1 Small ball probabilities of non-Gaussian processes

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

**Proposition 1** Let \( B_1 \) and \( B_2 \) be separable Banach spaces, and \( f(\cdot) : B_2 \rightarrow B_1 \) be a function such that for any \( u \in B_2 \), there exist constants \( r, s > 0 \), which may depend on \( u \), such that for any \( v \in B_2 \) sufficiently close to \( u \), we have \( r\|v - u\| \leq \|f(v) - f(u)\| \leq s\|v - u\| \). If \( T \) and \( G \) are random elements with \( T = f(G) \), and the small ball probability of \( G \) satisfies the bounds described in (9) in the main paper, then similar bounds also hold for \( T \).

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

\[
\mathbb{P} [s\|G - g\| \leq h] \leq \mathbb{P} [\|f(G) - f(g)\| \leq h] \leq \mathbb{P} [r\|G - g\| \leq h]
\]

\[
\iff \phi_G \left( g, \frac{h}{s} \right) \leq \phi_T(x, h) \leq \phi_G \left( g, \frac{h}{r} \right). \tag{1}
\]

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

$$C_1 h^{t_1} \exp \left[ -C_2 (1/h)^{t_2} (\log(1/h))^{t_3} \right] \leq \phi_{\mathbf{G}}(\mathbf{g}, h) \leq C_3 h^{t_4} \exp \left[ -C_4 (1/h)^{t_5} (\log(1/h))^{t_6} \right]$$

as $h \to 0^+$. 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 $\mathbf{T} = \mathbf{G}/U$, 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

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

Let $0 \leq U \leq u_0$ 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

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

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

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

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

$$
\mathbb{E}[\phi_G(tU, hU)] \\
\geq \mathbb{E}[C'_1 \exp[-(1/2)\|tU\|^2](hU)^{t_1}\exp[-m(hU)]] \\
\geq \mathbb{E}[C'_1 \exp[-(1/2)U^2\|t\|^2](hU)^{t_1}\exp[-m(hU)]\mathbb{I}(U \geq l_0)] \\
\geq \mathbb{E}[C'_1 \exp[-(1/2)\|t\|^2](hU)^{t_1}\exp[-m(hU)]\mathbb{I}(U \geq l_0)] \\
\geq C'_1 \exp[-(1/2)\|t\|^2]l_0^{t_1}h^{t_1}\exp[-m(hl_0)] \\
= C'_1 \exp[-(1/2)\|t\|^2]l_0^{t_1}h^{t_1} \left[- \left( \frac{1}{l_0} \right)^{t_2} \left( 1 - \frac{\log l_0}{\log h} \right)^{t_2} \left( \frac{1}{h} \right)^{t_2} \left( \log \frac{1}{h} \right)^{t_5} \right] \\
\geq C'_1 \exp[-(1/2)\|t\|^2]l_0^{t_1}h^{t_1} \left[- 2 \left( \frac{1}{l_0} \right)^{t_2} \left( \frac{1}{h} \right)^{t_2} \left( \log \frac{1}{h} \right)^{t_5} \right] \\
\geq C'_1 \exp[-(1/2)\|t\|^2]l_0^{t_1}h^{t_1} \left[- \left( \frac{1}{l_0} \right)^{t_2} \left( \frac{1}{h} \right)^{t_2} \left( \log \frac{1}{h} \right)^{t_5} \right]
$$

(4)

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

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

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

**Proof** We have

$$
\phi_T(t,h) = \mathbb{P} \left[ \left\| G - t\sqrt{\chi^2/k} \right\| \leq h\sqrt{\chi^2/k} \right] \\
= \mathbb{E} \left[ \mathbb{P} \left[ \left\| G - t\sqrt{\chi^2/k} \right\| \leq h\sqrt{\chi^2/k} \bigg\vert \chi \right] \right] \\
= \frac{1}{2^k \Gamma(\frac{k}{2})} \int_0^\infty \phi_G \left( t\sqrt{\frac{u}{k}}, h\sqrt{\frac{u}{k}} \right) e^{-\frac{u}{2}} u^{k-1} du. 
$$

(5)

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

$$
t_5 = 1 + \frac{t_2}{2}. 
$$

(6)

Define

$$
U(h) = (m_1(h))^{\frac{1}{t_2}}.
$$

(7)
Clearly, \( U(h) \to \infty \) as \( h \to 0^+ \). Also,

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

(8)

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

\[
\phi_G \left( t \sqrt{\frac{u}{k}}, h \sqrt{\frac{u}{k}} \right) \leq C_3 \left( h \sqrt{\frac{u}{k}} \right)^{t_4} \exp \left[ -C_4 m_1 \left( h \sqrt{\frac{u}{k}} \right) \right] \quad \text{(9)}
\]

\[
= \frac{C_3}{k^{t_4}} \left( h \sqrt{\frac{u}{k}} \right)^{t_4} \exp \left[ -C_4 k^{t_4} u^{-\frac{t_4}{2}} \left( \frac{1}{h} \right)^{t_2} \left( \log \frac{1}{h} \right)^{t_3} \left( 1 + \frac{\log \sqrt{k}}{\log \frac{1}{h}} - \frac{\log \sqrt{u}}{\log \frac{1}{h}} \right)^{t_3} \right] \quad \text{(9)}
\]

\[
\leq \frac{C_3}{k^{t_4}} \left( h \sqrt{\frac{u}{k}} \right)^{t_4} \exp \left[ -C_4 k^{t_4} \left( \frac{1}{t_2 + 2} \right)^{t_3} \left( m_1(h) \right)^{\frac{1}{t_3}} \right],
\]

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 \( h \),

\[
\phi_T(t, h) = \frac{1}{2^{t_4} \Gamma \left( \frac{t_4}{2} \right)} \int_{0}^{U(h)} \phi_G \left( t \sqrt{\frac{u}{k}}, h \sqrt{\frac{u}{k}} \right) e^{-\frac{t_4}{2} u^{-\frac{1}{2}}} du + \frac{1}{2^{t_4} \Gamma \left( \frac{t_4}{2} \right)} \int_{U(h)}^{\infty} \phi_G \left( t \sqrt{\frac{u}{k}}, h \sqrt{\frac{u}{k}} \right) e^{-\frac{t_4}{2} u^{-\frac{1}{2}}} du
\]

\[
< \frac{1}{2^{t_4} \Gamma \left( \frac{t_4}{2} \right)} C_3 \int_{0}^{U(h)} h^{t_4} \exp \left[ -C_4 k^{t_4} \left( \frac{1}{t_2 + 2} \right)^{t_3} \left( m_1(h) \right)^{\frac{1}{t_3}} \right] e^{-\frac{t_4}{2} u^{-\frac{1}{2}}} du + \frac{1}{2^{t_4} \Gamma \left( \frac{t_4}{2} \right)} \int_{U(h)}^{\infty} \exp \left[ -\frac{1}{4} U(h) \right] e^{-\frac{t_4}{2} u^{-\frac{1}{2}}} du
\]

\[
< \frac{1}{2^{t_4} \Gamma \left( \frac{t_4}{2} \right)} C_3 \left[ \int_{0}^{\infty} e^{-\frac{t_4}{2} u^{-\frac{1}{2}} - 1} du \right] h^{t_4} \exp \left[ -C_4 k^{t_4} \left( \frac{1}{t_2 + 2} \right)^{t_3} \left( m_1(h) \right)^{\frac{1}{t_3}} \right]
\]

\[
+ \frac{1}{2^{t_4} \Gamma \left( \frac{t_4}{2} \right)} \left[ \int_{0}^{\infty} e^{-\frac{t_4}{2} u^{-\frac{1}{2}} - 1} du \right] \exp \left[ -\frac{1}{4} (m_1(h))^{\frac{1}{t_3}} \right]
\]

\[
= \left( \Gamma \left( \frac{t_4}{2} \right) \right)^{\frac{3}{2}} C_3 \left[ \int_{0}^{\infty} e^{-\frac{t_4}{2} u^{-\frac{1}{2}} - 1} du \right] h^{t_4} \exp \left[ -C_4 k^{t_4} \left( \frac{1}{t_2 + 2} \right)^{t_3} \left( m_1(h) \right)^{\frac{1}{t_3}} \right] + 2^{\frac{t_4}{2}} \exp \left[ -\frac{1}{4} (m_1(h))^{\frac{1}{t_3}} \right]
\]
We now proceed to find a lower bound for $\phi_T(t, h)$. From (9) in the main paper, (6), (7) and (8), we get that for all sufficiently small $h$ and for any $U(h) \leq u \leq 2U(h)$,

$$\phi_G \left( \sqrt[4]{u}, h \sqrt[4]{u} \right)$$

$$\geq C'_1 \exp \left[ -\frac{1}{2} \left\| t \sqrt[4]{u} \right\| ^2 \left( h \sqrt[4]{u} \right) ^{t_1} \exp \left[ -C_2 m_1 \left( h \sqrt[4]{u} \right) \right] \right]$$

$$= \frac{C'_1}{k^{1/2}} u \frac{2}{k} h^{t_1} \exp \left[ -u \left\| t \right\| ^2 - C_2 k^{1/2} u - \frac{1}{2} \left( \frac{1}{h} \right) ^{t_2} \left( \log \frac{1}{h} \right) ^{t_3} \left( 1 + \log \frac{\sqrt{k}}{\log \frac{1}{h}} - \frac{\log \sqrt{u}}{\log \frac{1}{h}} \right) \right]$$

$$\geq \frac{C'_1}{k^{1/2}} u \frac{2}{k} h^{t_1} \exp \left[ -\left( \frac{\left\| t \right\| ^2}{k} + C_2 k^{1/2} \left( \frac{2}{t_5} \right) ^{t_3} \left( m_1(h) \right) ^{1/2} \right) \right]$$

$$= \frac{C'_1}{k^{1/2}} u \frac{2}{k} h^{t_1} \exp \left[ -\left( \frac{\left\| t \right\| ^2}{k} + C_2 k^{1/2} \left( \frac{2}{t_5} \right) ^{t_3} \left( m_1(h) \right) ^{1/2} \right) \right], \tag{11}$$

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

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

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

$$\phi_T(t, h)$$

$$\geq \frac{1}{2^{1/2} F \left( \frac{1}{2} \right) \left( 2U(h) \right) \int _{U(h)} h^{t_1} \exp \left[ -\left( \frac{\left\| t \right\| ^2}{k} + C_2 k^{1/2} \left( \frac{2}{t_5} \right) ^{t_3} \left( m_1(h) \right) ^{1/2} \right) \right] e^{-\frac{7}{2} u \frac{1}{k} - 1} du$$

$$\geq \frac{1}{2^{1/2} F \left( \frac{1}{2} \right) k^{1/2}} \left[ \int _{U(h)} e^{-\frac{7}{2} u \frac{1}{k} - 1} du \right] h^{t_1} \exp \left[ -\left( \frac{\left\| t \right\| ^2}{k} + C_2 k^{1/2} \left( \frac{2}{t_5} \right) ^{t_3} \left( m_1(h) \right) ^{1/2} \right) \right]$$

$$\geq \frac{1}{2^{1/2} F \left( \frac{1}{2} \right) k^{1/2}} \left[ \int _{U(h)} e^{-\frac{7}{2} u \frac{1}{k} - 1} du \right] h^{t_1} \exp \left[ -\left( \frac{\left\| t \right\| ^2}{k} + C_2 k^{1/2} \left( \frac{2}{t_5} \right) ^{t_3} \left( m_1(h) \right) ^{1/2} \right) \right]$$

$$= \frac{1}{2^{1/2} F \left( \frac{1}{2} \right) k^{1/2}} \left( U(h) \right) ^{t_1} h^{t_1} \exp \left[ -\left( \frac{1}{2} + \frac{\left\| t \right\| ^2}{k} + C_2 k^{1/2} \left( \frac{2}{t_5} \right) ^{t_3} \left( m_1(h) \right) ^{1/2} \right) \right]. \tag{12}$$
So, from (10) and (12), we have for all sufficiently small $h$,
\[ h^{t_1} \exp \left[ -u_1 \left( \frac{1}{h} \right)^{t_2} \left( \log \frac{1}{h} \right)^{t_3} \right] < \phi_T(t, h) < u_2 \exp \left[ -u_3 \left( \frac{1}{h} \right)^{t_2} \left( \log \frac{1}{h} \right)^{t_3} \right], \]
where
\[ u_1 = \left( \frac{1}{2} + \frac{\|t\|^2}{k} + C_2 k^{\frac{t_3}{2}} \left( \frac{2}{t_3} \right)^{t_3} \right), \quad u_2 = \left( \frac{\Gamma \left( \frac{t_3 + 2}{2} \right)}{\Gamma \left( \frac{t_3}{2} \right)} \right) \left( \frac{2}{k} \right)^{t_3} C_3 + 2^\frac{t_3}{2} \]
and $u_3 = \min \left\{ C_4 k^{\frac{t_3}{2}} \left( \frac{1}{t_2 + 2} \right)^{t_3}, \frac{1}{4} \right\}$.

\[ \square \]

2 Results required to prove Theorem 7

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

**Proof** Since $V_n = o_P(1)$ as $n \to \infty$, for any $a > 0$ and any $\epsilon > 0$,
\[ \mathbb{P}[U_n > a + V_n] \geq \mathbb{P}[U_n > 2a \text{ and } V_n < (a/2)] \]
\[ \geq \mathbb{P}[U_n > 2a] - \mathbb{P}[V_n > (a/2)] \]
\[ > \mathbb{P}[U_n > 2a] - \epsilon \]
for all sufficiently large $n$, which completes the proof. \[ \square \]

**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 $h_n^{\frac{1}{2}} \phi(x, h_n) \to 0$ as $n \to \infty$. Then, there exist $c > 0$ and $\delta > 0$ such that
\[ \mathbb{P} \left[ \left( m^{-1} \log n \right)^{-\beta} \| \hat{\Theta}_n(x) - \Theta(x) \| > c \right] > \delta \]
for all sufficiently large $n$. 

Proof Recall from subsection 4.1 in the main paper that $B_n(x) = \hat{B}_n(x) + \tilde{B}_n(x)$, where $\tilde{B}_n(x) = o_P(h_n^2)$, and $\hat{B}_n(x)$ is a non-random quantity. So, from (3) in the main paper and condition B(iii), we have

$$\hat{\Theta}_n(x) - \Theta(x) = \hat{B}_n(x) + V_n(x) + Q_n(x), \quad (13)$$

where $Q_n(x) = R_n(x) + \tilde{R}_n(x) = o_P(\max \{ h_n^2, [n\phi(x, h_n)]^{-1/2} \})$ as $n \to \infty$.

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

$$|\phi_i(v)| \leq \|v\|. \quad (14)$$

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

$$[n\phi(x, h_n)]^{-1/2}[E_n^{(2)}(x)]^{-1/2} E_n^{(1)}(x) \hat{\phi}_i(V_n(x)) \to Z \quad (15)$$

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

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

$$h_n^2 n\phi(x, h_n) \to 0 \text{ as } n \to \infty. \quad (16)$$

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

$$[n\phi(x, h_n)]^{-1/2} > h_n^2 \geq c_1 (m^{-1}(\log n))^{\beta}$$

$$\implies (m^{-1}(\log n))^{-\beta} [n\phi(x, h_n)]^{-1/2} > c_1^{\beta}, \quad (17)$$

where $c_1 > 0$ is a constant. Since $Q_n(x) = o_P(\max \{ h_n^2, [n\phi(x, h_n)]^{-1/2} \})$ as $n \to \infty$, from (16), we have $Q_n(x) = o_P([n\phi(x, h_n)]^{-1/2})$ as $n \to \infty$.

Further, from B(i), we get that $h_n^{-\beta} \hat{B}_n(x)$ is bounded, and hence from (16), we have $[n\phi(x, h_n)]^{1/2} \hat{B}_n(x) \to 0$ as $n \to \infty$. Therefore,

$$[n\phi(x, h_n)]^{1/2} \left[ \| \hat{B}_n(x) \| + \| Q_n(x) \| \right] = o_P(1) \quad (18)$$

as $n \to 0$. Take

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

where $Z$ 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 $n$,

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

$$\geq \mathbb{P} \left[ \frac{[n\phi(x, h_n)]^{1/2} \left[ \| V_n(x) \| - \| \hat{B}_n(x) \| - \| Q_n(x) \| \right]}{(m^{-1}(\log n))^{\beta} [n\phi(x, h_n)]^{1/2}} > c \right]$$
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 $h_n^{2\beta}n\phi(x, h_n) \to \infty$ as $n \to \infty$. Then, there exist $c > 0$ and $\delta > 0$ such that

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

for all sufficiently large $n$.

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

$$h_n^{2\beta}n\phi(x, h_n) \to \infty \text{ as } n \to \infty. \quad (19)$$

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

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

as $n \to \infty$, which implies $h_n^{-\beta}V_n(x) = o_P(1)$ as $n \to \infty$. Therefore,

$$h_n^{-\beta} \left[ \|V_n(x)\| + \|Q_n(x)\| \right] = o_P(1) \quad (20)$$

as $n \to \infty$. Note that we have chosen $\Theta(x)$ 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 $n$,

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

where $b_1$ is a constant. Take

$$c = \frac{b_1c_1^\beta}{4} \quad \text{and} \quad \delta = \frac{1}{2}.$$

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

$$P \left[ \left( m^{-1} \log n \right)^{-\beta} \| \hat{\Theta}_n(x) - \Theta(x) \| > c \right]$$
and (22), we get that for all sufficiently large \( n \) constants described in A(i). So, from (8) in the main paper, Lemma 1, (14),

\[
\sup_{n} \left[ \frac{h_n^{-\beta} \left( \| \hat{B}_n(x) \| - \| V_n(x) \| - \| Q_n(x) \| \right)}{(m^{-1}(\log n))^\beta h_n^{-\beta}} \right] > c
\]

\[
\geq \mathbb{P} \left[ h_n^{-\beta} \| \hat{B}_n(x) \| > c_1^{-\beta} + h_n^{-\beta} \left( \| V_n(x) \| + \| Q_n(x) \| \right) \right]
\]

\[
\geq \mathbb{P} \left[ h_n^{-\beta} \| \hat{B}_n(x) \| > 2c_1^{-\beta} - \frac{1}{4} \right]
\]

\[
= \mathbb{P} \left[ h_n^{-\beta} \| \hat{B}_n(x) \| > \frac{b_1}{2} - \frac{1}{4} = \frac{3}{4} > \delta. \right]
\]

\[\square\]

**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 \( 0 < \epsilon_1 < h_n^{-\beta} n\phi(x, h_n) < \epsilon_2 \) for all sufficiently large \( n \) and some \( \epsilon_1 \) and \( \epsilon_2 \). Then, there exist \( c > 0 \) and \( \delta > 0 \) such that

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

for all sufficiently large \( n \).

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

\[
0 < \epsilon_1 < h_n^{-\beta} n\phi(x, h_n) < \epsilon_2
\]

(22) for all sufficiently large \( n \) and some \( \epsilon_1 \) and \( \epsilon_2 \). From (40) in the main paper and (22), we get that for all sufficiently large \( n \),

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

(23)

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

\[
\max \left\{ h_n^{-\beta}, \left[ n\phi(x, h_n) \right]^{-1/2} \right\} \leq \max \{ \sqrt{\epsilon_2}, 1 \} \left[ n\phi(x, h_n) \right]^{-1/2}
\]

\[
\Rightarrow \left[ n\phi(x, h_n) \right]^{1/2} \| Q_n(x) \| = o_p(1)
\]

(24)

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

\[
\| n\phi(x, h_n) \|^{1/2} \| \hat{B}_n(x) \| \leq \| n\phi(x, h_n) \|^{1/2} h_n^\beta \| \hat{B}_n(x) \| \leq \sqrt{\epsilon_2} \| L_x \| \| b_F \|
\]

(25)

for all sufficiently large \( n \). Take

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

where \( Z \) 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),
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

\[
0 < c_1 \leq (h_n^{(b)})^{2\beta}n\phi(x, h_n^{(b)}) \leq c_2 < \infty
\]  

(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

\[
(h_n^{(b)})^{2\beta+t_1}nC_1 \exp\left[-m(h_n^{(b)})\right] \leq c_2
\]

and

\[
c_1 \leq (h_n^{(b)})^{2\beta+t_1}nC_3 \exp\left[-(C_4/C_2)m(h_n^{(b)})\right]
\]

\[
\Rightarrow (h_n^{(b)})^{2\beta+t_1}n \exp\left[-m(h_n^{(b)})\right] \leq \frac{c_2}{C_1}
\]

and

\[
\frac{C_1}{C_3} \leq (h_n^{(b)})^{2\beta+t_1}n \exp\left[-(C_4/C_2)m(h_n^{(b)})\right]
\]
\[ \log \frac{\hat{c}}{m(h_n^{(b)})} \geq -(2\beta + t_2) \log \frac{1}{h_n^{(b)}} + \frac{\log n}{m(h_n^{(b)})} - 1 \leq \frac{\log \hat{c}}{m(h_n^{(b)})}, \]

and
\[ \log \frac{\hat{c}}{m(h_n^{(b)})} \leq -(2\beta + t_4) \log \frac{1}{h_n^{(b)}} + \frac{\log n}{m(h_n^{(b)})} - \frac{C_4}{C_2}, \tag{27} \]

for all sufficiently large \( n \). When \( t_2 > 0 \) in (9), we have
\[ \log \frac{\hat{c}}{m(h_n^{(b)})} \to 0, \quad \frac{\log \hat{c}}{m(h_n^{(b)})} \to 0, \]
\[ \log \frac{\hat{c}}{m(h_n^{(b)})} \to 0 \quad \text{and} \quad \frac{\log n}{m(h_n^{(b)})} \to 0 \]
as \( n \to \infty \). Therefore, given any \( \epsilon > 0 \), from (27), we have for all sufficiently large \( n \),
\[ \frac{\log n}{m(h_n^{(b)})} \leq 1 + \epsilon \quad \text{and} \quad \frac{C_4}{C_2} - \epsilon \leq \frac{\log n}{m(h_n^{(b)})} \]
\[ \implies \frac{\log n}{1 + \epsilon} \leq m(h_n^{(b)}) \leq \frac{\log n}{(C_4/C_2) - \epsilon} \]
\[ \implies m^{-1} \left( \frac{\log n}{1 + \epsilon} \right) \geq h_n^{(b)} \geq m^{-1} \left( \frac{\log n}{(C_4/C_2) - \epsilon} \right). \tag{28} \]

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 \( n \),
\[ m^{-1} \left( \frac{\log n}{1 + \epsilon} \right) \geq h_n^{(op)} \geq m^{-1} \left( \frac{\log n}{1 - \epsilon} \right). \tag{29} \]

Since \( m(h) \) is strictly monotone decreasing function for \( h \in (0,1) \) and \( m(h) \to \infty \) as \( h \to 0^+ \), \( m^{-1}(u) \) is well-defined for all \( u > 1 \) and \( m^{-1}(u) \to 0^+ \) as \( u \to \infty \). Given \( \epsilon > 0 \), we have
\[ m \left( e^{-\frac{1}{t_2} (1 + \epsilon) h} \right) = cm(h) \frac{1}{(1 + \epsilon)^{t_2}} \left( 1 - \frac{\log(1 + \epsilon)}{\log \frac{1}{h}} + \frac{1}{t_2 \log \frac{1}{h}} \right)^{t_2}, \]
\[ m \left( e^{-\frac{1}{t_2} (1 - \epsilon) h} \right) = cm(h) \frac{1}{(1 - \epsilon)^{t_2}} \left( 1 - \frac{\log(1 - \epsilon)}{\log \frac{1}{h}} + \frac{1}{t_2 \log \frac{1}{h}} \right)^{t_2}. \]

For sufficiently small \( h > 0 \), we have
\[ \frac{1}{(1 + \epsilon)^{t_2}} \left( 1 - \frac{\log(1 + \epsilon)}{\log \frac{1}{h}} + \frac{1}{t_2 \log \frac{1}{h}} \right)^{t_2} < 1 < \frac{1}{(1 - \epsilon)^{t_2}} \left( 1 - \frac{\log(1 - \epsilon)}{\log \frac{1}{h}} + \frac{1}{t_2 \log \frac{1}{h}} \right)^{t_2}, \]
which implies
\[ m \left( e^{-\frac{1}{t_2} (1 + \epsilon) h} \right) < cm(h) < m \left( e^{-\frac{1}{t_2} (1 - \epsilon) h} \right). \]
for all sufficiently small \( h > 0 \). Hence, for all sufficiently large \( u \), we have

\[
m \left( e^{-\frac{1}{2} t} (1 + \epsilon) m^{-1}(u) \right) < cu < m \left( e^{-\frac{1}{2} t} (1 - \epsilon) m^{-1}(u) \right)
\]

\[
\Rightarrow c^{-\frac{1}{2} t} (1 - \epsilon) < \frac{m^{-1}(cu)}{m^{-1}(u)} < c^{-\frac{1}{2} t} (1 + \epsilon).
\] (30)

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

\[
m^{-1}(cu) \rightarrow c^{-\frac{1}{2} t}
\] (31)
as \( u \rightarrow \infty \). 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 \( \Theta_n^{(op)}(x) \) be as defined in Theorem 9. Then, under the conditions in Theorem 9,

\[
(h_n^{(op)})^{-\beta} \left\| \hat{\Theta}_n^{(op)}(x) - \Theta(x) \right\| = o_p(1) \quad \text{as } n \rightarrow \infty,
\]

and (\( h_n^{(op)} \))\(^{-2\beta} \mathbb{E} \left\| \hat{\Theta}_n^{(op)}(x) - \Theta(x) \right\|^2 \rightarrow 0 \quad \text{as } n \rightarrow \infty.

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

\[
(h_n^{(op)})^{-\beta} \left\| \hat{\Theta}_n^{(op)}(x) - \Theta(x) \right\| = o_p(1) \quad \text{as } n \rightarrow \infty.
\] (32)

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

\[
(d(x, z))^{-\beta} \left\| F(z) - F(x) \right\| \rightarrow 0 \quad \text{as } d(x, z) \rightarrow 0.
\]

Consequently,

\[
(h_n^{(op)})^{-\beta} \left\| B_n^{(op)}(x) \right\| = o_p(1) \quad \text{as } n \rightarrow \infty,
\] (33)

and (\( h_n^{(op)} \))\(^{-2\beta} \mathbb{E} \left\| B_n^{(op)}(x) \right\|^2 \rightarrow 0 \quad \text{as } n \rightarrow \infty.

(34)

From Theorem 3 and (32), we have

\[
(h_n^{(op)})^{-2\beta} \mathbb{E} \left\| V_n^{(op)}(x) \right\|^2
\]

\[
= \left( (h_n^{(op)})^{-2\beta} \left( n\phi(x, h_n^{(op)}) \right)^{-1} \right) \mathbb{E} \left\| V_n^{(op)}(x) \right\|^2 \rightarrow 0
\] (35)
as $n \to \infty$, and from (35) and the Markov inequality, we get
\[(h_n^{(op)})^{-\beta} \|V_n^{(op)}(x)\| = o_p(1) \text{ as } n \to \infty.\] (36)

From condition B(iii) and (32), we have
\[(h_n^{(op)})^{-\beta} \|R_n^{(op)}(x)\| = o_p(1) \text{ as } n \to \infty.\] (37)

Since $\|\tilde{\phi}_n\| = 1$, when $E[\|R_n(x)\|^2] = o(\delta_n^2)$ as $n \to \infty$, from (32), we have
\[(h_n^{(op)})^{-2\beta}E[\|R_n^{(op)}(x)\|^2] \to 0 \text{ as } n \to \infty.\] (38)

Therefore, from (33), (36) and (37), we have
\[(h_n^{(op)})^{-\beta} \|\hat{\Theta}^{(op)}_n(x) - \Theta(x)\|
\leq (h_n^{(op)})^{-\beta} \|B_n^{(op)}(x)\| + (h_n^{(op)})^{-\beta} \|V_n^{(op)}(x)\| + (h_n^{(op)})^{-\beta} \|R_n^{(op)}(x)\|
= o_p(1) \text{ as } n \to \infty.

Further, from (34), (35) and (38), we have
\[(h_n^{(op)})^{-2\beta}E[\|\hat{\Theta}^{(op)}_n(x) - \Theta(x)\|^2
\leq 3(h_n^{(op)})^{-2\beta}E[\|B_n^{(op)}(x)\|^2 + 3(h_n^{(op)})^{-2\beta}E[\|V_n^{(op)}(x)\|^2 + 3(h_n^{(op)})^{-2\beta}E[\|R_n^{(op)}(x)\|^2
\to 0 \text{ as } n \to \infty.\]

\[\Box\]

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

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

for all sufficiently large $n$. Further,
\[(h_n^{(b)})^{-2\beta}E[\|\hat{\Theta}_n^{(b)}(x) - \Theta(x)\|^2 \text{ is bounded away from } 0 \text{ as } n \to \infty.\]

Proof Let $h_n^{(b)}$ satisfy (26). Since $F(\cdot) \in F(x, \beta_1, G)$ for some $\beta_1 > \beta$, we have
\[(d(x, z))^{-\beta} \|F(z) - F(x)\| \to 0 \text{ as } d(x, z) \to 0.\] (39)

Consequently,
\[(h_n^{(b)})^{-\beta} \|B_n^{(b)}(x)\| = o_p(1) \text{ as } n \to \infty.\] (40)
Let $Z$ be the normal random variable described in (15). Given any $\epsilon > 0$, there exists $\delta > 0$ such that

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

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

$$\mathbb{P} \left[ \left\| (h_n^{(b)})^{-\beta} V_n^{(b)}(x) \right\| > 2\delta \right]$$

$$= \mathbb{P} \left[ \left\| (h_n^{(b)})^{-\beta} (n\phi(x, h_n^{(b)}))^{-1/2} \left( n\phi(x, h_n^{(b)}) \right)^{1/2} V_n^{(b)}(x) \right\| > 2\delta \right]$$

$$\geq \mathbb{P} \left[ \left\| n\phi(x, h_n^{(b)})^{1/2} V_n^{(b)}(x) \right\| > 2\delta \sqrt{c_2} \right]$$

$$= \mathbb{P} \left[ \left\| n\phi(x, h_n^{(b)})^{1/2} \tilde{\phi}_n(V_n^{(b)}(x)) \right\| > 2\delta \sqrt{c_2} \right]$$

$$\geq \mathbb{P} \left[ \left\| n\phi(x, h_n^{(b)})^{1/2} |E_n^{(2)}(x)|^{-1/2} E_n^{(1)}(x) \tilde{\phi}_n(V_n^{(b)}(x)) \right\| > 2\delta \sqrt{c_2 l^{-1} L} \right]$$

$$> 1 - \epsilon$$

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

$$(h_n^{(b)})^{-\beta} R_n^{(b)}(x) = o_p(1) \quad \text{as } n \to \infty.$$  \hspace{1cm} (43)

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

$$\mathbb{P} \left[ (h_n^{(b)})^{-\beta} \left\| \tilde{\phi}_n^{(b)}(x) - \Theta(x) \right\| > \delta \right]$$

$$\geq \mathbb{P} \left[ (h_n^{(b)})^{-\beta} \left\| V_n^{(b)}(x) \right\| - (h_n^{(b)})^{-\beta} \left\| B_n^{(b)}(x) \right\| - (h_n^{(b)})^{-\beta} \left\| R_n^{(b)}(x) \right\| > \delta \right]$$

$$= \mathbb{P} \left[ (h_n^{(b)})^{-\beta} \left\| V_n^{(b)}(x) \right\| > \delta + (h_n^{(b)})^{-\beta} \left\| B_n^{(b)}(x) \right\| + (h_n^{(b)})^{-\beta} \left\| R_n^{(b)}(x) \right\| \right]$$

$$> \mathbb{P} \left[ (h_n^{(b)})^{-\beta} \left\| V_n^{(b)}(x) \right\| > 2\delta \right] > 1 - \epsilon$$

for all sufficiently large $n$.

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

$$\mathbb{E} \left[ \left\| \tilde{\phi}_n^{(b)}(x) - \Theta(x) \right\|^2 \right]$$

$$= \mathbb{E} \left[ \left\| B_n^{(b)}(x) + V_n^{(b)}(x) + R_n^{(b)}(x) \right\|^2 \right]$$

$$\geq \mathbb{E} \left[ (\tilde{\phi}_n(B_n^{(b)}(x)))^2 \right] + \mathbb{E} \left[ (\tilde{\phi}_n(V_n^{(b)}(x)))^2 \right] + \mathbb{E} \left[ (\tilde{\phi}_n(R_n^{(b)}(x)))^2 \right]$$
From (39), we have for some constant $c_6$, Further, since $\|\hat{\phi}_n\| = 1$, from Theorem 3 and (26), we have for all sufficiently large $n$ and for some constant $c_6$. Further, since $\|\hat{\phi}_n\| = 1$, from Theorem 3 and (26), we have for some constant $c_7 > 0$ and for all sufficiently large $n$. Since $\|\hat{\phi}_n\| = 1$, when $\mathbb{E}[\|R_n(x)\|^2] = o(\tilde{a}_n^2)$ as $n \to \infty$, from (26) and (32), we have as $n \to \infty$. Therefore, from (44), (45), (46), (47) and (48), we have 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

\[
\hat{D}_n(x, h) = \frac{1}{1 + \epsilon_0} \sigma^2 \zeta_n \frac{\log n}{n \phi(x, h)}.
\]

\[
\hat{C}_n(x, h) = \max_{h' \in \mathbb{H}_n} \left( \left\| \Phi_n(x, h') - \Phi_n(x, \max\{h, h'\}) \right\|^2 - \hat{D}_n(x, h') \right).
\]

Then,

\[
C_n(x, h) \leq \hat{C}_n(x, h) + \max_{h' \in \mathbb{H}_n} \left( \hat{D}_n(x, h') - D_n(x, h') \right) + \hat{D}_n(x, h).
\]

Proof The proof is straightforward from the definitions of $C_n(x, h)$, $D_n(x, h)$, $\hat{C}_n(x, h)$ and $\hat{D}_n(x, h)$. $\square$

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

\[
E \left[ \max_{h' \in \mathbb{H}_n} \left( \hat{D}_n(x, h') - D_n(x, h') \right) \right] < \frac{1}{n^2},
\]

and

\[
E [D_n(x, h)] \leq 3 \hat{D}_n(x, h) + \frac{3 \zeta_0 \sigma^2}{n^2}.
\]

Proof Define the event

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

where $\epsilon_0$ 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 $n \geq n_1$,

\[
\mathbb{P}[(U(x))^c] = \mathbb{P} \left[ \bigcup_{h' \in \mathbb{H}_n} \left\{ \left| \frac{\hat{\phi}(x, h') - \phi(x, h)}{\phi(x, h')} \right| \geq \epsilon_0 \phi(x, h') \right\} \right]
\]

\[
\leq \sum_{h' \in \mathbb{H}_n} \mathbb{P} \left[ \sum_{i=1}^{n} \left| d(x, X_i) \leq h' \right| \phi(x, h') \right] \geq \epsilon_0 n \phi(x, h') \right]
\]

\[
< 2 \sum_{h' \in \mathbb{H}_n} \exp \left[ -4 \log n \right] \leq \frac{2}{n^3}.
\] (49)

Note that

\[
E \left[ \max_{h' \in \mathbb{H}_n} \left( \hat{D}_n(x, h') - D_n(x, h') \right) \right]
\]

\[
= E \left[ \max_{h' \in \mathbb{H}_n} \left( \hat{D}_n(x, h') - D_n(x, h') \right) + I(U(x)) \right]
\]
When \( I(U(x)) = 1 \), we have

\[
(1 - \epsilon_0) \phi(x, h') < \hat{\phi}(x, h') < (1 + \epsilon_0) \phi(x, h') \quad \text{for all } h' \in \mathbb{H}_n
\]

\[
\iff \frac{1}{(1 + \epsilon_0)} \phi(x, h') < \frac{1}{\hat{\phi}(x, h')} < \frac{1}{(1 - \epsilon_0) \phi(x, h')} \quad \text{for all } h' \in \mathbb{H}_n
\]

\[\Rightarrow \max_{h' \in \mathbb{H}_n} \left( \hat{D}_n(x, h') - D_n(x, h') \right) + I(U(x)) = 0\]

\[\Rightarrow \mathbb{E} \left[ \max_{h' \in \mathbb{H}_n} \left( \hat{D}_n(x, h') - D_n(x, h') \right) + I(U(x)) \right] = 0. \quad (52)\]

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 \( n \geq \max \{ n_1, n_2 \} \),

\[
\mathbb{E} \left[ \max_{h' \in \mathbb{H}_n} \left( \hat{D}_n(x, h') - D_n(x, h') \right) + I(U(x))^c \right]
\]

\[\leq \sum_{h' \in \mathbb{H}_n} \mathbb{E} \left[ \left( \hat{D}_n(x, h') - D_n(x, h') \right) + I(U(x))^c \right]
\]

\[\leq \sum_{h' \in \mathbb{H}_n} \hat{D}_n(x, h') \mathbb{P} [(U(x))^c]
\]

\[= \sum_{h' \in \mathbb{H}_n} \frac{1}{(1 + \epsilon_0)} \sigma^2 \zeta_n \frac{\log n}{n \phi(x, h')} \mathbb{P} [(U(x))^c] < 2 \zeta_0 \sigma^2 \frac{1}{\log n} \frac{1}{n^2}. \quad (53)\]

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 \( n \geq \max \{ n_1, n_2, n_3 \} \),

\[
\mathbb{E} \left[ \max_{h' \in \mathbb{H}_n} \left( D_n(x, h') - \hat{D}_n(x, h') \right) \right] < \frac{1}{n^2}. \quad (54)\]

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

\[
\mathbb{E} [D_n(x, h)] = \mathbb{E} [D_n(x, h) I(U(x))] + \mathbb{E} [D_n(x, h) I((U(x))^c)]
\]

\[\leq \frac{(1 + \epsilon_0)}{(1 - \epsilon_0)} \hat{D}_n(x, h) + \sigma^2 \zeta_n n \mathbb{P} [(U(x))^c] \]

\[< 3 \hat{D}_n(x, h) + \frac{3 \zeta_0 \sigma^2}{n^2}. \quad (55)\]

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

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

\[
\mathbb{P} \left[ \sum_{i=1}^{n} \log (G(Y_i) - E[G(Y_i) | X_i]) \frac{K(h^{-1}d(x, X_i))}{n \mathbb{E}[K(h^{-1}d(x, X))]} > y \right] \leq n^{-3}
\]
for all \( h' \in \mathbb{R}_n \). Further, given any \( c_1 > 0, c_2 > 0 \) and any \( 0 < \epsilon < 1 \), we have, for all sufficiently large \( n \),

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

for all \( h' \in \mathbb{R}_n \) 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

\[
\mathbb{E}[\|\xi\|^m] \leq (m! / 2) b_n^2 H^{m-2}
\]

for all integers \( m \geq 2 \). Let

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

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

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

Now, we choose

\[
\xi_i = L_x (G(Y_i) - \mathbb{E}[G(Y_i) | X_i]) \frac{K(h'^{-1}d(x, X_i))}{n\mathbb{E}[K(h'^{-1}d(x, X))]}\]

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

\[
\mathbb{E}[\|\xi_1 + \cdots + \xi_n\|]
\]

\[
= \mathbb{E} \left[ \left\| \sum_{i=1}^{n} L_x (G(Y_i) - \mathbb{E}[G(Y_i) | X_i]) \frac{K(h'^{-1}d(x, X_i))}{n\mathbb{E}[K(h'^{-1}d(x, X))]} \right\| \right]
\]

\[
\leq \left[ \mathbb{E} \left[ \left\| \sum_{i=1}^{n} L_x (G(Y_i) - \mathbb{E}[G(Y_i) | X_i]) \frac{K(h'^{-1}d(x, X_i))}{n\mathbb{E}[K(h'^{-1}d(x, X))]} \right\|^2 \right] \right]^{1/2}
\]

\[
\leq \left[ \epsilon \sum_{i=1}^{n} \mathbb{E} \left[ \left\| L_x (G(Y_i) - \mathbb{E}[G(Y_i) | X_i]) \frac{K(h'^{-1}d(x, X_i))}{n\mathbb{E}[K(h'^{-1}d(x, X))]} \right\|^2 \right] \right]^{1/2}
\]

\[
= \epsilon \sum_{i=1}^{n} \mathbb{E} \left[ \left\| L_x (G(Y_i) - \mathbb{E}[G(Y_i) | X_i]) \frac{K(h'^{-1}d(x, X_i))}{n\mathbb{E}[K(h'^{-1}d(x, X))]} \right\|^2 \right]^{1/2}
\]
\[ \sqrt{c} \frac{\sigma L}{l^{\sqrt{n} \phi(x, h')}} = \beta_n, \]

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

\[ E\|\xi_i\|^m = E\left\| l_{\infty}(G(Y_i) - E[G(Y_i) | X_i]) K(h'^{-1}d(x, X_i)) \right\|^m \]

\[ \leq \frac{m!}{2} \left( \frac{\sigma L}{ln\phi(x, h')} \right)^{m-2} \frac{\sigma^2 L^2}{l^2 n^2 \phi(x, h')} , \]

and we can take

\[ U^2_n = \frac{\sigma^2 L^2}{l^2 n \phi(x, h')} \quad \text{and} \quad H = \frac{\sigma L}{ln\phi(x, h')} . \]

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

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

for all sufficiently large \( n \) and for all \( h' \in H_n \). Also,

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

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

\[ P \left[ \left( \sum_{i=1}^{n} l_{\infty}(G(Y_i) - E[G(Y_i) | X_i]) K(h'^{-1}d(x, X_i)) \right)^m \right] > y \]

\[ \leq \exp \left[ -\frac{(yl\log n - \sqrt{c} \sigma L)^2}{8\sigma^2 L^2 + 4yl\sigma L} \right] < \exp [-3 \log n] = n^{-3} . \]

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

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

for all sufficiently large \( n \) and all \( h' \in H_n \). Also, given any \( 0 < \epsilon < 1 \), we have, for all sufficiently large \( n \),

\[ \epsilon \frac{c_2 \sqrt{c_1 D_n(x, h')} + t}{U_n} \geq \epsilon \sqrt{\log n} \left( \frac{lc_2 \sqrt{c_1 d^2 \zeta_n}}{\sigma L} \right) > \sqrt{c} \]

\[ \implies \left( \frac{c_2 \sqrt{c_1 D_n(x, h')} + t}{U_n} - \sqrt{c} \right)^2 > (1 - \epsilon)^2 \frac{c_1 D_n(x, h')}{{U_n}^2} + t \]
for all \( h' \in \mathbb{H}_n \) and all \( t \geq 0 \). Now,
\[
\left( \frac{c_2 \sqrt{c_1 D_n(x, h') + t}}{U_n} - \frac{\beta_n}{U_n} \right) \frac{H}{2U_n} \leq c_2 \sqrt{c_1 D_n(x, h') + t} \frac{H}{2U_n} < c_2 \sqrt{c_1 D_n(x, h') + t} \frac{l}{\sigma L}
\]
for all \( h' \in \mathbb{H}_n \) and all \( t \geq 0 \). So, from (56), we get that for all sufficiently large \( n \),
\[
\mathbb{P} \left[ \sum_{i=1}^{n} \mathbb{E} \left( (G(Y_i) - \mathbb{E}[G(Y_i) | X_i]) K(h^{-1}d(x, X_i)) \right) > c_2 \sqrt{c_1 D_n(x, h') + t} \right]
\]
\[
\leq \exp \left[ \frac{(1 - \epsilon)^2 c_2^2 n \phi(x, h') (c_1 D_n(x, h') + t)}{8 \sigma L^2 \left( \sigma + c_2 \sqrt{c_1 D_n(x, h') + t} \right)} \right]
\]
\[
\leq \exp \left[ \frac{(1 - \epsilon)^2 c_2^2 n \phi(x, h') c_2^2 (c_1 D_n(x, h') + t)}{16 \sigma^2 L^2} \right]
\]
\[
+ \exp \left[ \frac{(1 - \epsilon)^2 c_2^2 n \phi(x, h') c_2 \sqrt{c_1 D_n(x, h') + t}}{16 \sigma L^2} \right]
\]
for all \( h' \in \mathbb{H}_n \) and for all \( t \geq 0 \).

\[\square\]

\textbf{Lemma 11} Let \( \hat{C}_n(x, h) \) 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 \( n \geq N_2 \),
\[
\hat{C}_n(x, h) \leq M_1 h^{2\beta} + 24 \max_{h' \in \mathbb{H}_n, h' \leq h} \left( \left\| V_n(x, h') \right\|^2 - \frac{\tilde{D}_n(x, h')}{24} \right) + 12 \max_{h' \in \mathbb{H}_n, h' \leq h} \left( \left\| R_n(x, h') \right\|^2 - \left( M_h^{2\beta} + \left\| V_n(x, h') \right\|^2 \right) \right)
\]
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
\[
\mathbb{P} \left[ \max_{h' \in \mathbb{H}_n, h' \leq h} \left( \left\| R_n(x, h') \right\|^2 - \left( M_h^{2\beta} + \left\| V_n(x, h') \right\|^2 \right) \right) > \frac{1}{n^2} \right] \leq 2n^{-2},
\]
and
\[
\mathbb{E} \left[ \max_{h' \in \mathbb{H}_n, h' \leq h} \left( \left\| V_n(x, h') \right\|^2 - \frac{\tilde{D}_n(x, h')}{24} \right) \right] < \frac{1}{n}.
\]
\textbf{Proof} Note that
\[
\hat{C}_n(x, h) = \max_{h' \in \mathbb{H}_n, h' \leq h} \left( \left\| \tilde{C}_n(x, h') - \tilde{C}_n(x, h) \right\|^2 - \tilde{D}_n(x, h') \right) + \max_{h' \in \mathbb{H}_n, h' \leq h} \left( 2 \left\| \tilde{C}_n(x, h') - \tilde{C}_n(x, h) \right\|^2 + 2 \left\| \tilde{C}_n(x, h) - \tilde{C}_n(x, h') \right\|^2 - \tilde{D}_n(x, h') \right)
\]
\[ \leq 2 \max_{h' \in \mathbb{H}_n, h' \leq h} \left( \left\| \hat{\Theta}_n(x, h') - \Theta(x) \right\|^2 - \frac{\tilde{D}_n(x, h')}{4} \right) + \\
+ 2 \max_{h' \in \mathbb{H}_n, h' \leq h} \left( \left\| \hat{\Theta}_n(x, h) - \Theta(x) \right\|^2 - \frac{\tilde{D}_n(x, h')}{4} \right) + \\
\leq 4 \max_{h' \in \mathbb{H}_n, h' \leq h} \left( \left\| \hat{\Theta}_n(x, h') - \Theta(x) \right\|^2 - \frac{\tilde{D}_n(x, h')}{4} \right) \]  

(57)

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

\[
\max_{h' \in \mathbb{H}_n, h' \leq h} \left( \left\| \hat{\Theta}_n(x, h') - \Theta(x) \right\|^2 - \frac{\tilde{D}_n(x, h')}{4} \right) + \\
\leq 3 \max_{h' \in \mathbb{H}_n, h' \leq h} \left( \| B_n(x, h') \|^2 + \| V_n(x, h') \|^2 + \| R_n(x, h') \|^2 - \frac{\tilde{D}_n(x, h')}{12} \right) + \\
\leq 3 \max_{h' \in \mathbb{H}_n, h' \leq h} \left( \| B_n(x, h') \|^2 + Mh^{2\beta} \right) \\
+ 6 \max_{h' \in \mathbb{H}_n, h' \leq h} \left( \| V_n(x, h') \|^2 - \frac{\tilde{D}_n(x, h')}{24} \right) + \\
+ 3 \max_{h' \in \mathbb{H}_n, h' \leq h} \left( \| R_n(x, h') \|^2 - \left( Mh^{2\beta} + \| V_n(x, h') \|^2 \right) \right) \]  

(58)

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

\[
\max_{h' \in \mathbb{H}_n, h' \leq h} \left( \| B_n(x, h') \|^2 + Mh^{2\beta} \right) \leq M_1 h^{2\beta} \]  

(59)

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

Next, define the event

\[
S(x, h') = \left\{ \frac{1}{n} \sum_{i=1}^{n} K(h^{-1}d(x, X_i)) > (1 - \epsilon_0) \right\},
\]

where $\epsilon_0$ is the number described in Lemma 8. From assumption D(ii) and the fact that $\max\{h' \mid h' \in \mathbb{H}_n\} \rightarrow 0$ as $n \rightarrow \infty$, we have for all sufficiently large $n$,

\[
P \left[ \max_{h' \in \mathbb{H}_n, h' \leq h} \left( \| R_n(x, h') \|^2 - \left( Mh^{2\beta} + \| V_n(x, h') \|^2 \right) \right) > \frac{1}{n^2} \right] \\
\leq \sum_{h' \in \mathbb{H}_n} P \left[ \left( \| R_n(x, h') \|^2 - \left( Mh^{2\beta} + \| V_n(x, h') \|^2 \right) \right) > 0 \right] \\
\leq \sum_{h' \in \mathbb{H}_n} P \left[ \| R_n(x, h') \|^2 > Mh^{2\beta} + \| V_n(x, h') \|^2 \right] \]
Since $B$ for all sufficiently large $n$

and all $h' \in H_n$

also, from Lemma 10, we get

\[
\sum_{h' \in H_n} \mathbb{P}[\left\{S(x, h')\right\}] = \sum_{h' \in H_n} \mathbb{P}\left[\left\{1 - \frac{K(h^{-1}d(x, X_i))}{nE[K(h^{-1}d(x, X)]]} \right\} \geq \epsilon_0\right]
\]

\[
\leq \sum_{h' \in H_n} \exp\{-3\log n\} \leq n^{-2}.
\] (61)

Also, from Lemma 10, we get

\[
\sum_{h' \in H_n} \mathbb{P}[\left\{V_n(x, h')\right\}] \geq \frac{1}{2} \sum_{i=1}^{n} \left\{1 - \frac{K(h^{-1}d(x, X_i))}{nE[K(h^{-1}d(x, X)]]} \right\} \geq (1 - \epsilon_0)\epsilon_2
\]

\[
\leq n^{-2}
\] (62)

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

\[
\mathbb{P}\left[\max_{h' \in H_n, h' \leq n} \left(\left|R_n(x, h')\right|^2 - \left(Mh'^2 + \left\|V_n(x, h')\right\|^2\right)\right) > \frac{1}{n^2}\right] \leq 2n^{-2}
\] (63)

for all sufficiently large $n$ and all $h \in H_n$. Next,

\[
\mathbb{E}\left[\max_{h' \in H_n, h' \leq n} \left(\left\|V_n(x, h')\right\|^2 - \frac{\hat{D}_n(x, h')}{24}\right)\right]
\]

\[
\leq \sum_{h' \in H_n} \mathbb{E}\left[\left(\left\|V_n(x, h')\right\|^2 - \frac{\hat{D}_n(x, h')}{24}\right)\right]
\]

\[
\leq \sum_{h' \in H_n} \mathbb{E}\left[\left(\left\|V_n(x, h')\right\|^2 \mathbb{I}(S(x, h')) - \frac{\hat{D}_n(x, h')}{24}\right)\right]
\]

\[
+ \sum_{h' \in H_n} \mathbb{E}\left[\left\|V_n(x, h')\right\|^2 \mathbb{I}((S(x, h'))^c)\right].
\] (64)

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

\[
\sum_{h' \in H_n} \mathbb{E}\left[\left\|V_n(x, h')\right\|^2 \mathbb{I}((S(x, h'))^c)\right]
\]
\[
\begin{align*}
\sum_{h' \in \mathbb{H}_n} \mathbb{E} \left[ \frac{\sum_{i=1}^{n} L_\infty (G(Y_i) - \mathbb{E}(G(Y_i) | X_i)) K(h'^{-1}d(x, X_i))}{(\sum_{i=1}^{n} K(h'^{-1}d(x, X_i)))^2} \right] I((S(x, h'))^c) \\
\leq \sum_{h' \in \mathbb{H}_n} c \mathbb{E} \left[ \frac{\sum_{i=1}^{n} K^2(h'^{-1}d(x, X_i))}{(\sum_{i=1}^{n} K(h'^{-1}d(x, X_i)))^2} I((S(x, h'))^c) \right] \\
\leq c \sigma^2 \sum_{h' \in \mathbb{H}_n} \mathbb{P}[(S(x, h'))^c] \leq c \sigma^2 n^{-2} 
\end{align*}
\]

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 \(n\),

\[
\begin{align*}
\sum_{h' \in \mathbb{H}_n} \mathbb{E} \left[ \left( \left( \frac{V_n(x, h')}{24} \right) + \hat{D}_n(x, h') \right) I((S(x, h'))^c) \right] - \frac{\hat{D}_n(x, h')}{24} \\
= \sum_{h' \in \mathbb{H}_n} \int_{0}^{\infty} \mathbb{P} \left[ \left( \left( \frac{V_n(x, h')}{24} \right) + \hat{D}_n(x, h') \right) I((S(x, h'))^c) \right] dt \\
= \sum_{h' \in \mathbb{H}_n} \int_{0}^{\infty} \mathbb{P} \left[ \left( \left( \frac{V_n(x, h')}{24} \right) + \hat{D}_n(x, h') \right) \right] dt \\
\leq \sum_{h' \in \mathbb{H}_n} \int_{0}^{\infty} \exp \left[ -\left( \frac{(1 - \epsilon_0)^3 L^2 n \phi(x, h')}{16 \sigma^2 L^2} \right) \right] dt \\
+ \sum_{h' \in \mathbb{H}_n} \int_{0}^{\infty} \exp \left[ -\left( \frac{(1 - \epsilon_0)^3 L^2 n \phi(x, h')}{16 \sigma^2 L^2} \right) \right] dt. 
\end{align*}
\]

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

\[
\begin{align*}
\sum_{h' \in \mathbb{H}_n} \int_{0}^{\infty} \exp \left[ -\left( \frac{(1 - \epsilon_0)^3 L^2 n \phi(x, h')}{16 \sigma^2 L^2} \right) \right] dt \\
= 2 \sum_{h' \in \mathbb{H}_n} \int_{\sqrt{\frac{1}{24} D_n(x, h')}}^{\infty} \exp \left[ -\left( \frac{(1 - \epsilon_0)^3 L^2 n \phi(x, h')}{16 \sigma^2 L^2} \right) \right] ds \\
< \frac{1}{n \log n} 
\end{align*}
\]

for all sufficiently large \(n\). Next, we take

\[
\epsilon_0 \geq 768 \left( \frac{1 + \epsilon_0}{1} \right)^2 \frac{L^2}{(1 - \epsilon_0)^4 \sqrt{2}}.
\]
Since $\zeta_n \to \zeta_0$ as $n \to \infty$, we have

$$\zeta_n > 768 \frac{(1 + \epsilon_0) L^2}{(1 - \epsilon_0)^2 L}$$  \hspace{1cm} (69)$$

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

$$\sum_{h' \in H_n} \int_0^\infty \exp \left[ - \frac{(1 - \epsilon_0)^4 I^2 n \phi(x, h')}{16 \sigma^2 L^2} \left( \frac{1}{24} D_n(x, h') + t \right) \right] dt$$

$$= \sum_{h' \in H_n} \int_0^\infty \exp \left[ - \frac{(1 - \epsilon_0)^4 I^2 n \phi(x, h')}{16 \sigma^2 L^2} \right] ds$$

$$= \sum_{h' \in H_n} \frac{16 \sigma^2 L^2}{(1 - \epsilon_0)^4 I^2 n \phi(x, h')} \exp \left[ - \frac{(1 - \epsilon_0)^4 I^2 n \phi(x, h')} {16 \sigma^2 L^2} \frac{1}{24} D_n(x, h') \right]$$

$$= \sum_{h' \in H_n} \frac{16 \sigma^2 L^2}{(1 - \epsilon_0)^4 I^2 n \phi(x, h')} \exp \left[ - \frac{1}{768} \frac{(1 - \epsilon_0)^4 I^2}{(1 + \epsilon_0) L^2 \zeta_n (2 \log n)} \right]$$

$$\leq \frac{16 \sigma^2 L^2}{(1 - \epsilon)^4 I^2 (\log n)^2 n^{-1}} < \frac{1}{n \log n}$$  \hspace{1cm} (70)$$

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

$$\sum_{h' \in H_n} E \left[ \left( \| V_n(x, h') \|^2 I(S(x, h')) - \frac{D_n(x, h')} {24} \right) \right] < \frac{2}{n \log n}$$  \hspace{1cm} (71)$$

for all sufficiently large $n$. Therefore, from (64), (65) and (71), we get that for all sufficiently large $n$ and all $h \in H_n$,

$$E \left[ \max_{h' \in H_n, h' \leq h} \left( \| V_n(x, h') \|^2 - \frac{D_n(x, h')} {24} \right) \right] < \frac{1}{n}$$  \hspace{1cm} (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 $h \in H_n$. Hence, the proof is complete from (57), (58), (59), (63) and (72). \qed

From (68), we see that $\zeta_0$ depends on the choice of $\epsilon_0$, and it increases with an increase in the value of $\epsilon_0$. Taking $\epsilon_0 = 0.1$ we see that

$$\zeta_0 = 1500 \frac{L^2}{l^2}$$  \hspace{1cm} (73)$$

satisfies (68). Taking smaller values of $\epsilon_0$, we can further decrease the value of $\zeta_0$, but it cannot be less than 768 in view of (68).

References