Moderate deviations for deconvolution kernel density estimators with ordinary smooth measurement errors

Weixing Song
Department of Statistics, Kansas State University, Manhattan, 108F, Dickens Hall KS, 66502, United States

Abstract
In this paper, we establish the pointwise and uniform moderate deviations limit results for the deconvolution kernel density estimator in the errors-in-variables model, when the measurement error possesses an ordinary smooth distribution. The results are similar to the moderate deviations theorems for the classical kernel density estimators, but a factor related to the ordinary smooth order is needed to account for the measurement errors.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction
In many practical situations, the random variable $X$ cannot be observed directly, instead, a surrogate $Y$ can be obtained which is related to $X$ in the manner $Y = X + \varepsilon$, where $\varepsilon$ is the measurement error with a known density function, $X$ and $\varepsilon$ are independent. This is the so-called errors-in-variables or the convolution model. The errors-in-variables modeling is widely used in microfluorimetry, nutrition study, electrophoresis, biostatistics, and some other areas. The investigations on its theories and applications have received much attention in the past decades. A typical example of the errors-in-variables model is given in Fan (1991) about the AIDS study, where $Y$ may be the time from some starting point to the time that symptoms appear, $\varepsilon$ may be the time from the start point to the time that infection occurs, and $X$ is the time from the occurrence of infection to the time of symptoms. A comprehensive introduction to the errors-in-variables model, as well as many other real data examples, can be found in Fuller (1987), Carroll et al. (1995) and the references therein.

The problem of interest in errors-in-variables model is to estimate the density function of $X$. Because $X$ is not observable and its density is unknown, the kernel type estimator cannot be applied here. The most commonly used estimator for the density $f_X$ of $X$ is the deconvolution kernel density estimator, which is based on the additive structure of $X$ and $\varepsilon$. It is well known that this additivity structure, and the independence of $X$ and $\varepsilon$ imply that $\phi_Y(t) = \phi_X(t)\phi_{\varepsilon}(t)$ for $t \in \mathbb{R}$, where $\phi(t)$ denotes the characteristic function of a random variable which is specified by the subscript. Since the density function of $\varepsilon$, denoted by $f_{\varepsilon}$, is known hence $\phi_{\varepsilon}$ is known, and $Y$ is observable hence $\phi_Y$ can be estimated, so the inverse Fourier transformation on $\hat{\phi}_Y(t)/\phi_{\varepsilon}(t)$ can provide us an estimator for $f_X$, where $\hat{\phi}_Y(t)$ is an estimate of the characteristic function of $Y$. Using a kernel function $K$ with bandwidth $h_n$, Stefanski and Carroll (1991) consider the following so-called deconvolution kernel density estimator of $X$:

$$\hat{f}_n(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-itx) \frac{\phi_X(th_n)\hat{\phi}_n(t)}{\phi_{\varepsilon}(t)} \, dt, \quad i = \sqrt{-1},$$

(1.1)

where $\phi(t)$ is the characteristic function of some kernel function $K$, $\hat{\phi}_n(t)$ is the empirical characteristic function of the sample $Y_1, \ldots, Y_n$ defined by $\hat{\phi}_n(t) = n^{-1} \sum_{i=1}^{n} \exp(-itY_i)$, and $h_n$ has the same meaning as the bandwidth in the classical
kernel density estimators. Let
\[ g_n(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-ity) \frac{\phi_k(t)}{\phi_k(t/h_n)} \, dt, \]
then (1.1) has the following form which resembles the classical kernel density estimator,
\[ \hat{f}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h_n} g_n \left( \frac{x - Y_i}{h_n} \right). \]
g_n is significantly different from the classical kernel functions in that \( g_n \) might be a complex function, and it depends on the bandwidth \( h_n \), hence on the sample size.

The large sample behavior of \( \hat{f}_n(x) \) strongly depends on the smoothness of the distribution of the measurement error \( \varepsilon \).

Using the terms in Fan and Truong (1993), a distribution is called ordinary smooth if the tails of its characteristic function decay to 0 at an algebraic rate; it is called supersmooth if its characteristic function has tails approaching 0 exponentially fast. It is known that the local and global rates of convergence of the sequences of deconvolution kernel density estimators are slower than that of the classical kernel density estimators. Moreover, the convergence rate is much slower in the supersmooth cases than in the ordinary smooth cases. Related works on the consistency and the asymptotic normality of the deconvolution kernel density estimator \( \hat{f}_n(x) \) can be found in Carroll and Hall (1988), Devroye (1989), Fan (1991), and Fan and Liu (1997) and the references therein. But the literature seems scant in the discussion of the moderate and large deviation results for the deconvolution kernel density estimators, while there are many such works on the classical kernel estimators, see Louani (1998) and Gao (2003) for more details. The current paper will try to fill this void partly. In particular, we will focus on the moderate deviation limit theorems when the measurement error has an ordinary smooth distribution. Further research will be conducted in the future for the large deviation results for the ordinary smooth measurement models, and the moderate and large deviation results for the supersmooth cases. Another interesting topic will be the moderate and large deviation results for the nonparametric prediction in measurement error regression models based on the work of Carroll et al. (2009). Since a parametric rate \( n^{-1/2} \) can be achieved for the nonparametric prediction, so the deviation results should be similar to the nonparametric prediction in the classical regression models.

The paper is organized as follows. Section 2 will present the necessary assumptions for the theory, the pointwise and uniform moderate deviation results will be also introduced; the proofs of the main results will be provided in Section 3.

2. Assumptions and main results

The ordinary smooth condition on \( \varepsilon \) often takes the following form
\[(C1) \ t^\beta \phi_\varepsilon(t) \to c, \ t^{\beta+1} \phi_\varepsilon'(t) \to -\beta c, \text{ as } t \to +\infty, \text{ with some constant } c \neq 0 \text{ and } \beta \geq 0. \text{ Moreover, } \phi_\varepsilon(t) \neq 0 \text{ for all } t, \]
where \( \beta \) is called the smooth order.

For example, the Gamma density function \( \alpha^p x^{p-1} \exp(-\alpha x) I(x > 0)/\Gamma(p) \) is ordinary smooth with \( \beta = p \), and the double exponential density function \( \exp(-|x|)/2 \) is ordinary smooth with \( \beta = 2 \).

The following condition is needed to guarantee the boundedness of the density function \( Y \), which is similar to condition (A1) in Gao (2003).

\[(C2) \ \text{The density function of } \varepsilon \text{ is continuous and } f_\varepsilon(t) \to 0 \text{ as } t \to \infty. \]

Obviously, the above mentioned Gamma and double exponential distributions satisfy (C2).

As for the kernel function, we shall assume
\[(C3) \ \phi_k(t) \text{ is a symmetric function, having second bounded integrable derivatives.} \]
\[(C4) \ \int_{-\infty}^{\infty} |\phi_k(t)| + |\phi_k'(t)| |t|^{\beta} \, dt < \infty, \int_{-\infty}^{\infty} |t|^{2\beta} |\phi_k(t)|^2 \, dt < \infty. \]

The following condition will be imposed on the density function of \( X \).

\[(C5) \ \text{The second derivative of the unknown } f_k \text{ exists and is continuous.} \]

For the bandwidth \( h_n \), we shall assume
\[(C6) \ nh_n \to \infty, \quad \frac{b_n}{nh_n} \to 0, \quad \frac{nh_n \log h_n}{b_n^2} \to 0. \]

Similar to Giné and Guillou (2001) and Gao (2003), the following condition is needed when using Giné and Guillou (2001)'s exponential inequality to prove the uniform moderate deviation result,
\[(C7) \ g_n(y) \text{ is a bounded, square integrable function in the linear span (the set of finite linear combinations) of functions } q \geq 0 \text{ satisfying the following property: the subgraph of } q, \{(s, u): q(s) \geq u\}, \text{ can be represented as a finite number of Boolean operations among sets of the form } \{(s, u): w(s, u) \geq \psi(u)\}, \text{ where } w \text{ is a polynomial on } \mathbb{R}^2 \text{ and } \psi \text{ is an arbitrary real function.} \]
As an example, if $\varepsilon$ has a double exponential distribution $f_\varepsilon(x) = (\sqrt{2}\sigma_0^2) \exp(-\sqrt{2}|x|/\sigma_0)$ with the characteristic function $\phi_{\varepsilon}(t) = (1 + \sigma_0^2 t^2/2)^{-1}$, and $K$ is further chosen to be the Gaussian kernel $K(x) = (\sqrt{2\pi})^{-1} \exp(-x^2/2)$, then one can show that

$$g_n(y) = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{y^2}{2} \right) \left[ 1 - \frac{\sigma_0^2}{2h_n^2} (y^2 - 1) \right].$$

Easy to see that $g_n(y)$ is a bounded real function of bounded variation, hence $g_n(y)$ satisfies (C7).

Under above conditions, we can show the following pointwise moderate deviation limit result for the estimator $\hat{f}_n(x)$.

**Theorem 2.1.** Suppose (C1)–(C6) hold. Then for any $x \in \mathbb{R}$, any closed set $F \subset \mathbb{R}$, any open set $G \subset \mathbb{R}$, we have

$$\lim_{n \to \infty} \sup_{\lambda > 0} \frac{nh_n}{b_n^2} \log P \left( \frac{nh_n^{\beta+1}}{b_n} \left[ \hat{f}_n(x) - \hat{E}_n f \right] \right) \leq - \inf_{x \in F} I_x(\lambda),$$

$$\lim_{n \to \infty} \inf_{\lambda \in G} \frac{nh_n}{b_n^2} \log P \left( \frac{nh_n^{\beta+1}}{b_n} \left[ \hat{f}_n(x) - \hat{E}_n f \right] \right) \geq - \inf_{\lambda \in G} I_x(\lambda),$$

where

$$I_x(t) = \frac{\lambda^2}{2G(x)} - G(x) = \frac{f_\varepsilon(x)}{2\pi |c|^2} \int_{-\infty}^{\infty} |t|^{2\beta} |\phi_X(t)|^2 dt.$$

and $f_\varepsilon$ is the density function of $Y$.

Furthermore, we also have the following uniform moderate deviation limit theorem for the estimator $\hat{f}_n(x)$.

**Theorem 2.2.** If assumptions (C1)–(C7) hold, then for any $\lambda > 0$,

$$\lim_{n \to \infty} \frac{nh_n}{b_n^2} \log P \left( \frac{nh_n^{\beta+1}}{b_n} \left\| \hat{f}_n - \hat{E}_n f \right\|_\infty > \lambda \right) = -I(\lambda),$$

where $\| \cdot \|_\infty$ denotes the supreme norm of a function and

$$I(\lambda) = \lambda^2 \frac{2}{\|f_\varepsilon\|_\infty L}, \quad L = \frac{1}{2\pi |c|^2} \int_{-\infty}^{\infty} |t|^{2\beta} |\phi_X(t)|^2 dt.$$

If we further assume

(C8) $\phi_X(t) = 1 + O(|t|^2)$ as $t \to 0$,

(C9) $nh_n^{\beta+1} / b_n \to 0$ as $n \to \infty$,

(C10) $f_\varepsilon(x)$ and its second derivative is bounded,

then the following corollary holds.

**Corollary 2.1.** Suppose (C1)–(C10) hold, then for any $x \in \mathbb{R}$, any closed set $F \subset \mathbb{R}$, any open set $G \subset \mathbb{R}$,

$$\lim_{n \to \infty} \sup_{\lambda > 0} \frac{nh_n}{b_n^2} \log P \left( \frac{nh_n^{\beta+1}}{b_n} \left[ \hat{f}_n(x) - f_\varepsilon(x) \right] \right) \leq - \inf_{x \in F} I_x(\lambda),$$

$$\lim_{n \to \infty} \inf_{\lambda \in G} \frac{nh_n}{b_n^2} \log P \left( \frac{nh_n^{\beta+1}}{b_n} \left[ \hat{f}_n(x) - f_\varepsilon(x) \right] \right) \geq - \inf_{\lambda \in G} I_x(\lambda),$$

and for any $\lambda > 0$,

$$\lim_{n \to \infty} \frac{nh_n}{b_n^2} \log P \left( \frac{nh_n^{\beta+1}}{b_n} \left\| \hat{f}_n - f_\varepsilon(x) \right\|_\infty > \lambda \right) = -I(\lambda),$$

where $I_x(\lambda)$ and $I(\lambda)$ are the same as in Theorems 2.1 and 2.2, respectively.

Condition (C8) is also used in Fan (1991). A typical example of $K$ satisfying (C8) is the standard normal density function. Condition (C9) guarantees the bias of $\hat{f}_n(x)$ converges to 0 sufficiently fast, while the condition (C10) ensures this convergence is uniform in $x \in \mathbb{R}$.
3. The proofs of main results

Proof of Theorem 2.1. From Fan (1991), conditions (C1), (C3)–(C5) imply
\[ |h_n^\delta g_n(y)| \leq B(y) \overset{\text{def}}{=} \min(C_1, C_2/y), \]
where \( C_1, C_2 \) are two constants, and
\[
h_n^\delta g_n(y) \rightarrow C(y) \overset{\text{def}}{=} \frac{1}{2\pi c} \int_{-\infty}^{\infty} \exp(-|ty|) t^\delta \phi_k(t)\,dt
\]
as \( n \to \infty \). For convenience, let \( Z_n = h_n^\delta g_n((x - Y_i)/h_n) \). Note that assumption (C1) implies that the density function \( f_Y(\cdot) \) of \( Y \) is bounded. Therefore, for any \( k \geq 3 \), we have
\[
E|Z_n|^k = \int_{-\infty}^{\infty} h_n^\delta g_n(x-y) \left| \int \frac{f_Y(y)}{h_n} \right|^k \, dy = \int_{-\infty}^{\infty} h_n^\delta g_n(x) \left| \int f_Y(y) \right|^k \, dy \leq h_n \| f_Y \|_{\infty} \cdot \int_{-\infty}^{\infty} |h_n^\delta g_n(x)|^k \, dy \leq h_n \| f_Y \|_{\infty} \int_{-\infty}^{\infty} |B(y)|^k \, dy.
\]
To proceed, we need an upper bound for the above integration. Let \( d = \max\{C_1, C_2\} \). Then
\[
\int_{-\infty}^{\infty} |B(y)|^k \, dy = 2d^k \int_{0}^{\infty} [1 + 1/y^k] \, dy = 2d^k \int_{0}^{1} \, dy + 2d^k \int_{1}^{\infty} \frac{1}{y^k} \, dy = 2d^k \left[ 1 + \frac{1}{k - 1} \right] \leq 3d^k.
\]
Therefore, \( E|Z_n| - EZ_n|^k \leq 3 \cdot 2^k d^k h \| f_Y \|_{\infty} \) and
\[
\sum_{k=3}^{\infty} \frac{1}{k!} \left| \frac{tb_n}{nh_n} (Z_n - EZ_n) \right|^k \leq 3h_n \| f_Y \|_{\infty} \sum_{k=3}^{\infty} \frac{1}{k!} \frac{2^k d^k t^k b_n^k}{n^k h_n^k}
\]
and
\[
\sum_{k=3}^{\infty} \frac{1}{k!} \left| \frac{tb_n}{nh_n} (Z_n - EZ_n) \right|^k = O \left( \left( \frac{2tb_n}{nh_n} \right)^3 h_n \right).
\]
By assumption (C6), we have
\[
\Phi_n(x) = E \left( \exp\left[ \frac{tb_n}{nh_n} (f_Y(x) - E\hat{f}_n(x)) \right] \right)
\]
\[
= E \left( \exp \left[ \frac{tb_n}{nh_n} \sum_{i=1}^{n} (Z_i - EZ_i) \right] \right) = \left[ E \left( \exp \left[ \frac{tb_n}{nh_n} (Z_n - EZ_n) \right] \right) \right]^n
\]
\[
= \left[ 1 + \frac{tb_n}{nh_n} E(Z_n - EZ_n) + \frac{t^2b_n^2}{2n^2 h_n^2} E(Z_n - EZ_n)^2 + O \left( \left( \frac{2tb_n}{nh_n} \right)^3 h_n \right) \right]^n.
\]
By (3.1) and (3.2), we can obtain
\[
h_n^{-1} E(Z_n - EZ_n)^2 \overset{\text{def}}{=} \frac{f_Y(x)}{2\pi c|t|^2} \int_{-\infty}^{\infty} |t|^{2\delta} |\phi_k(t)|^2 \, dt \overset{\text{def}}{=} G(x).
\]
Therefore, we have
\[
\Phi_n(x) = \left[ 1 + \frac{t^2b_n^2}{2n^2 h_n^2} G(x)(1 + o(1)) + O \left( \frac{2tb_n}{nh_n} \right)^3 h_n \right]^n.
\]
Finally, by letting \( n \to \infty \),
\[
\Phi_n(x) = \lim_{n \to \infty} \frac{nh_n}{b_n^2} \log \Phi_n(x) = \frac{t^2}{2} G(x),
\]
and the dual of \( \Phi_n(x) \) is given by
\[
I_\lambda(t) = \sup_{t \in \mathbb{R}} \{ t\lambda - \Phi_n(t) \} = \frac{\lambda^2}{2G(x)}.
\]
Note that \( \Phi_n(x) \) is differentiable with respect to \( t \in \mathbb{R} \), from Gärnter–Ellis Theorem, we have the desired result. \( \square \)
To show the uniform moderate deviation result, we need an exponential inequality for the empirical process from Giné and Guillou (2001). For the sake of completeness, we state the exponential inequality below.

**Lemma 3.1.** Let $\mathcal{F}$ be a measurable uniformly bounded VC class of functions, and let $\sigma^2$ and $U$ be any numbers such that $\sigma^2 \geq \sup_{f \in \mathcal{F}} \Var(f)$, $U \geq \sup_{f \in \mathcal{F}} \|f\|_\infty$, and $0 < \sigma \leq U/2$. Then, there exist constants $C$ and $L$ depending only on the characteristic of the class $\mathcal{F}$, such that the inequality

$$P \left( \left\| \sum_{i=1}^n (f(\xi_i) - Ef(\xi_i)) \right\|_\infty > t \right) \leq L \exp \left[ -t \frac{U}{L(\sqrt{n} \sigma + U \sqrt{\log(U/\sigma)}^2)} \right]$$

holds for all

$$t \geq C \left( U \log \frac{U}{\sigma} + \sqrt{\frac{n}{\sigma}} \sqrt{\log \frac{U}{\sigma}} \right).$$

More about the VC (Vapnik–Chervonenkis) class of functions, please see Pollard (1984). We also need the following lemmas which are similar to the lemmas 2.4 and 2.5 in Gao (2003).

**Lemma 3.2.** Let assumptions (C1)–(C7) hold. For any $0 < \delta < 1$, let $l_{n,k}$, $k = 1, 2, \ldots, l_n$, be $l_n$ intervals with length $\delta h_n$, such that $\{l_{n,k}, k = 1, \ldots, l_n\}$ is a covering of $[-h_n^{-1}, h_n]$ and $l_n \leq 2h_n^2 + 1$. Let $x_{n,k} \in l_{n,k}$, $1 \leq k \leq l_n$, $n \geq 1$. Then for any $\varepsilon > 0$,

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \frac{n h_n}{b_n^2} \log P \left( b_n \sup_{1 \leq k \leq l_n, x \in l_{n,k}} \left| \hat{f}_{n,k}(x) - Ef_{n,k}(x) \right| \geq \varepsilon \right) = -\infty,$$

where $\hat{f}_{n,k}(x) = \hat{f}_n(x) - \hat{f}_n(x_{n,k})$.

**Proof.** By (C7), the class of functions $\mathcal{F} = \{h_n^2 g_n((x - \cdot)/h_n), x \in \mathbb{R}, h_n \neq 0\}$ is a bounded measurable VC class of functions, so the following classes of functions

$$\mathcal{F}_{n,k} = \left\{h_n^2 g_n \left( \frac{x - \cdot}{h_n} \right) - h_n^2 g_n \left( \frac{x_{n,k} - \cdot}{h_n} \right) : x \in l_{n,k} \right\}, \quad k = 1, \ldots, l_n; \quad n \geq 1$$

are measurable VC classes of functions. Moreover, there is a common VC characteristic for all these classes that does not depend on $k$ and $n$. Note that

$$\int \left[ h_n^2 g_n \left( \frac{x-y}{h_n} \right) - h_n^2 g_n \left( \frac{x_{n,k}-y}{h_n} \right) \right]^2 f(y) dy = h_n \int \left[ h_n^2 g_n(z) - h_n^2 g_n \left( \frac{x_{n,k}-z}{h_n} \right) \right]^2 f(y) (y-hz) dz \leq 2\|f\|_\infty^2 h_n.$$ 

Now let us take $U_n = 2\|h_n^2 g_n(y)\|_\infty$, $\mathcal{F} = \mathcal{F}_{n,k}$, and $\sigma_n^2 = 2\|f\|_\infty^2 h_n$. Then for $n$ large enough, we have $\sigma_n \leq U_n/2$. Too see this point, note that, from (3.1), $U_n = 2\|h_n^2 g_n(y)\|_\infty = O(1)$, and $\sigma_n = O(h_n^{1/2})$. We also have $\sqrt{n} \sigma_n \geq U_n/\sqrt{\log(U_n/\sigma_n)}$ because of $h_n \to 0$ and $n h_n^2 / (1/h_n^2) \to \infty$. Moreover, from (C6), one has

$$\lim_{n \to \infty} \sup_{U_n \to \infty} \frac{U_n \log(U_n/\sigma_n) + \sqrt{n} \sigma_n \log(U_n/\sigma_n)}{b_n} \leq \limsup_{n \to \infty} \frac{2 \sqrt{n} \sigma_n \log(U_n/\sigma_n)}{b_n} = 0.$$ 

Consequently, there exists a positive integer number $n_0$ such that for any $n \geq n_0$,

$$b_n \geq C \left( U_n \log(U_n/\sigma_n) + \sqrt{n} \sigma_n \log(U_n/\sigma_n) \right)$$

and $\sqrt{n} \sigma_n + U_n \log(U_n/\sigma_n) \leq 2 \sqrt{2n} h_n \|f\|_\infty$.

Now applying Giné and Guillou (2001)'s lemma, for $n \geq n_0$, one has

$$P \left( b_n^{1/2} \sup_{1 \leq k \leq l_n, x \in l_{n,k}} \left| \hat{f}_{n,k}(x) - Ef_{n,k}(x) \right| \geq \varepsilon \right) \leq U_n \exp \left\{ -\frac{b_n \varepsilon U_n}{4Lnh_n \|f\|_\infty} \right\}.$$
Therefore,
\[
\lim_{\delta \to 0} \lim_{n \to \infty} \sup_{n} \frac{n h_n^{\beta+1}}{b_n^2} \log P \left( \frac{n h_n^{\beta+1}}{b_n} \sup_{1 \leq k \leq n} \sup_{x \in \left[ -b_n^{-1}, b_n^{-1} \right]} |\hat{f}(x) - \hat{g}(x)| \geq \varepsilon \right)
\leq \lim_{n \to \infty} \frac{n h_n^{\beta+1}}{b_n^2} \log LL_n - \lim_{n \to \infty} \frac{n h_n^{\beta+1}}{2L b_n} \log \left( 1 + \frac{U_n b_n e}{4 n h L L_n} \|f_L\|_{\infty} \right)
\leq -\frac{\varepsilon^2}{8 L^2 n \|f_L\|_{\infty}}.
\]

Letting \(\eta \to 0\) implies the desired result. \(\square\)

**Lemma 3.3.** Assume (C1)–(C7) hold. Then for any \(\varepsilon > 0\),
\[
\lim_{n \to \infty} \frac{n h_n^{\beta+1}}{b_n^2} \log P \left( \frac{n h_n^{\beta+1}}{b_n} \sup_{x \in \left[ -b_n^{-1}, b_n^{-1} \right]} |\hat{f}(x) - \hat{g}(x)| \geq \varepsilon \right) = -\infty.
\]

**Proof.** Set
\[
\mathcal{F}_n = \left\{ h_n^{\beta} g_n \left( \frac{x}{h_n} \right) : x \notin [-h_n^{-1}, h_n^{-1}] \right\}.
\]

Then \(\mathcal{F}_n, n \geq 1\) are measurable VC classes of functions with VC characteristic \((A, v)\) that does not depend on \(n\). Note that (C1) implies \(f_L(y) \to 0\) as \(y \to \infty\). So, for any \(\eta \in (0, \varepsilon)\), there exists an \(n_0 > 0\) such that for any \(n \geq n_0\),
\[
\sup_{x \notin [-h_n^{-1}, h_n^{-1}]} h_n^2 \int \frac{g_n^2}{h_n^2} \left( \frac{x - y}{h_n} \right) f_L(y) dy \leq \eta h_n.
\]

Now let \(U_n = \|h_n^{\beta} g_n(x)\|_{\infty}, \mathcal{F} = \mathcal{F}_n\), and \(\sigma_n^2 = \eta h_n\). Then for \(n\) large enough, \(\sigma_n \leq U_n/2\). \(\sqrt{n} \sigma_n \geq U_n \sqrt{\log(U_n/\sigma_n)}\). From (C6), one has,
\[
\lim_{n \to \infty} \frac{U_n \log(U_n/\sigma_n) + \sqrt{n} \sigma_n \log(U_n/\sigma_n)}{b_n} \leq \lim_{n \to \infty} \frac{2 \sqrt{n} \sigma_n \log(U_n/\sigma_n)}{b_n} = 0.
\]

Consequently, there exists a positive integer number \(n_0\) such that for any \(n \geq n_0\),
\[
b_n \geq C \left( \frac{U_n \log(U_n/\sigma_n) + \sqrt{n} \sigma_n \log(U_n/\sigma_n)}{b_n} \right)
\]
and \(\sqrt{n} \sigma_n + U_n \sqrt{\log(U_n/\sigma_n)} \leq 2 \eta h_n\). Now applying Giné and Guillou’s lemma, for \(n \geq n_0\), one has
\[
P \left( \frac{n h_n^{\beta+1}}{b_n} \sup_{x \in [-b_n^{-1}, b_n^{-1}]} |\hat{f}(x) - \hat{g}(x)| \geq \varepsilon \right) \leq L \exp \left\{ -\frac{b_n e \varepsilon}{4 U_n \log \left( 1 + \frac{U_n b_n e}{4 n h L L_n} \|f_L\|_{\infty} \right)} \right\}.
\]

Therefore,
\[
\lim_{n \to \infty} \frac{n h_n^{\beta+1}}{b_n^2} \log P \left( \frac{n h_n^{\beta+1}}{b_n} \sup_{x \in [-h_n^{-1}, h_n^{-1}]} |\hat{f}(x) - \hat{g}(x)| \geq \varepsilon \right) \leq \lim_{n \to \infty} \frac{n h_n^{\beta+1}}{b_n^2} \log L - \lim_{n \to \infty} \frac{n h_n^{\beta+1}}{L U_n b_n} \log \left( 1 + \frac{U_n b_n e}{4 n h L L_n} \right) \leq -\frac{\varepsilon^2}{4 L n \|f_L\|_{\infty}}.
\]

Letting \(\eta \to 0\) implies the desired result. \(\square\)

**Proof of Theorem 2.2.** The lower bound is easy. In fact, for any \(x \in \mathbb{R}\), from Theorem 2.1 we have
\[
\liminf_{n \to \infty} \frac{n h_n^{\beta+1}}{b_n^2} \log P \left( \frac{n h_n^{\beta+1}}{b_n} \|\hat{f}(x) - \hat{g}(x)\|_{\infty} \geq \lambda \right) \geq \liminf_{n \to \infty} \frac{n h_n^{\beta+1}}{b_n^2} \log P \left( \frac{n h_n^{\beta+1}}{b_n} \|\hat{f}(x) - \hat{g}(x)\| \geq \lambda \right) \geq -I_\lambda(\lambda).
\]

Hence
\[
\liminf_{n \to \infty} \frac{n h_n^{\beta+1}}{b_n^2} \log P \left( \frac{n h_n^{\beta+1}}{b_n} \|\hat{f}(x) - \hat{g}(x)\|_{\infty} \geq \lambda \right) \geq -I(\lambda).
\]
To show the opposite inequality, note that \( \| \hat{f}_n - E(\hat{f}_n) \|_\infty \) equals to
\[
\max \left\{ \sup_{x \in [-h_n^{-1}, h_n^{-1}]} |\hat{f}_n(x) - \hat{E}_n(x)|, \sup_{x \notin [-h_n^{-1}, h_n^{-1}]} |\hat{f}_n(x) - E_n(x)| \right\}.
\]
and \( \sup_{x \in [-h_n^{-1}, h_n^{-1}]} |\hat{f}_n(x) - \hat{E}_n(x)| \) is bounded above by
\[
\max \left\{ \sup_{1 \leq k \leq b_n} |\hat{f}_{n,k}(x) - \hat{E}_{n,k}(x)| + |\hat{f}_n(z_{n,k}) - E_{n,k}(z_{n,k})| \right\}.
\]

By Lemmas 3.2 and 3.3 and using inequality: max\( \log a, \log b \) \( \leq \log(a + b) \leq \log 2 + \max(\log a, \log b), a \geq 0, b \geq 0, \) we have that for any \( 0 < \epsilon < \lambda, \)
\[
\lim_{n \to \infty} \sup_{n \geq 3} \frac{nh_n}{b_n^2} \log P \left( \frac{nh_n^{\beta + 1}}{b_n} \|\hat{f}_n - E_{\hat{f}_n}\|_\infty > \lambda \right)
\leq \lim_{n \to \infty} \sup_{n \geq 3} \frac{nh_n}{b_n^2} \log P \left( \frac{nh_n^{\beta + 1}}{b_n} \sup_{x \in [-h_n^{-1}, h_n^{-1}]} |\hat{f}_n(x) - \hat{E}_n(x)| > \lambda \right)
\leq \lim_{n \to \infty} \sup_{n \geq 3} \frac{nh_n}{b_n^2} \log \left[ P \left( \frac{nh_n^{\beta + 1}}{b_n} \max_{1 \leq k \leq b_n} |\hat{f}_{n,k}(x) - \hat{E}_{n,k}(x)| \geq \lambda - \epsilon \right) \right]
\leq \lim_{n \to \infty} \sup_{n \geq 3} \frac{nh_n}{b_n^2} \log P \left( \frac{nh_n^{\beta + 1}}{b_n} \max_{1 \leq k \leq b_n} |\hat{f}_{n,k}(x) - \hat{E}_{n,k}(x)| > \lambda - \epsilon \right).
\]
By Chebyshev’s inequality,
\[
P \left( \frac{nh_n^{\beta + 1}}{b_n} \max_{1 \leq k \leq b_n} |\hat{f}_{n,k}(x) - \hat{E}_{n,k}(x)| \geq \lambda - \epsilon \right)
\leq \frac{1}{\max_{1 \leq k \leq b_n} \left\{ \exp \left( \frac{b_n^2}{nh_n} \frac{h_n}{(\lambda - \epsilon)t} \Phi_{x_n}(t) \right) \right\}}
\leq \frac{1}{\max_{1 \leq k \leq b_n} \left\{ \exp \left( \frac{b_n^2}{nh_n} \frac{h_n}{(\lambda - \epsilon)t} \right) \right\}} \left[ 1 + \frac{nh_n^2}{b_n^2} \max_{x \in \mathbb{R}} \Phi_{x_n}(t) (1 + o(1)) + O \left( \frac{2t b_n}{nh_n} \right) \right].
\]
Therefore,
\[
\lim_{n \to \infty} \sup_{n \geq 3} \frac{nh_n}{b_n^2} \log P \left( \frac{nh_n^{\beta + 1}}{b_n} \|\hat{f}_n - E_{\hat{f}_n}\|_\infty > \lambda \right) \leq - \left( \frac{\lambda - \epsilon}{t} \right) \max_{x \in \mathbb{R}} \Phi_{x_n}(t).
\]
The desired result follows from the fact that the last inequality holds for arbitrary \( 0 < \epsilon < \lambda \) and any \( t > 0. \)

**Proof of Corollary 2.1.** To show the first two claims one only has to investigate the term \( \Phi_{x_n}(t). \) Now \( \Phi_{x_n}(t) = E \exp(t b_nh_n^2 [\hat{f}_n(x) - f_X(x)]), \) it can be rewritten as
\[
\Phi_{x_n}(t) = E \exp(t b_nh_n^2 [\hat{f}_n(x) - \hat{E}_n(x)]) \cdot \exp(t b_nh_n^2 [E_{\hat{f}_n}(x) - f_X(x)]).
\]
So, it suffices to show that
\[
\frac{nh_n}{b_n^2} \cdot t b_nh_n^2 [E_{\hat{f}_n}(x) - f_X(x)] \to 0
\]
as \( n \to \infty. \) The conditions (C8) and (C10) imply
\[
E_{\hat{f}_n}(x) - f_X(x) = O(h_n^2) \quad \text{uniformly in } x \in \mathbb{R}, \tag{3.3}
\]
so the desired result follows from (C9).

The third claim follows from similar argument, hence omitted here for brevity. \( \square \)
References