}{k} + C_2 k^{\frac{t_2}{2}}\left(\frac{2}{t_5}\right)^{t_3}\right)(m_1(h))^{\frac{1}{t_5}}\right] \\
&= \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} \frac{C_1'}{k^{\frac{t_1}{2}}} (U(h))^{\frac{t_1+k}{2}} h^{t_1} \exp\left[-\left(\frac{1}{2} + \frac{\|\mathbf{t}\|^2}{k} + C_2 k^{\frac{t_2}{2}}\left(\frac{2}{t_5}\right)^{t_3}\right)(m_1(h))^{\frac{1}{t_5}}\right] \\
&> h^{t_1} \exp\left[-\left(\frac{1}{2} + \frac{\|\mathbf{t}\|^2}{k} + C_2 k^{\frac{t_2}{2}}\left(\frac{2}{t_5}\right)^{t_3}\right)\left(\frac{1}{h}\right)^{\frac{t_2}{t_5}}\left(\log\frac{1}{h}\right)^{\frac{t_3}{t_5}}\right].
\end{aligned} \tag{12}$$

So, from (10) and (12), we have for all sufficiently small h,

$$h^{t_1} \exp\left[-u_1 \left(\frac{1}{h}\right)^{\frac{t_2}{t_5}} \left(\log \frac{1}{h}\right)^{\frac{t_3}{t_5}}\right] < \phi_{\mathbf{T}}(\mathbf{t}, h) < u_2 \exp\left[-u_3 \left(\frac{1}{h}\right)^{\frac{t_2}{t_5}} \left(\log \frac{1}{h}\right)^{\frac{t_3}{t_5}}\right],$$

where

$$u_1 = \left(\frac{1}{2} + \frac{\|\mathbf{t}\|^2}{k} + C_2 k^{\frac{t_2}{2}} \left(\frac{2}{t_5}\right)^{t_3}\right), \quad u_2 = \left(\frac{\Gamma\left(\frac{t_4+k}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} \left(\frac{2}{k}\right)^{\frac{t_4}{2}} C_3 + 2^{\frac{k}{2}}\right)$$

and $u_3 = \min\left\{C_4 k^{\frac{t_2}{2}} \left(\frac{1}{t_2+2}\right)^{t_3}, \frac{1}{4}\right\}$.

□

2 Results required to prove Theorem 7

Lemma 1 *Let $\{\mathbf{U}_n\}$ be a sequence of real random variables and let $\{\mathbf{V}_n\}$ be another sequence of positive random variables with $\mathbf{V}_n = o_{\mathbb{P}}(1)$ as $n \longrightarrow \infty$. Then, for any $a > 0$ and any $\epsilon > 0$, $\mathbb{P}[\mathbf{U}_n > a + \mathbf{V}_n] > \mathbb{P}[\mathbf{U}_n > 2a] - \epsilon$ for all sufficiently large n.*

Proof Since $\mathbf{V}_n = o_{\mathbb{P}}(1)$ as $n \longrightarrow \infty$, for any $a > 0$ and any $\epsilon > 0$,

$$\begin{aligned}
\mathbb{P}[\mathbf{U}_n > a + \mathbf{V}_n] &\geq \mathbb{P}[\mathbf{U}_n > 2a \text{ and } \mathbf{V}_n < (a/2)] \\
&\geq \mathbb{P}[\mathbf{U}_n > 2a] - \mathbb{P}[\mathbf{V}_n > (a/2)] \\
&> \mathbb{P}[\mathbf{U}_n > 2a] - \epsilon
\end{aligned}$$

for all sufficiently large n, which completes the proof. □

Lemma 2 *Suppose that in (9) in the main paper, we have either $t_2 > 0$, or $t_3 > 1$ with $C_2 = C_4$, the kernel $K(\cdot)$ satisfies A(i), and the decomposition (3) in the main paper along with conditions B(i)–B(iii), C(i) and C(ii) are satisfied. Consider a bandwidth sequence $\{h_n\}$ that satisfies A(ii) and $h_n^{2\beta} n\phi(\mathbf{x}, h_n) \longrightarrow 0$ as $n \longrightarrow \infty$. Then, there exist $c > 0$ and $\delta > 0$ such that*

$$\mathbb{P}\left[\left(m^{-1}(\log n)\right)^{-\beta} \left\|\widehat{\Theta}_n(\mathbf{x}) - \Theta(\mathbf{x})\right\| > c\right] > \delta$$

for all sufficiently large n.

Proof Recall from subsection 4.1 in the main paper that $B_n(\mathbf{x}) = \tilde{B}_n(\mathbf{x}) + \tilde{R}_n(\mathbf{x})$, where $\tilde{R}_n(\mathbf{x}) = o_{\mathbb{P}}(h_n^{\beta})$, and $\tilde{B}_n(\mathbf{x})$ is a non-random quantity. So, from (3) in the main paper and condition B(iii), we have

$$\widehat{\Theta}_n(\mathbf{x}) - \Theta(\mathbf{x}) = \tilde{B}_n(\mathbf{x}) + V_n(\mathbf{x}) + Q_n(\mathbf{x}), \tag{13}$$

where $Q_n(\mathbf{x}) = R_n(\mathbf{x}) + \tilde{R}_n(\mathbf{x}) = o_{\mathbb{P}}\big(\max\big\{h_n^{\beta}, \big[n\phi(\mathbf{x}, h_n)\big]^{-1/2}\big\}\big)$ as $n \longrightarrow \infty$.

Recall the projection functional $\tilde{\phi}_i(\cdot)$ defined in subsection 4.1 and the positive integer i_0 mentioned in condition C(ii). Note that $\|\tilde{\phi}_{i_0}\| = 1$. So, for all $\mathbf{v} \in \mathcal{B}$,

$$|\tilde{\phi}_{i_0}(\mathbf{v})| \leq \|v\|. \tag{14}$$

Using A(i), A(ii), B(ii), C(ii) and arguments similar to those in Theorem 4, we get

$$[n\phi(\mathbf{x}, h_n)]^{1/2}[E_n^{(2)}(\mathbf{x})]^{-1/2}E_n^{(1)}(\mathbf{x})\tilde{\phi}_{i_0}(V_n(\mathbf{x})) \longrightarrow \mathbf{Z} \tag{15}$$

in distribution as $n \longrightarrow \infty$, where $\mathbf{Z}$ follows a normal distribution with mean zero and variance $\mathbb{V}(\mathbf{x}) > 0$.

Next, consider $\{h_n\}$ that satisfies A(ii) and

$$h_n^{2\beta} n\phi(\mathbf{x}, h_n) \longrightarrow 0 \text{ as } n \longrightarrow \infty. \tag{16}$$

From (40) in the main paper and (16), we get that for all sufficiently large n,

$$\begin{aligned} &[n\phi(\mathbf{x}, h_n)]^{-1/2} > h_n^{\beta} > c_1^{\beta}\left(m^{-1}(\log n)\right)^{\beta} \\ \implies &\left(m^{-1}(\log n)\right)^{-\beta}[n\phi(\mathbf{x}, h_n)]^{-1/2} > c_1^{\beta}, \end{aligned} \tag{17}$$

where $c_1 > 0$ is a constant. Since $Q_n(\mathbf{x}) = o_{\mathbb{P}}\big(\max\big\{h_n^{\beta}, \big[n\phi(\mathbf{x}, h_n)\big]^{-1/2}\big\}\big)$ as $n \longrightarrow \infty$, from (16), we have $Q_n(\mathbf{x}) = o_{\mathbb{P}}\big(\big[n\phi(\mathbf{x}, h_n)\big]^{-1/2}\big)$ as $n \longrightarrow \infty$. Further, from B(i), we get that $h_n^{-\beta}\tilde{B}_n(\mathbf{x})$ is bounded, and hence from (16), we have $[n\phi(\mathbf{x}, h_n)]^{1/2}\tilde{B}_n(\mathbf{x}) \longrightarrow \mathbf{0}$ as $n \longrightarrow \infty$. Therefore,

$$[n\phi(\mathbf{x}, h_n)]^{1/2}\left[\big\|\tilde{B}_n(\mathbf{x})\big\| + \big\|Q_n(\mathbf{x})\big\|\right] = o_{\mathbb{P}}(1) \tag{18}$$

as $n \longrightarrow 0$. Take

$$c = \frac{lc_1^{\beta}}{2L} \quad \text{and} \quad \delta = \frac{1}{2}\mathbb{P}[|\mathbf{Z}| > 1],$$

where $\mathbf{Z}$ is the normal random variable described in (15). So, from (8) in the main paper, Lemma 1, (14), (15), (17), (18) and the triangle inequality, we have for all sufficiently large n,

$$\begin{aligned} &\mathbb{P}\big[\left(m^{-1}(\log n)\right)^{-\beta}\big\|\widehat{\Theta}_n(\mathbf{x}) - \Theta(\mathbf{x})\big\| > c\big] \\ &\geq \mathbb{P}\left[\frac{[n\phi(\mathbf{x}, h_n)]^{1/2}\left[\big\|V_n(\mathbf{x})\big\| - \big\|\tilde{B}_n(\mathbf{x})\big\| - \big\|Q_n(\mathbf{x})\big\|\right]}{\left(m^{-1}(\log n)\right)^{\beta}[n\phi(\mathbf{x}, h_n)]^{1/2}} > c\right] \end{aligned}$$

$$\geq \mathbb{P}\left[[n\phi(\mathbf{x},h_n)]^{1/2}\left|\tilde{\phi}_{i_0}(V_n(\mathbf{x}))\right| > cc_1^{-\beta} + [n\phi(\mathbf{x},h_n)]^{1/2}\left[\left\|\tilde{B}_n(\mathbf{x})\right\| + \left\|Q_n(\mathbf{x})\right\|\right]\right]$$
$$\geq \mathbb{P}\left[[n\phi(\mathbf{x},h_n)]^{1/2}\left|\tilde{\phi}_{i_0}(V_n(\mathbf{x}))\right| > 2cc_1^{-\beta}\right] - \frac{\delta}{2}$$
$$\geq \mathbb{P}\left[\left|[n\phi(\mathbf{x},h_n)]^{1/2}[E_n^{(2)}(\mathbf{x})]^{-1/2}E_n^{(1)}(\mathbf{x})\tilde{\phi}_{i_0}(V_n(\mathbf{x}))\right| > 2cc_1^{-\beta}\frac{L}{l}\right] - \frac{\delta}{2}$$
$$= \mathbb{P}\left[\left|[n\phi(\mathbf{x},h_n)]^{1/2}[E_n^{(2)}(\mathbf{x})]^{-1/2}E_n^{(1)}(\mathbf{x})\tilde{\phi}_{i_0}(V_n(\mathbf{x}))\right| > 1\right] - \frac{\delta}{2} > \delta.$$

□

Lemma 3 *Suppose that in* (9) *in the main paper, we have either* $t_2 > 0$, *or* $t_3 > 1$ *with* $C_2 = C_4$, *the kernel* $K(\cdot)$ *satisfies A(i), and the decomposition* (3) *in the main paper along with conditions B(i)–B(iii), C(i) and C(ii) are satisfied. Consider a bandwidth sequence* $\{h_n\}$ *that satisfies A(ii) and* $h_n^{2\beta} n\phi(\mathbf{x}, h_n) \longrightarrow \infty$ *as* $n \longrightarrow \infty$. *Then, there exist* $c > 0$ *and* $\delta > 0$ *such that*

$$\mathbb{P}\left[\left(m^{-1}(\log n)\right)^{-\beta}\left\|\widehat{\Theta}_n(\mathbf{x}) - \Theta(\mathbf{x})\right\| > c\right] > \delta$$

for all sufficiently large n.

Proof Consider $\{h_n\}$ that satisfies A(ii) and

$$h_n^{2\beta} n\phi(\mathbf{x}, h_n) \longrightarrow \infty \text{ as } n \longrightarrow \infty. \tag{19}$$

Let $Q_n(\mathbf{x})$ be as defined in (13). Since $Q_n(\mathbf{x}) = o_{\mathbb{P}}\big(\max\{h_n^{\beta}, [n\phi(\mathbf{x},h_n)]^{-1/2}\}\big)$ as $n \longrightarrow \infty$, from (19), we have $Q_n(\mathbf{x}) = o_{\mathbb{P}}\big(h_n^{\beta}\big)$ as $n \longrightarrow \infty$. Further, from Theorem 3 in the main paper and (19), we get

$$h_n^{-2\beta}\mathbb{E}[\|V_n(\mathbf{x})\|^2] = h_n^{-2\beta}\,[n\phi(\mathbf{x},h_n)]^{-1}\, n\phi(\mathbf{x},h_n)\mathbb{E}[\|V_n(\mathbf{x})\|^2] \longrightarrow 0$$

as $n \longrightarrow \infty$, which implies $h_n^{-\beta} V_n(\mathbf{x}) = o_{\mathbb{P}}\big(1\big)$ as $n \longrightarrow \infty$. Therefore,

$$h_n^{-\beta}\left[\left\|V_n(\mathbf{x})\right\| + \left\|Q_n(\mathbf{x})\right\|\right] = o_{\mathbb{P}}\big(1\big) \tag{20}$$

as $n \longrightarrow \infty$. Note that we have chosen $\Theta(\mathbf{x})$ satisfying C(i), so that for any kernel $K(\cdot)$ satisfying A(i) and any sequence of bandwidths $\{h_n\}$ satisfying A(ii), we have for all sufficiently large n,

$$h_n^{-\beta}\|\tilde{B}_n(\mathbf{x})\| \geq b_1 > 0, \tag{21}$$

where b_1 is a constant. Take

$$c = \frac{b_1 c_1^{\beta}}{4} \quad \text{and} \quad \delta = \frac{1}{2}.$$

Then, from (40) in the main paper, Lemma 1, (20), (21) and the triangle inequality, we have for all sufficiently large n,

$$\mathbb{P}\big[\left(m^{-1}(\log n)\right)^{-\beta}\left\|\widehat{\Theta}_n(\mathbf{x}) - \Theta(\mathbf{x})\right\| > c\big]$$

$$
\begin{aligned}
&\geq \mathbb{P}\left[\frac{h_n^{-\beta}\left[\|\tilde{B}_n(\mathbf{x})\| - \|V_n(\mathbf{x})\| - \|Q_n(\mathbf{x})\|\right]}{(m^{-1}(\log n))^{\beta}\, h_n^{-\beta}} > c\right] \\
&\geq \mathbb{P}\left[h_n^{-\beta}\|\tilde{B}_n(\mathbf{x})\| > cc_1^{-\beta} + h_n^{-\beta}\left[\|V_n(\mathbf{x})\| + \|Q_n(\mathbf{x})\|\right]\right] \\
&\geq \mathbb{P}\left[h_n^{-\beta}\|\tilde{B}_n(\mathbf{x})\| > 2cc_1^{-\beta}\right] - \frac{1}{4} \\
&= \mathbb{P}\left[h_n^{-\beta}\|\tilde{B}_n(\mathbf{x})\| > \frac{b_1}{2}\right] - \frac{1}{4} = \frac{3}{4} > \delta.
\end{aligned}
$$

□

Lemma 4 *Suppose that in* (9) *in the main paper, we have either* $t_2 > 0$, *or* $t_3 > 1$ *with* $C_2 = C_4$, *the kernel* $K(\cdot)$ *satisfies A(i), and the decomposition* (3) *in the main paper along with conditions B(i)–B(iii), C(i) and C(ii) are satisfied. Consider a bandwidth sequence* $\{h_n\}$ *that satisfies A(ii), and* $0 < \epsilon_1 < h_n^{2\beta} n\phi(\mathbf{x}, h_n) < \epsilon_2$ *for all sufficiently large* n *and some* ϵ_1 *and* ϵ_2. *Then, there exist* $c > 0$ *and* $\delta > 0$ *such that*

$$
\mathbb{P}\left[\left(m^{-1}(\log n)\right)^{-\beta}\left\|\widehat{\Theta}_n(\mathbf{x}) - \Theta(\mathbf{x})\right\| > c\right] > \delta
$$

for all sufficiently large n.

Proof Consider $\{h_n\}$ that satisfies A(ii) and

$$
0 < \epsilon_1 < h_n^{2\beta} n\phi(\mathbf{x}, h_n) < \epsilon_2 \tag{22}
$$

for all sufficiently large n and some ϵ_1 and ϵ_2. From (40) in the main paper and (22), we get that for all sufficiently large n,

$$
\left(m^{-1}(\log n)\right)^{\beta}[n\phi(\mathbf{x}, h_n)]^{1/2} < \frac{[n\phi(\mathbf{x}, h_n)]^{1/2} h_n^{\beta}}{c_1^{\beta}} < \frac{\sqrt{\epsilon_2}}{c_1^{\beta}}, \tag{23}
$$

where $c_1 > 0$ is a constant. Let $Q_n(\mathbf{x})$ be as defined in (13). Since $Q_n(\mathbf{x}) = o_{\mathbb{P}}\left(\max\left\{h_n^{\beta}, [n\phi(\mathbf{x}, h_n)]^{-1/2}\right\}\right)$ as $n \longrightarrow \infty$, from (22), we have

$$
\begin{aligned}
&\max\left\{h_n^{\beta}, [n\phi(\mathbf{x}, h_n)]^{-1/2}\right\} \leq \max\{\sqrt{\epsilon_2}, 1\}[n\phi(\mathbf{x}, h_n)]^{-1/2} \\
\Longrightarrow\ &[n\phi(\mathbf{x}, h_n)]^{1/2}\|Q_n(\mathbf{x})\| = o_{\mathbb{P}}(1)
\end{aligned} \tag{24}
$$

as $n \longrightarrow \infty$. From A(ii), B(i) and (22), we get

$$
[n\phi(\mathbf{x}, h_n)]^{1/2}\|\tilde{B}_n(\mathbf{x})\| \leq [n\phi(\mathbf{x}, h_n)]^{1/2} h_n^{\beta} h_n^{-\beta}\|\tilde{B}_n(\mathbf{x})\| \leq \sqrt{\epsilon_2}\|\mathbb{L}_{\mathbf{x}}\| b_F \tag{25}
$$

for all sufficiently large n. Take

$$
c = \frac{c_1^{\beta} l}{2\sqrt{\epsilon_2} L} \quad \text{and} \quad \delta = \frac{1}{2}\mathbb{P}\left[|\mathbf{Z}| > 1 + \sqrt{\epsilon_2}\frac{L}{l}\|\mathbb{L}_{\mathbf{x}}\| b_F\right],
$$

where $\mathbf{Z}$ is the normal random variable described in (15), and l and L are the constants described in A(i). So, from (8) in the main paper, Lemma 1, (14), (15), (23), (24), (25) and the triangle inequality, we have for all sufficiently large n,

$$\begin{aligned}
&\mathbb{P}\big[\left(m^{-1}(\log n)\right)^{-\beta} \|\widehat{\Theta}_n(\mathbf{x}) - \Theta(\mathbf{x})\| > c\big] \\
&\geq \mathbb{P}\left[\frac{[n\phi(\mathbf{x}, h_n)]^{1/2}\left[\|V_n(\mathbf{x})\| - \|\tilde{B}_n(\mathbf{x})\| - \|Q_n(\mathbf{x})\|\right]}{(m^{-1}(\log n))^{\beta}\, [n\phi(\mathbf{x}, h_n)]^{1/2}} > c \right] \\
&\geq \mathbb{P}\left[[n\phi(\mathbf{x}, h_n)]^{1/2} \left|\tilde{\phi}_{i_0}\left(V_n(\mathbf{x})\right)\right| > c\sqrt{\epsilon_2} c_1^{-\beta} + \sqrt{\epsilon_2}\|\mathbb{L}_{\mathbf{x}}\| b_F + [n\phi(\mathbf{x}, h_n)]^{1/2}\|Q_n(\mathbf{x})\| \right] \\
&\geq \mathbb{P}\left[[n\phi(\mathbf{x}, h_n)]^{1/2}[E_n^{(2)}(\mathbf{x})]^{-1/2} E_n^{(1)}(\mathbf{x}) \left|\tilde{\phi}_{i_0}\left(V_n(\mathbf{x})\right)\right| > \frac{2\sqrt{\epsilon_2}L}{c_1^{\beta} l} c + \sqrt{\epsilon_2}\frac{L}{l}\|\mathbb{L}_{\mathbf{x}}\| b_F \right] - \frac{\delta}{2} \\
&\geq \mathbb{P}\left[\left| [n\phi(\mathbf{x}, h_n)]^{1/2}[E_n^{(2)}(\mathbf{x})]^{-1/2} E_n^{(1)}(\mathbf{x}) \tilde{\phi}_{i_0}\left(V_n(\mathbf{x})\right)\right| > 1 + \sqrt{\epsilon_2}\frac{L}{l}\|\mathbb{L}_{\mathbf{x}}\| b_F \right] - \frac{\delta}{2} \\
&> \delta.
\end{aligned}$$

□

3 Results required to prove Theorem 9

Lemma 5 *Suppose assumptions A(i) and A(ii) are satisfied. Let $\{h_n^{(b)}\}$ be a sequence of bandwidths that satisfies A(ii) and balances the bias and the variance so that*

$$0 < c_1 \leq (h_n^{(b)})^{2\beta} n\phi(\mathbf{x}, h_n^{(b)}) \leq c_2 < \infty \tag{26}$$

for all sufficiently large n, where c_1, c_2 are some constants. Also, let $\{h_n^{(op)}\}$ denote the sequence of optimum bandwidths minimizing (25) *in the proof of Theorem 5 in the main paper. Assume that $t_2 > 0$ in the bounds on the small ball probability of the covariate in* (9) *in the main paper. Then,*

$$0 < c_3 \leq \frac{h_n^{(b)}}{h_n^{(op)}} \leq c_4 < \infty$$

for all sufficiently large n, where c_3, c_4 are some constants.

Proof Recall from (10) in the main paper that $m(h) = C_2(1/h)^{t_2}(\log(1/h))^{t_3}$ for $0 < h < 1$. From (9) in the main paper and (26), we have

$$\begin{aligned}
&(h_n^{(b)})^{2\beta+t_1} n C_1 \exp\left[-m(h_n^{(b)})\right] \leq c_2 \\
&\text{and} \quad c_1 \leq (h_n^{(b)})^{2\beta+t_4} n C_3 \exp\left[-(C_4/C_2) m(h_n^{(b)})\right] \\
\Longrightarrow\ &(h_n^{(b)})^{2\beta+t_1} n \exp\left[-m(h_n^{(b)})\right] \leq \frac{c_2}{C_1} \\
&\text{and} \quad \frac{c_1}{C_3} \leq (h_n^{(b)})^{2\beta+t_4} n \exp\left[-(C_4/C_2) m(h_n^{(b)})\right]
\end{aligned}$$

$$\Longrightarrow \frac{-(2\beta+t_1)\log\frac{1}{h_n^{(b)}}}{m(h_n^{(b)})}+\frac{\log n}{m(h_n^{(b)})}-1\le\frac{\log\frac{c_2}{C_1}}{m(h_n^{(b)})}$$

$$\text{and}\quad \frac{\log\frac{c_1}{C_3}}{m(h_n^{(b)})}\le\frac{-(2\beta+t_4)\log\frac{1}{h_n^{(b)}}}{m(h_n^{(b)})}+\frac{\log n}{m(h_n^{(b)})}-\frac{C_4}{C_2} \tag{27}$$

for all sufficiently large n. When $t_2>0$ in (9), we have

$$\frac{-(2\beta+t_1)\log\frac{1}{h_n^{(b)}}}{m(h_n^{(b)})}\longrightarrow 0,\quad \frac{\log\frac{c_2}{C_1}}{m(h_n^{(b)})}\longrightarrow 0,$$

$$\frac{\log\frac{c_1}{C_3}}{m(h_n^{(b)})}\longrightarrow 0\quad\text{and}\quad\frac{-(2\beta+t_4)\log\frac{1}{h_n^{(b)}}}{m(h_n^{(b)})}\longrightarrow 0$$

as $n\longrightarrow\infty$. Therefore, given any $\epsilon>0$, from (27), we have for all sufficiently large n,

$$\frac{\log n}{m(h_n^{(b)})}\le 1+\epsilon\quad\text{and}\quad\frac{C_4}{C_2}-\epsilon\le\frac{\log n}{m(h_n^{(b)})}$$

$$\Longrightarrow\frac{\log n}{1+\epsilon}\le m(h_n^{(b)})\le\frac{\log n}{(C_4/C_2)-\epsilon}$$

$$\Longrightarrow m^{-1}\left(\frac{\log n}{1+\epsilon}\right)\ge h_n^{(b)}\ge m^{-1}\left(\frac{\log n}{(C_4/C_2)-\epsilon}\right). \tag{28}$$

Next, we consider our optimum bandwidth $h_n^{(op)}$. From (34) in the proof of Theorem 5 in the main paper, we have, given any $\epsilon>0$ and for all sufficiently large n,

$$m^{-1}\left(\frac{\log n}{1+\epsilon}\right)\ge h_n^{(op)}\ge m^{-1}\left(\frac{\log n}{1-\epsilon}\right). \tag{29}$$

Since $m(h)$ is strictly monotone decreasing function for $h\in(0,1)$ and $m(h)\longrightarrow\infty$ as $h\longrightarrow 0^+$, $m^{-1}(u)$ is well-defined for all $u>1$ and $m^{-1}(u)\longrightarrow 0^+$ as $u\longrightarrow\infty$. Given $\epsilon>0$, we have

$$m\left(c^{-\frac{1}{t_2}}(1+\epsilon)h\right)=cm(h)\frac{1}{(1+\epsilon)^{t_2}}\left(1-\frac{\log(1+\epsilon)}{\log\frac{1}{h}}+\frac{1}{t_2}\frac{\log c}{\log\frac{1}{h}}\right)^{t_3},$$

$$m\left(c^{-\frac{1}{t_2}}(1-\epsilon)h\right)=cm(h)\frac{1}{(1-\epsilon)^{t_2}}\left(1-\frac{\log(1-\epsilon)}{\log\frac{1}{h}}+\frac{1}{t_2}\frac{\log c}{\log\frac{1}{h}}\right)^{t_3}.$$

For sufficiently small $h>0$, we have

$$\frac{1}{(1+\epsilon)^{t_2}}\left(1-\frac{\log(1+\epsilon)}{\log\frac{1}{h}}+\frac{1}{t_2}\frac{\log c}{\log\frac{1}{h}}\right)^{t_3}<1<\frac{1}{(1-\epsilon)^{t_2}}\left(1-\frac{\log(1-\epsilon)}{\log\frac{1}{h}}+\frac{1}{t_2}\frac{\log c}{\log\frac{1}{h}}\right)^{t_3},$$

which implies

$$m\left(c^{-\frac{1}{t_2}}(1+\epsilon)h\right)<cm(h)<m\left(c^{-\frac{1}{t_2}}(1-\epsilon)h\right)$$

for all sufficiently small $h > 0$. Hence, for all sufficiently large u, we have

$$m\left(c^{-\frac{1}{t_2}}(1+\epsilon)m^{-1}(u)\right) < cu < m\left(c^{-\frac{1}{t_2}}(1-\epsilon)m^{-1}(u)\right)$$
$$\implies c^{-\frac{1}{t_2}}(1-\epsilon) < \frac{m^{-1}(cu)}{m^{-1}(u)} < c^{-\frac{1}{t_2}}(1+\epsilon). \quad (30)$$

From (30), we get that for any $c > 0$,

$$\frac{m^{-1}(cu)}{m^{-1}(u)} \longrightarrow c^{-\frac{1}{t_2}} \quad (31)$$

as $u \longrightarrow \infty$. Therefore, using (28), (29) and (31), we have

$$0 < c_3 \leq \frac{h_n^{(b)}}{h_n^{(op)}} \leq c_4 < \infty$$

for all sufficiently large n, where c_3, c_4 are some constants. □

Lemma 6 *We denote our optimum bandwidth minimizing* (25) *in the proof of Theorem 5 as* $h_n^{(op)}$. *Let* $\widehat{\Theta}_n^{(op)}(\mathbf{x})$ *be as defined in Theorem 9. Then, under the conditions in Theorem 9,*

$$(h_n^{(op)})^{-\beta}\left\|\widehat{\Theta}_n^{(op)}(\mathbf{x}) - \Theta(\mathbf{x})\right\| = o_{\mathbb{P}}(1) \quad \textit{as } n \longrightarrow \infty,$$
$$\textit{and} \quad (h_n^{(op)})^{-2\beta}\mathbb{E}\left\|\widehat{\Theta}_n^{(op)}(\mathbf{x}) - \Theta(\mathbf{x})\right\|^2 \longrightarrow 0 \quad \textit{as } n \longrightarrow \infty.$$

Proof From (29) in the proof of Theorem 5 and the lower bound of $\phi(\mathbf{x}, h)$ in (9) in the main paper, we get

$$(h_n^{(op)})^{2\beta} n\phi(\mathbf{x}, h_n^{(op)}) \longrightarrow \infty \text{ as } n \longrightarrow \infty. \quad (32)$$

Since $F(\cdot) \in \mathcal{F}(\mathbf{x}, \beta_1, \mathcal{G})$ for some $\beta_1 > \beta$, we have

$$(d(\mathbf{x}, \mathbf{z}))^{-\beta}\|F(\mathbf{z}) - F(\mathbf{x})\| \longrightarrow 0 \quad \text{as } d(\mathbf{x}, \mathbf{z}) \longrightarrow 0.$$

Consequently,

$$(h_n^{(op)})^{-\beta}\left\|B_n^{(op)}(\mathbf{x})\right\| = o_{\mathbb{P}}(1) \quad \text{as } n \longrightarrow \infty, \quad (33)$$
$$\text{and } (h_n^{(op)})^{-2\beta}\mathbb{E}\left\|B_n^{(op)}(\mathbf{x})\right\|^2 \longrightarrow 0 \quad \text{as } n \longrightarrow \infty. \quad (34)$$

From Theorem 3 and (32), we have

$$(h_n^{(op)})^{-2\beta}\mathbb{E}\left\|V_n^{(op)}(\mathbf{x})\right\|^2$$
$$= \left((h_n^{(op)})^{-2\beta}\left(n\phi(\mathbf{x}, h_n^{(op)})\right)^{-1}\right) n\phi(\mathbf{x}, h_n^{(op)})\mathbb{E}\left\|V_n^{(op)}(\mathbf{x})\right\|^2 \longrightarrow 0 \quad (35)$$

as $n \longrightarrow \infty$, and from (35) and the Markov inequality, we get

$$(h_n^{(op)})^{-\beta} \left\| V_n^{(op)}(\mathbf{x}) \right\| = o_{\mathbb{P}}(1) \text{ as } n \longrightarrow \infty. \quad (36)$$

From condition B(iii) and (32), we have

$$(h_n^{(op)})^{-\beta} \left\| R_n^{(op)}(\mathbf{x}) \right\| = o_{\mathbb{P}}(1) \quad \text{as } n \longrightarrow \infty. \quad (37)$$

Since $\|\tilde{\phi}_{i_0}\| = 1$, when $\mathbb{E}[\|R_n(\mathbf{x})\|^2] = o(\delta_n^2)$ as $n \longrightarrow \infty$, from (32), we have

$$(h_n^{(op)})^{-2\beta} \mathbb{E} \left\| R_n^{(op)}(\mathbf{x}) \right\|^2 \longrightarrow 0 \quad \text{as } n \longrightarrow \infty. \quad (38)$$

Therefore, from (33), (36) and (37), we have

$$\begin{aligned}
&(h_n^{(op)})^{-\beta} \left\| \widehat{\Theta}_n^{(op)}(\mathbf{x}) - \Theta(\mathbf{x}) \right\| \\
&\leq (h_n^{(op)})^{-\beta} \left\| B_n^{(op)}(\mathbf{x}) \right\| + (h_n^{(op)})^{-\beta} \left\| V_n^{(op)}(\mathbf{x}) \right\| + (h_n^{(op)})^{-\beta} \left\| R_n^{(op)}(\mathbf{x}) \right\| \\
&= o_{\mathbb{P}}(1) \quad \text{as } n \longrightarrow \infty.
\end{aligned}$$

Further, from (34), (35) and (38), we have

$$\begin{aligned}
&(h_n^{(op)})^{-2\beta} \mathbb{E} \left\| \widehat{\Theta}_n^{(op)}(\mathbf{x}) - \Theta(\mathbf{x}) \right\|^2 \\
&\leq 3(h_n^{(op)})^{-2\beta} \mathbb{E} \left\| B_n^{(op)}(\mathbf{x}) \right\|^2 + 3(h_n^{(op)})^{-2\beta} \mathbb{E} \left\| V_n^{(op)}(\mathbf{x}) \right\|^2 + 3(h_n^{(op)})^{-2\beta} \mathbb{E} \left\| R_n^{(op)}(\mathbf{x}) \right\|^2 \\
&\longrightarrow 0 \quad \text{as } n \longrightarrow \infty.
\end{aligned}$$

□

Lemma 7 *Let $h_n^{(b)}$ and $\widehat{\Theta}_n^{(b)}(\mathbf{x})$ be as defined in Theorem 9. Then, under the conditions in Theorem 9, given any $\epsilon > 0$, there is $\delta > 0$ such that*

$$\mathbb{P}\left[(h_n^{(b)})^{-\beta} \left\| \widehat{\Theta}_n^{(b)}(\mathbf{x}) - \Theta(\mathbf{x}) \right\| > \delta \right] > 1 - \epsilon$$

for all sufficiently large n. Further,

$$(h_n^{(b)})^{-2\beta} \mathbb{E} \left\| \widehat{\Theta}_n^{(b)}(\mathbf{x}) - \Theta(\mathbf{x}) \right\|^2 \text{ is bounded away from 0 as } n \longrightarrow \infty.$$

Proof Let $h_n^{(b)}$ satisfy (26). Since $F(\cdot) \in \mathcal{F}(\mathbf{x}, \beta_1, \mathcal{G})$ for some $\beta_1 > \beta$, we have

$$(d(\mathbf{x}, \mathbf{z}))^{-\beta} \|F(\mathbf{z}) - F(\mathbf{x})\| \longrightarrow 0 \quad \text{as } d(\mathbf{x}, \mathbf{z}) \longrightarrow 0. \quad (39)$$

Consequently,

$$(h_n^{(b)})^{-\beta} \left\| B_n^{(b)}(\mathbf{x}) \right\| = o_{\mathbb{P}}(1) \quad \text{as } n \longrightarrow \infty. \quad (40)$$

Let $\mathbf{Z}$ be the normal random variable described in (15). Given any $\epsilon > 0$, there exists $\delta > 0$ such that

$$\mathbb{P}\left[|\mathbf{Z}| > 2\delta\sqrt{c_2}l^{-1}L\right] > 1-\epsilon, \tag{41}$$

where c_2 is a constant described in (26), and l, L are constants described in assumption A(i). Hence, from (8) in the main paper, (15), (26) and (41), we have

$$\begin{aligned}
&\mathbb{P}\left[\left\|(h_n^{(b)})^{-\beta}V_n^{(b)}(\mathbf{x})\right\| > 2\delta\right] \\
&= \mathbb{P}\left[\left((h_n^{(b)})^{-\beta}\left(n\phi(\mathbf{x},h_n^{(b)})\right)^{-1/2}\right)\left\|\left(n\phi(\mathbf{x},h_n^{(b)})\right)^{1/2}V_n^{(b)}(\mathbf{x})\right\| > 2\delta\right] \\
&\geq \mathbb{P}\left[\left\|\left(n\phi(\mathbf{x},h_n^{(b)})\right)^{1/2}V_n^{(b)}(\mathbf{x})\right\| > 2\delta\sqrt{c_2}\right] \\
&\geq \mathbb{P}\left[\left|\left(n\phi(\mathbf{x},h_n^{(b)})\right)^{1/2}\tilde{\phi}_{i_0}(V_n^{(b)}(\mathbf{x}))\right| > 2\delta\sqrt{c_2}\right] \\
&\geq \mathbb{P}\left[\left|\left(n\phi(\mathbf{x},h_n^{(b)})\right)^{1/2}[E_n^{(2)}(\mathbf{x})]^{-1/2}E_n^{(1)}(\mathbf{x})\tilde{\phi}_{i_0}(V_n^{(b)}(\mathbf{x}))\right| > 2\delta\sqrt{c_2}l^{-1}L\right] \\
&> 1-\epsilon
\end{aligned} \tag{42}$$

for all sufficiently large n. From condition B(iii) and (26), we have

$$(h_n^{(b)})^{-\beta}\left\|R_n^{(b)}(\mathbf{x})\right\| = o_{\mathbb{P}}(1) \quad \text{as } n \longrightarrow \infty. \tag{43}$$

Therefore, from Lemma 1, (40), (42) and (43), we have

$$\begin{aligned}
&\mathbb{P}\left[(h_n^{(b)})^{-\beta}\left\|\widehat{\Theta}_n^{(b)}(\mathbf{x}) - \Theta(\mathbf{x})\right\| > \delta\right] \\
&\geq \mathbb{P}\left[(h_n^{(b)})^{-\beta}\left\|V_n^{(b)}(\mathbf{x})\right\| - (h_n^{(b)})^{-\beta}\left\|B_n^{(b)}(\mathbf{x})\right\| - (h_n^{(b)})^{-\beta}\left\|R_n^{(b)}(\mathbf{x})\right\| > \delta\right] \\
&= \mathbb{P}\left[(h_n^{(b)})^{-\beta}\left\|V_n^{(b)}(\mathbf{x})\right\| > \delta + (h_n^{(b)})^{-\beta}\left\|B_n^{(b)}(\mathbf{x})\right\| + (h_n^{(b)})^{-\beta}\left\|R_n^{(b)}(\mathbf{x})\right\|\right] \\
&> \mathbb{P}\left[(h_n^{(b)})^{-\beta}\left\|V_n^{(b)}(\mathbf{x})\right\| > 2\delta\right] > 1-\epsilon
\end{aligned}$$

for all sufficiently large n.

We proceed to prove the second part of the lemma. Since $|\tilde{\phi}_{i_0}(\mathbf{v})| \leq \|\mathbf{v}\|$ for any $\mathbf{v}$, from an application of the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&\mathbb{E}\left\|\widehat{\Theta}_n^{(b)}(\mathbf{x}) - \Theta(\mathbf{x})\right\|^2 \\
&= \mathbb{E}\left\|B_n^{(b)}(\mathbf{x}) + V_n^{(b)}(\mathbf{x}) + R_n^{(b)}(\mathbf{x})\right\|^2 \\
&\geq \mathbb{E}\left[\tilde{\phi}_{i_0}(B_n^{(b)}(\mathbf{x})) + \tilde{\phi}_{i_0}(V_n^{(b)}(\mathbf{x})) + \tilde{\phi}_{i_0}(R_n^{(b)}(\mathbf{x}))\right]^2 \\
&= \mathbb{E}\left[\left(\tilde{\phi}_{i_0}(B_n^{(b)}(\mathbf{x}))\right)^2\right] + \mathbb{E}\left[\left(\tilde{\phi}_{i_0}(V_n^{(b)}(\mathbf{x}))\right)^2\right] + \mathbb{E}\left[\left(\tilde{\phi}_{i_0}(R_n^{(b)}(\mathbf{x}))\right)^2\right]
\end{aligned}$$

$$
\begin{aligned}
&+ 2\mathbb{E}\left[\tilde{\phi}_{i_0}(R_n^{(b)}(\mathbf{x}))\left(\tilde{\phi}_{i_0}(B_n^{(b)}(\mathbf{x})) + \tilde{\phi}_{i_0}(V_n^{(b)}(\mathbf{x}))\right)\right] \\
&\geq \mathbb{E}\left[\left(\tilde{\phi}_{i_0}(B_n^{(b)}(\mathbf{x}))\right)^2\right] + \mathbb{E}\left[\left(\tilde{\phi}_{i_0}(V_n^{(b)}(\mathbf{x}))\right)^2\right] + \mathbb{E}\left[\left(\tilde{\phi}_{i_0}(R_n^{(b)}(\mathbf{x}))\right)^2\right] \\
&- 2\left[\mathbb{E}\left[\left(\tilde{\phi}_{i_0}(R_n^{(b)}(\mathbf{x}))\right)^2\right]\right]^{1/2}\left[\mathbb{E}\left[\left(\tilde{\phi}_{i_0}(B_n^{(b)}(\mathbf{x}))\right)^2\right] + \mathbb{E}\left[\left(\tilde{\phi}_{i_0}(V_n^{(b)}(\mathbf{x}))\right)^2\right]\right]^{1/2}.
\end{aligned} \tag{44}
$$

From (39), we have

$$
(h_n^{(b)})^{-2\beta}\mathbb{E}\left[\left(\tilde{\phi}_{i_0}(B_n^{(b)}(\mathbf{x}))\right)^2\right] \leq (h_n^{(b)})^{-2\beta}\mathbb{E}\left\|B_n^{(b)}(\mathbf{x})\right\|^2 \longrightarrow 0 \tag{45}
$$

as $n \longrightarrow \infty$. From (8) in the main paper, (15) and (26), we have

$$
\begin{aligned}
&(h_n^{(b)})^{-2\beta}\mathbb{E}\left[\left(\tilde{\phi}_{i_0}(V_n^{(b)}(\mathbf{x}))\right)^2\right] \\
&= \left((h_n^{(b)})^{-2\beta}\left(n\phi(\mathbf{x}, h_n^{(b)})\right)^{-1}\right) n\phi(\mathbf{x}, h_n^{(b)})\mathbb{E}\left[\left(\tilde{\phi}_{i_0}(V_n^{(b)}(\mathbf{x}))\right)^2\right] \\
&\geq \frac{1}{c_2}\,\mathbb{P}\left[\left|\left(n\phi(\mathbf{x}, h_n^{(b)})\right)^{1/2}[E_n^{(2)}(\mathbf{x})]^{-1/2}E_n^{(1)}(\mathbf{x})\tilde{\phi}_{i_0}(V_n^{(b)}(\mathbf{x}))\right| > l^{-1}L\right] \\
&> c_6 > 0
\end{aligned} \tag{46}
$$

for all sufficiently large n and for some constant c_6. Further, since $\|\tilde{\phi}_{i_0}\| = 1$, from Theorem 3 and (26), we have

$$
\begin{aligned}
&(h_n^{(b)})^{-2\beta}\mathbb{E}\left[\left(\tilde{\phi}_{i_0}(V_n^{(b)}(\mathbf{x}))\right)^2\right] \\
&\leq \left((h_n^{(b)})^{-2\beta}\left(n\phi(\mathbf{x}, h_n^{(b)})\right)^{-1}\right) n\phi(\mathbf{x}, h_n^{(b)})\mathbb{E}\left\|V_n^{(b)}(\mathbf{x})\right\|^2 \leq \frac{c_7}{c_1}
\end{aligned} \tag{47}
$$

for some constant $c_7 > 0$ and for all sufficiently large n. Since $\|\tilde{\phi}_{i_0}\| = 1$, when $\mathbb{E}[\|R_n(\mathbf{x})\|^2] = o(\delta_n^2)$ as $n \longrightarrow \infty$, from (26) and (32), we have

$$
(h_n^{(b)})^{-2\beta}\mathbb{E}\left[\left(\tilde{\phi}_{i_0}(R_n^{(b)}(\mathbf{x}))\right)^2\right] \leq (h_n^{(b)})^{-2\beta}\mathbb{E}\left\|R_n^{(b)}(\mathbf{x})\right\|^2 \longrightarrow 0 \tag{48}
$$

as $n \longrightarrow \infty$.

Therefore, from (44), (45), (46), (47) and (48), we have

$$
(h_n^{(b)})^{-2\beta}\mathbb{E}\left\|\widehat{\Theta}_n^{(b)}(\mathbf{x}) - \Theta(\mathbf{x})\right\|^2 \geq \frac{c_6}{2} > 0
$$

for all sufficiently large n. □

4 Results required to prove Theorem 10

Lemma 8 *Let $0 < \epsilon_0 < 0.5$ be fixed. For $h \in \mathbb{H}_n$, define*

$$\tilde{D}_n(\mathbf{x}, h) = \frac{1}{(1+\epsilon_0)} \sigma^2 \zeta_n \frac{\log n}{n\phi(\mathbf{x}, h)},$$
$$\tilde{C}_n(\mathbf{x}, h) = \max_{h' \in \mathbb{H}_n} \left(\left\| \widehat{\Theta}_n(\mathbf{x}, h') - \widehat{\Theta}_n(\mathbf{x}, \max\{h, h'\}) \right\|^2 - \tilde{D}_n(\mathbf{x}, h') \right)_+ .$$

Then,

$$C_n(\mathbf{x}, h) \leq \tilde{C}_n(\mathbf{x}, h) + \max_{h' \in \mathbb{H}_n} \left(\tilde{D}_n(\mathbf{x}, h') - D_n(\mathbf{x}, h') \right)_+ .$$

Proof The proof is straight forward from the definitions of $C_n(\mathbf{x}, h)$, $D_n(\mathbf{x}, h)$, $\tilde{C}_n(\mathbf{x}, h)$ and $\tilde{D}_n(\mathbf{x}, h)$. □

Lemma 9 *Let $\tilde{D}_n(\mathbf{x}, h)$ be as defined in Lemma 8, where $h \in \mathbb{H}_n$. Then, there exists a positive integer N_1 such that for all $n \geq N_1$,*

$$\mathbb{E}\left[\max_{h' \in \mathbb{H}_n} \left(\tilde{D}_n(\mathbf{x}, h') - D_n(\mathbf{x}, h') \right)_+ \right] < \frac{1}{n^2},$$
$$\text{and} \quad \mathbb{E}\left[D_n(\mathbf{x}, h) \right] \leq 3\tilde{D}_n(\mathbf{x}, h) + \frac{3\zeta_0 \sigma^2}{n^2}.$$

Proof Define the event

$$\mathbb{U}(\mathbf{x}) = \bigcap_{h' \in \mathbb{H}_n} \left\{ \left| \frac{\widehat{\phi}(\mathbf{x}, h')}{\phi(\mathbf{x}, h')} - 1 \right| < \epsilon_0 \right\},$$

where ϵ_0 is as in Lemma 8. Since the cardinality of $\mathbb{H}_n$ is at most n, from an application of the Bernstein inequality, we get that there exists an integer n_1 such that for all $n \geq n_1$,

$$\begin{aligned} \mathbb{P}\left[(\mathbb{U}(\mathbf{x}))^c \right] &= \mathbb{P}\left[\bigcup_{h' \in \mathbb{H}_n} \left\{ \left| \widehat{\phi}(\mathbf{x}, h') - \phi(\mathbf{x}, h') \right| \geq \epsilon_0 \phi(\mathbf{x}, h') \right\} \right] \\ &\leq \sum_{h' \in \mathbb{H}_n} \mathbb{P}\left[\left| \sum_{i=1}^{n} \left[\mathbb{I}\left(d(\mathbf{x}, \mathbf{X}_i) \leq h' \right) - \phi(\mathbf{x}, h') \right] \right| \geq \epsilon_0 n \phi(\mathbf{x}, h') \right] \\ &< 2 \sum_{h' \in \mathbb{H}_n} \exp\left[-4 \log n \right] \leq \frac{2}{n^3}. \end{aligned} \tag{49}$$

Note that

$$\begin{aligned} &\mathbb{E}\left[\max_{h' \in \mathbb{H}_n} \left(\tilde{D}_n(\mathbf{x}, h') - D_n(\mathbf{x}, h') \right)_+ \right] \\ &= \mathbb{E}\left[\max_{h' \in \mathbb{H}_n} \left(\tilde{D}_n(\mathbf{x}, h') - D_n(\mathbf{x}, h') \right)_+ \mathbb{I}(\mathbb{U}(\mathbf{x})) \right] \end{aligned}$$

$$+\mathbb{E}\left[\max_{h'\in\mathbb{H}_n}\left(\tilde{D}_n(\mathbf{x},h')-D_n(\mathbf{x},h')\right)_+\mathbb{I}\left((\mathbb{U}(\mathbf{x}))^c\right)\right]. \tag{50}$$

When $\mathbb{I}(\mathbb{U}(\mathbf{x}))=1$, we have

$$\begin{aligned}
&(1-\epsilon_0)\phi(\mathbf{x},h')<\widehat{\phi}(\mathbf{x},h')<(1+\epsilon_0)\phi(\mathbf{x},h') \quad \text{for all } h'\in\mathbb{H}_n\\
\Longleftrightarrow\ &\frac{1}{(1+\epsilon_0)}\frac{1}{\phi(\mathbf{x},h')}<\frac{1}{\widehat{\phi}(\mathbf{x},h')}<\frac{1}{(1-\epsilon_0)}\frac{1}{\phi(\mathbf{x},h')} \quad \text{for all } h'\in\mathbb{H}_n \qquad (51)\\
\Longrightarrow\ &\max_{h'\in\mathbb{H}_n}\left(\tilde{D}_n(\mathbf{x},h')-D_n(\mathbf{x},h')\right)_+\mathbb{I}(\mathbb{U}(\mathbf{x}))=0\\
\Longrightarrow\ &\mathbb{E}\left[\max_{h'\in\mathbb{H}_n}\left(\tilde{D}_n(\mathbf{x},h')-D_n(\mathbf{x},h')\right)_+\mathbb{I}(\mathbb{U}(\mathbf{x}))\right]=0. \qquad (52)
\end{aligned}$$

Let n_2 be a positive integer such that for all $n\geq n_2$, $\zeta_n\leq(1+\epsilon_0)\zeta_0$. So, from (49), we get that for all $n\geq\max\{n_1,n_2\}$,

$$\begin{aligned}
&\mathbb{E}\left[\max_{h'\in\mathbb{H}_n}\left(\tilde{D}_n(\mathbf{x},h')-D_n(\mathbf{x},h')\right)_+\mathbb{I}\left((\mathbb{U}(\mathbf{x}))^c\right)\right]\\
&\leq\sum_{h'\in\mathbb{H}_n}\mathbb{E}\left[\left(\tilde{D}_n(\mathbf{x},h')-D_n(\mathbf{x},h')\right)_+\mathbb{I}\left((\mathbb{U}(\mathbf{x}))^c\right)\right]\\
&\leq\sum_{h'\in\mathbb{H}_n}\tilde{D}_n(\mathbf{x},h')\mathbb{P}\left[(\mathbb{U}(\mathbf{x}))^c\right]\\
&=\sum_{h'\in\mathbb{H}_n}\frac{1}{(1+\epsilon_0)}\sigma^2\zeta_n\frac{\log n}{n\phi(\mathbf{x},h')}\mathbb{P}\left[(\mathbb{U}(\mathbf{x}))^c\right]<2\zeta_0\sigma^2\frac{1}{\log n}\frac{1}{n^2}. \qquad (53)
\end{aligned}$$

Let $n_3=\min\{n\,|\,\log n>(2/(1+\epsilon_0))\sigma^2\zeta_0\}$. Then, from (50), (52) and (53), we get that for all $n\geq\max\{n_1,n_2,n_3\}$,

$$\mathbb{E}\left[\max_{h'\in\mathbb{H}_n}\left(D_n(\mathbf{x},h')-\tilde{D}_n(\mathbf{x},h')\right)_+\right]<\frac{1}{n^2}. \tag{54}$$

Next, from (49) and (51), we have for all $n\geq n_1$,

$$\begin{aligned}
\mathbb{E}\left[D_n(\mathbf{x},h)\right]&=\mathbb{E}\left[D_n(\mathbf{x},h)\mathbb{I}\left(\mathbb{U}(\mathbf{x})\right)\right]+\mathbb{E}\left[D_n(\mathbf{x},h)\mathbb{I}\left((\mathbb{U}(\mathbf{x}))^c\right)\right]\\
&\leq\frac{(1+\epsilon_0)}{(1-\epsilon_0)}\tilde{D}_n(\mathbf{x},h)+\sigma^2\zeta_n n\mathbb{P}\left[(\mathbb{U}(\mathbf{x}))^c\right]\\
&<3\tilde{D}_n(\mathbf{x},h)+\frac{3\zeta_0\sigma^2}{n^2}. \qquad (55)
\end{aligned}$$

Taking $N_1=\max\{n_1,n_2,n_3\}$, the proof is complete from (54) and (55). □

Lemma 10 *Let the assumptions of Theorem 10 be satisfied. Let $y>0$. We have for all sufficiently large n,*

$$\mathbb{P}\left[\left\|\sum_{i=1}^{n}\mathbb{L}_{\mathbf{x}}\left(G(\mathbf{Y}_i)-\mathbb{E}[G(\mathbf{Y}_i)\,|\,\mathbf{X}_i]\right)\frac{K(h'^{-1}d(\mathbf{x},\mathbf{X}_i))}{n\mathbb{E}[K(h'^{-1}d(\mathbf{x},\mathbf{X}))]}\right\|>y\right]\leq n^{-3}$$

for all $h' \in \mathbb{H}_n$. Further, given any $c_1 > 0$, $c_2 > 0$ and any $0 < \epsilon < 1$, we have, for all sufficiently large n,

$$\mathbb{P}\left[\left\|\sum_{i=1}^{n} \mathbb{L}_{\mathbf{x}}\left(G(\mathbf{Y}_i) - \mathbb{E}[G(\mathbf{Y}_i) \,|\, \mathbf{X}_i]\right) \frac{K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))}{n\mathbb{E}[K(h'^{-1}d(\mathbf{x}, \mathbf{X}))]}\right\| > c_2\sqrt{c_1 D_n(\mathbf{x}, h') + t}\right]$$
$$\leq \exp\left[-\frac{(1-\epsilon)^2 l^2 n\phi(\mathbf{x}, h') c_2^2 (c_1 D_n(\mathbf{x}, h') + t)}{16\sigma^2 L^2}\right]$$
$$+ \exp\left[-\frac{(1-\epsilon)^2 l^2 n\phi(\mathbf{x}, h') c_2 \sqrt{c_1 D_n(\mathbf{x}, h') + t}}{16\sigma L^2}\right].$$

for all $h' \in \mathbb{H}_n$ and all $t \geq 0$.

Proof We use the following result from Yurinskiĭ (1976): Let $\xi_1, \cdots, \xi_n \in \mathcal{B}$ be independent random elements with

$$\mathbb{E}\|\xi_j\|^m \leq (m!/2) b_j^2 H^{m-2}$$

for all integers $m \geq 2$. Let

$$\beta_n \geq \mathbb{E}\|\xi_1 + \cdots + \xi_n\|, \qquad U_n^2 = b_1^2 + \cdots + b_n^2.$$

If $\bar{u} = u - (\beta_n/U_n) > 0$, then

$$\mathbb{P}[\|\xi_1 + \cdots + \xi_n\| \geq uU_n] \leq \exp\left[-\frac{\bar{u}^2}{8(1 + (\bar{u}H/2U_n))}\right]. \tag{56}$$

Now, we choose

$$\xi_i = \mathbb{L}_{\mathbf{x}}\left(G(\mathbf{Y}_i) - \mathbb{E}[G(\mathbf{Y}_i) \,|\, \mathbf{X}_i]\right) \frac{K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))}{n\mathbb{E}[K(h'^{-1}d(\mathbf{x}, \mathbf{X}))]}$$

for $i = 1, \cdots, n$. Since $\mathcal{B}$ is a type 2 Banach space, from D(i), we have

$$\mathbb{E}\|\xi_1 + \cdots + \xi_n\|$$
$$= \mathbb{E}\left\|\sum_{i=1}^{n} \mathbb{L}_{\mathbf{x}}\left(G(\mathbf{Y}_i) - \mathbb{E}[G(\mathbf{Y}_i) \,|\, \mathbf{X}_i]\right) \frac{K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))}{n\mathbb{E}[K(h'^{-1}d(\mathbf{x}, \mathbf{X}))]}\right\|$$
$$\leq \left[\mathbb{E}\left\|\sum_{i=1}^{n} \mathbb{L}_{\mathbf{x}}\left(G(\mathbf{Y}_i) - \mathbb{E}[G(\mathbf{Y}_i) \,|\, \mathbf{X}_i]\right) \frac{K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))}{n\mathbb{E}[K(h'^{-1}d(\mathbf{x}, \mathbf{X}))]}\right\|^2\right]^{\frac{1}{2}}$$
$$\leq \left[c\sum_{i=1}^{n} \mathbb{E}\left\|\mathbb{L}_{\mathbf{x}}\left(G(\mathbf{Y}_i) - \mathbb{E}[G(\mathbf{Y}_i) \,|\, \mathbf{X}_i]\right) \frac{K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))}{n\mathbb{E}[K(h'^{-1}d(\mathbf{x}, \mathbf{X}))]}\right\|^2\right]^{\frac{1}{2}}$$
$$= \sqrt{c}\left[\sum_{i=1}^{n} \mathbb{E}\left[\mathbb{E}\left[\|\mathbb{L}_{\mathbf{x}}\left(G(\mathbf{Y}_i) - \mathbb{E}[G(\mathbf{Y}_i) \,|\, \mathbf{X}_i]\right)\|^2 \,\middle|\, \mathbf{X}_i\right] \frac{K^2(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))}{(n\mathbb{E}[K(h'^{-1}d(\mathbf{x}, \mathbf{X}))])^2}\right]\right]^{\frac{1}{2}}$$

$$\leq \sqrt{c}\frac{\sigma L}{l\sqrt{n\phi(\mathbf{x}, h')}} = \beta_n,$$

where c is a positive constant. Also, again using D(i), we get

$$\begin{aligned}\mathbb{E}\|\xi_i\|^m &= \mathbb{E}\left\|\mathbb{L}_{\mathbf{x}}\left(G(\mathbf{Y}_i) - \mathbb{E}[G(\mathbf{Y}_i)\,|\,\mathbf{X}_i]\right)\frac{K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))}{n\mathbb{E}[K(h'^{-1}d(\mathbf{x}, \mathbf{X}))]}\right\|^m \\ &\leq \frac{m!}{2}\left(\frac{\sigma L}{ln\phi(\mathbf{x}, h')}\right)^{m-2}\frac{\sigma^2 L^2}{l^2 n^2\phi(\mathbf{x}, h')},\end{aligned}$$

and we can take

$$U_n^2 = \frac{\sigma^2 L^2}{l^2 n\phi(\mathbf{x}, h')} \quad \text{and} \quad H = \frac{\sigma L}{ln\phi(\mathbf{x}, h')}.$$

So, $(\beta_n/U_n) = \sqrt{c}$. Now,

$$\frac{y}{U_n} - \frac{\beta_n}{U_n} = \frac{yl\sqrt{n\phi(\mathbf{x}, h')}}{\sigma L} - \sqrt{c} \geq \frac{yl\log n}{\sigma L} - \sqrt{c} > 0$$

for all sufficiently large n and for all $h' \in \mathbb{H}_n$. Also,

$$\left(\frac{y}{U_n} - \frac{\beta_n}{U_n}\right)\frac{H}{2U_n} = \left(\frac{yl\sqrt{n\phi(\mathbf{x}, h')}}{\sigma L} - \sqrt{c}\right)\frac{1}{2\sqrt{n\phi(\mathbf{x}, h')}} < \frac{yl}{2\sigma L}.$$

So, from (56), we get that for all sufficiently large n (depending on y),

$$\begin{aligned}&\mathbb{P}\left[\left\|\sum_{i=1}^{n}\mathbb{L}_{\mathbf{x}}\left(G(\mathbf{Y}_i) - \mathbb{E}[G(\mathbf{Y}_i)\,|\,\mathbf{X}_i]\right)\frac{K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))}{n\mathbb{E}[K(h'^{-1}d(\mathbf{x}, \mathbf{X}))]}\right\| > y\right] \\ &< \exp\left[-\frac{\left(yl\log n - \sqrt{c}\sigma L\right)^2}{8\sigma^2 L^2 + 4yl\sigma L}\right] < \exp\left[-3\log n\right] = n^{-3}.\end{aligned}$$

For the next part in the statement of this lemma, we have

$$\min_{t\geq 0}\frac{c_2\sqrt{c_1 D_n(\mathbf{x}, h') + t}}{U_n} \geq \sqrt{\log n}\frac{lc_2\sqrt{c_1\frac{2}{3}\sigma^2\zeta_n}}{\sigma L} > \sqrt{c} = \frac{\beta_n}{U_n}$$

for all sufficiently large n and all $h' \in \mathbb{H}_n$. Also, given any $0 < \epsilon < 1$, we have, for all sufficiently large n,

$$\begin{aligned}&\epsilon\frac{c_2\sqrt{c_1 D_n(\mathbf{x}, h') + t}}{U_n} \geq \epsilon\sqrt{\log n}\left(\frac{lc_2\sqrt{c_1\frac{2}{3}\sigma^2\zeta_n}}{\sigma L}\right) > \sqrt{c} \\ \implies &\left(\frac{c_2\sqrt{c_1 D_n(\mathbf{x}, h') + t}}{U_n} - \sqrt{c}\right)^2 > (1-\epsilon)^2 c_2^2\frac{c_1 D_n(\mathbf{x}, h') + t}{U_n^2}\end{aligned}$$

for all $h' \in \mathbb{H}_n$ and all $t \geq 0$. Now,

$$\left(\frac{c_2\sqrt{c_1 D_n(\mathbf{x}, h') + t}}{U_n} - \frac{\beta_n}{U_n}\right)\frac{H}{2U_n} \leq c_2\sqrt{c_1 D_n(\mathbf{x}, h') + t}\frac{H}{2U_n^2} < c_2\sqrt{c_1 D_n(\mathbf{x}, h') + t}\frac{l}{\sigma L}$$

for all $h' \in \mathbb{H}_n$ and all $t \geq 0$. So, from (56), we get that for all sufficiently large n,

$$\begin{aligned}
&\mathbb{P}\left[\left\|\sum_{i=1}^{n} \mathbb{L}_{\mathbf{x}}\left(G(\mathbf{Y}_i) - \mathbb{E}[G(\mathbf{Y}_i) \mid \mathbf{X}_i]\right)\frac{K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))}{n\mathbb{E}[K(h'^{-1}d(\mathbf{x}, \mathbf{X}))]}\right\| > c_2\sqrt{c_1 D_n(\mathbf{x}, h') + t}\right] \\
&\leq \exp\left[-\frac{(1-\epsilon)^2 c_2^2 l^2 n\phi(\mathbf{x}, h')(c_1 D_n(\mathbf{x}, h') + t)}{8\sigma L^2\left(\sigma + c_2\sqrt{c_1 D_n(\mathbf{x}, h') + t}\right)}\right] \\
&\leq \exp\left[-\frac{(1-\epsilon)^2 l^2 n\phi(\mathbf{x}, h') c_2^2(c_1 D_n(\mathbf{x}, h') + t)}{16\sigma^2 L^2}\right] \\
&\quad + \exp\left[-\frac{(1-\epsilon)^2 l^2 n\phi(\mathbf{x}, h') c_2\sqrt{c_1 D_n(\mathbf{x}, h') + t}}{16\sigma L^2}\right]
\end{aligned}$$

for all $h' \in \mathbb{H}_n$ and for all $t \geq 0$. □

Lemma 11 *Let $\tilde{C}_n(\mathbf{x}, h)$ be as defined in Lemma 8, where $h \in \mathbb{H}_n$. Let the assumptions in Theorem 10 be satisfied. Then, there exists an integer N_2 such that for all $n \geq N_2$,*

$$\begin{aligned}
\tilde{C}_n(\mathbf{x}, h) \leq M_1 h^{2\beta} &+ 24 \max_{h' \in \mathbb{H}_n,\, h' \leq h}\left(\|V_n(\mathbf{x}, h')\|^2 - \frac{\tilde{D}_n(\mathbf{x}, h')}{24}\right)_+ \\
&+ 12 \max_{h' \in \mathbb{H}_n,\, h' \leq h}\left(\|R_n(\mathbf{x}, h')\|^2 - \left(M h'^{2\beta} + \|V_n(\mathbf{x}, h')\|^2\right)\right)_+
\end{aligned}$$

for all $h \in \mathbb{H}_n$, where $M_1 > 0$ is some constant. Further, for all $n \geq N_2$ and all $h \in \mathbb{H}_n$, we have

$$\mathbb{P}\left[\max_{h' \in \mathbb{H}_n,\, h' \leq h}\left(\|R_n(\mathbf{x}, h')\|^2 - \left(M h'^{2\beta} + \|V_n(\mathbf{x}, h')\|^2\right)\right)_+ > \frac{1}{n^2}\right] \leq 2n^{-2},$$

$$\text{and} \quad \mathbb{E}\left[\max_{h' \in \mathbb{H}_n,\, h' \leq h}\left(\|V_n(\mathbf{x}, h')\|^2 - \frac{\tilde{D}_n(\mathbf{x}, h')}{24}\right)_+\right] < \frac{1}{n}.$$

Proof Note that

$$\begin{aligned}
&\tilde{C}_n(\mathbf{x}, h) \\
&= \max_{h' \in \mathbb{H}_n,\, h' \leq h}\left(\left\|\widehat{\Theta}_n(\mathbf{x}, h') - \widehat{\Theta}_n(\mathbf{x}, h)\right\|^2 - \tilde{D}_n(\mathbf{x}, h')\right)_+ \\
&\leq \max_{h' \in \mathbb{H}_n,\, h' \leq h}\left(2\left\|\widehat{\Theta}_n(\mathbf{x}, h') - \Theta(\mathbf{x})\right\|^2 + 2\left\|\widehat{\Theta}_n(\mathbf{x}, h) - \Theta(\mathbf{x})\right\|^2 - \tilde{D}_n(\mathbf{x}, h')\right)_+
\end{aligned}$$

$$
\begin{aligned}
&\le 2 \max_{h' \in \mathbb{H}_n,\, h' \le h} \left(\left\| \widehat{\Theta}_n(\mathbf{x}, h') - \Theta(\mathbf{x}) \right\|^2 - \frac{\tilde{D}_n(\mathbf{x}, h')}{4} \right)_+ \\
&\quad + 2 \max_{h' \in \mathbb{H}_n,\, h' \le h} \left(\left\| \widehat{\Theta}_n(\mathbf{x}, h) - \Theta(\mathbf{x}) \right\|^2 - \frac{\tilde{D}_n(\mathbf{x}, h')}{4} \right)_+ \\
&\le 4 \max_{h' \in \mathbb{H}_n,\, h' \le h} \left(\left\| \widehat{\Theta}_n(\mathbf{x}, h') - \Theta(\mathbf{x}) \right\|^2 - \frac{\tilde{D}_n(\mathbf{x}, h')}{4} \right)_+
\end{aligned} \tag{57}
$$

since $\tilde{D}_n(\mathbf{x}, h') \ge \tilde{D}_n(\mathbf{x}, h)$ for $h' \le h$. From (3) in the main paper, we have

$$
\begin{aligned}
&\max_{h' \in \mathbb{H}_n,\, h' \le h} \left(\left\| \widehat{\Theta}_n(\mathbf{x}, h') - \Theta(\mathbf{x}) \right\|^2 - \frac{\tilde{D}_n(\mathbf{x}, h')}{4} \right)_+ \\
&\le 3 \max_{h' \in \mathbb{H}_n,\, h' \le h} \left(\|B_n(\mathbf{x}, h')\|^2 + \|V_n(\mathbf{x}, h')\|^2 + \|R_n(\mathbf{x}, h')\|^2 - \frac{\tilde{D}_n(\mathbf{x}, h')}{12} \right)_+ \\
&\le 3 \max_{h' \in \mathbb{H}_n,\, h' \le h} \left(\|B_n(\mathbf{x}, h')\|^2 + M h'^{2\beta} \right) \\
&\quad + 6 \max_{h' \in \mathbb{H}_n,\, h' \le h} \left(\|V_n(\mathbf{x}, h')\|^2 - \frac{\tilde{D}_n(\mathbf{x}, h')}{24} \right)_+ \\
&\quad + 3 \max_{h' \in \mathbb{H}_n,\, h' \le h} \left(\|R_n(\mathbf{x}, h')\|^2 - \left(M h'^{2\beta} + \|V_n(\mathbf{x}, h')\|^2 \right) \right)_+ .
\end{aligned} \tag{58}
$$

From assumption B(i) and the fact that $\max\{h' \mid h' \in \mathbb{H}_n\} \longrightarrow 0$ as $n \longrightarrow \infty$, we get that for all sufficiently large n,

$$
\max_{h' \in \mathbb{H}_n,\, h' \le h} \left(\|B_n(\mathbf{x}, h')\|^2 + M h'^{2\beta} \right) \le M_1 h^{2\beta} \tag{59}
$$

for all $h \in \mathbb{H}_n$, where $M_1 > 0$ is a constant.

Next, define the event

$$
\mathbb{S}(\mathbf{x}, h') = \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{K(h'^{-1} d(\mathbf{x}, \mathbf{X}_i))}{\mathbb{E}\left[K(h'^{-1} d(\mathbf{x}, \mathbf{X}))\right]} > (1 - \epsilon_0) \right\},
$$

where ϵ_0 is the number described in Lemma 8. From assumption D(ii) and the fact that $\max\{h' \mid h' \in \mathbb{H}_n\} \longrightarrow 0$ as $n \longrightarrow \infty$, we have for all sufficiently large n,

$$
\begin{aligned}
&\mathbb{P}\left[\max_{h' \in \mathbb{H}_n,\, h' \le h} \left(\|R_n(\mathbf{x}, h')\|^2 - \left(M h'^{2\beta} + \|V_n(\mathbf{x}, h')\|^2 \right) \right)_+ > \frac{1}{n^2} \right] \\
&\le \sum_{h' \in \mathbb{H}_n} \mathbb{P}\left[\left(\|R_n(\mathbf{x}, h')\|^2 - \left(M h'^{2\beta} + \|V_n(\mathbf{x}, h')\|^2 \right) \right)_+ > 0 \right] \\
&\le \sum_{h' \in \mathbb{H}_n} \mathbb{P}\left[\|R_n(\mathbf{x}, h')\|^2 > M h'^{2\beta} + \|V_n(\mathbf{x}, h')\|^2 \right]
\end{aligned}
$$

$$\leq \sum_{h' \in \mathbb{H}_n} \mathbb{P}\left[\|V_n(\mathbf{x}, h')\| > \epsilon_2\right]$$
$$\leq \sum_{h' \in \mathbb{H}_n} \mathbb{P}\left[\|V_n(\mathbf{x}, h')\| > \epsilon_2 \text{ and } \mathbb{I}(\mathbb{S}(\mathbf{x}, h')) = 1\right] + \sum_{h' \in \mathbb{H}_n} \mathbb{P}\left[(\mathbb{S}(\mathbf{x}, h'))^c\right]. \quad (60)$$

Now, using assumption A(i), the fact that $n\phi(\mathbf{x}, h') \geq (\log n)^2$ for all $h' \in \mathbb{H}_n$ and the Bernstein inequality, we get that for all sufficiently large n,

$$\sum_{h' \in \mathbb{H}_n} \mathbb{P}\left[(\mathbb{S}(\mathbf{x}, h'))^c\right] = \sum_{h' \in \mathbb{H}_n} \mathbb{P}\left[\frac{1}{n}\sum_{i=1}^{n}\left[1 - \frac{K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))}{\mathbb{E}\left[K(h'^{-1}d(\mathbf{x}, \mathbf{X}))\right]}\right] \geq \epsilon_0\right]$$
$$\leq \sum_{h' \in \mathbb{H}_n} \exp\left[-3\log n\right] \leq n^{-2}. \quad (61)$$

Also, from Lemma 10, we get

$$\sum_{h' \in \mathbb{H}_n} \mathbb{P}\left[\|V_n(\mathbf{x}, h')\| > \epsilon_2 \text{ and } \mathbb{I}(\mathbb{S}(\mathbf{x}, h')) = 1\right]$$
$$\leq \sum_{h' \in \mathbb{H}_n} \mathbb{P}\left[\left\|\sum_{i=1}^{n} \mathbb{L}_{\mathbf{x}}\left(G(\mathbf{Y}_i) - \mathbb{E}[G(\mathbf{Y}_i) \mid \mathbf{X}_i]\right) \frac{K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))}{n\mathbb{E}[K(h'^{-1}d(\mathbf{x}, \mathbf{X}))]}\right\| > (1-\epsilon_0)\epsilon_2\right]$$
$$\leq n^{-2} \quad (62)$$

for all sufficiently large n. Hence, from (60), (61) and (62), we have

$$\mathbb{P}\left[\max_{h' \in \mathbb{H}_n,\, h' \leq h}\left(\|R_n(\mathbf{x}, h')\|^2 - \left(Mh'^{2\beta} + \|V_n(\mathbf{x}, h')\|^2\right)\right)_+ > \frac{1}{n^2}\right] \leq 2n^{-2} \quad (63)$$

for all sufficiently large n and all $h \in \mathbb{H}_n$. Next,

$$\mathbb{E}\left[\max_{h' \in \mathbb{H}_n,\, h' \leq h}\left(\|V_n(\mathbf{x}, h')\|^2 - \frac{\tilde{D}_n(\mathbf{x}, h')}{24}\right)_+\right]$$
$$\leq \sum_{h' \in \mathbb{H}_n} \mathbb{E}\left[\left(\|V_n(\mathbf{x}, h')\|^2 - \frac{\tilde{D}_n(\mathbf{x}, h')}{24}\right)_+\right]$$
$$\leq \sum_{h' \in \mathbb{H}_n} \mathbb{E}\left[\left(\|V_n(\mathbf{x}, h')\|^2 \,\mathbb{I}(\mathbb{S}(\mathbf{x}, h')) - \frac{\tilde{D}_n(\mathbf{x}, h')}{24}\right)_+\right]$$
$$+ \sum_{h' \in \mathbb{H}_n} \mathbb{E}\left[\|V_n(\mathbf{x}, h')\|^2 \,\mathbb{I}\left((\mathbb{S}(\mathbf{x}, h'))^c\right)\right]. \quad (64)$$

Since $\mathcal{B}$ is a type 2 Banach space, from D(i) and (61), we have

$$\sum_{h' \in \mathbb{H}_n} \mathbb{E}\left[\|V_n(\mathbf{x}, h')\|^2 \,\mathbb{I}\left((\mathbb{S}(\mathbf{x}, h'))^c\right)\right]$$

$$
\begin{aligned}
&= \sum_{h' \in \mathbb{H}_n} \mathbb{E}\left[\frac{\mathbb{E}\left[\left\|\sum_{i=1}^n \mathbb{L}_{\mathbf{x}}\left(G(\mathbf{Y}_i) - \mathbb{E}[G(\mathbf{Y}_i) \,|\, \mathbf{X}_i]\right) K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))\right\|^2 \,\middle|\, \mathbf{X}_1, \cdots, \mathbf{X}_n\right]}{\left(\sum_{i=1}^n K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))\right)^2} \mathbb{I}\left((\mathbb{S}(\mathbf{x}, h'))^c\right)\right] \\
&\leq \sum_{h' \in \mathbb{H}_n} \mathbb{E}\left[\frac{c\sum_{i=1}^n \mathbb{E}\left[\left\|\mathbb{L}_{\mathbf{x}}\left(G(\mathbf{Y}_i) - \mathbb{E}[G(\mathbf{Y}_i) \,|\, \mathbf{X}_i]\right)\right\|^2 \,\middle|\, \mathbf{X}_i\right] K^2(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))}{\left(\sum_{i=1}^n K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))\right)^2} \mathbb{I}\left((\mathbb{S}(\mathbf{x}, h'))^c\right)\right] \\
&\leq c\sigma^2 \sum_{h' \in \mathbb{H}_n} \mathbb{E}\left[\frac{\sum_{i=1}^n K^2(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))}{\left(\sum_{i=1}^n K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))\right)^2} \mathbb{I}\left((\mathbb{S}(\mathbf{x}, h'))^c\right)\right] \\
&\leq c\sigma^2 \sum_{h' \in \mathbb{H}_n} \mathbb{P}\left[(\mathbb{S}(\mathbf{x}, h'))^c\right] \leq c\sigma^2 n^{-2} \qquad (65)
\end{aligned}
$$

for all sufficiently large n, where $c > 0$ is a constant. On the other hand, taking $\epsilon = \epsilon_0$ in Lemma 10, we have for all sufficiently large n,

$$
\begin{aligned}
&\sum_{h' \in \mathbb{H}_n} \mathbb{E}\left[\left(\|V_n(\mathbf{x}, h')\|^2 \mathbb{I}(\mathbb{S}(\mathbf{x}, h')) - \frac{\tilde{D}_n(\mathbf{x}, h')}{24}\right)_+\right] \\
&= \sum_{h' \in \mathbb{H}_n} \int_0^\infty \mathbb{P}\left[\left(\|V_n(\mathbf{x}, h')\|^2 \mathbb{I}(\mathbb{S}(\mathbf{x}, h')) - \frac{\tilde{D}_n(\mathbf{x}, h')}{24}\right)_+ \geq t\right] dt \\
&= \sum_{h' \in \mathbb{H}_n} \int_0^\infty \mathbb{P}\left[\|V_n(\mathbf{x}, h')\| \, \mathbb{I}(\mathbb{S}(\mathbf{x}, h')) \geq \sqrt{\frac{\tilde{D}_n(\mathbf{x}, h')}{24} + t}\right] dt \\
&\leq \sum_{h' \in \mathbb{H}_n} \int_0^\infty \mathbb{P}\left[\left\|\sum_{i=1}^n \mathbb{L}_{\mathbf{x}}\left(G(\mathbf{Y}_i) - \mathbb{E}[G(\mathbf{Y}_i) \,|\, \mathbf{X}_i]\right) \frac{K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i))}{n\mathbb{E}\left[K(h'^{-1}d(\mathbf{x}, \mathbf{X}_i)\right]}\right\| \geq (1-\epsilon_0)\sqrt{\frac{\tilde{D}_n(\mathbf{x}, h')}{24} + t}\right] dt \\
&\leq \sum_{h' \in \mathbb{H}_n} \int_0^\infty \exp\left[-\frac{(1-\epsilon_0)^4 l^2 n\phi(\mathbf{x}, h')}{16\sigma^2 L^2}\left(\frac{1}{24}D_n(\mathbf{x}, h') + t\right)\right] dt \\
&\quad + \sum_{h' \in \mathbb{H}_n} \int_0^\infty \exp\left[-\frac{(1-\epsilon_0)^3 l^2 n\phi(\mathbf{x}, h')}{16\sigma L^2}\sqrt{\frac{1}{24}D_n(\mathbf{x}, h') + t}\right] dt. \qquad (66)
\end{aligned}
$$

Now, for the second term on the right hand side of (66), we have

$$
\begin{aligned}
&\sum_{h' \in \mathbb{H}_n} \int_0^\infty \exp\left[-\frac{(1-\epsilon_0)^3 l^2 n\phi(\mathbf{x}, h')}{16\sigma L^2}\sqrt{\frac{1}{24}D_n(\mathbf{x}, h') + t}\right] dt \\
&= 2\sum_{h' \in \mathbb{H}_n} \int_{\sqrt{\frac{1}{24}D_n(\mathbf{x}, h')}}^\infty \exp\left[-\frac{(1-\epsilon_0)^3 l^2 n\phi(\mathbf{x}, h')}{16\sigma L^2} s\right] s ds \\
&< \frac{1}{n\log n} \qquad (67)
\end{aligned}
$$

for all sufficiently large n. Next, we take

$$
\zeta_0 \geq 768\frac{(1+\epsilon_0)^2}{(1-\epsilon_0)^4}\frac{L^2}{l^2}. \qquad (68)
$$

Since $\zeta_n \longrightarrow \zeta_0$ as $n \longrightarrow \infty$, we have

$$\zeta_n > 768 \frac{(1+\epsilon_0)}{(1-\epsilon_0)^4} \frac{L^2}{l^2} \tag{69}$$

for all sufficiently large n. Consequently, for the first term on the right hand side of (66), we have from (69),

$$\begin{aligned}
&\sum_{h' \in \mathbb{H}_n} \int_0^\infty \exp\left[-\frac{(1-\epsilon_0)^4 l^2 n\phi(\mathbf{x},h')}{16\sigma^2 L^2}\left(\frac{1}{24} D_n(\mathbf{x},h') + t\right)\right] dt \\
&= \sum_{h' \in \mathbb{H}_n} \int_{\frac{1}{24} D_n(\mathbf{x},h')}^\infty \exp\left[-\frac{(1-\epsilon_0)^4 l^2 n\phi(\mathbf{x},h')}{16\sigma^2 L^2} s\right] ds \\
&= \sum_{h' \in \mathbb{H}_n} \frac{16\sigma^2 L^2}{(1-\epsilon_0)^4 l^2 n\phi(\mathbf{x},h')} \exp\left[-\frac{(1-\epsilon_0)^4 l^2 n\phi(\mathbf{x},h')}{16\sigma^2 L^2} \frac{1}{24} D_n(\mathbf{x},h')\right] \\
&= \sum_{h' \in \mathbb{H}_n} \frac{16\sigma^2 L^2}{(1-\epsilon_0)^4 l^2 n\phi(\mathbf{x},h')} \exp\left[-\frac{1}{768}\frac{(1-\epsilon_0)^4}{(1+\epsilon_0)}\frac{l^2}{L^2}\zeta_n (2\log n)\right] \\
&\leq \frac{16\sigma^2 L^2}{(1-\epsilon)^4 l^2 (\log n)^2} n^{-1} < \frac{1}{n\log n}
\end{aligned} \tag{70}$$

for all sufficiently large n. Hence, from (66), (67) and (70), we have

$$\sum_{h' \in \mathbb{H}_n} \mathbb{E}\left[\left(\|V_n(\mathbf{x},h')\|^2\, \mathbb{I}(\mathbb{S}(\mathbf{x},h')) - \frac{\tilde{D}_n(\mathbf{x},h')}{24}\right)_+\right] < \frac{2}{n\log n} \tag{71}$$

for all sufficiently large n. Therefore, from (64), (65) and (71), we get that for all sufficiently large n and all $h \in \mathbb{H}_n$,

$$\mathbb{E}\left[\max_{h' \in \mathbb{H}_n,\, h' \leq h}\left(\|V_n(\mathbf{x},h')\|^2 - \frac{\tilde{D}_n(\mathbf{x},h')}{24}\right)_+\right] < \frac{1}{n}. \tag{72}$$

We choose an integer N_2 large enough such that the assertions in (59), (63) and (72) are satisfied for all $n \geq N_2$ and all $h \in \mathbb{H}_n$. Hence, the proof is complete from (57), (58), (59), (63) and (72). □

From (68), we see that ζ_0 depends on the choice of ϵ_0, and it increases with an increase in the value of ϵ_0. Taking $\epsilon_0 = 0.1$ we see that

$$\zeta_0 = 1500 \frac{L^2}{l^2} \tag{73}$$

satisfies (68). Taking smaller values of ϵ_0, we can further decrease the value of ζ_0, but it cannot be less than 768 in view of (68).

References

Yurinskiĭ, V. (1976). Exponential inequalities for sums of random vectors. *Journal of Multivariate Analysis*, 6(4):473–499.