T E C H N I C A L

R E P O R T

0325

BOUNDARY ESTIMATION AND ESTIMATION OF DISCONTINUITY POINTS IN DECONVOLUTION PROBLEMS

A. DELAIGLE and I. GIJBELS

I A P S T A T I S T I C S N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

Boundary estimation and estimation of discontinuity points in deconvolution problems*

A. Delaigle and I. Gijbels

Institut de Statistique, Université catholique de Louvain, 20 Voie du Roman Pays, Louvain-la-Neuve, Belgium.

Abstract

This paper studies estimation of the boundary of the support of a density function when only a contaminated sample from the density is available. Estimation of the boundary of the support is a first necessary step when estimating a density with support different from the whole real line, since then modifications of the usual kernel type estimators are needed for consistent estimation of the density at the endpoints of its support. Apart from this, boundary estimation is also of interest on its own, since it is the basic problem in, for example, frontier estimation in efficiency analysis in econometrics. The method proposed in this paper can also be used for estimating locations of discontinuity points of a density in the same deconvolution context. We establish the asymptotic law of the proposed estimator as well as approximate expressions for its mean squared error, and this for various types of error densities. These expressions then serve to discuss rates of convergence of the estimator and also shed some light on theoretical choices of the bandwidth parameter involved. An illustration of the method is given.

Key words and phrases: asymptotic law, boundary points, deconvolution problem, density estimation, diagnostic function, discontinuity points, endpoints, rates of convergence.

1 Introduction

In this paper, we consider kernel estimation of boundary points and discontinuity points of a density from a sample that has been contaminated by random noise. The contamination problem is often referred to as a deconvolution problem and has applications in many different fields such as chemistry or public health. In this context a so-called deconvolution kernel density estimator has been proposed in the literature. See, for example, Carroll and Hall (1988) and Stefanski and Carroll (1990). When the support of the density is

* This research was supported by 'Projet d'Actions de Recherche Concertées' (No. 98/03–217) from the Belgian government. Financial support from the IAP research network nr P5/24 of the Belgian State (Federal Office for Scientific, Technical and Cultural Affairs) is also gratefully acknowledged.

not the whole real line, the deconvolution kernel density estimator is not consistent at the boundary points: one has to modify this estimator and take the boundary points into account, such as is done in, for example, Zhang and Karunamuni (2000). When these points are unknown, it is necessary to estimate them. Similar problems arise when the density has discontinuity points.

A second interest in the estimation of boundary points comes from the problem of economic frontier estimation, which has applications in many different fields such as for example public services or banks. A typical example of frontier estimation is when one tries to determine the maximal level of output that can be produced for a given level of input. This frontier (or production) function allows to investigate efficiencies of firms and other instances. Many different methods have been developed to estimate a frontier, but most of them rely on the fact that the observations do not contain any measurement error. In the more realistic setup of stochastic frontier estimation, where the data are contaminated by noise, these methods do generally not provide consistent estimators of the frontier and other methods have to be applied.

Some authors have already turned their attention to the boundary estimation problem. In the specific context of frontier estimation, Kneip and Simar (1996) propose, under quite restrictive assumptions, a method to estimate a stochastic frontier from panel data. Hall and Simar (2002) propose a kernel method to estimate a stochastic frontier that is consistent when the error variance is small. We propose a deconvolution kernel method to estimate a boundary point, that is consistent whatever the error variance. The idea of the method is to estimate the boundary point by the maximiser of a certain diagnostic function. This is related to procedures used in the error-free case to estimate discontinuity points of a density or a regression function.

Similarly as for density estimation in the deconvolution context, the behaviour of our estimator depends strongly on the type of error that contaminates the data. See for example Fan (1991c), who considers two classes of error densities: the ordinary smooth and the supersmooth error densities. We prove consistency of the proposed boundary estimation method for both classes of error densities. We also discuss theoretical choices of bandwidth parameters.

This paper is organized as follows. In Section 2 we present the general problem of boundary point estimation. The estimation procedure is introduced in Section 3, where we also provide a brief discussion on other existing methods. In Section 4 we establish the asymptotic law for the estimator and deduce approximate expressions for its bias and variance, and this for the ordinary smooth and the supersmooth error cases. The proofs of the results are given in Section 5 (for ordinary smooth errors) and in Section 6 (for supersmooth errors).

2 Deconvolution and boundary points

Suppose we are interested in a density f_X, but we observe an i.i.d. sample $Y_1, \ldots, Y_n$ from the density f_Y, where

$$Y_i = X_i + Z_i \qquad\qquad i = 1, \ldots, n\,, \tag{2.1}$$

and where for all i, Z_i is a r.v. independent of X_i, of known continuous density f_Z, representing the error in the data, and X_i is a r.v. of density f_X. The case where f_Z is totally or partially unknown may also be considered, if further observations, such as for example a sample from f_Z itself, are available. See for example Barry and Diggle (1995), Neumann (1997a) and Li and Vuong (1998).

In the case where the density f_X is continuous, a so-called deconvolving kernel density estimator of f_X has been proposed. Consider a kernel function K and a smoothing parameter $h = h_n > 0$, depending on n, called the bandwidth. The deconvolving kernel estimator of f_X is then defined by

$$\widehat{f}_X(x; h) = \frac{1}{nh} \sum_{j=1}^{n} K_Z\Big(\frac{x - Y_j}{h}\Big),$$

where $K_Z(u) = (2\pi)^{-1} \int e^{-itu} \varphi_K(t)/\varphi_Z(t/h)\, dt$, with φ_L the Fourier transform of a function or a random variable L. See Carroll and Hall (1988) and Stefanski and Carroll (1990) for an introduction to this estimator. Throughout this paper, we consider real and symmetric kernel functions K. Note that this guarantees that K_Z is real and symmetric.

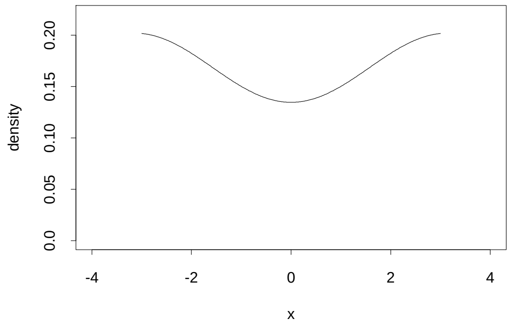

Figure 2.1: *A density with bounded support* $[-3, 3]$.

In this paper, we consider the case where the uncontaminated density f_X has one or two finite boundary points and f_X is not continuous in these points. Figure 2.1 shows an example of such a density f_X, with support $[-3, 3]$.

In the case of a single left (respectively right) boundary point τ, the density f_X is supported on an interval $[\tau, +\infty[$ (respectively $]-\infty, \tau]$) with $|\tau| < \infty$, and may be decomposed into

$$f_X = r_X + d\, I_{[\tau, +\infty[}\,, \tag{2.2}$$

with $d > 0$ (respectively $d < 0$) where r_X is continuous on the whole real line and

$$d = \lim_{\substack{x \to \tau \\ >}} f_X(x) - \lim_{\substack{x \to \tau \\ <}} f_X(x) = f_X(\tau^+) - f_X(\tau^-)$$

is the size of the discontinuity. Here we used the standard notations $g(a^+) = \lim_{x \underset{>}{\to} a} g(x)$ for the right-hand limit, and $g(a^-) = \lim_{x \underset{<}{\to} a} g(x)$ for the left-hand limit of a function g at a point $a \in I\!\!R$. In the case of two boundary points τ_1 and τ_2, when f_X is supported on the compact interval $[\tau_1, \tau_2]$, we can model this as

$$f_X = r_X + d_1 I_{[\tau_1,+\infty[} + d_2 I_{[\tau_2,+\infty[},$$

where r_X is continuous on the whole real line,

$$d_1 = \lim_{x \underset{>}{\to} \tau_1} f_X(x) - \lim_{x \underset{<}{\to} \tau_1} f_X(x) = \lim_{x \underset{>}{\to} \tau_1} f_X(x) = f_X(\tau_1^+) > 0$$

is the size of the left discontinuity, and

$$d_2 = \lim_{x \underset{>}{\to} \tau_2} f_X(x) - \lim_{x \underset{<}{\to} \tau_2} f_X(x) = - \lim_{x \underset{<}{\to} \tau_2} f_X(x) = -f_X(\tau_2^-) < 0$$

is the size of the right discontinuity.

3 The estimation method

When the data of interest are observed directly (i.e. without error), a simple and consistent approach for estimating the boundary points is to estimate the left endpoint of the support by the smallest observation, and the right endpoint by the largest observation. In the case of measurement error however, these simple estimators $\min(Y_1, \ldots, Y_n)$ and $\max(Y_1, \ldots, Y_n)$ will not be consistent estimators of the boundary points of f_X but rather of those of f_Y. More precisely, if f_X is supported on $[\tau_1, \tau_2]$ and f_Z is supported on $[a, b]$ say, then f_Y is supported on $[\tau_1 + a, \tau_2 + b]$, as can be seen easily from the fact that under model (2.1), $f_Y = f_X * f_Z$ where $*$ denotes the deconvolution product, i.e. $f_Y(x) = \int f_X(x-y) f_Z(y) dy$. So, $\min(Y_1, \ldots, Y_n)$ and $\max(Y_1, \ldots, Y_n)$ are consistent estimates of $\tau_1 + a$ and $\tau_2 + b$ respectively. This is illustrated in the left panel of Figure 3.2 which shows results from 100 simulations of samples of sizes $n = 100$ and $n = 1000$ from the density presented in Figure 2.1, contaminated with Laplace error (supported on $I\!\!R$). Presented are the 100 estimated values of the boundary points -3 and 3. The index of the simulation is presented on the horizontal axes.

The method we propose uses the fact that a boundary point is a particular discontinuity point of the density. The idea is then to use methods that exist in the error-free case to detect a discontinuity point, and to adapt them to the case of boundary point estimation with contaminated data. We focus on kernel methods. In the error-free case, several such methods to detect a discontinuity point have been proposed. They are all based on the following basic idea: estimate a discontinuity point by the maximizer of an appropriate diagnostic function. Chu and Cheng (1996) choose as diagnostic function the difference of two kernel density estimators. Couallier (1999, 2000) uses the derivative of a kernel density estimator. See Wu and Chu (1993), Gijbels, Hall and Kneip (1999), Gijbels and Goderniaux (2001) and Goderniaux (2001), among others, for similar methods in the regression context.

We propose a diagnostic function based on derivative estimation, and for a single boundary point τ, we define the estimator of τ by

$$\widehat{\tau} = \mathrm{argmax}_x \, |\widehat{J}(x)|, \tag{3.1}$$

where the diagnostic function $\widehat{J}(x)$ is given by

$$\widehat{J}(x) = \frac{1}{nh} \sum_{i=1}^{n} K_Z' \Big(\frac{x - Y_i}{h} \Big).$$

Hence our diagnostic function $\widehat{J}$ is equal to h times the derivative of the deconvolution kernel density estimator of f_X (the multiplicative factor h is only introduced for convenience). The kernel estimate of f_X is a smooth function. In the current context, it is to be expected that this estimate will be continuous but with large derivatives when approaching the endpoints (large positive derivative for a left endpoint and large negative derivative for a right endpoint). This is the motivation behind the estimator (3.1).

The consistency of the method will be established in Section 4, but we already give some heuristic arguments here. Suppose that the variance of $\widehat{J}(x)$ tends to zero as n tends to infinity. Then asymptotically $\widehat{J}(x)$ behaves like $\mathrm{E}\,\widehat{J}(x)$, where, under sufficient regularity conditions, we have

$$\mathrm{E}\,\widehat{J}(x) = d\, K\Big(\frac{x - \tau}{h} \Big) + O(h). \tag{3.2}$$

See Lemma 5.3.

If we choose a kernel K which is a symmetric function, with a global maximum in zero, we see that, asymptotically, expression (3.2) is maximized (for $d > 0$) or minimized (for $d < 0$) at $x = \tau$, which implies that asymptotically, $\mathrm{argmax}_{x\in I\!R}\, \widehat{J}(x) = \tau$ (if $d > 0$) or $\mathrm{argmin}_{x\in I\!R}\, \widehat{J}(x) = \tau$ (if $d < 0$). Hence in practice our estimator can be written as

$$\widehat{\tau} = \mathrm{argmax}_x \, \widehat{J}(x), \;\; (\text{ respectively } \widehat{\tau} = \mathrm{argmin}_x \, \widehat{J}(x)),$$

for a left (respectively right) boundary point τ.

Note that the situation where f_X has a discontinuity point can also be modelled via (2.2), with $d > 0$ or $d < 0$ the size of the discontinuity. Hence the above method may be applied for the estimation of a discontinuity point as well, but in that case one has no information on the sign of the discontinuity and has to stick to the original definition (3.1).

In the case of two boundary points τ_1 and τ_2, we define, using similar ideas, the estimators

$$\widehat{\tau_1} = \mathrm{argmax}_x \, \widehat{J}(x) \;\text{ and } \widehat{\tau_2} = \mathrm{argmin}_x \, \widehat{J}(x).$$

To illustrate the method, Figure 3.1 presents, for increasing bandwidths, the diagnostic function $\widehat{J}$ for a sample of size $n = 100$ from the density of Figure 2.1 contaminated by

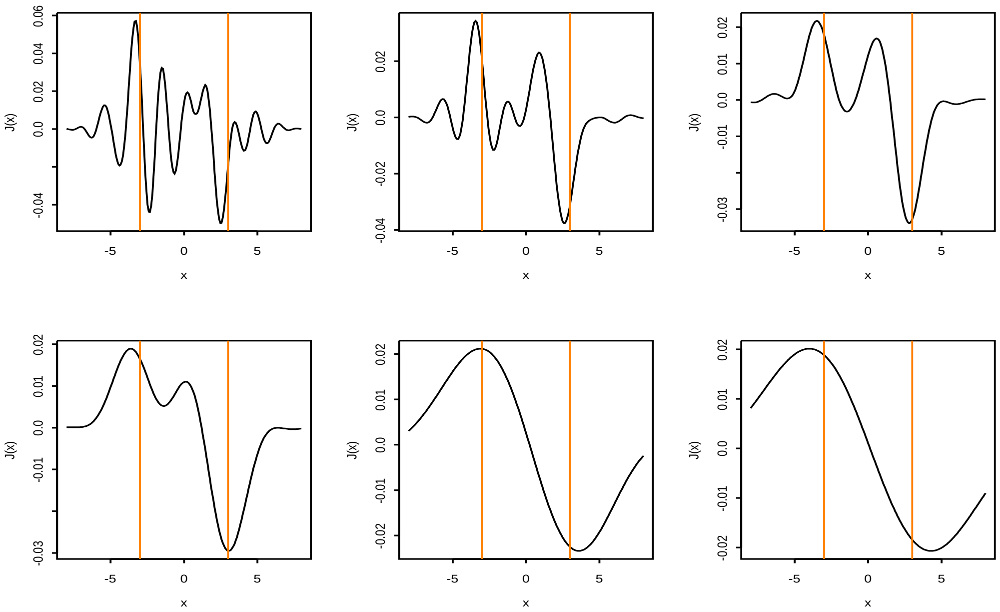

Figure 3.1: *Typical shape of $\widehat{J}(x)$ for a sample of size $n = 100$ from the density in Figure 2.1. contaminated with a Laplace error, for increasing bandwidths (from the left to the right and from the top to the bottom).*

a Laplace error. The actual endpoints (-3 and 3) are indicated by vertical lines. We see that the diagnostic function is indeed maximized at points close to -3 and minimized at points close to 3. From this figure it is also clear that choosing an appropriate bandwidth for the diagnostic function $\widehat{J}$ will be an important practical issue. This issue together with other practical implementation issues of the method and illustrations on its use will be the subject of a forthcoming paper by the same authors. See Delaigle and Gijbels (2003b).

In Figure 3.2 (right panel) we illustrate the performance of the proposed method (with an appropriate data-driven choice of the bandwidth) for the density in Figure 2.1 contaminated with a Laplace error. As can be seen from this figure, the newly proposed method appears as being consistent and of good finite sample performance, in contrast to the naive estimators $\min(Y_1, \ldots, Y_n)$ and $\max(Y_1, \ldots, Y_n)$.

In Section 4, we will see that, as for density estimation from contaminated data, the asymptotic properties of our estimator depend strongly on the error distribution. As in Fan (1991c) we classify the errors in two categories, the supersmooth distributions and the ordinary smooth distributions.

Definition 1. *The distribution of a random variable Z is said to be*

(i) supersmooth of order β if its characteristic function $\varphi_Z(t)$ satisfies:

$$d_0|t|^{\beta_0}\exp(-|t|^{\beta}/\gamma) \leq |\varphi_Z(t)| \leq d_1|t|^{\beta_1}\exp(-|t|^{\beta}/\gamma) \quad as\ t \to \infty,$$

for some positive constants d_0, d_1, β, γ and constants β_0 and β_1;

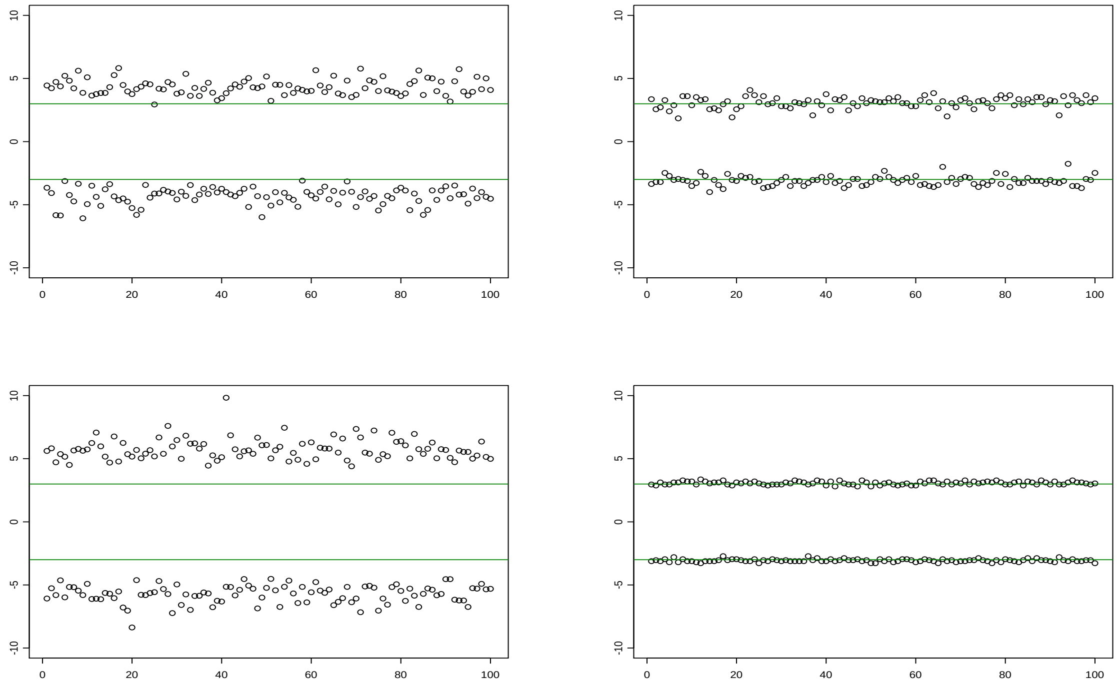

Figure 3.2: *Estimation of boundary points* -3 *and 3. Results from 100 simulations of samples of sizes* $n = 100$ *(upper panels) and* $n = 1000$ *(lower panels) from the density in Figure 2.1, contaminated with a Laplace error. Left panels: estimators* $\min(Y_1, \ldots, Y_n)$ *and* $\max(Y_1, \ldots, Y_n)$*; Right panels: proposed estimation method.*

(ii) ordinary smooth of order β *if its characteristic function* $\varphi_Z(t)$ *satisfies:*

$$d_0|t|^{-\beta} \leq |\varphi_Z(t)| \leq d_1|t|^{-\beta} \quad \textit{as } t \to \infty,$$

for some positive constants d_0, d_1 *and* β*.*

We will see that for supersmooth error densities (e.g. normal and Cauchy densities) the rate of convergence of the estimator is logarithmic, and hence very slow. By contrast this rate is much faster (algebraic) for ordinary smooth error densities (e.g. gamma and Laplace densities). The same distinction shows up when considering deconvolving kernel density estimation (see for example Fan (1991c)).

Other estimation methods

Under different conditions, alternative ways to estimate a boundary or a discontinuity point in the case of contaminated data have been proposed in the literature. Hall and Simar (2002) consider a kernel type estimator of a boundary point which is consistent when the error variance is small: they assume that the error variance (the variance of Z) depends on the sample size n and tends to zero as n tends to infinity.

Neumann (1997b) restricts to the case of ordinary smooth error densities, and studies optimal rates of convergence for an estimate of a discontinuity point. He considers a diagnostic function, similar to ours, defined by $\widehat{J}(x) = n^{-1} \sum_{i=1}^n K_Z\big(\frac{Y_i - x}{h}\big)$, with K a kernel that satisfies some quite particular conditions and depends on two bandwidths h and g. More precisely, he assumes that for all x and $\theta \in I\!R$, $K\big(\frac{x-\theta}{h}; g\big) = \frac{1}{h}\int \rho\big(\frac{y-\theta}{h}\big)\mu\big(\frac{x-y}{g}\big)\,dy$ where μ is the Haar wavelet and ρ is a continuous non-negative symmetric function with compact support $[-1, 1]$, satisfying $||\varphi_\rho(t)(1 + |t|)^{\beta+1}||_{L_p} < \infty$ for $p = 1, 2$ and $|\varphi_\rho(t)| \leq C(1 + |t|)^{-\beta}$ (here $\| \cdot \|$ denotes the usual L_p norm). Neumann (1997b) establishes optimal rates of convergence of the estimator $\widehat{\tau}$ to the discontinuity point τ. Although in theory this method reaches optimal rates of convergence, there are some questions arising for the practical implementation of it. For example: how to choose the bandwidths h and g in practice? In many cases, our estimator, which uses a single bandwidth and a simpler class of kernels, has a good finite sample performance (see Delaigle and Gijbels (2003b)), and also deals with the more difficult supersmooth error case.

4 Asymptotic law and consistency of the estimator

Consider τ, a single boundary or discontinuity point of a density f_X. In this section we show that the estimator for τ introduced in the previous section is a consistent estimator, and establish its asymptotic law. We assume that τ lies in a compact interval $[A, B]$, and thus our estimator is defined as $\widehat{\tau} = \text{argmax}_{x\in[A,B]} |\widehat{J}(x)|$. We introduce the following notations. Let $I_n = \{x \in [A, B] : |x - \tau| > h^{1+\delta}\}$, with $\delta > 0$, and let E_n be the set of endpoints of a partition of the interval $[A, B]$ in n^{1+q} intervals of equal size, i.e. $E_n = \{z_0, \ldots , z_{n^{1+q}}\}$, where $z_0 = A < z_1 < \ldots < z_{n^{1+q}} = B$, and $z_{j+1} - z_j = (B - A)/n^{1+q}$ for $j = 0, \ldots , n^{1+q} - 1$. For a square integrable function f, let also $R(f)$ denote $\int f^2(x)\,dx$. Throughout this paper we make the assumption that for all $t \in I\!R$, $\varphi_Z(t) \neq 0$ and $\|\varphi_K\|_\infty < \infty$. We will treat the ordinary smooth and the supersmooth error cases separately.

4.1 Ordinary smooth error case

If the error is ordinary smooth of order β, we assume that φ_K is supported on $[-B, B]$, with $0 < B < \infty$. We also introduce the following assumptions.

Condition A:

(A1) K is a symmetric, three times continuously differentiable, kth order kernel ($k \geq 2$ even), such that $\|K\|_\infty = K(0) > \max_{x\neq 0} |K(x)|$, $K''(0) < 0$, $\int |u|^{k+3}|K''(u)|\,du < \infty$, and $\int |t|^{2r+2\beta}|\varphi_K(t)|^2\,dt < \infty$, $\int |uK^{(r)}(u)|\,du < \infty$ for $r \leq 3$;

(A2) r_X is Lipschitz continuous with Lipschitz constant L;

(A3) f_Y is differentiable on $I\!R \setminus \{\tau\}$ such that $\sup_{x\in I\!R\setminus\{\tau\}} |f_Y^{(\ell)}(x)| < \infty$ for $\ell = 0, 1$;

(A4) K_Z is four times continuously differentiable such that, for $r \le 4$, $\int |K_Z^{(r)}(u)|^2\,du \sim h^{-2\beta}$, $\|K_Z^{(r)}\|_\infty = O(h^{-\beta})$, $\int |K_Z''(u)|\,du = O(h^{-\beta})$ and $\int |u|\cdot|K_Z''(u)|^2\,du = O(h^{-2\beta})$;

(A5) $h \to 0$ as $n \to \infty$, such that, for some $p \in I\!N_0 = I\!N \setminus \{0\}$,

$$nh^{\frac{2+2\delta+\beta}{1+q}} \to \infty \qquad \text{and} \qquad \sum_{n=1}^{\infty} n^{1+q-p}h^{-4\delta p-p-2\beta p} < \infty,$$

with $0 < \delta < 1/2$ and $q \in I\!N_0$ as above;

(A6) r_X is l times continuously differentiable on $I\!R$ with $0 \le l \le k$ (with k the order of the kernel) and r_X is $l+2+\rho$ times continuously differentiable on $I\!R \setminus \{\tau\}$, where $\rho = \mathbf{1}\{l \text{ is odd}\}$, such that $\|r_X^{(j)}\|_\infty < \infty$ for $j = 0, \dots, l$, and $\sup_{x \in I\!R\setminus\{\tau\}} |r_X^{(j)}(x)| < \infty$ for $j = l+1, \dots, l+2+\rho$;

(A6') r_X is l times continuously differentiable on $I\!R$ with $l > k$ and r_X is $k+3$ times continuously differentiable on $I\!R \setminus \{\tau\}$ with $\|r_X^{(j)}\|_\infty < \infty$ for $j = 0, \dots, k+1$ and $\sup_{x \in I\!R\setminus\{\tau\}} |r_X^{(j)}(x)| < \infty$ for $j = k+2, k+3$.

Although these conditions look rather involved, they are quite common in density deconvolution problems. A discussion on these conditions is provided in Section 4.3 which illustrates this statement.

Under the conditions stated, we prove the convergence in law of the estimator to a normal density. From the asymptotic law, we deduce an approximate expression for the Mean Squared Error (MSE) of the estimator $\widehat{\tau}$ and discuss rates of convergence of the estimator to its target point τ. The proofs of the results are deferred to Section 5.

The asymptotic law of the estimator is described in the following theorem, which necessitates the following notations. For a positive integer k_2, defined in Theorem 4.1, let

$$\begin{aligned} B_\tau &= \frac{f_Y(\tau^+) + f_Y(\tau^-)}{2}\,, \\ D_{\tau,j} &= \frac{(-1)^{k_2+1}}{(k_2+1)!}[r_X^{(k_2+1)}(\tau^+) + (-1)^j\, r_X^{(k_2+1)}(\tau^-)] \int_{-\infty}^{0} u^{k_2+1}K''(u)\,du \quad j = 1, 2\,. \end{aligned} \tag{4.1}$$

Theorem 4.1. *Suppose the error is ordinary smooth of order β and assume Conditions (A1)–(A5). Then we have,*

(i) if Condition (A6) holds, and $h = o(n^{-\frac{1}{2\beta+2k_1+7}})$, with $k_1 = l+\rho-1$, then

$$\sqrt{\frac{n}{h}}\Big[\frac{\tau - \widehat{\tau}}{\sqrt{R(K_Z'')}} - \frac{h^{k_2+2}D_{\tau,j}}{dK''(0)\sqrt{R(K_Z'')}}\Big] \xrightarrow{\text{L}} N\Big(0; \frac{B_\tau}{d^2\{K''(0)\}^2}\Big), \tag{4.2}$$

with $k_2 = l+\rho$ and $j = 1$, and where $R(K_Z'') = \int (K_Z''(u))^2 du$;

(ii) if Condition (A6') holds, and $h = o(n^{-\frac{1}{2\beta+2k_1+7}})$, *with* $k_1 = k$, *then (4.2) holds with* $k_2 = k+1$ *and* $j = 2$.

From this theorem, we deduce the following corollary, which provides an *approximate* mean squared error ($\mathcal{A}$MSE) of the estimator $\widehat{\tau}$ of τ.

Corollary 4.1. *Suppose the error is ordinary smooth of order* β, *and assume Conditions (A1)–(A5). Then we have,*

(i) if Condition (A6) holds, and $h = o(n^{-\frac{1}{2\beta+2k_1+7}})$, *with* $k_1 = l+\rho-1$, *then*

$$\mathcal{A}\mathrm{MSE}[\widehat{\tau}] = \frac{h^{2k_2+4}D_{\tau,j}^2}{d^2\{K''(0)\}^2} + \frac{hR(K_Z'')B_\tau}{nd^2\{K''(0)\}^2}, \tag{4.3}$$

with $k_2 = l+\rho$ *and* $j = 1$;

(ii) if Condition (A6') holds, and $h = o(n^{-\frac{1}{2\beta+2k_1+7}})$, *with* $k_1 = k$, *then (4.3) holds with* $k_2 = k+1$ *and* $j = 2$.

Remark 1. Note that in the particular case of no noise in the data (i.e. $\beta = 0$), $f_Y = f_X$ and B_τ is equal to $[f_X(\tau^+) + f_X(\tau^-)]/2$. In the case where r_X is not differentiable in τ (i.e. $l = \rho = 0$), it is easy to check that $D_{\tau,1} = K(0)[r_X'(\tau^+) - r_X'(\tau^-)]$. See the proof of Theorem 4.1.

From Theorem 4.1, we deduce that

$$\sqrt{\frac{n}{h}}\Big[\frac{\tau-\widehat{\tau}}{\sqrt{R(K_Z'')}} - \frac{h^{k_2+2}D_{\tau,j}}{dK''(0)\sqrt{R(K_Z'')}}\Big] = O_{\mathrm{P}}(1), \qquad \text{for } j = 1 \text{ and } 2,$$

where $k_2 = l+\rho$ (for $j = 1$) or $k_2 = k+1$ (for $j = 2$), and thus

$$\tau - \widehat{\tau} = O_{\mathrm{P}}\Big(\sqrt{\frac{hR(K_Z'')}{n}}\Big) + \frac{h^{k_2+2}D_{\tau,j}}{dK''(0)} = O_{\mathrm{P}}\Big(\sqrt{\frac{hR(K_Z'')}{n}}\Big) + O(h^{k_2+2}). \tag{4.4}$$

Under the conditions of the theorem, we know that $R(K_Z'') \sim h^{-2\beta}$ (by condition (A4) and since K_Z is real). Hence the term $hR(K_Z'')$ is of order $h^{1-2\beta}$. To obtain rates of convergence of the estimator, we need to investigate both terms at the right-hand side of expression (4.4). The appearance of the term $hR(K_Z'') \sim h^{1-2\beta}$ immediately leads to making the distinction between the case where the exponent $1-2\beta$ is ≥ 0 and the case where $1-2\beta < 0$. Similarly, this distinction is coming up when looking at $\mathcal{A}$MSE in (4.3).

If $0 \leq \beta \leq 1/2$ (which includes the error-free case where $\beta = 0$) and hence $1-2\beta \geq 0$, minimization of (4.4) with respect to h leads to choosing h as small as possible, since in this case h is in the numerator of both terms. Under the conditions of the theorem, we can take for example $h \sim n^{-\frac{1}{2\beta+1}+\epsilon}$ with $\epsilon > 0$ as small as we want (see Proposition 4.3). Since in this case the second term of (4.4) is negligible compared to the first term, we can deduce that $\tau - \widehat{\tau} = O_{\mathrm{P}}(n^{-\frac{1}{2\beta+1}+\epsilon})$, with $\epsilon > 0$ as small as we want, and thus a rate as close as we want to $n^{-\frac{1}{2\beta+1}}$.

For his estimator, involving the particular construction of the kernel and two bandwidths, Neumann (1997b) obtains a rate of convergence of the MSE of order $n^{-\frac{2}{2\beta+1}}$, which is slightly better than the above rate.

For $\beta > 1/2$, or equivalently $1 - 2\beta < 0$, we see that the optimal bandwidth is the balancing bandwidth (i.e. the bandwidth which makes the two terms of (4.4) of the same order). In case (i) of the theorem (with $D_{\tau,1} \neq 0$), this bandwidth satisfies $h^{2l+4+2\rho} \sim \frac{1}{nh^{2\beta-1}}$, i.e. $h \sim n^{-\frac{1}{2\beta+2l+3+2\rho}}$. From (4.4), we conclude then that $\tau - \widehat{\tau} = O_{\mathrm{P}}(n^{-\frac{l+2+\rho}{2\beta+2l+3+2\rho}})$. In case (ii), if $D_{\tau,2} \neq 0$, the bandwidth satisfies $h^{2k+6} \sim \frac{1}{nh^{2\beta-1}}$, i.e. $h \sim n^{-\frac{1}{2\beta+2k+5}}$ and $\tau - \widehat{\tau} = O_{\mathrm{P}}(n^{-\frac{k+3}{2\beta+2k+5}})$. If $D_{\tau,2} = 0$, one has to go one or several steps further in the Taylor expansions used in the proofs, until finding a non-zero leading term. Note that such a higher order Taylor expansion will not always lead to an explicit expression for the bias part (for example if $r_X \in C_\infty$ such that $r_X^{(j)}(\tau) = 0$ for all $j \in I\!N_0$). A similar remark can be made for the case $D_{\tau,1} = 0$.

Neumann (1997b) does not consider the particular case where r_X is differentiable and gives a rate of convergence for the MSE of the estimator of order $n^{-\frac{4}{2\beta+3}}$ for the large class of functions f_X discontinuous in τ. For the case of a density f_X such that r_X is not differentiable in τ we obtain the same rate of convergence. In the case where r_X is differentiable in τ, we obtain faster rates. For example, for a second order kernel (i.e. $k = 2$), if r_X is $l = 0 < k$ times differentiable, the $\mathcal{A}$MSE of $\widehat{\tau}$ tends to zero at a rate $n^{-\frac{4}{2\beta+3}}$ for both our method and Neumann's method. For the same kernel, if r_X is $l = 1 < k$ (respectively $l = 3 > k$) times differentiable, the $\mathcal{A}$MSE of $\widehat{\tau}$ tends to zero at a rate $n^{-\frac{8}{2\beta+7}}$ (respectively $n^{-\frac{10}{2\beta+9}}$) for our method, which is faster than the rate $n^{-\frac{4}{2\beta+3}}$.

4.2 Supersmooth error case

If the error is supersmooth of order β with β_0 and β_1 as in Definition 1, we assume that φ_K is supported on $[-B, B]$, with $0 < B < \infty$. For simplicity we prove the results for $B = 1$. We also introduce the following conditions.

Condition B:

(B1)–(B3): the same as Conditions (A1)–(A3), but where the condition on φ_K changes into: $\int |t|^{2r-2\beta_0} |\varphi_K(t)|^2 \, dt < \infty$;

(B4) K_Z is four times continuously differentiable such that, for $r \leq 4$, $\|K_Z^{(r)}\|_\infty = O\big(h^{\beta_0} \exp(h^{-\beta}/\gamma)\big)$, $R(K_Z^{(r)}) = O\big(h^{2\beta_0} \exp(2h^{-\beta}/\gamma)\big)$, and $\int |K_Z''(u)| \, du = O\big(h^{\beta_3} \exp(h^{-\beta}/\gamma)\big)$, with β_3 a real constant;

(B5) $f_Y(\tau^-) > 0$ and for n large enough,

$$h^{-1}\int_{-\infty}^{\infty}\{K_Z''(y)\}^2 f_Y(\tau - hy)\,dy \geq c_5 f_Y(\tau^-)h^{c_6}\exp\big(\frac{2}{h^\beta\gamma} - \frac{4\beta b_n}{h^\beta\gamma}\big) \qquad (4.5)$$

where $b = h^{\beta/(4+10)}$, c_5 is a positive constant and c_6 is a constant depending on β_0.

(B6)–(B6'): the same as Conditions (A6)–(A6').

As in the ordinary smooth error case, these conditions may look rather involved, but as illustrated in Section 4.3 they are among rather commonly-assumed conditions for this (difficult) class of error densities.

If the error is supersmooth of order β and if Condition B is satisfied, we obtain the same asymptotic law of the estimator as in the ordinary smooth error case. From there, we deduce the same approximate expression for the MSE of $\widehat{\tau}$. However the rates of convergence of the estimator $\widehat{\tau}$ to its target point τ are different.

The asymptotic law of the estimator $\widehat{\tau}$ is provided in the following theorem. Note that in the supersmooth error case the study of the variance of the random quantity $h\widehat{J}'(\tau)$ is more tedious than in the ordinary smooth error case.

Theorem 4.2. *Suppose the error is supersmooth of order $\beta > 0$ and $h = d\cdot(2/\gamma)^{1/\beta}(\ln n)^{-1/\beta}$ with $d > 1$. Further assume that Conditions (B1)–(B5) hold. Then we have*

(i) if Condition (B6) holds, then

$$\frac{\tau - \widehat{\tau}}{h\sqrt{\mathrm{Var}(h\widehat{J}'(\tau))}} - \frac{\mathrm{E}[h\widehat{J}'(\tau)]}{dK''(0)\sqrt{\mathrm{Var}(h\widehat{J}'(\tau))}} \xrightarrow{\mathrm{L}} N\Big(0;\frac{1}{d^2\{K''(0)\}^2}\Big), \qquad (4.6)$$

with $\mathrm{E}[h\widehat{J}'(\tau)] = h^{l+1+\rho}D_{\tau,1} + O(h^{l+\rho+2})$*, and where* $\mathrm{Var}(h\widehat{J}'(\tau)) = (nh)^{-1}\int\big\{K_Z''(u)\big\}^2 f_Y(\tau - hu)\,du + O(n^{-1})$*;*

(ii) if Condition (B6') holds, then (4.6) holds with $\mathrm{E}[h\widehat{J}'(\tau)] = h^{k+2}D_{\tau,2} + O(h^{k+3})$.

Further, if $\int|u|\cdot|K_Z''(u)|^2\,du = O(R(K_Z''))$, we can write

$$\sqrt{nh}\Big[\frac{\tau - \widehat{\tau}}{h\sqrt{R(K_Z'')}} - \frac{\mathrm{E}[h\widehat{J}'(\tau)]}{dK''(0)\sqrt{R(K_Z'')}}\Big] \xrightarrow{\mathrm{L}} N\Big(0;\frac{B_\tau}{d^2\{K''(0)\}^2}\Big).$$

The following corollary establishes an approximate mean squared error of the estimator $\widehat{\tau}$.

Corollary 4.2. *Suppose the error is supersmooth of order $\beta > 0$ and $h = d\cdot(2/\gamma)^{1/\beta}(\ln n)^{-1/\beta}$ with $d > 1$. Further assume that Conditions (B1)–(B5) hold. Then we have* (*i*) *if Condition (B6) holds, then*

$$\mathcal{A}\mathrm{MSE}[\widehat{\tau}] = \frac{h^{2k_2+4}D^2_{\tau,j}}{d^2\{K''(0)\}^2} + \frac{h^2\,\mathrm{Var}(h\widehat{J}'(\tau))}{d^2\{K''(0)\}^2}, \tag{4.7}$$

with $j = 1$ and $k_2 = l + \rho$;

(*ii*) *if Condition (B6') holds, then (4.7) holds with $j = 2$ and $k_2 = k + 1$.*

If we further assume that $\int |u| \cdot |K_Z''(u)|^2\,du = O(R(K_Z''))$, we have

$$\mathcal{A}\mathrm{MSE}[\widehat{\tau}] = \frac{h^{2k_2+4}D^2_{\tau,j}}{d^2\{K''(0)\}^2} + \frac{hR\big(K_Z''\big)B_\tau}{nd^2\{K''(0)\}^2}.$$

Under the conditions of the theorem, we have $R(K_Z'') = O\big(h^{2\beta_0}\exp(2h^{-\beta}/\gamma)\big)$. With the condition imposed on the bandwidth, the first term of (4.4) is thus of order $O_{\mathrm{P}}\big((\ln n)^{-(\beta_0+1/2)/\beta} n^{\frac{d-\beta}{2}-\frac{1}{2}}\big)$, which is negligible compared to the second term. Hence we obtain, from (4.4), that in case (*i*), $\widehat{\tau} - \tau = O_{\mathrm{P}}\big((\ln n)^{-(l+2+\rho)/\beta}\big)$, and in case (*ii*), $\widehat{\tau} - \tau = O_{\mathrm{P}}\big((\ln n)^{-(k+3)/\beta}\big)$. So, in both cases, we obtain much slower rates of convergence than in the ordinary smooth error case.

4.3 Discussion on the conditions

We now provide some discussion on Conditions A and B introduced in Sections 4.1 and 4.2. We first remark that it is not difficult to find a kernel that satisfies Condition (A1). See for example Delaigle and Gijbels (2003a). Condition (A2) simply imposes a certain regularity on the continuous part of the density f_X of X. Condition (A3) is not very restrictive either and is satisfied in many cases. For example it is satisfied as soon as the contaminating density f_Z is bounded differentiable on $\mathbb{R}$, since we have $f_Y(x) = f_X * f_Z(x)$ and $f_Y'(x) = f_X * f_Z'(x)$. Note that we do not impose f_Y to be differentiable in τ. As a matter of fact, doing so we include the error-free case where $f_Y = f_X$ (and hence τ is a discontinuity point of f_Y).

Although it may look quite overwhelming at first sight, Condition (A4) is standard in deconvolution problems. See for example Fan (1991a,b,c) and the discussion below. A surprising particularity of this condition is that the asymptotic behaviour of the derivatives of the pseudo kernel K_Z is the same whatever the order of the derivative. Note in particular also that the supremum of these derivatives tends to infinity as $n \to \infty$. This partly explains the considerable difficulty in proving results in the deconvolution context (in comparison and in contrast to the classical error-free density estimation setup).

With a little more effort, it is possible to translate the set of conditions (A4) or (B4) into simple conditions on the kernel K and the density f_Z. The next two propositions provide such possible sets of simple conditions in the case of an ordinary smooth error density (Proposition 4.1) or a supersmooth error density (Proposition 4.2). See Delaigle (2003) for proofs of these two propositions.

Proposition 4.1. *Consider an ordinary smooth error density f_Z. Let $r \in I\!N$. If $\varphi_Z(t) \neq 0$ for all t, and if φ_K is supported on $[-B, B]$ with $0 < B < \infty$, we have*

(i) $R(K_Z^{(r)}(\cdot; h)) = (2\pi)^{-1} \int t^{2r} |\varphi_K(t)|^2 |\varphi_Z(t/h)|^{-2} dt.$

(ii) For j and $k \in I\!N$, if $\|\varphi_K\|_\infty < \infty$, we have, for n large enough,

$$c_1(j)\, h^{-k\beta} \leq \int |t|^j |\varphi_K(t)|^k |\varphi_Z(t/h)|^{-k}\, dt \leq c_2(j)\, h^{-k\beta},$$

where $$c_1(j) = (2d_1)^{-k} \int_{\{B/2<|t|\leq B\}} |t|^{j+k\beta} |\varphi_K(t)|^k\, dt$$

and $$c_2(j) = 2(2/d_0)^k \int |t|^{j+k\beta} |\varphi_K(t)|^k\, dt.$$

(iii) If $\|\varphi_K\|_\infty < \infty$, we have, for n large enough,

$$\|K_Z^{(r)}\|_\infty \leq 2(d_0\pi)^{-1} h^{-\beta} \int |t|^{r+\beta} |\varphi_K(t)|\, dt.$$

(iv) If φ_K and φ_Z are twice continuously differentiable such that $B_K \stackrel{def}{=} \max_{0\leq j\leq 2} \|\varphi_K^{(j)}\|_\infty < \infty$ and $\max_{0\leq j\leq 2} \int |t|^{r+\beta+j-2} |\varphi_K^{(j)}(t)|\, dt < \infty$, $B_Z = \max_{0\leq j\leq 2} \|\varphi_Z^{(j)}\|_\infty < \infty$ and

$$d_{0,j} |t|^{-\beta-j} \leq |\varphi_Z^{(j)}(t)| \leq d_{1,j} |t|^{-\beta-j} \quad \text{as } |t| \to \infty, \text{ for } j = 0, 1, 2 \tag{4.8}$$

for some positive constants $d_{0,j}, d_{1,j}$, β, then, for $r = 0$ and $\beta \geq 1$, or for $r \geq 1$, we have

$$\int |K_Z^{(r)}(x; h)|\, dx = O(h^{-\beta}) \text{ and } \int |x| \cdot |K_Z^{(r)}(x; h)|^2\, dx = O(h^{-2\beta}).$$

Remark 2. Result *(i)* in Proposition 4.1 in fact also holds for supersmooth error densities.

Proposition 4.2. *Consider a supersmooth error density f_Z. Let $r \in I\!N$. If $\varphi_Z(t) \neq 0$ for all t, and if φ_K is supported on $[-B, B]$ with $0 < B < \infty$ and $\|\varphi_K\|_\infty < \infty$, we have*

(i) for j and $k \in I\!N$, if $\int |t|^{j-k\beta_0} |\varphi_K(t)|^k\, dt < \infty$, we have, for n large enough,

$$\int |t|^j |\varphi_K(t)|^k |\varphi_Z(t/h)|^{-k}\, dt \leq \frac{2^{k+1}}{d_0^k} h^{k\beta_0} \exp(k|B/h|^\beta/\gamma) \int |t|^{j-k\beta_0} |\varphi_K(t)|^k\, dt.$$

(ii) If $\int |t|^{r-\beta_0} |\varphi_K(t)|\, dt < \infty$, we have, for n large enough,

$$\|K_Z^{(r)}\|_\infty \leq \frac{2h^{\beta_0}}{\pi d_0} \exp(|B/h|^\beta/\gamma) \int |t|^{r-\beta_0} |\varphi_K(t)|\, dt.$$

(iii) *If φ_K and φ_Z are twice continuously differentiable such that* $B_K \stackrel{def}{=} \max_{0\le j\le 2}\|\varphi_K^{(j)}\|_\infty < \infty$, $\sup_{0\le j\le 2}\int |t|^{-\beta_0}|\varphi_K^{(j)}(t)|\,dt < \infty$, $\sup_{j=0,1}\int |t|^{\beta_1-2\beta_0-\beta+1}|\varphi_K^{(j)}(t)|\,dt < \infty$, $B_Z = \max_{0\le j\le 2}\|\varphi_Z^{(j)}\|_\infty < \infty$ *and, for* $j = 0, 1, 2$,

$$d_{0,j}|t|^{\beta_0+j\beta-j}\exp(-|t|^\beta/\gamma) \le |\varphi_Z^{(j)}(t)| \le d_{1,j}|t|^{\beta_1+j\beta-j}\exp(-|t|^\beta/\gamma), \tag{4.9}$$

as $|t| \to \infty$, *for some positive constants* $d_{0,j}, d_{1,j}, \beta, \gamma$, *and constants* β_0 *and* β_1, *then, for* $r \ge 0$ *and* $\beta > 0$, *we have*

$$\int |K_Z^{(r)}(x;h)|\,dx = O\Big(h^{-\beta_3}\exp(|B/h|^\beta/\gamma)\Big),$$

where $\beta_3 = 2\beta - 3\beta_0 + 2\beta_1$.

Note that Condition (4.8) is nothing but the extension of the definition of an ordinary smooth error density towards the derivatives of the characteristic function. It is satisfied by many ordinary smooth error densities. Similarly, Condition (4.9) is an extension of the definition of a supersmooth error density.

The next proposition tells us when Condition (A5) is satisfied when considering a bandwidth of the form $n^{-1/\alpha}$. This is the type of bandwidth that one usually encounters in deconvolution problems with an ordinary smooth error (see Fan (1991a,b,c)). For a proof of the proposition, see Section 5.4.

Proposition 4.3. *Suppose that* $h \sim n^{-1/\alpha}$, *with* $\alpha > 2\beta + 1$. *Then for all* $0 < \delta < \min(\frac{\alpha-1-2\beta}{4}, 1/2)$, *Condition (A5) is satisfied as soon as we take* $q > (\beta + 4 - \alpha)/\alpha$ *and* $p > \frac{1+\alpha+\alpha q}{\alpha-2\beta-1-4\delta}$.

The next proposition established in Fan (1991a) gives conditions under which (4.5) of Condition (B5) is satisfied. It can be checked in the proof given by Fan (1991a) that the proposition holds under the following (less restrictive) set of conditions.

Proposition 4.4 (Fan (1991a)). *Suppose that the error density is supersmooth of order* $\beta > 0$. *Suppose also that* $\varphi_Z(t) = R_Z(t) + i\,I_Z(t)$, *where* $I_Z(t)$ *and* $R_Z(t) \in I\!R$, *with either* $I_Z(t) = o(R_Z(t))$ *or* $R_Z(t) = o(I_Z(t))$ *as* $t \to \infty$, *and* $\varphi_K(t)$ *is symmetric with support* $[-1, 1]$, *such that* $\varphi_K(t) > c.(1-t)^{r+3}$, *for* $t \in [1-\eta, 1[$ *for some* $\eta > 0$ *and* $c > 0$. *Then, if* f_Y *is continuous and differentiable on* $I\!R \setminus \{x\}$, $f_Y(x^-) > 0$ *and* $\sup_{z\ne x}|f_Y^{(j)}(z)| < \infty$ *for* $j = 0, 1$, *we have, for* n *sufficiently large,*

$$h^{-1}\int_{-\infty}^{\infty}\{K_Z^{(r)}(y)\}^2 f_Y(x-hy)\,dy \ge c_5 f_Y(x^-)h^{c_6}\exp\big(\frac{2}{h^\beta\gamma} - \frac{4\beta b_n}{h^\beta\gamma}\big),$$

where $b = h^{\beta/(2r+10)}$, c_5 *is a positive constant depending on* r *and* c_6 *is a constant depending on* β_0.

Note that the condition on φ_Z in Proposition 4.4 essentially means that the tail of φ_Z is purely real or purely imaginary. Such a condition is easy to check. For example when f_Z is symmetric, then φ_Z is purely real. The condition on φ_K in Proposition 4.4 is also easily verified.

Conditions (A6) and (A6') are only introduced to distinguish the case where the smoothness of r_X is larger than that of K from the case where the smoothness of K is larger than that of r_X.

5 Proofs for the ordinary smooth error case

In this section we prove the results of Section 4.1 regarding ordinary smooth error densities of order β.

When considering a Taylor expansion of r_X around a boundary or a discontinuity point τ, we will consider the left-hand and right-hand derivatives $r_X^{(j)}(\tau^-) = \lim_{x \underset{<}{\to} \tau} r_X^{(j)}(x)$ and $r_X^{(j)}(\tau^+) = \lim_{x \underset{>}{\to} \tau} r_X^{(j)}(x)$. Suppose that r_X is $\ell + 1$ times continuously differentiable on $I\!R \setminus \{\tau\}$, and let $c < \tau$. We can apply an ℓth order Taylor expansion of r_X around c for all $u < 0$:

$$r_X(c+u) = r_X(c) + ur'_X(c) + \ldots + \frac{u^\ell}{\ell!} r_X^{(\ell)}(c) + R_\ell(c),$$

where $R_\ell(c) = \dfrac{u^{\ell+1}}{(\ell+1)!} r_X^{(\ell+1)}(c+\theta u)$, with $0 < \theta < 1$, is the remainder term in the Taylor expansion. If we take the limit $c \underset{<}{\to} \tau$ we get

$$r_X(\tau+u) = r_X(\tau) + ur'_X(\tau^-) + \ldots + \frac{u^\ell}{\ell!} r_X^{(\ell)}(\tau^-) + R_\ell(\tau),$$

where $R_\ell(\tau) = \dfrac{u^{\ell+1}}{(\ell+1)!} r_X^{(\ell+1)}(\tau+\theta u)$, with $0 < \theta < 1$. Similarly, for $u > 0$, we get

$$r_X(\tau+u) = r_X(\tau) + ur'_X(\tau^+) + \ldots + \frac{u^\ell}{\ell!} r_X^{(\ell)}(\tau^+) + R_\ell(\tau).$$

5.1 Intermediate results

We start by giving some intermediate results that are necessary to prove Theorem 4.1 and Corollary 4.1. The proofs of these two results are given in Section 5.2 and the intermediate results of this section are proved in Section 5.3. Among the conditions used to prove these intermediate results (the lemmas and propositions) several resemble, but depend on the mth derivative of K and K_Z, for different values of m. For simplicity of presentation, we write these conditions for a general mth order derivative (with m denoted as a superscript in the notations for the conditions), and refer to them in the text by giving the adequate m value. Note that these conditions are all part of Condition A for given values of m.

Condition C:

($\mathbf{C}_1^m$) K is m times continuously differentiable such that $\lim_{|x|\to\infty} K^{(m-1)}(x) = 0$ and φ_K is supported on $[-B, B]$ with $0 < B < \infty$;

($\mathbf{C}_2^m$) K_Z is m times continuously differentiable;

($\mathbf{C}_3^m$) $\int |u| \cdot |K_Z^{(m)}(u)|^2\, du = O(h^{-2\beta})$;

($\mathbf{C}_4^m$) $\int |K_Z^{(m)}(u)|\, du = O(h^{-\beta})$;

($\mathbf{C}_5^m$) $\|K_Z^{(m)}\|_\infty = O(h^{-\beta})$;

($\mathbf{C}_6^m$) $\int |K_Z^{(m)}(u)|^2\, du \sim h^{-2\beta}$;

($\mathbf{C}_7^m$) $\int |uK^{(m)}(u)|\, du < \infty$.

The following lemma is a generalization of a result of Stefanski and Carroll (1990). It implies that the deconvolution kernel estimator of $f_X^{(r)}$ has the same bias as the kernel estimator of $f_X^{(r)}$ in the error-free case. See Delaigle and Gijbels (2002) for a proof.

Lemma 5.1. *Let $r \geq 0$. If K is a kth order symmetric kernel, r times differentiable such that φ_K is supported on $[-B, B]$, with $0 < B < \infty$, we have*

$$\mathrm{E}\Big[K_Z^{(r)}\Big(\frac{x-Y}{h}\Big)\Big] = \mathrm{E}\Big[K^{(r)}\Big(\frac{x-X}{h}\Big)\Big]. \tag{5.1}$$

Lemma 5.2 below describes the asymptotic behaviour of the quantity $E[h\widehat{J}'(\tau)]$.

Lemma 5.2. *Assume Conditions* (C_1^2) *and* (C_2^2)*, and let K be a kth order symmetric kernel ($k \geq 2$ even), such that $\int |u|^{k+3}|K''(u)|\, du < \infty$, and K has a maximum in 0. Then we have*

(*i*) *if Condition* (A6) *holds, then*

$$\mathrm{E}\Big[\frac{1}{nh}\sum_{i=1}^n K_Z''\Big(\frac{\tau - Y_i}{h}\Big)\Big] = h^{k_2+1} D_{\tau,j} + O(h^{k_1+3}), \tag{5.2}$$

with $j = 1$, $k_2 = l + \rho$ and $k_1 = l + \rho - 1$;

(*ii*) *if Condition* (A6') *holds, then (5.2) holds with $j = 2$, $k_2 = k + 1$ and $k_1 = k$.*

The next lemma gives the asymptotic behaviour of $E[h^r \widehat{J}^{(r)}(x)]$, for any x, and any $r \geq 0$. Note though that for the case $r = 1$, the following lemma would give as first term $dK'((x-\tau)/h))$ which equals 0 for $x = \tau$.

Lemma 5.3. *Let $r \geq 0$. Under* (A2), (C_1^{r+1}), (C_2^{r+1}) *and* (C_7^{r+1})*, we have, for all x,*

$$\mathrm{E}\Big[h^r \widehat{J}^{(r)}(x)\Big] = dK^{(r)}\Big(\frac{x-\tau}{h}\Big) + O(h)\,,$$

uniformly in x.

Lemma 5.4 allows to study the behaviour of the variance of the random quantity $h^{r-1}\widehat{J}^{(r-1)}(\tau)$.

Lemma 5.4. *Let $r \geq 0$. Under Conditions* (A2), (A3), (C_1^{r+1}), (C_2^{r+1}), (C_3^r), (C_7^{r+1}) *and if $K_Z^{(r)}$ is symmetric, we have*

$$\mathrm{Var}\left[h^{-1}K_Z^{(r)}\Big(\frac{\tau - Y}{h}\Big)\right] = h^{-1}\frac{f_Y(\tau^+) + f_Y(\tau^-)}{2} \cdot R(K_Z^{(r)}) + O(h^{-2\beta}). \tag{5.3}$$

The next lemma proves the convergence in law of the rth derivative of the deconvolution kernel density estimator evaluated at the point τ. This result was already established by Fan (1991a), but under different conditions (in particular, under the condition that the density f_X is continuous and differentiable).

Lemma 5.5. *Let $r \geq 0$. Under Conditions* (A2), (A3), (C_1^{r+1}), (C_2^{r+1}), (C_3^r), (C_4^r), (C_5^r), (C_6^r), (C_7^{r+1}), *and if $K_Z^{(r)}$ is symmetric and $nh \to \infty$ as $n \to \infty$, we have*

$$\frac{\frac{1}{nh}\sum_{i=1}^{n} K_Z^{(r)}\big(\frac{\tau - Y_i}{h}\big) - \mathrm{E}\left[\frac{1}{nh}\sum_{i=1}^{n} K_Z^{(r)}\big(\frac{\tau - Y_i}{h}\big)\right]}{\sqrt{\mathrm{Var}\left[\frac{1}{nh}\sum_{i=1}^{n} K_Z^{(r)}\big(\frac{\tau - Y_i}{h}\big)\right]}} \xrightarrow{\mathrm{L}} N(0;1). \tag{5.4}$$

Lemma 5.6 (L_p distance between $\widehat{J}^{(r)}$ and $\mathrm{E}\,\widehat{J}^{(r)}$). *Let $r \geq 0$. Under* (A2), (C_1^{r+1}), (C_2^{r+1}), (C_5^{r+1}), (C_6^{r+1}), (C_7^{r+1}), *if $\int t^{2r+2+2\beta}|\varphi_K(t)|^2\,dt < \infty$, $\sup_{x\in I\!R\setminus\{\tau\}}|f_Y(x)| < \infty$ and if $nh \to \infty$ as $n \to \infty$, we have, for all $p \in I\!N_0$ and for n large enough,*

$$\mathrm{E}[h^r\widehat{J}^{(r)}(x) - h^r\,\mathrm{E}\,\widehat{J}^{(r)}(x)]^{2p} \leq 2n^{-p}h^{-p-2\beta p} \cdot \Big\{2\,c_\beta(K) \cdot \sup_{x\in I\!R\setminus\{\tau\}}|f_Y(x)|\,(2/d_0)^2\Big\}^p,$$

where $c_\beta(K) = \pi^{-1}\int t^{2r+2+2\beta}|\varphi_K(t)|\,dt$.

Lemma 5.7. *Let $r \geq 0$. Assume Conditions* (A2), (A5), (C_1^{r+1}), (C_2^{r+2}), (C_5^{r+1}), (C_5^{r+2}), (C_6^{r+1}) *and* (C_7^{r+1}). *Further assume that $\|K^{(r+1)}\|_\infty < \infty$, $\int t^{2r+2+2\beta}|\varphi_K(t)|^2\,dt < \infty$ and* $\sup_{x\in I\!R\setminus\{\tau\}}|f_Y(x)| < \infty$. *Then, if $nh \to \infty$ as $n \to \infty$, we have for all $\epsilon > 0$*

$$\sum_{n=1}^{\infty} P\Big(h^{-2\delta}\sup_{x\in[A,B]}|h^r\widehat{J}^{(r)}(x) - h^r\,\mathrm{E}\,\widehat{J}^{(r)}(x)| > \epsilon\Big) < \infty\,, \tag{5.5}$$

with $\delta > 0$ as in Condition (A5).

Lemma 5.8. *Suppose that* $\widehat{\tau} = \tau + O(h^{1+\eta})$ *a.s., with* $\eta > 0$. *Let* $r \geq 0$. *Assume Conditions* (A2)*,* (A5)*,* (C$_1^3$)*,* (C$_2^4$)*,* (C$_5^3$)*,* (C$_5^4$)*,* (C$_6^3$)*,* (C$_7^3$)*, and suppose that* $\|K^{(3)}\|_\infty < \infty$, $\sup_{x\in I\!R\setminus\{\tau\}} |f_Y(x)| < \infty$ *and* $\int t^{6+2\beta}|\varphi_K(t)|^2\,dt < \infty$. *Then, if* $nh \to \infty$ *as* $n \to \infty$*, we have*

$$h^2\widehat{J}''(\xi) \xrightarrow{a.s} dK''(0), \tag{5.6}$$

for any ξ *between* τ *and* $\widehat{\tau}$.

Proposition 5.2 below shows that under the conditions of the theorem, the condition $\widehat{\tau} - \tau = O(h^{1+\eta})$ a.s. of Lemma 5.8 is satisfied, with η corresponding to δ appearing in Condition (A5). To prove this we also need Proposition 5.1. Since these two results are of some independent interest, we state them as propositions instead of lemmas.

Proposition 5.1. *Under Conditions* (A1) *to* (A5)*, we have*

$$\sum_{n=1}^{\infty} P(\sup_{x\in I_n} |\widehat{J}(x)| \geq |\widehat{J}(\tau)|) < \infty. \tag{5.7}$$

Proposition 5.2. *Under Conditions* (A1) *to* (A5)*, we have*

$$\widehat{\tau} - \tau = O(h^{1+\delta}) \quad a.s.\ , \tag{5.8}$$

with $\delta > 0$ *as in Condition (A5).*

To prove these two key results, we need the following extra lemmas. Recall the notation δ appearing in the definition of the partition $I_n = \{x \in [A,B] : |x-\tau| > h^{1+\delta}\}$.

Lemma 5.9. *Under Conditions* (A2)*,* (C$_1^1$)*,* (C$_2^1$) *and* (C$_7^1$)*, and if* K *is a symmetric kernel maximized in 0 such that* $\|K\|_\infty = K(0)$*, we have, for* n *large enough and* $\delta > 0$*,*

$$\sup_{x\in I_n} |\,\mathrm{E}\,\widehat{J}(x)| = |d|K(h^\delta) + O(h).$$

Lemma 5.10. *Under Conditions* (A2)*,* (C$_1^1$)*,* (C$_2^1$) *and* (C$_7^1$)*, and if* K *is a symmetric kernel maximized in 0 with* $\|K\|_\infty = K(0)$*, and* K *is 3 times continuously differentiable with* $\|K^{(3)}\|_\infty < \infty$ *we have, for* n *large enough, and* $\delta > 0$

$$\sup_{x\in I_n} |\,\mathrm{E}\,\widehat{J}(x)| - |\,\mathrm{E}\,\widehat{J}(\tau)| = \frac{|d|}{2}\,h^{2\delta}K''(0) + o(h^{2\delta}) + O(h).$$

5.2 Proofs of the asymptotic law and the corollary

Using the intermediate results of Section 5.1, we can prove Theorem 4.1 and Corollary 4.1.

Proof of Theorem 4.1.

By applying a Taylor expansion of $\widehat{J}'$ around τ, we can write

$$0 = \widehat{J}'(\widehat{\tau}) = \widehat{J}'(\tau) + (\widehat{\tau} - \tau)\widehat{J}''(\xi),$$

where ξ lies between τ and $\widehat{\tau}$. Thus we have

$$\tau - \widehat{\tau} = \frac{\widehat{J}'(\tau)}{\widehat{J}''(\xi)}, \tag{5.9}$$

where by Lemma 5.8, $\widehat{J}''(\xi)$ is almost surely different from zero as $n \to \infty$, since $K''(0) < 0$. Under the conditions of the theorem, we have $\widehat{\tau} - \tau = O(h^{1+\delta})$ a.s. (see Proposition 5.2). Hence the conditions of Lemma 5.8 are all satisfied.

Below, we search for the asymptotic law of $\widehat{J}'(\tau)/\widehat{J}''(\xi)$ and deduce from there the asymptotic law of $\tau - \widehat{\tau}$. From Lemma 5.5, for $r = 2$, we know that

$$\frac{h\widehat{J}'(\tau) - \mathrm{E}[h\widehat{J}'(\tau)]}{\sqrt{\mathrm{Var}(h\widehat{J}'(\tau))}} = \frac{\frac{1}{nh}\sum_{i=1}^{n} K_Z''\left(\frac{\tau - Y_i}{h}\right) - \mathrm{E}\left[\frac{1}{nh}\sum_{i=1}^{n} K_Z''\left(\frac{\tau - Y_i}{h}\right)\right]}{\sqrt{\mathrm{Var}\left[\frac{1}{nh}\sum_{i=1}^{n} K_Z''\left(\frac{\tau - Y_i}{h}\right)\right]}} \xrightarrow{\mathrm{L}} N(0;1)$$

where, by Lemma 5.4,

$$\mathrm{Var}\left[\frac{1}{nh}\sum_{i=1}^{n} K_Z''\Big(\frac{\tau - Y_i}{h}\Big)\right] = \frac{1}{n}\,\mathrm{Var}\left(h^{-1}K_Z''\Big(\frac{\tau - Y}{h}\Big)\right) = \frac{B_\tau}{nh}\,R\big(K_Z''\big) + O(n^{-1}h^{-2\beta}),$$

with $B_\tau = \frac{f_Y(\tau^+) + f_Y(\tau^-)}{2}$, and where $R(K_Z'') \sim h^{-2\beta}$ (see Condition (A4)). We now distinguish between cases (i) and (ii) of Theorem 4.1.

Case (i)

From Lemma 5.2 (i), we know that $\mathrm{E}[h\widehat{J}'(\tau)] = h^{l+1+\rho}D_{\tau,1} + O(h^{l+\rho+2})$. Hence we have

$$\frac{h\widehat{J}'(\tau) - h^{l+1+\rho}D_{\tau,1} + O(h^{l+\rho+2})}{\sqrt{\frac{B_\tau}{nh}\,R\big(K_Z''\big) + O(n^{-1}h^{-2\beta})}} \xrightarrow{\mathrm{L}} N(0;1)$$

Now

$$\frac{h\widehat{J}'(\tau) - h^{l+1+\rho}D_{\tau,1} + O(h^{l+\rho+2})}{\sqrt{\frac{B_\tau}{nh}\,R\big(K_Z''\big) + O(n^{-1}h^{-2\beta})}} = \frac{h\widehat{J}'(\tau) - h^{l+1+\rho}D_{\tau,1}}{\sqrt{\frac{B_\tau}{nh}\,R\big(K_Z''\big)}} \cdot a_n + b_n,$$

where

$$a_n = \frac{\sqrt{\frac{B_\tau}{nh}\, R\big(K_Z''\big)}}{\sqrt{\frac{B_\tau}{nh}\, R\big(K_Z''\big) + O(n^{-1}h^{-2\beta})}} \quad \text{and} \quad b_n = \frac{O(h^{l+\rho+2})}{\sqrt{\frac{B_\tau}{nh}\, R\big(K_Z''\big) + O(n^{-1}h^{-2\beta})}} \,.$$

Since $R(K_Z'') \sim h^{-2\beta}$, we have $a_n \to 1$ and $b_n \to 0$ as $n \to \infty$, and we deduce that

$$\frac{h\widehat{J}'(\tau) - h^{l+1+\rho} D_{\tau,1}}{\sqrt{\frac{B_\tau}{nh}\, R\big(K_Z''\big)}} \xrightarrow{\mathrm{L}} N(0;1)$$

i.e.

$$\sqrt{nh}\, \frac{h\widehat{J}'(\tau) - h^{l+1+\rho} D_{\tau,1}}{\sqrt{R\big(K_Z''\big)}} \xrightarrow{\mathrm{L}} N(0; B_\tau)$$

and thus by Lemma 5.8,

$$\sqrt{nh}\, \frac{h\widehat{J}'(\tau) - h^{l+1+\rho} D_{\tau,1}}{h^2 \widehat{J}''(\xi) \sqrt{R\big(K_Z''\big)}} \xrightarrow{\mathrm{L}} N\Big(0; \frac{B_\tau}{d^2\{K''(0)\}^2}\Big)$$

i.e.

$$\sqrt{\frac{n}{h}} \Big[\frac{\widehat{J}'(\tau)}{\widehat{J}''(\xi)\sqrt{R\big(K_Z''\big)}} - \frac{h^{l+\rho} D_{\tau,1}}{\widehat{J}''(\xi)\sqrt{R\big(K_Z''\big)}} \Big] \xrightarrow{\mathrm{L}} N\Big(0; \frac{B_\tau}{d^2\{K''(0)\}^2}\Big)$$

or again, by (5.9),

$$\sqrt{\frac{n}{h}} \Big[\frac{\tau - \widehat{\tau}}{\sqrt{R\big(K_Z''\big)}} - \frac{h^{l+\rho} D_{\tau,1}}{\widehat{J}''(\xi)\sqrt{R\big(K_Z''\big)}} \Big] \xrightarrow{\mathrm{L}} N\Big(0; \frac{B_\tau}{d^2\{K''(0)\}^2}\Big)$$

which, by Lemma 5.8, implies that

$$\sqrt{\frac{n}{h}} \Big[\frac{\tau - \widehat{\tau}}{\sqrt{R\big(K_Z''\big)}} - \frac{h^{l+2+\rho} D_{\tau,1}}{dK''(0)\sqrt{R\big(K_Z''\big)}} \Big] \xrightarrow{\mathrm{L}} N\Big(0; \frac{B_\tau}{d^2\{K''(0)\}^2}\Big).$$

Note that in the particular case where $l = 0$ (and hence $\rho = 0$), the quantity $D_{\tau,1}$ defined in (4.1) reduces to

$$D_{\tau,1} = -[r_X'(\tau^+) - r_X'(\tau^-)] \int_{-\infty}^{0} uK''(u)\, du \,,$$

where further

$$\int_{-\infty}^{0} uK''(u)\, du = [uK'(u)]_{-\infty}^{0} - \int_{-\infty}^{0} K'(u)\, du = -[K]_{-\infty}^{0} = -K(0),$$

by Condition (A1).

Case (ii)

From Lemma 5.2 (ii), we know that $\mathrm{E}[h\widehat{J}'(\tau)] = h^{k+2}D_{\tau,2} + O(h^{k+3})$. Hence, proceeding as for the case (i), we deduce that

$$\sqrt{\frac{n}{h}}\Big[\frac{\tau-\widehat{\tau}}{\sqrt{R\big(K''_Z\big)}} - \frac{h^{k+3}D_{\tau,2}}{dK''(0)\sqrt{R\big(K''_Z\big)}}\Big] \xrightarrow{\text{L}} N\Big(0; \frac{B_\tau}{d^2\{K''(0)\}^2}\Big).$$

$\square$

Proof of Corollary 4.1. From Theorem 4.1, an approximate expression for the expectation ($\mathcal{A}$E) may be derived as follows

$$\mathcal{A}\mathrm{E}\Big[\frac{\tau-\widehat{\tau}}{\sqrt{R\big(K''_Z\big)}} - \frac{h^{k_2+2}D_{\tau,j}}{dK''(0)\sqrt{R\big(K''_Z\big)}}\Big] = o(\sqrt{h/n}).$$

Hence, we obtain

$$\mathcal{A}\mathrm{E}\Big[\frac{\tau-\widehat{\tau}}{\sqrt{R\big(K''_Z\big)}}\Big] = \frac{h^{k_2+2}D_{\tau,j}}{dK''(0)\sqrt{R\big(K''_Z\big)}} + o(\sqrt{h/n})$$

and thus

$$\mathcal{A}\mathrm{E}\big[\tau-\widehat{\tau}\big] = \frac{h^{k_2+2}D_{\tau,j}}{dK''(0)} + o(\sqrt{hR\big(K''_Z\big)/n}).$$

Using Condition (A4), we obtain

$$\mathcal{A}\mathrm{E}\big[\widehat{\tau}\big] = \tau - \frac{h^{k_2+2}D_{\tau,j}}{dK''(0)} + o(n^{-1/2}h^{(1-2\beta)/2}).$$

An approximate expression for the variance ($\mathcal{A}$Var) is given by

$$\frac{n}{h}\,\mathcal{A}\mathrm{Var}\Big[\frac{\tau-\widehat{\tau}}{\sqrt{R\big(K''_Z\big)}} - \frac{h^{k_2+2}D_{\tau,j}}{dK''(0)\sqrt{R\big(K''_Z\big)}}\Big] = \frac{B_\tau}{d^2\{K''(0)\}^2}(1+o(1)),$$

i.e.

$$\mathcal{A}\mathrm{Var}[\widehat{\tau}] = \frac{hR\big(K''_Z\big)B_\tau}{nd^2\{K''(0)\}^2}(1+o(1)).$$

Finally an approximate mean squared error may be written as

$$\mathcal{A}\mathrm{MSE}[\widehat{\tau}] = \frac{h^{2k_2+4}D^2_{\tau,j}}{d^2\{K''(0)\}^2} + \frac{hR\big(K''_Z\big)B_\tau}{nd^2\{K''(0)\}^2}.$$

$\square$

5.3 Proofs of the intermediate results

In this section we prove the intermediate results of Section 5.1.

Proof of Lemma 5.2. From Lemma 5.1, we can write

$$
\begin{aligned}
\mathrm{E}\Big[\frac{1}{nh}\sum_{i=1}^{n} K_Z''\Big(\frac{\tau - Y_i}{h}\Big)\Big] =& \frac{1}{h}\,\mathrm{E}\Big[K''\Big(\frac{\tau - X}{h}\Big)\Big] \\
=& \int K''(u) f_X(\tau - hu)\,du \\
=& \int K''(u) r_X(\tau - hu)\,du + d\int K''(u) I_{[\tau,+\infty[}(\tau - hu)\,du \\
=& \int K''(u) r_X(\tau - hu)\,du + d\int_{-\infty}^{0} K''(u)\,du \\
=& \int K''(u) r_X(\tau - hu)\,du + dK'(0) \\
=& \int K''(u) r_X(\tau - hu)\,du \\
=& \int_{-\infty}^{0} K''(u) r_X(\tau - hu)\,du + \int_{0}^{+\infty} K''(u) r_X(\tau - hu)\,du\ ,
\end{aligned}
$$

since K has a maximum in zero.

To study the behaviour of this term, we make Taylor expansions of r_X around τ^+ or τ^-. Therefore we need to know which moments of K'' exist and are equal to 0. Since K is a kth order kernel, we have

$$\int K(u)\,du = 1, \int uK(u)\,du = 0, \dots, \int u^{k-1}K(u)\,du = 0,$$

and $0 < |\int u^k K(u)\,du| < \infty$. This implies that $\lim_{u\to\pm\infty} u^j K(u) = 0$, for $j = 0, \dots, k$, and thus for $j = 0, \dots, k$

$$\int u^j K'(u)\,du = \Big[u^j K(u)\Big]_{-\infty}^{\infty} - j\int u^{j-1}K(u)\cdot \mathbf{1}\{j > 0\}\,du = 0,$$

which implies that $\lim_{u\to\pm\infty} u^j K'(u) = 0$, for $j = 0, \dots, k$ and thus for $j = 0, \dots, k$ we also have

$$\int u^j K''(u)\,du = \Big[u^j K'(u)\Big]_{-\infty}^{\infty} - j\int u^{j-1}K'(u)\cdot \mathbf{1}\{j > 0\}\,du = 0. \tag{5.10}$$

The conditions of the lemma imply that $|\int u^{k+1}K''(u)\,du| < \infty$ and K'' is symmetric. Hence, since k is even, we also have $\int u^{k+1}K''(u)\,du = 0$.

Now we treat the cases (i) and (ii) separately.

Case (i)

For positive u we make a Taylor expansion of r_X of order $l+1+\rho$ around τ^-:

$$
\begin{aligned}
r_X(\tau-hu) &= \sum_{j=0}^{l+1+\rho} \frac{(-hu)^j}{j!} r_X^{(j)}(\tau^-) + \frac{(-hu)^{l+2+\rho}}{(l+2+\rho)!} r_X^{(l+2+\rho)}(\tau-\theta_1 hu) \\
&= \sum_{j=0}^{l} \frac{(-hu)^j}{j!} r_X^{(j)}(\tau) + \sum_{j=l+1}^{l+1+\rho} \frac{(-hu)^j}{j!} r_X^{(j)}(\tau^-) + \frac{(-hu)^{l+2+\rho}}{(l+2+\rho)!} r_X^{(l+2+\rho)}(\tau-\theta_1 hu),
\end{aligned}
$$

where $0<\theta_1<1$.

For negative u, we make a Taylor expansion of r_X of order $l+1+\rho$ around τ^+:

$$
r_X(\tau-hu) = \sum_{j=0}^{l} \frac{(-hu)^j}{j!} r_X^{(j)}(\tau) + \sum_{j=l+1}^{l+1+\rho} \frac{(-hu)^j}{j!} r_X^{(j)}(\tau^+) + \frac{(-hu)^{l+2+\rho}}{(l+2+\rho)!} r_X^{(l+2+\rho)}(\tau-\theta_2 hu),
$$

where $0<\theta_2<1$. We then obtain

$$
\begin{aligned}
&\mathrm{E}\Big[\frac{1}{nh}\sum_{i=1}^{n} K_Z''\Big(\frac{\tau-Y_i}{h}\Big)\Big] \\
&= \int_{-\infty}^{0} K''(u) r_X(\tau-hu)\,du + \int_{0}^{+\infty} K''(u) r_X(\tau-hu)\,du \\
&= \sum_{j=0}^{l} \frac{(-h)^j}{j!} r_X^{(j)}(\tau) \int_{-\infty}^{+\infty} u^j K''(u)\,du + \sum_{j=l+1}^{l+1+\rho} \frac{(-h)^j}{j!} r_X^{(j)}(\tau^+) \int_{-\infty}^{0} u^j K''(u)\,du \\
&\quad + \sum_{j=l+1}^{l+1+\rho} \frac{(-h)^j}{j!} r_X^{(j)}(\tau^-) \int_{0}^{+\infty} u^j K''(u)\,du \\
&\quad + \frac{(-h)^{l+2+\rho}}{(l+2+\rho)!} \int_{-\infty}^{0} u^{l+2+\rho} r_X^{(l+2+\rho)}(\tau-\theta_2 hu) K''(u)\,du \\
&\quad + \frac{(-h)^{l+2+\rho}}{(l+2+\rho)!} \int_{0}^{+\infty} u^{l+2+\rho} r_X^{(l+2+\rho)}(\tau-\theta_1 hu) K''(u)\,du.
\end{aligned}
$$

Now if l is odd, $l+1$ is even and $u^{l+1}K''(u)$ is symmetric. Thus in this case,

$$
\int_{-\infty}^{0} u^{l+1} K''(u)\,du = \int_{0}^{+\infty} u^{l+1} K''(u)\,du = \frac{1}{2}\int_{-\infty}^{+\infty} u^{l+1} K''(u)\,du = 0,
$$

by expression (5.10), since $l+1\le k+1$. If l is even, then $\rho=0$ and $l+1+\rho=l+1$. Hence in both cases, we have

$$
\begin{aligned}
&\sum_{j=l+1}^{l+1+\rho} \frac{(-h)^j}{j!} r_X^{(j)}(\tau^+) \int_{-\infty}^{0} u^j K''(u)\,du + \sum_{j=l+1}^{l+1+\rho} \frac{(-h)^j}{j!} r_X^{(j)}(\tau^-) \int_{0}^{+\infty} u^j K''(u)\,du \\
&= \frac{(-h)^{l+1+\rho}}{(l+1+\rho)!} r_X^{(l+1+\rho)}(\tau^+) \int_{-\infty}^{0} u^{l+1+\rho} K''(u)\,du
\end{aligned}
$$

$$+\frac{(-h)^{l+1+\rho}}{(l+1+\rho)!}r_X^{(l+1+\rho)}(\tau^-)\int_0^{+\infty}u^{l+1+\rho}K''(u)\,du.$$

Since $l+1+\rho$ is always odd, we further have

$$\int_0^{+\infty}u^{l+1+\rho}K''(u)\,du=-\int_{-\infty}^0 u^{l+1+\rho}K''(u)\,du.$$

We conclude that

$$\begin{aligned}&\mathrm{E}\Big[\frac{1}{nh}\sum_{i=1}^n K_Z''\Big(\frac{\tau-Y_i}{h}\Big)\Big]\\ &\quad=\frac{(-1)^{l+1+\rho}h^{l+1+\rho}}{(l+1+\rho)!}[r_X^{(l+1+\rho)}(\tau^+)-r_X^{(l+1+\rho)}(\tau^-)]\int_{-\infty}^0 u^{l+1+\rho}K''(u)\,du+O(h^{l+2+\rho}).\end{aligned}$$

Case (ii)

We proceed as in (i) and make Taylor expansions of r_X of order $k+2$ around τ^+ and τ^-. Since r_X is at least $k+1$ times differentiable in τ, we have, for $0\le j\le k+1$, $r_X^{(j)}(\tau^+)=r_X^{(j)}(\tau^-)=r_X^{(j)}(\tau)$. Hence we can write

$$\begin{aligned}&\mathrm{E}\Big[\frac{1}{nh}\sum_{i=1}^n K_Z''\Big(\frac{\tau-Y_i}{h}\Big)\Big]\\ &=\int_{-\infty}^0 K''(u)r_X(\tau-hu)\,du+\int_0^{+\infty}K''(u)r_X(\tau-hu)\,du\\ &=\sum_{j=0}^{k+1}\frac{(-h)^j}{j!}r_X^{(j)}(\tau)\int_{-\infty}^{+\infty}u^jK''(u)\,du+\frac{h^{k+2}}{(k+2)!}r_X^{(k+2)}(\tau^+)\int_{-\infty}^0 u^{k+2}K''(u)\,du\\ &\quad+\frac{h^{k+2}}{(k+2)!}r_X^{(k+2)}(\tau^-)\int_0^{+\infty}u^{k+2}K''(u)\,du+\frac{(-h)^{k+3}}{(k+3)!}\int_{-\infty}^0 u^{k+3}r_X^{(k+3)}(\tau-\theta_1hu)K''(u)\,du\\ &\quad+\frac{(-h)^{k+3}}{(k+3)!}\int_0^{+\infty}u^{k+3}r_X^{(k+3)}(\tau-\theta_2hu)K''(u)\,du\\ &=\frac{h^{k+2}}{(k+2)!}[r_X^{(k+2)}(\tau^+)+r_X^{(k+2)}(\tau^-)]\int_{-\infty}^0 u^{k+2}K''(u)\,du+O(h^{k+3}),\end{aligned}$$

where $0<\theta_1,\theta_2<1$, and since the symmetry of $u^{k+2}K''(u)$ implies that

$$\int_{-\infty}^0 u^{k+2}K''(u)\,du=\int_0^{+\infty}u^{k+2}K''(u)\,du=\frac{1}{2}\int_{-\infty}^{+\infty}u^{k+2}K''(u)\,du.$$

$\square$

Proof of Lemma 5.3. We want to prove that

$$\mathrm{E}\Big[h^r\widehat{J}^{(r)}(x)\Big]=dK^{(r)}\Big(\frac{x-\tau}{h}\Big)+O(h).$$

By Lemma 5.1, we have

$$\begin{aligned}
\mathrm{E}\left[h^r \widehat{\mathcal{J}}^{(r)}(x)\right] =& \mathrm{E}\left[\frac{1}{nh}\sum_{i=1}^{n} K_Z^{(r+1)}\Big(\frac{x-Y_i}{h}\Big)\right] \\
=& \frac{1}{h}\mathrm{E}\left[K^{(r+1)}\Big(\frac{x-X}{h}\Big)\right] \\
=& \int K^{(r+1)}(u) f_X(x-hu)\, du \\
=& \int K^{(r+1)}(u) r_X(x-hu)\, du + d\int K^{(r+1)}(u) I_{[\tau,+\infty[}(x-hu)\, du \\
=& \int K^{(r+1)}(u) r_X(x-hu)\, du + d\int_{-\infty}^{\frac{x-\tau}{h}} K^{(r+1)}(u)\, du.
\end{aligned}$$

By Lipschitz continuity of r_X, we can write

$$\begin{aligned}
\Big|\int K^{(r+1)}(u)[r_X(x-hu) - r_X(x)]\, du\Big| \leq & \int |K^{(r+1)}(u)| \cdot |r_X(x-hu) - r_X(x)|\, du \\
=& Lh\int |uK^{(r+1)}(u)|\, du \\
=& O(h)\ ,
\end{aligned}$$

uniformly in x (since the upperbound does not depend on x). Thus

$$\int K^{(r+1)}(u) r_X(x-hu)\, du = r_X(x)\int K^{(r+1)}(u)\, du + O(h) = O(h),$$

since $\int K^{(r+1)}(u)\, du = [K^{(r)}]_{-\infty}^{+\infty} = 0$. Finally we have

$$\mathrm{E}\left[h^r \widehat{\mathcal{J}}^{(r)}(x)\right] = d\int_{-\infty}^{\frac{x-\tau}{h}} K^{(r+1)}(u)\, du + O(h) = dK^{(r)}\Big(\frac{x-\tau}{h}\Big) + O(h).$$

$\square$

Proof of Lemma 5.4. We have

$$\begin{aligned}
\mathrm{Var}\left[h^{-1}K_Z^{(r)}\Big(\frac{\tau-Y}{h}\Big)\right] =& h^{-2}\,\mathrm{Var}\left(K_Z^{(r)}\Big(\frac{\tau-Y}{h}\Big)\right) \\
=& h^{-2}\,\mathrm{E}\left[\Big\{K_Z^{(r)}\Big(\frac{\tau-Y}{h}\Big)\Big\}^2\right] - h^{-2}\Big\{\mathrm{E}\left[K_Z^{(r)}\Big(\frac{\tau-Y}{h}\Big)\right]\Big\}^2 \\
=& h^{-2}\,\mathrm{E}\left[\Big\{K_Z^{(r)}\Big(\frac{\tau-Y}{h}\Big)\Big\}^2\right] + O(1) \\
=& h^{-1}\int \big\{K_Z^{(r)}(u)\big\}^2 f_Y(\tau-hu)\, du + O(1), \qquad (5.11)
\end{aligned}$$

where we used Lemma 5.3. Now, by a first order Taylor expansion of f_Y around τ^+ or τ^-, we can further write

$$\mathrm{Var}\left[h^{-1}K_Z^{(r)}\Big(\frac{\tau-Y}{h}\Big)\right] = h^{-1}\int_{-\infty}^{0} \big\{K_Z^{(r)}(u)\big\}^2 [f_Y(\tau^+) - huf_Y'(\tau-\theta_1 hu)]\, du$$

$$
\begin{aligned}
&+ h^{-1}\int_0^{+\infty} \{K_Z^{(r)}(u)\}^2 [f_Y(\tau^-) - huf_Y'(\tau - \theta_2 hu)]\, du + O(1) \\
=& h^{-1} f_Y(\tau^+) \int_{-\infty}^0 \{K_Z^{(r)}(u)\}^2\, du - \int_{-\infty}^0 u\{K_Z^{(r)}(u)\}^2 f_Y'(\tau - \theta_1 hu)\, du \\
&+ h^{-1} f_Y(\tau^-) \int_0^{+\infty} \{K_Z^{(r)}(u)\}^2\, du - \int_0^{+\infty} u\{K_Z^{(r)}(u)\}^2 f_Y'(\tau - \theta_2 hu)\, du + O(1) \\
=& h^{-1}\frac{f_Y(\tau^+) + f_Y(\tau^-)}{2} \int \{K_Z^{(r)}(u)\}^2\, du - \int_{-\infty}^0 u\{K_Z^{(r)}(u)\}^2 f_Y'(\tau - \theta_1 hu)\, du \\
&- \int_0^{+\infty} u\{K_Z^{(r)}(u)\}^2 f_Y'(\tau - \theta_2 hu)\, du + O(1),
\end{aligned}
$$

with θ_1 and θ_2 between 0 and 1, where we used the symmetry of $K_Z^{(r)}$. Now

$$
\Big| \int_{\mathcal{I}} u\{K_Z^{(r)}(u)\}^2 f_Y'(\tau - \theta_j hu)\, du \Big| \leq \sup_{x \in I\!R\setminus\{\tau\}} |f_Y'(x)| \int_{\mathcal{I}} |u| \cdot |K_Z^{(r)}(u)|^2\, du = O(h^{-2\beta}),
$$

where $\mathcal{I} =]-\infty, 0[$ and $j = 1$ or $\mathcal{I} =]0, +\infty[$ and $j = 2$, and the proof is completed. $\square$

Proof of Lemma 5.5. Denote $h^{-1}K_Z^{(r)}\big(\frac{\tau - Y_i}{h}\big)$ by $Y_{n,i}, i = 1, \ldots, n$. Note that for all $i \neq j$, $Y_{n,i} \perp\!\!\!\perp Y_{n,j}$. We will use the central limit theorem for triangular arrays of random variables. See for example Serfling (1980), page 32. Since in our case the variables $Y_{n,i}, i = 1, \ldots, n$ are i.i.d., we need to verify if the following Lyapounov condition is satisfied: for some $\eta > 0$,

$$
\lim_{n\to\infty} \frac{n\,\mathrm{E}\,|Y_{n,1} - \mathrm{E}(Y_{n,1})|^{2+\eta}}{(n\,\mathrm{Var}(Y_{n,1}))^{(2+\eta)/2}} = 0. \tag{5.12}
$$

We will do so by studying the behaviour of the numerator and the denominator separately.

Numerator:

Remember Minkowski's inequality: $\{\mathrm{E}\,|X + Y|^p\}^{1/p} \leq \{\mathrm{E}\,|X|^p\}^{1/p} + \{\mathrm{E}\,|Y|^p\}^{1/p}$. If we apply this inequality with $X = Y_{n,1}$, $Y = -\mathrm{E}(Y_{n,1})$ and $p = 2 + \eta$, we find

$$
\begin{aligned}
\Big\{\mathrm{E}\,|Y_{n,1} - \mathrm{E}(Y_{n,1})|^{2+\eta}\Big\}^{\frac{1}{2+\eta}} \leq& \Big\{\mathrm{E}\,|Y_{n,1}|^{2+\eta}\Big\}^{\frac{1}{2+\eta}} + \Big\{\mathrm{E}\,|-\mathrm{E}(Y_{n,1})|^{2+\eta}\Big\}^{\frac{1}{2+\eta}} \\
=& \Big\{\mathrm{E}\,|Y_{n,1}|^{2+\eta}\Big\}^{\frac{1}{2+\eta}} + |\,\mathrm{E}(Y_{n,1})| \\
\leq& \Big\{\mathrm{E}\,|Y_{n,1}|^{2+\eta}\Big\}^{\frac{1}{2+\eta}} + \mathrm{E}\,|Y_{n,1}|\,,
\end{aligned}
$$

which is equivalent to

$$
\mathrm{E}\,|Y_{n,1} - \mathrm{E}(Y_{n,1})|^{2+\eta} \leq \Big(\big\{\mathrm{E}\,|Y_{n,1}|^{2+\eta}\big\}^{\frac{1}{2+\eta}} + \mathrm{E}\,|Y_{n,1}|\Big)^{2+\eta}.
$$

Now if $|K_Z^{(r)}|$ is integrable and f_Y is bounded on $\mathbb{R} \setminus \{\tau\}$, we have, using the symmetry of $K_Z^{(r)}$,

$$\begin{aligned}
\mathrm{E}\,|Y_{n,1}| =& h^{-1}\,\mathrm{E}\left|K_Z^{(r)}\Big(\frac{\tau - Y_1}{h}\Big)\right| \\
=& \int |K_Z^{(r)}(u)| f_Y(\tau - hu)\,du \\
=& \int_{-\infty}^{0} |K_Z^{(r)}(u)| f_Y(\tau - hu)\,du + \int_0^{+\infty} |K_Z^{(r)}(u)| f_Y(\tau - hu)\,du \\
\leq& \sup_{x \in \mathbb{R}\setminus\{\tau\}} |f_Y(x)| \int |K_Z^{(r)}(u)|\,du \\
=& O(h^{-\beta})\ ,
\end{aligned}$$

and

$$\begin{aligned}
h^{1+\eta}\,\mathrm{E}\,|Y_{n,1}|^{2+\eta} =& h^{-1}\,\mathrm{E}\left|K_Z^{(r)}\Big(\frac{\tau - Y_1}{h}\Big)\right|^{2+\eta} \\
=& \int |K_Z^{(r)}(u)|^{2+\eta} f_Y(\tau - hu)\,du \\
\leq& ||K_Z^{(r)}||_\infty^{1+\eta} \cdot \sup_{x \in \mathbb{R}\setminus\{\tau\}} |f_Y(x)| \int |K_Z^{(r)}(u)|\,du \\
=& O(h^{-\beta(2+\eta)}).
\end{aligned}$$

We conclude that

$$\begin{aligned}
h^{1+\eta}\,\mathrm{E}\,|Y_{n,1} - \mathrm{E}(Y_{n,1})|^{2+\eta} \leq& \Big(\big\{h^{1+\eta}\,\mathrm{E}\,|Y_{n,1}|^{2+\eta}\big\}^{\frac{1}{2+\eta}} + h^{\frac{1+\eta}{2+\eta}}\,\mathrm{E}\,|Y_{n,1}|\Big)^{2+\eta} \\
=& \Big(O(h^{-\beta}) + h^{\frac{1+\eta}{2+\eta}} O(h^{-\beta})\Big)^{2+\eta} \\
=& O(h^{-\beta(2+\eta)}).
\end{aligned}$$

Denominator:

By Lemma 5.4 and Condition (C_6^r), we have

$$\mathrm{Var}(Y_{n,1}) = h^{-1}\frac{f_Y(\tau^+) + f_Y(\tau^-)}{2} R(K_Z^{(r)}) + O(h^{-2\beta}) \sim h^{-2\beta-1}.$$

We now are ready to verify the Lyapounov condition in (5.12). From the above considerations, we find

$$\lim_{n\to\infty} \frac{n\,\mathrm{E}\,|Y_{n,1} - \mathrm{E}(Y_{n,1})|^{2+\eta}}{(n\,\mathrm{Var}(Y_{n,1}))^{(2+\eta)/2}} = \lim_{n\to\infty} \frac{1}{(nh)^{\eta/2}} \frac{h^{1+\eta}\,\mathrm{E}\,|Y_{n,1} - \mathrm{E}(Y_{n,1})|^{2+\eta}}{(h\,\mathrm{Var}(Y_{n,1}))^{(2+\eta)/2}} = 0,$$

since $h^{-\beta(2+\eta)} = o\big((nh)^{\eta/2} h^{-\beta(2+\eta)}\big)$, and hence the numerator is negligible compared to the denominator. This concludes the proof.

□

Proof of Lemma 5.6. We want to find an upper bound for $\mathrm{E}[h^r \widehat{J}^{(r)}(x) - h^r \mathrm{E}\, \widehat{J}^{(r)}(x)]^{2p}$. By definition, we have $h^r \widehat{J}^{(r)}(x) = \dfrac{1}{nh}\displaystyle\sum_{i=1}^{n} K_Z^{(r+1)}\Big(\frac{x-Y_i}{h}\Big)$. Thus

$$
\begin{aligned}
&\mathrm{E}[h^r \widehat{J}^{(r)}(x) - h^r \mathrm{E}\, \widehat{J}^{(r)}(x)]^{2p} \\
=&\frac{1}{n^{2p}h^{2p}} \mathrm{E}\Big[\sum_{i=1}^{n}\Big\{K_Z^{(r+1)}\Big(\frac{x-Y_i}{h}\Big) - \mathrm{E}\, K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big\}\Big]^{2p} \\
=&\frac{1}{n^{2p}h^{2p}} \mathrm{E} \sum_{i_1,i_2,\dots i_{2p}=1}^{n} \Big[K_Z^{(r+1)}\Big(\frac{x-Y_{i_1}}{h}\Big) - \mathrm{E}\, K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big] \\
&\times \Big[K_Z^{(r+1)}\Big(\frac{x-Y_{i_2}}{h}\Big) - \mathrm{E}\, K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big] \dots \Big[K_Z^{(r+1)}\Big(\frac{x-Y_{i_{2p}}}{h}\Big) - \mathrm{E}\, K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big].
\end{aligned}
$$

If one of the indices $i_1, \dots, i_{2p}$ in the sum, say the index i_j, is different from all the other indices, then the corresponding term, $\big[K_Z^{(r+1)}\big(\frac{x-Y_{i_j}}{h}\big) - \mathrm{E}\, K_Z^{(r+1)}\big(\frac{x-Y_1}{h}\big)\big]$, has an expectation equal to zero. This implies that the only terms in the sum that can lead to a non-zero contribution are the terms for which all the indices appear at least two times. This means that we only have to consider the terms of the sum where at most p different indices appear and each index i_k appears $l_{i_k} \geq 2$ times. Therefore, we can write

$$
\begin{aligned}
&\mathrm{E}[h^r \widehat{J}^{(r)}(x) - h^r \mathrm{E}\, \widehat{J}^{(r)}(x)]^{2p} \\
=&\frac{1}{n^{2p}h^{2p}} \sum_{k=1}^{p} \sum_{i_1\neq i_2 \neq \dots \neq i_k} \sum_{\substack{l_{i_1}, l_{i_2}, \dots l_{i_k} \geq 2 \\ l_{i_1} + \dots + l_{i_k} = 2p}} \mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_{i_1}}{h}\Big) - \mathrm{E}\, K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^{l_{i_1}} \\
&\times \mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_{i_2}}{h}\Big) - \mathrm{E}\, K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^{l_{i_2}} \\
&\times \dots \times \mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_{i_k}}{h}\Big) - \mathrm{E}\, K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^{l_{i_k}} \\
=&\frac{1}{n^{2p}h^{2p}} \sum_{i=1}^{n} \mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_i}{h}\Big) - \mathrm{E}\, K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^{2p} \\
&+ \frac{1}{n^{2p}h^{2p}} \sum_{i=1}^{n} \sum_{j\neq i} \sum_{l_i=2}^{2p-2} \mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_i}{h}\Big) - \mathrm{E}\, K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^{l_i} \\
&\times \mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_j}{h}\Big) - \mathrm{E}\, K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^{2p-l_i} \\
&+ \dots + \frac{1}{n^{2p}h^{2p}} \sum_{i_1\neq i_2 \neq \dots \neq i_p} \mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_{i_1}}{h}\Big) - \mathrm{E}\, K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^{2} \\
&\times \mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_{i_2}}{h}\Big) - \mathrm{E}\, K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^{2} \\
&\times \dots \times \mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_{i_p}}{h}\Big) - \mathrm{E}\, K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^{2}. \qquad (5.13)
\end{aligned}
$$

We have, for all $l \geq 2$,

$$\begin{aligned}
\mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_i}{h}\Big) - \mathrm{E}\,K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^l \\
&= \sum_{j=0}^{l}(-1)^j C_l^j\,\mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_i}{h}\Big)\Big]^j\Big[\mathrm{E}\,K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^{l-j}, \qquad (5.14)
\end{aligned}$$

where C_l^j denotes the binomial coefficient $\begin{pmatrix} l \\ j \end{pmatrix}$. We now first investigate the behaviour of all expectations appearing in the above sum. By Lemma 5.3, we have that

$$\mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big] = dhK^{(r)}\Big(\frac{x-\tau}{h}\Big) + O(h^2) = O(h).$$

Further, for all $j \geq 2$, we find

$$\begin{aligned}
\mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_i}{h}\Big)\Big]^j &= \int \Big\{K_Z^{(r+1)}\Big(\frac{x-y}{h}\Big)\Big\}^j f_Y(y)\,dy \\
&= h\int \big\{K_Z^{(r+1)}(u)\big\}^j f_Y(x-hu)\,du \\
&= h\int_{-\infty}^{\frac{x-\tau}{h}} \big\{K_Z^{(r+1)}(u)\big\}^j f_Y(x-hu)\,du \\
&\quad + h\int_{\frac{x-\tau}{h}}^{\infty} \big\{K_Z^{(r+1)}(u)\big\}^j f_Y(x-hu)\,du,
\end{aligned}$$

where

$$\begin{aligned}
h\Big|\int_{\mathcal{I}} \big\{K_Z^{(r+1)}(u)\big\}^j f_Y(x-hu)\,du\Big| &\leq h\,\|K_Z^{(r+1)}\|_\infty^{j-2} \sup_{x\in I\!R\setminus\{\tau\}} |f_Y(x)| \cdot \int |K_Z^{(r+1)}(u)|^2\,du \\
&= O(h^{1-j\beta}),
\end{aligned}$$

for $\mathcal{I} =]-\infty, \frac{x-\tau}{h}[$ or $\mathcal{I} =]\frac{x-\tau}{h}, +\infty[$. From expression (5.14) we then get

$$\begin{aligned}
\mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_i}{h}\Big) - \mathrm{E}\,K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^l \\
&= O(h^l) + \sum_{j=2}^{l}(-1)^j C_l^j\,\mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_i}{h}\Big)\Big]^j\Big[\mathrm{E}\,K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^{l-j} \\
&= O(h^l) + \sum_{j=2}^{l} O(h^{1-\beta j})\cdot O(h^{l-j}) \\
&= O(h^{1-l\beta}).
\end{aligned}$$

Finally we get

$$\begin{aligned}\mathrm{E}[h^r \widehat{J}^{(r)}(x) - h^r \,\mathrm{E}\, \widehat{J}^{(r)}(x)]^{2p} =& O\Big(\frac{nh^{1-2\beta p}}{n^{2p}h^{2p}}\Big) + O\Big(\frac{n^2h^{2-2\beta p}}{n^{2p}h^{2p}}\Big)\ldots + O\Big(\frac{n^p h^{p-2\beta p}}{n^{2p}h^{2p}}\Big)\\ =& O\Big(n^{-p}h^{-p-2\beta p}\Big),\end{aligned}$$

since $nh \to \infty$ as $n \to \infty$.

From the above calculations, we also see that we can write the following

$$\begin{aligned}\mathrm{E}[h^r \widehat{J}^{(r)}(x) - & h^r \,\mathrm{E}\, \widehat{J}^{(r)}(x)]^{2p}\\ =& \frac{n\cdot(n-1)\cdot\ldots\cdot(n-p+1)}{n^{2p}h^{2p}}\\ &\times\Big\{\mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big) - \mathrm{E}\,K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^2\Big\}^p\cdot(1+o(1))\\ =& \frac{1}{n^p h^{2p}}\Big\{\mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big) - \mathrm{E}\,K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^2\Big\}^p\cdot(1+o(1))\\ =& \frac{1}{n^p h^{2p}}\Big\{\mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^2\Big\}^p\cdot(1+o(1))\\ \le& \frac{1}{n^p h^{2p}}\Big\{2h\sup_{x\in I\!R\setminus\{\tau\}}|f_Y(x)|\cdot R(K_Z^{(r+1)})\Big\}^p\cdot(1+o(1))\\ \le& \frac{1}{n^p h^{2p}}\Big\{2c_\beta(K)h\sup_{x\in I\!R\setminus\{\tau\}}|f_Y(x)|\cdot(2/d_0)^2h^{-2\beta}\Big\}^p\cdot(1+o(1))\\ =& n^{-p}h^{-p-2\beta p}\cdot\Big\{2c_\beta(K)\sup_{x\in I\!R\setminus\{\tau\}}|f_Y(x)|\cdot(2/d_0)^2\Big\}^p\cdot(1+o(1)),\end{aligned}$$

where $c_\beta(K) = \pi^{-1}\int t^{2r+2+2\beta}|\varphi_K(t)|^2\,dt$, and where we used Proposition 4.1 (i) and (ii). This implies that, for n large enough, we have

$$\mathrm{E}[h^r \widehat{J}^{(r)}(x) - h^r \,\mathrm{E}\, \widehat{J}^{(r)}(x)]^{2p} \le 2n^{-p}h^{-p-2\beta p}\cdot\Big\{2c_\beta(K)\sup_{x\in I\!R\setminus\{\tau\}}|f_Y(x)|\cdot(2/d_0)^2\Big\}^p.$$

□

Proof of Lemma 5.7. We want to prove that

$$\sum_{n=1}^{\infty} P\Big(h^{-2\delta}\sup_{x\in[A,B]}|h^r\widehat{J}^{(r)}(x) - h^r\,\mathrm{E}\,\widehat{J}^{(r)}(x)| > \epsilon\Big) < \infty.$$

Let E_n be as defined on page 8. For each x in $[A,B]$ there exists at least one point z in E_n such that $|x-z| \le (B-A)n^{-(1+q)}$. Denote that point by $z(x)$. For all $\omega\in\Omega$, we have

$$\begin{aligned}\sup_{x\in[A,B]}|h^r\widehat{J}^{(r)}(x) - h^r\,\mathrm{E}\,\widehat{J}^{(r)}(x)| \le& \sup_{x\in[A,B]}|h^r\widehat{J}^{(r)}(x) - h^r\widehat{J}^{(r)}(z(x))|\\ &+\sup_{x\in[A,B]}|h^r\widehat{J}^{(r)}(z(x)) - h^r\,\mathrm{E}\,\widehat{J}^{(r)}(z(x))|\\ &+\sup_{x\in[A,B]}|h^r\,\mathrm{E}\,\widehat{J}^{(r)}(z(x)) - h^r\,\mathrm{E}\,\widehat{J}^{(r)}(x)|\end{aligned}$$

$$\equiv S_{1,n} + S_{2,n} + S_{3,n}, \tag{5.15}$$

where to simplify the notations, we do not indicate specifically the dependence of the random variables on ω. We will treat these three terms separately.

For the first term, note that for all $\omega \in \Omega$ and for all $x \in [A, B]$ we have (by the mean-value theorem)

$$|h^r \widehat{J}^{(r)}(x) - h^r \widehat{J}^{(r)}(z(x))| \leq h^r |\widehat{J}^{(r+1)}(\xi)| \cdot |x - z(x)|$$

where ξ lies between x and $z(x)$. Now under the conditions of the lemma, we have

$$|h^r \widehat{J}^{(r+1)}(\xi)| = \Big| \frac{1}{nh^2} \sum_{i=1}^{n} K_Z^{(r+2)} \Big(\frac{\xi - Y_i}{h} \Big) \Big| \leq h^{-2} \|K_Z^{(r+2)}\|_\infty \leq c_1 h^{-2-\beta},$$

with c_1 a positive constant independent of ω and n. We conclude that

$$\sup_{x \in [A,B]} |h^r \widehat{J}^{(r)}(x) - h^r \widehat{J}^{(r)}(z(x))| \leq c_1 \cdot (B - A)\, h^{-2-\beta} n^{-1-q}.$$

Thus

$$h^{-2\delta} S_{1,n} = h^{-2\delta} \sup_{x \in [A,B]} |h^r \widehat{J}^{(r)}(x) - h^r \widehat{J}^{(r)}(z(x))| \leq c_1 \cdot (B - A)\, h^{-2-2\delta-\beta} n^{-1-q}.$$

For handling the second term in (5.15), note that we have, for all $\omega \in \Omega$,

$$\sup_{x \in [A,B]} |h^r \widehat{J}^{(r)}(z(x)) - h^r \,\mathrm{E}\, \widehat{J}^{(r)}(z(x))| \leq \sup_{z \in E_n} |h^r \widehat{J}^{(r)}(z) - h^r \,\mathrm{E}\, \widehat{J}^{(r)}(z)|.$$

Hence for all $\epsilon > 0$ and $\ell \geq 1$, we can write

$$\begin{aligned}
\sum_{n=1}^{\ell} P\Big(h^{-2\delta} S_{2,n} > \epsilon\Big) &\leq \sum_{n=1}^{\ell} P\Big(h^{-2\delta} \sup_{z \in E_n} |h^r \widehat{J}^{(r)}(z) - h^r \,\mathrm{E}\, \widehat{J}^{(r)}(z)| > \epsilon\Big) \\
&= \sum_{n=1}^{\ell} P\Big(h^{-2\delta} \max_{z \in E_n} |h^r \widehat{J}^{(r)}(z) - h^r \,\mathrm{E}\, \widehat{J}^{(r)}(z)| > \epsilon\Big) \\
&= \sum_{n=1}^{\ell} P\Big(\Big\{h^{-2\delta} |h^r \widehat{J}^{(r)}(z_0) - h^r \,\mathrm{E}\, \widehat{J}^{(r)}(z_0)| > \epsilon\Big\} \cup \ldots \\
&\qquad \cup \Big\{h^{-2\delta} |h^r \widehat{J}^{(r)}(z_{n^{1+q}}) - h^r \,\mathrm{E}\, \widehat{J}^{(r)}(z_{n^{1+q}})| > \epsilon\Big\}\Big) \\
&\leq \sum_{n=1}^{\ell} \Big[P\Big(h^{-2\delta} |h^r \widehat{J}^{(r)}(z_0) - h^r \,\mathrm{E}\, \widehat{J}^{(r)}(z_0)| > \epsilon\Big) \\
&\qquad + \ldots + P\Big(h^{-2\delta} |h^r \widehat{J}^{(r)}(z_{n^{1+q}}) - h^r \,\mathrm{E}\, \widehat{J}^{(r)}(z_{n^{1+q}})| > \epsilon\Big)\Big].
\end{aligned}$$

Now by Chebychev's Inequality and Lemma 5.6, we have that, for any $z \in I\!R$, and n large enough (say $n \geq M$)

$$P\Big(h^{-2\delta}|h^r \widehat{J}^{(r)}(z) - h^r \operatorname{E} \widehat{J}^{(r)}(z)| > \epsilon\Big) \leq \frac{\operatorname{E}[h^r \widehat{J}^{(r)}(z) - h^r \operatorname{E} \widehat{J}^{(r)}(z)]^{2p}}{\epsilon^{2p}/h^{-4\delta p}} \leq c_2 n^{-p} h^{-p-2\beta p} h^{-4\delta p},$$

where c_2 is independent of n. Thus we have, for $\ell \geq M$,

$$\begin{aligned}
&\sum_{n=1}^{\ell} P\Big(h^{-2\delta} \sup_{z \in E_n} |h^r \widehat{J}^{(r)}(z) - h^r \operatorname{E} \widehat{J}^{(r)}(z)| > \epsilon\Big) \\
&\qquad \leq (M-1) + \sum_{n=M}^{\ell} P\Big(h^{-2\delta} \sup_{z \in E_n} |h^r \widehat{J}^{(r)}(z) - h^r \operatorname{E} \widehat{J}^{(r)}(z)| > \epsilon\Big) \\
&\qquad \leq (M-1) + c_2 \sum_{n=M}^{\ell} (n^{1+q}+1) n^{-p} h^{-4\delta p - p - 2\beta p}.
\end{aligned}$$

We deduce that

$$\begin{aligned}
\lim_{\ell \to \infty} \sum_{n=1}^{\ell} P\Big(h^{-2\delta} S_{2,n} > \epsilon\Big) &\leq \lim_{\ell \to \infty} \sum_{n=1}^{\ell} P\Big(h^{-2\delta} \sup_{z \in E_n} |h^r \widehat{J}^{(r)}(z) - h^r \operatorname{E} \widehat{J}^{(r)}(z)| > \epsilon\Big) \\
&\leq (M-1) + c_2 \lim_{\ell \to \infty} \sum_{n=M}^{\ell} (n^{1+q}+1) n^{-p} h^{-4\delta p - p - 2\beta p} \\
&= (M-1) + c_2 \sum_{n=M}^{\infty} n^{1+q} n^{-p} h^{-4\delta p - p - 2\beta p} \\
&\quad + c_2 \sum_{n=M}^{\infty} n^{-p} h^{-4\delta p - p - 2\beta p} \\
&< \infty,
\end{aligned}$$

by Condition (A5). We conclude that

$$\sum_{n=1}^{\infty} P\Big(h^{-2\delta} S_{2,n} > \epsilon\Big) = \lim_{\ell \to \infty} \sum_{n=1}^{\ell} P\Big(h^{-2\delta} S_{2,n} > \epsilon\Big) < \infty. \tag{5.16}$$

For the third term in (5.15), we have

$$\begin{aligned}
|h^r \operatorname{E} \widehat{J}^{(r)}(z(x)) - h^r \operatorname{E} \widehat{J}^{(r)}(x)| &= \Big| \int K^{(r+1)}(u)[f_X(z(x)-hu) - f_X(x-hu)]\, du \Big| \\
&\leq \Big| \int K^{(r+1)}(u)[r_X(z(x)-hu) - r_X(x-hu)]\, du \Big| \\
&\quad + \Big| \int K^{(r+1)}(u)[dI_{[\tau,+\infty[}(z(x)-hu) - dI_{[\tau,+\infty[}(x-hu)]\, du \Big|.
\end{aligned} \tag{5.17}$$

By Lipschitz continuity of r_X we know that

$$|r_X(z(x) - hu) - r_X(x - hu)| \leq L|z(x) - hu - (x - hu)| = L|z(x) - x|$$

and thus

$$\begin{aligned}\Big|\int K^{(r+1)}(u)[r_X(z(x) - hu) - r_X(x - hu)]\, du\Big| & \\ &\leq \int |K^{(r+1)}(u)| \cdot |r_X(z(x) - hu) - r_X(x - hu)|\, du \\ &\leq L|z(x) - x| \int |K^{(r+1)}(u)|\, du \\ &\leq L \cdot (B - A)\, n^{-1-q} \int |K^{(r+1)}(u)|\, du.\end{aligned}$$

For the second integral on the right-hand side of inequality (5.17), note that $I_{[\tau,+\infty[}(z(x)-hu) = 1$ if $u \leq \frac{z(x)-\tau}{h}$ and 0 otherwise. Similarly $I_{[\tau,+\infty[}(x - hu) = 1$ if $u \leq \frac{x-\tau}{h}$ and 0 otherwise. Thus if $m(x)$ denotes $\min\left(\frac{z(x)-\tau}{h}, \frac{x-\tau}{h}\right)$ and $M(x)$ denotes $\max\left(\frac{z(x)-\tau}{h}, \frac{x-\tau}{h}\right)$, we can write $I_{[\tau,+\infty[}(z(x) - hu) - I_{[\tau,+\infty[}(x - hu) = \pm 1_{\{u\in]m(x),M(x)]\}}$, and thus

$$\begin{aligned}\Big|\int K^{(r+1)}(u)[dI_{[\tau,+\infty[}(z(x) - hu) - dI_{[\tau,+\infty[}(x - hu)]\, du\Big| & \\ &\leq |d| \cdot \|K^{(r+1)}\|_\infty \cdot \int |I_{[\tau,+\infty[}(z(x) - hu) - I_{[\tau,+\infty[}(x - hu)|\, du \\ &= |d| \cdot \|K^{(r+1)}\|_\infty \cdot (M(x) - m(x)).\end{aligned}$$

Under the conditions on K, and noting that $M(x) - m(x) = |z(x) - x|/h \leq (B - A)\, n^{-1-q}/h$, we conclude that, for all $x \in [A, B]$,

$$\begin{aligned}|h^r \, \mathrm{E}\, \widehat{J}^{(r)}(z(x)) - h^r \, \mathrm{E}\, \widehat{J}^{(r)}(x)| \leq& L \cdot (B - A)\, n^{-1-q} \int |K^{(r+1)}(u)|\, du \\ &+ |d| \cdot (B - A) \cdot \|K^{(r+1)}\|_\infty n^{-1-q} h^{-1} \\ \leq& c_3 n^{-1-q} h^{-1},\end{aligned}$$

with c_3 a positive constant independent of ω and n. Finally we obtain

$$h^{-2\delta} S_{3,n} = h^{-2\delta} \sup_{x\in[A,B]} |h^r \, \mathrm{E}\, \widehat{J}^{(r)}(z(x)) - h^r \, \mathrm{E}\, \widehat{J}^{(r)}(x)| \leq c_3 n^{-1-q} h^{-1-2\delta}.$$

Let ϵ be any positive real number. Since we have shown that, for all $\omega \in \Omega$,

$$h^{-2\delta} S_{1,n} + h^{-2\delta} S_{3,n} \leq c_1\, (B - A) h^{-2-2\delta-\beta} n^{-1-q} + c_3 n^{-1-q} h^{-1-2\delta},$$

which, under Condition (A5), tends to zero as $n \to \infty$, we have, for n large enough (say $n \geq M$)

$$h^{-2\delta} S_{1,n} + h^{-2\delta} S_{3,n} \leq \epsilon/2.$$

Thus for $n \geq M$ we can write

$$\begin{aligned} P\Big(h^{-2\delta}S_{1,n} + h^{-2\delta}S_{2,n} + h^{-2\delta}S_{3,n} > \epsilon\Big) =& P\Big(h^{-2\delta}S_{2,n} > \epsilon - h^{-2\delta}S_{1,n} - h^{-2\delta}S_{3,n}\Big) \\ \leq& P\Big(h^{-2\delta}S_{2,n} > \epsilon/2\Big), \end{aligned}$$

and therefore

$$\begin{aligned} \sum_{n=1}^{\infty} P\Big(h^{-2\delta}S_{1,n} + h^{-2\delta}S_{2,n} + h^{-2\delta}S_{3,n} > \epsilon\Big) =& \sum_{n=1}^{M-1} P\Big(h^{-2\delta}S_{1,n} + h^{-2\delta}S_{2,n} + h^{-2\delta}S_{3,n} > \epsilon\Big) \\ &+ \sum_{n=M}^{\infty} P\Big(h^{-2\delta}S_{1,n} + h^{-2\delta}S_{2,n} + h^{-2\delta}S_{3,n} > \epsilon\Big) \\ \leq& (M-1) + \sum_{n=M}^{\infty} P\Big(h^{-2\delta}S_{2,n} > \epsilon/2\Big) \\ <& \infty, \end{aligned}$$

by statement (5.16).

□

Proof of Lemma 5.8. We have

$$\begin{aligned} |h^2\widehat{J}''(\xi) - dK''(0)| \leq& |h^2\widehat{J}''(\xi) - h^2\,\mathrm{E}\,\widehat{J}''(\tau)| + |h^2\,\mathrm{E}\,\widehat{J}''(\tau) - dK''(0)| \\ \leq& \sup_{x\in[\tau\wedge\widehat{\tau},\tau\vee\widehat{\tau}]} |h^2\widehat{J}''(x) - h^2\,\mathrm{E}\,\widehat{J}''(\tau)| + |h^2\,\mathrm{E}\,\widehat{J}''(\tau) - dK''(0)| \\ \leq& \sup_{x\in[\tau\wedge\widehat{\tau},\tau\vee\widehat{\tau}]} |h^2\widehat{J}''(x) - h^2\,\mathrm{E}\,\widehat{J}''(x)| + \sup_{x\in[\tau\wedge\widehat{\tau},\tau\vee\widehat{\tau}]} |h^2\,\mathrm{E}\,\widehat{J}''(x) - h^2\,\mathrm{E}\,\widehat{J}''(\tau)| \\ &+ |\,\mathrm{E}\,h^2\widehat{J}''(\tau) - dK''(0)| \\ \leq& \sup_{x\in[A,B]} |h^2\widehat{J}''(x) - h^2\,\mathrm{E}\,\widehat{J}''(x)| + \sup_{x\in[\tau\wedge\widehat{\tau},\tau\vee\widehat{\tau}]} |h^2\,\mathrm{E}\,\widehat{J}''(x) - h^2\,\mathrm{E}\,\widehat{J}''(\tau)| \\ &+ |\,\mathrm{E}\,h^2\widehat{J}''(\tau) - dK''(0)| \\ \equiv& T_{1,n} + T_{2,n} + T_{3,n}. \end{aligned}$$

Now, by Lemma 5.7, we have that, for all $\epsilon > 0$,

$$\sum_{n=1}^{\infty} P\Big(h^{-2\delta} \sup_{x\in[A,B]} |h^2\widehat{J}''(x) - h^2\,\mathrm{E}\,\widehat{J}''(x)| > \epsilon\Big) < \infty,$$

i.e.

$$\forall \epsilon > 0: \quad P\Big(h^{-2\delta}T_{1,n} > \epsilon \text{ i.o.}\Big) = 0\ ,$$

where i.o. stands for infinitely often. This implies that $h^{-2\delta}T_{1,n} \xrightarrow{a.s.} 0$, and thus $T_{1,n} \xrightarrow{a.s.} 0$.

For the second term, by Lemma 5.3 for $r = 2$, we have, for all $x \in [A, B]$,

$$h^2\,\mathrm{E}\,\widehat{J}''(x) - h^2\,\mathrm{E}\,\widehat{J}''(\tau) = dK''\Big(\frac{x-\tau}{h}\Big) - dK''(0) + O(h),$$

where the remainder term is uniform in x and where, by Taylor, we can write

$$K''\Big(\frac{x-\tau}{h}\Big) = K''(0) + \frac{x-\tau}{h}K^{(3)}(\theta),$$

with θ between 0 and $(x-\tau)/h$. Thus for all $x \in [A, B]$, we have

$$h^2\,\mathrm{E}\,\widehat{J}''(x) - h^2\,\mathrm{E}\,\widehat{J}''(\tau) = d\frac{x-\tau}{h}K^{(3)}(\theta) + O(h).$$

In particular, since $[\tau \wedge \widehat{\tau}, \tau \vee \widehat{\tau}] \subset [A, B]$ we can write

$$\begin{aligned}
\sup_{x\in[\tau\wedge\widehat{\tau},\tau\vee\widehat{\tau}]} |h^2\,\mathrm{E}\,\widehat{J}''(x) - h^2\,\mathrm{E}\,\widehat{J}''(\tau)| &\leq h^{-1}|d|\cdot\|K^{(3)}\|_\infty \cdot \sup_{x\in[\tau\wedge\widehat{\tau},\tau\vee\widehat{\tau}]} |x-\tau| + O(h)\\
&= h^{-1}|d|\cdot\|K^{(3)}\|_\infty\cdot|\widehat{\tau}-\tau| + O(h)\\
&= O(h^\eta) + O(h),
\end{aligned}$$

where the last equality holds almost surely. This proves that $T_{2,n} \xrightarrow{a.s} 0$.

For the third term, we know from Lemma 5.3 that it tends to 0.

In conclusion, we have $T_{1,n} \xrightarrow{a.s.} 0$, $T_{2,n} \xrightarrow{a.s} 0$ and $T_{3,n} \to 0$. This implies that $T_{1,n} + T_{2,n} + T_{3,n} \xrightarrow{a.s.} 0$

□

Proof of Proposition 5.1. We have

$$\begin{aligned}
\sup_{x\in I_n} |\widehat{J}(x)| - |\widehat{J}(\tau)| &\leq \sup_{x\in I_n} |\widehat{J}(x) - \mathrm{E}\,\widehat{J}(x)| + \sup_{x\in I_n} |\,\mathrm{E}\,\widehat{J}(x)| - |\widehat{J}(\tau)|\\
&= \sup_{x\in I_n} |\widehat{J}(x) - \mathrm{E}\,\widehat{J}(x)| + \sup_{x\in I_n} |\,\mathrm{E}\,\widehat{J}(x)| - |\widehat{J}(\tau) - \mathrm{E}\,\widehat{J}(\tau) + \mathrm{E}\,\widehat{J}(\tau)|\\
&\leq \sup_{x\in I_n} |\widehat{J}(x) - \mathrm{E}\,\widehat{J}(x)| + \sup_{x\in I_n} |\,\mathrm{E}\,\widehat{J}(x)| + |\widehat{J}(\tau) - \mathrm{E}\,\widehat{J}(\tau)| - |\,\mathrm{E}\,\widehat{J}(\tau)|.
\end{aligned}$$

Now, by Lemma 5.10 ,we have, for n large enough,

$$\sup_{x\in I_n} |\,\mathrm{E}\,\widehat{J}(x)| - |\,\mathrm{E}\,\widehat{J}(\tau)| = \frac{|d|}{2}\,h^{2\delta}K''(0) + o(h^{2\delta}),$$

for $0 < \delta < 1/2$ and as in Condition (A5). Since $K''(0) < 0$, for n large enough, say $n \geq M$, we have

$$\sup_{x\in I_n} |\,\mathrm{E}\,\widehat{J}(x)| - |\,\mathrm{E}\,\widehat{J}(\tau)| \leq \frac{|d|}{4}\,h^{2\delta}K''(0).$$

We find

$$
\begin{aligned}
&\sum_{n=1}^{\infty} P\big(\sup_{x\in I_n}|\widehat{J}(x)|-|\widehat{J}(\tau)|\geq 0\big)\\
&\quad\leq\sum_{n=1}^{\infty} P\big(\sup_{x\in I_n}|\widehat{J}(x)-\mathrm{E}\,\widehat{J}(x)|+\sup_{x\in I_n}|\,\mathrm{E}\,\widehat{J}(x)|-|\,\mathrm{E}\,\widehat{J}(\tau)|+|\widehat{J}(\tau)-\mathrm{E}\,\widehat{J}(\tau)|\geq 0\big)\\
&\quad\leq M+\sum_{n=M}^{\infty} P\big(\sup_{x\in I_n}|\widehat{J}(x)-\mathrm{E}\,\widehat{J}(x)|+\sup_{x\in I_n}|\,\mathrm{E}\,\widehat{J}(x)|-|\,\mathrm{E}\,\widehat{J}(\tau)|+|\widehat{J}(\tau)-\mathrm{E}\,\widehat{J}(\tau)|\geq 0\big)\\
&\quad\leq M+\sum_{n=M}^{\infty} P\big(\sup_{x\in I_n}|\widehat{J}(x)-\mathrm{E}\,\widehat{J}(x)|+|\widehat{J}(\tau)-\mathrm{E}\,\widehat{J}(\tau)|\geq -\frac{|d|}{4}h^{2\delta}K''(0)\big)\\
&\quad= M+\sum_{n=M}^{\infty} P\big(h^{-2\delta}[\sup_{x\in I_n}|\widehat{J}(x)-\mathrm{E}\,\widehat{J}(x)|+|\widehat{J}(\tau)-\mathrm{E}\,\widehat{J}(\tau)|]\geq -\frac{|d|}{4}K''(0)\big)\\
&\quad\leq M+\sum_{n=M}^{\infty} P\big(h^{-2\delta}\sup_{x\in I_n}|\widehat{J}(x)-\mathrm{E}\,\widehat{J}(x)|\geq -\frac{|d|}{8}K''(0)\big)\\
&\qquad\qquad+\sum_{n=M}^{\infty} P\big(h^{-2\delta}|\widehat{J}(\tau)-\mathrm{E}\,\widehat{J}(\tau)|\geq -\frac{|d|}{8}K''(0)\big)\\
&\quad\leq M+\sum_{n=M}^{\infty} P\big(h^{-2\delta}\sup_{x\in I_n}|\widehat{J}(x)-\mathrm{E}\,\widehat{J}(x)|> -\frac{|d|}{9}K''(0)\big)\\
&\qquad\qquad+\sum_{n=M}^{\infty} P\big(h^{-2\delta}|\widehat{J}(\tau)-\mathrm{E}\,\widehat{J}(\tau)|> -\frac{|d|}{9}K''(0)\big)\\
&\quad<\infty,
\end{aligned}
$$

since $-|d|K''(0)>0$, and by Lemma 5.7. $\square$

Proof of Proposition 5.2. By definition, we have $\widehat{\tau}=\underset{x\in[A,B]}{\operatorname{argmax}}|\widehat{J}(x)|$. Hence

$$\forall x\in[A,B],\quad |\widehat{J}(\widehat{\tau})|\geq|\widehat{J}(x)|.$$

In particular, if $\widehat{\tau}\in I_n$, we have

$$|\widehat{J}(\widehat{\tau})|=\max_{x\in[A,B]}|\widehat{J}(x)|=\sup_{x\in I_n}|\widehat{J}(x)|\geq|\widehat{J}(\tau)|$$

In other words, we have

$$\big\{\omega:\widehat{\tau}\in I_n\big\}\subset\Big\{\omega:\sup_{x\in I_n}|\widehat{J}(x)|\geq|\widehat{J}(\tau)|\Big\},$$

or again

$$P\big(\widehat{\tau}\in I_n\big)\leq P\Big(\sup_{x\in I_n}|\widehat{J}(x)|\geq|\widehat{J}(\tau)|\Big).$$

From Proposition 5.1, we know that $\sum_{n=1}^{\infty} P(\sup_{x\in I_n} |\widehat{J}(x)| \geq |\widehat{J}(\tau)|) < \infty$. Hence we can write

$$
\begin{aligned}
\sum_{n=1}^{\infty} P(\sup_{x\in I_n} |\widehat{J}(x)| \geq |\widehat{J}(\tau)|) < \infty \Longrightarrow & \sum_{n=1}^{\infty} P(\widehat{\tau} \in I_n) < \infty \\
\Longleftrightarrow & \sum_{n=1}^{\infty} P(|\widehat{\tau} - \tau| > h^{1+\delta}) < \infty \\
\Longleftrightarrow & \sum_{n=1}^{\infty} P\Big(\frac{|\widehat{\tau} - \tau|}{h^{1+\delta}} > 1\Big) < \infty \\
\Longrightarrow & P\Big(\Big\{\frac{|\widehat{\tau} - \tau|}{h^{1+\delta}} > 1\Big\} \text{ i.o.}\Big) = 0 \\
\Longleftrightarrow & P\Big(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \Big\{\frac{|\widehat{\tau}_m - \tau|}{h_m^{1+\delta}} > 1\Big\}\Big) = 0 \\
\Longleftrightarrow & P\Big(\bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \Big\{\frac{|\widehat{\tau}_m - \tau|}{h_m^{1+\delta}} \leq 1\Big\}\Big) = 1,
\end{aligned}
$$

by using the Borel-Cantelli lemma. Now define $A = \Big\{w \ : \ \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \{|\widehat{\tau}_m - \tau|/h_m^{1+\delta} \leq 1\}\Big\}$. We have just proven that $P(A) = 1$. We can write

$$
(\forall \omega \in A)(\exists n_\omega \in I\!N)(\forall m \geq n_\omega): \quad \frac{|\widehat{\tau}_m - \tau|}{h_m^{1+\delta}} \leq 1.
$$

Hence $\forall \omega \in A \ : \ \lim_{n\to\infty} \frac{|\widehat{\tau}_n - \tau|}{h_n^{1+\delta}} \leq 1$, and thus $\widehat{\tau} - \tau = O(h^{1+\delta})$ almost surely.

$\square$

Proof of Lemma 5.9. We want to prove that

$$
\sup_{x\in I_n} |\operatorname{E} \widehat{J}(x)| = |d| K(h^\delta) + O(h).
$$

By Lemma 5.3 we know that $\operatorname{E} \widehat{J}(x) = dK\big(\frac{x-\tau}{h}\big) + O(h)$. Since we want to maximize the absolute value of this expectation, we want the sum to be as large as possible. The second term tends to zero as $n \to \infty$, but we will see that if $x \in I_n$ is well chosen, the first term does not tend to zero, and hence dominates this expression for n large enough. The aim is thus to try to maximize the first term.

Since K is maximum in zero, we want $K\big(\frac{x-\tau}{h}\big)$ to be as close as possible to $K(0)$. This is possible if we choose $x = \operatorname{argmin}_{y\in \bar{I}_n} |y - \tau| = \tau \pm h^{1+\delta}$, where $\bar{I}_n$ denotes the closure of I_n. As a matter of fact, since 0 is also the global maximiser of $|K|$, we have, for n large enough,

$$
\sup_{x\in I_n} \Big|K\Big(\frac{x-\tau}{h}\Big)\Big| = \sup_{x\in I_n} K\Big(\frac{x-\tau}{h}\Big) = \max_{x\in \bar{I}_n} K\Big(\frac{x-\tau}{h}\Big) = K(\pm h^\delta) = K(h^\delta),
$$

by the symmetry of K and where $K(h^\delta) \to K(0)$ as $n \to \infty$, since $h^\delta \to 0$ as $n \to \infty$ and K is continuous. We conclude that, for n large enough,

$$\begin{aligned}
\sup_{x \in I_n} |\operatorname{E} \widehat{J}(x)| - |d| K(h^\delta) =& \sup_{x \in I_n} |\operatorname{E} \widehat{J}(x)| - |d| \sup_{x \in I_n} |K\big(\frac{x-\tau}{h}\big)| \\
\leq& \sup_{x \in I_n} \big[|\operatorname{E} \widehat{J}(x)| - |d||K\big(\frac{x-\tau}{h}\big)|\big] \\
\leq& \sup_{x \in I_n} |\operatorname{E} \widehat{J}(x) - dK\big(\frac{x-\tau}{h}\big)| \\
=& O(h).
\end{aligned}$$

□

Proof of Lemma 5.10. We want to prove that

$$\sup_{x \in I_n} |\operatorname{E} \widehat{J}(x)| - |\operatorname{E} \widehat{J}(\tau)| = \frac{|d|}{2} h^{2\delta} K''(0) + o(h^{2\delta}) + O(h).$$

By the proof of Lemma 5.9, we know that, for n large enough,

$$\sup_{x \in I_n} |\operatorname{E} \widehat{J}(x)| = |d| K(h^\delta) + O(h). \tag{5.18}$$

From Lemma 5.3, we also know that we have

$$\operatorname{E} \widehat{J}(\tau) = d \cdot K(0) + O(h),$$

where, for n large enough, the sign of the whole expression is dictated by the sign of the first term, and thus

$$|\operatorname{E} \widehat{J}(\tau)| = |d| \cdot K(0) + O(h).$$

By a second order Taylor expansion of K around 0, we have

$$K(h^\delta) = K(0) + h^\delta K'(0) + \frac{h^{2\delta}}{2} K''(0) + o(h^{2\delta}) = K(0) + \frac{h^{2\delta}}{2} K''(0) + o(h^{2\delta}),$$

since $K'(0) = 0$ and $\|K^{(3)}\|_\infty < \infty$, and thus

$$\sup_{x \in I_n} |\operatorname{E} \widehat{J}(x)| - |\operatorname{E} \widehat{J}(\tau)| = \frac{|d|}{2} h^{2\delta} K''(0) + o(h^{2\delta}) + O(h)$$

which concludes the proof.

□

5.4 Proof of Proposition 4.3

Under the conditions of the proposition, we have

$$nh^{\frac{2+2\delta+\beta}{1+q}} \sim n \cdot n^{-\frac{2+2\delta+\beta}{\alpha(1+q)}} = n^{\frac{\alpha+\alpha q-2-2\delta-\beta}{\alpha(1+q)}},$$

where, under the conditions on q,

$$\alpha+\alpha q-2-2\delta-\beta > \alpha+\beta+4-\alpha-2-2\delta-\beta = 2-2\delta > 0,$$

since $\delta < 1$. This proves that $nh^{\frac{2+2\delta+\beta}{1+q}} \to \infty$.

We next check the second part of Condition (A5). We have $h \sim n^{-1/\alpha}$. This implies that $\exists\, 0 < c_1 \le c_2 < \infty$ such that $c_1 n^{-1/\alpha} \le h \le c_2 n^{-1/\alpha}$. We thus have

$$\sum_{n=1}^{\infty} n^{1+q-p} h^{-4\delta p-p-2\beta p} \le c \sum_{n=1}^{\infty} n^{1+q-p} n^{\frac{4\delta p+p+2\beta p}{\alpha}} = c \sum_{n=1}^{\infty} n^{\frac{\alpha+\alpha q-\alpha p+4\delta p+p+2\beta p}{\alpha}},$$

where $c = c_1^{-4\delta p-p-2\beta p}$ and

$$\begin{aligned}\alpha+\alpha q-\alpha p+4\delta p+p+2\beta p =& \alpha+\alpha q-p[\alpha-1-2\beta-4\delta]\\ <& \alpha+\alpha q-1-\alpha-\alpha q\\ =& -1,\end{aligned}$$

which proves that $\sum_{n=1}^{\infty} n^{1+q-p} h^{-4\delta p-p-2\beta p} < \infty$.

6 Proofs for the supersmooth error case

6.1 Intermediate results

We start by giving some results that will be necessary to prove the main results, Theorem 4.2 and Corollary 4.2. The proofs of these main results are given in Section 6.2. Section 6.3 contains the proofs of the intermediate results of this section. The lemmas below are an adaptation of Lemmas 5.5–5.8 of Section 5.1 to the supersmooth error case.

Similarly as for the ordinary smooth error case, we define conditions that depend on the mth derivative of K_Z, for different values of m.

Condition D:

($\mathbf{D}_1^m$) $\int |K_Z^{(m)}(u)|\, du = O\big(h^{\beta_3} \exp(h^{-\beta}/\gamma)\big)$, with β_3 a real constant;

($\mathbf{D}_2^m$) $\|K_Z^{(m)}\|_\infty = O\big(h^{\beta_0} \exp(h^{-\beta}/\gamma)\big)$;

($\mathbf{D}_3^m$) $R(K_Z^{(m)}) = \int \Big\{K_Z^{(m)}(u)\Big\}^2 du = O\big(h^{2\beta_0} \exp(2h^{-\beta}/\gamma)\big)$.

The next result has been established by Fan (1991a), but under conditions different from our (less restrictive) conditions. We provide a new proof of this lemma.

Lemma 6.1. *Let $r \geq 0$. Suppose that $h = d \cdot (2/\gamma)^{1/\beta}(\ln n)^{-1/\beta}$ with $d > 0$. Assume Conditions* (A2), (C_1^{r+1}), (C_2^{r+1}), (C_7^{r+1}), (D_1^{r}), (D_2^{r}). *If expression* (4.5) *is satisfied and* $\sup_{x\in I\!R\setminus\{\tau\}} |f_Y(x)| < \infty$, *then statement* (5.4) *in Lemma 5.5 holds.*

Lemma 6.2. *Let $r \geq 0$. Assume Conditions* (A2), (C_1^{r+1}), (C_2^{r+1}), (C_7^{r+1}), (D_2^{r+1}) *and* (D_3^{r+1}). *Suppose that* $\sup_{x\in I\!R\setminus\{\tau\}} |f_Y(x)| < \infty$ *and* $\int t^{2r+2-2\beta_0}|\varphi_K(t)|^2\, dt < \infty$. *Then, if $nh \to \infty$ as $n \to \infty$, we have, for all $p \in I\!N_0$ and for n large enough,*

$$\mathrm{E}[h^r \widehat{J}^{(r)}(x) - h^2\,\mathrm{E}\,\widehat{J}^{(r)}(x)]^{2p} \leq 2n^{-p}h^{-p+2\beta_0 p}\exp(2ph^{-\beta}/\gamma)\Big\{c_{\beta_0}(K) \sup_{x\in I\!R\setminus\{\tau\}} |f_Y(x)| \cdot \frac{8}{\pi d_{0,0}^2}\Big\}^p,$$

where $c_{\beta_0}(K) = \int |t|^{2r+2-2\beta_0}|\varphi_K(t)|^2\, dt$.

Lemma 6.3. *Let $r \geq 0$ and $h = d \cdot (2/\gamma)^{1/\beta}(\ln n)^{-1/\beta}$ with $d > 1$. Assume Conditions* (A2), (C_1^{r+1}), (C_2^{r+2}), (C_7^{r+1}), (D_2^{r+1}), (D_2^{r+2}) *and* (D_3^{r+1}). *Suppose that* $\|K^{(r+1)}\|_\infty < \infty$, $\sup_{x\in I\!R\setminus\{\tau\}} |f_Y(x)| < \infty$ *and* $\int t^{2r+2-2\beta_0}|\varphi_K(t)|^2\, dt < \infty$. *Then statement* (5.5) *in Lemma 5.7 holds for any $\delta > 0$.*

Lemma 6.4. *Suppose that $\widehat{\tau} = \tau + O(h^{1+\eta})$ a.s., with $\eta > 0$, and $h = d\cdot(2/\gamma)^{1/\beta}(\ln n)^{-1/\beta}$ with $d > 1$. Assume Conditions* (A2), (C_1^3), (C_2^4), (C_7^3), (D_2^3), (D_2^4), (D_3^3). *Suppose that* $\|K^{(3)}\|_\infty < \infty$, $\sup_{x\in I\!R\setminus\{\tau\}} |f_Y(x)| < \infty$ *and* $\int t^{6-2\beta_0}|\varphi_K(t)|^2\, dt < \infty$. *Then statement* (5.6) *in Lemma 5.8 holds.*

Propositions 6.2 and 6.1 correspond to Propositions 5.2 and 5.1 adapted to the supersmooth error case. Their proofs are similar to the ordinary smooth error case and hence are omitted.

Proposition 6.1. *Suppose that $h = d \cdot (2/\gamma)^{1/\beta}(\ln n)^{-1/\beta}$ with $d > 1$, and $0 < \delta < 1/2$. Then under Conditions* (B1) *to* (B4), *statement* (5.7) *of Proposition 5.1 holds.*

Proposition 6.2. *Suppose that $h = d \cdot (2/\gamma)^{1/\beta}(\ln n)^{-1/\beta}$ with $d > 1$, and $0 < \delta < 1/2$. Then under Conditions* (B1) *to* (B4), *statement* (5.8) *of Proposition 5.2 holds.*

6.2 Proof of the asymptotic law and the corollary

Proof of Theorem 4.2. As in the proof of Theorem 4.1, we have

$$\tau - \widehat{\tau} = \frac{\widehat{J}'(\tau)}{\widehat{J}''(\xi)},$$

where ξ lies between τ and $\widehat{\tau}$. From Lemma 6.1 for $r = 2$, we know that

$$\frac{h\widehat{J}'(\tau) - \mathrm{E}\, h\widehat{J}'(\tau)}{\sqrt{\mathrm{Var}(h\widehat{J}'(\tau))}} = \frac{\frac{1}{nh}\sum\limits_{i=1}^{n} K_Z''\big(\frac{\tau-Y_i}{h}\big) - \mathrm{E}\left[\frac{1}{nh}\sum\limits_{i=1}^{n} K_Z''\big(\frac{\tau-Y_i}{h}\big)\right]}{\sqrt{\mathrm{Var}\left[\frac{1}{nh}\sum\limits_{i=1}^{n} K_Z''\big(\frac{\tau-Y_i}{h}\big)\right]}} \xrightarrow{\mathrm{L}} N(0;1),$$

and thus by Lemma 6.4,

$$\frac{h\widehat{J}'(\tau) - \mathrm{E}[h\widehat{J}'(\tau)]}{h^2\widehat{J}''(\xi)\sqrt{\mathrm{Var}(h\widehat{J}'(\tau))}} \xrightarrow{\mathrm{L}} N\Big(0; \frac{1}{d^2\{K''(0)\}^2}\Big)$$

i.e.

$$\frac{\widehat{J}'(\tau)}{h\widehat{J}''(\xi)\sqrt{\mathrm{Var}(h\widehat{J}'(\tau))}} - \frac{\mathrm{E}[h\widehat{J}'(\tau)]}{h^2\widehat{J}''(\xi)\sqrt{\mathrm{Var}(h\widehat{J}'(\tau))}} \xrightarrow{\mathrm{L}} N\Big(0; \frac{1}{d^2\{K''(0)\}^2}\Big)$$

or again

$$\frac{\tau - \widehat{\tau}}{h\sqrt{\mathrm{Var}(h\widehat{J}'(\tau))}} - \frac{\mathrm{E}[h\widehat{J}'(\tau)]}{dK''(0)\sqrt{\mathrm{Var}(h\widehat{J}'(\tau))}} \xrightarrow{\mathrm{L}} N\Big(0; \frac{1}{d^2\{K''(0)\}^2}\Big),$$

where in Case (i), we have $\mathrm{E}[h\widehat{J}'(\tau)] = h^{l+1+\rho}D_{\tau,1} + O(h^{l+\rho+2})$, and in Case (ii), we have $\mathrm{E}[h\widehat{J}'(\tau)] = h^{k+2}D_{\tau,2} + O(h^{k+3})$.

We may further develop the variance term if we note that, as in the proof of Lemma 5.4, we have

$$\begin{aligned}
\mathrm{Var}(h\widehat{J}'(\tau)) =& \frac{1}{nh^2}\,\mathrm{Var}\left[K_Z''\Big(\frac{\tau - Y_1}{h}\Big)\right] \\
=& \frac{1}{nh}\int \big\{K_Z''(u)\big\}^2 f_Y(\tau - hu)\,du + O(n^{-1}) \\
=& \frac{1}{nh}\int_{-\infty}^{0} \big\{K_Z''(u)\big\}^2 [f_Y(\tau^+) - huf_Y'(\tau - h\theta_1 u)]\,du \\
&+ \frac{1}{nh}\int_{0}^{+\infty} \big\{K_Z''(u)\big\}^2 [f_Y(\tau^-) - huf_Y'(\tau - h\theta_2 u)]\,du + O(n^{-1}),
\end{aligned}$$

with $0 < \theta_1, \theta_2 < 1$. Under Condition (B3) and if $\int |u| \cdot |K_Z''(u)|^2\,du = O(R(K_Z''))$, we can then write

$$\Big|\int_{\mathcal{I}} u\big\{K_Z''(u)\big\}^2 f_Y'(\tau - h\theta_j u)]\,du\Big| \le \sup_{x \in I\!R\setminus\{\tau\}} |f_Y'(x)| \int_{\mathcal{I}} |u| \cdot |K_Z''(u)|^2\,du = O(R(K_Z'')),$$

where $\mathcal{I} =]-\infty, 0[$ and $j = 1$ or $\mathcal{I} =]0, +\infty[$ and $j = 2$. We conclude that

$$\begin{aligned}
\operatorname{Var}(h\widehat{J}'(\tau)) =& f_Y(\tau^+)\frac{1}{nh}\int_{-\infty}^{0}\left\{K_Z''(u)\right\}^2 du + f_Y(\tau^-)\frac{1}{nh}\int_{0}^{+\infty}\left\{K_Z''(u)\right\}^2 du + O(n^{-1}R(K_Z''))\\
=& \frac{B_\tau}{nh} R\big(K_Z''\big) + O(n^{-1}R(K_Z'')),
\end{aligned}$$

since K_Z is symmetric.

$\square$

Proof of Corollary 4.2. From

$$\frac{\tau - \widehat{\tau}}{h\sqrt{\operatorname{Var}(h\widehat{J}'(\tau))}} - \frac{\mathrm{E}[h\widehat{J}'(\tau)]}{dK''(0)\sqrt{\operatorname{Var}(h\widehat{J}'(\tau))}} \xrightarrow{\mathrm{L}} N\Big(0; \frac{1}{d^2\{K''(0)\}^2}\Big),$$

an approximate expression for the variance ($\mathcal{A}$Var) is given by

$$\mathcal{A}\operatorname{Var}\Big[\frac{\tau - \widehat{\tau}}{h\sqrt{\operatorname{Var}(h\widehat{J}'(\tau))}}\Big] = \frac{1}{d^2\{K''(0)\}^2},$$

i.e.

$$\mathcal{A}\operatorname{Var}\Big[\widehat{\tau}\Big] = \frac{h^2\operatorname{Var}(h\widehat{J}'(\tau))}{d^2\{K''(0)\}^2}.$$

An approximate expression for the expectation ($\mathcal{A}$E) may be derived as follows

$$\mathcal{A}\mathrm{E}\Big[\frac{\tau - \widehat{\tau}}{h} - \frac{\mathrm{E}[h\widehat{J}'(\tau)]}{dK''(0)}\Big] = o\Big(\sqrt{\operatorname{Var}(h\widehat{J}'(\tau))}\Big).$$

Hence, we obtain

$$\mathcal{A}\mathrm{E}[\widehat{\tau}] = \tau - \frac{h\,\mathrm{E}[h\widehat{J}'(\tau)]}{dK''(0)} + o\Big(h\sqrt{\operatorname{Var}(h\widehat{J}'(\tau))}\Big).$$

We deduce that

$$\mathcal{A}\mathrm{E}[\widehat{\tau}] = \tau - \frac{h^{k_2+2}D_{\tau,j}}{dK''(0)} + o\Big(h\sqrt{\operatorname{Var}(h\widehat{J}'(\tau))}\Big).$$

Finally an approximate mean squared error may be written as

$$\mathcal{A}\mathrm{MSE}[\widehat{\tau}] = \frac{h^{2k_2+4}D_{\tau,j}^2}{d^2\{K''(0)\}^2} + \frac{h^2\operatorname{Var}(h\widehat{J}'(\tau))}{d^2\{K''(0)\}^2}.$$

If we further assume that $\int |u|\cdot|K_Z''(u)|^2\,du = O(R(K_Z''))$, we have $\operatorname{Var}(h\widehat{J}'(\tau)) = \frac{B_\tau}{nh} R\big(K_Z''\big) + O(n^{-1}R(K_Z''))$ and we find

$$\mathcal{A}\mathrm{MSE}[\widehat{\tau}] = \frac{h^{2k_2+4}D_{\tau,j}^2}{d^2\{K''(0)\}^2} + \frac{hR\big(K_Z''\big)B_\tau}{nd^2\{K''(0)\}^2}.$$

$\square$

6.3 Proofs of the intermediate results

In this section we prove the intermediate results of Section 6.1.

Proof of Lemma 6.1. Denote $h^{-1}K_Z^{(r)}\big(\frac{\tau-Y_i}{h}\big)$ by $Y_{n,i}, i = 1, \ldots, n$, where, for all $i \neq j$, $Y_{n,i} \perp\!\!\!\perp Y_{n,j}$. We will use the central limit theorem for triangular arrays of random variables. As in the ordinary smooth error case, we need to verify the Lyapounov condition (5.12). We will start by investigating separately the behaviour of the numerator and denominator in (5.12).

Numerator:

We have $\mathrm{E}\,|Y_{n,1} - \mathrm{E}(Y_{n,1})|^{2+\eta} \leq \big(\big\{\mathrm{E}\,|Y_{n,1}|^{2+\eta}\big\}^{\frac{1}{2+\eta}} + \mathrm{E}\,|Y_{n,1}|\big)^{2+\eta}$. Let $\beta_4 = \min(\beta_0, \beta_3)$. Under the conditions of the lemma, we have

$$\mathrm{E}\,|Y_{n,1}| \leq \sup_{x \in I\!R\backslash\{\tau\}} |f_Y(x)| \int |K_Z^{(r)}(u)|\,du = O\Big(h^{\beta_4}\exp(h^{-\beta}/\gamma)\Big)$$

and

$$h^{1+\eta}\,\mathrm{E}\,|Y_{n,1}|^{2+\eta} \leq \|K_Z^{(r)}\|_\infty^{1+\eta} \cdot \sup_{x \in I\!R\backslash\{\tau\}} |f_Y(x)| \int |K_Z^{(r)}(u)|\,du = O\Big(h^{(2+\eta)\beta_4}\exp((2+\eta)h^{-\beta}/\gamma)\Big).$$

We conclude that

$$\begin{aligned} h^{1+\eta}\,\mathrm{E}\,|Y_{n,1} - \mathrm{E}(Y_{n,1})|^{2+\eta} =& \Big[O\Big(h^{\beta_4}\exp(h^{-\beta}/\gamma)\Big) + h^{\frac{1+\eta}{2+\eta}}O\Big(h^{\beta_4}\exp(h^{-\beta}/\gamma)\Big)\Big]^{2+\eta} \\ =& O\Big(h^{\beta_4(2+\eta)}\exp((2+\eta)h^{-\beta}/\gamma)\Big). \end{aligned} \tag{6.1}$$

Denominator:

From expression (4.5), we have, as in the proof of Lemma 5.4 (using Lemma 5.3)

$$\begin{aligned} \mathrm{Var}(Y_{n,1}) =& h^{-1}\int \Big\{K_Z^{(r)}(u)\Big\}^2 f_Y(\tau - hu)\,du - h^{-2}\Big\{\mathrm{E}\Big[K_Z^{(r)}\Big(\frac{\tau - Y_1}{h}\Big)\Big]\Big\}^2 \\ =& h^{-1}\int \Big\{K_Z^{(r)}(u)\Big\}^2 f_Y(\tau - hu)\,du - \Big\{dK^{(r-1)}(0) + O(h)\Big\}^2 \\ \geq& c_5\, f_Y(\tau^-)h^{c_6}\exp(\frac{2}{\gamma h^\beta} - \frac{4\beta b}{\gamma h^\beta}) - \Big\{dK^{(r-1)}(0) + O(h)\Big\}^2 \\ \geq& \frac{c_5}{2}\, f_Y(\tau^-)h^{c_6}\exp(\frac{2}{\gamma h^\beta} - \frac{4\beta b}{\gamma h^\beta}), \end{aligned}$$

for n large enough, where $b = h^{\beta/(2r+10)}$ and c_5 is a positive constant independent of n and c_6 is a constant independent of n.

We now are ready to verify Condition (5.12).

$$\lim_{n\to\infty} \frac{n\,\mathrm{E}\,|Y_{n,1} - \mathrm{E}(Y_{n,1})|^{2+\eta}}{(n\,\mathrm{Var}(Y_{n,1}))^{(2+\eta)/2}}$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \frac{1}{(nh)^{\eta/2}} \frac{h^{1+\eta}\,\mathrm{E}\,|Y_{n,1}-\mathrm{E}(Y_{n,1})|^{2+\eta}}{(h\,\mathrm{Var}(Y_{n,1}))^{(2+\eta)/2}} \\
&\leq \Big[\frac{2}{c_5 f_Y(\tau^-)}\Big]^{(2+\eta)/2} \lim_{n\to\infty} \frac{1}{(nh)^{\eta/2}} \frac{h^{1+\eta}\,\mathrm{E}\,|Y_{n,1}-\mathrm{E}(Y_{n,1})|^{2+\eta}}{\big[h^{c_6+1}\exp(\frac{2}{\gamma h^\beta}-\frac{4\beta b}{\gamma h^\beta})\big]^{(2+\eta)/2}} \\
&= c\cdot \lim_{n\to\infty}\left\{ \frac{1}{(nh)^{\eta/2}}\cdot \frac{h^{\beta_4(2+\eta)}\exp((2+\eta)h^{-\beta}/\gamma)}{\big[h^{c_6+1}\exp(\frac{2}{\gamma h^\beta}-\frac{4\beta b}{\gamma h^\beta})\big]^{(2+\eta)/2}}\cdot \frac{h^{1+\eta}\,\mathrm{E}\,|Y_{n,1}-\mathrm{E}(Y_{n,1})|^{2+\eta}}{h^{\beta_4(2+\eta)}\exp((2+\eta)h^{-\beta}/\gamma)}\right\} \\
&= c\cdot \lim_{n\to\infty}\left\{ n^{-\eta/2}h^{\beta_5}\exp\big(2(2+\eta)\beta b h^{-\beta}/\gamma\big)\cdot \frac{h^{1+\eta}\,\mathrm{E}\,|Y_{n,1}-\mathrm{E}(Y_{n,1})|^{2+\eta}}{h^{\beta_4(2+\eta)}\exp((2+\eta)h^{-\beta}/\gamma)}\right\},
\end{aligned}
$$

where $c = \big[\frac{2}{c_5 f_Y(\tau^-)}\big]^{(2+\eta)/2}$ and $\beta_5 = (2+\eta)\,\beta_4 - \frac{2+\eta}{2}\,c_6 - \eta - 1$. By the assumptions of the lemma, we have

$$
\exp\Big(2(2+\eta)\beta b h^{-\beta}/\gamma\Big) = \exp\Big((2+\eta)\beta b d^{-\beta}\ln n\Big) = \exp\Big(\ln n^{(2+\eta)\beta d^{-\beta} b}\Big) = n^{(2+\eta)\beta d^{-\beta} b},
$$

with $b = h^{\beta/(2r+10)}$ tending to zero as $n\to\infty$, and thus

$$
\exp\Big(2(2+\eta)\beta b h^{-\beta}/\gamma\Big)\cdot n^{-\eta/2}h^{\beta_5} \sim n^{-\eta/2+(2+\eta)\beta d^{-\beta} b}(\ln n)^{-\beta_5/\beta},
$$

which tends to 0 as $n\to\infty$ (even if $\beta_5 < 0$), since $b\to 0$ as $n\to\infty$ and $\eta > 0$. Hence, by (6.1), the Lyapounov condition (5.12) is satisfied. This concludes the proof. $\square$

Proof of Lemma 6.2. Proceeding as in the proof of Lemma 5.6, we need to investigate the behaviour of each term in (5.13) and (5.14). We first investigate the behaviour of all expectations appearing in (5.14). The term $\mathrm{E}\big[K_Z^{(r+1)}\big(\frac{x-Y_1}{h}\big)\big]$ is the same as in the proof of Lemma 5.6. For all $j\geq 2$, we find

$$
\mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_i}{h}\Big)\Big]^j = h\int_{-\infty}^{\frac{x-\tau}{h}} \{K_Z^{(r+1)}(u)\}^j f_Y(x-hu)\,du + h\int_{\frac{x-\tau}{h}}^{\infty} \{K_Z^{(r+1)}(u)\}^j f_Y(x-hu)\,du,
$$

where, under the conditions of the lemma,

$$
\begin{aligned}
h\Big|\int_{-\infty}^{\frac{x-\tau}{h}} \{K_Z^{(r+1)}(u)\}^j f_Y(x-hu)\,du\Big| &\leq h\,\|K_Z^{(r+1)}\|_\infty^{j-2} \sup_{x\in I\!R\setminus\{\tau\}} |f_Y(x)|\cdot \int |K_Z^{(r+1)}(u)|^2\,du \\
&= O\Big(h^{j\beta_0+1}\exp(jh^{-\beta}/\gamma)\Big),
\end{aligned}
$$

and the same bound holds for the second integral. From expression (5.14) we then get

$$
\begin{aligned}
&\mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_i}{h}\Big) - \mathrm{E}\,K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^l \\
&\qquad = O(h^l) + \sum_{j=2}^{l}(-1)^j C_l^j\,\mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_i}{h}\Big)\Big]^j\Big[\mathrm{E}\,K^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^{l-j}
\end{aligned}
$$

$$
\begin{aligned}
&= O(h^l) + \sum_{j=2}^{l} O\Big(h^{j\beta_0+1}\exp(jh^{-\beta}/\gamma)\Big)\cdot O(h^{l-j})\\
&= O\Big(h^{l\beta_0+1}\exp(lh^{-\beta}/\gamma)\Big).
\end{aligned}
$$

Finally we get

$$
\begin{aligned}
\mathrm{E}[h^r\widehat{J}^{(r)}(x) - h^r\,\mathrm{E}\,\widehat{J}^{(r)}(x)]^{2p} =& O\Big(\frac{nh^{2p\beta_0+1}\exp(2ph^{-\beta}/\gamma)}{n^{2p}h^{2p}}\Big) + O\Big(\frac{n^2h^{2p\beta_0+2}\exp(2ph^{-\beta}/\gamma)}{n^{2p}h^{2p}}\Big)\\
&+ \ldots + O\Big(\frac{n^p h^{2p\beta_0+p}\exp(2ph^{-\beta}/\gamma)}{n^{2p}h^{2p}}\Big)\\
=& O\Big(n^{-p}h^{-p+2\beta_0 p}\exp(2ph^{-\beta}/\gamma)\Big),
\end{aligned}
$$

since $nh \to \infty$ as $n \to \infty$.

From the above calculations we have, proceeding as in the proof of Lemma 5.6, and by Propositions 4.1 (i) and 4.2 (i),

$$
\begin{aligned}
&\mathrm{E}[h^r\widehat{J}^{(r)}(x) - h^r\,\mathrm{E}\,\widehat{J}^{(r)}(x)]^{2p}\\
&\quad= \frac{1}{n^p h^{2p}}\Big\{\mathrm{E}\Big[K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\Big]^2\Big\}^p\cdot(1+o(1))\\
&\quad\le \frac{1}{n^p h^{2p}}\Big\{2h\cdot\sup_{x\in I\!R\setminus\{\tau\}}|f_Y(x)|\cdot R(K_Z^{(r+1)})\Big\}^p\cdot(1+o(1))\\
&\quad\le \frac{1}{n^p h^{2p}}\Big\{2h\cdot\sup_{x\in I\!R\setminus\{\tau\}}|f_Y(x)|\cdot\frac{4}{\pi d_{0,0}^2}\;c_{\beta_0}(K)h^{2\beta_0}\exp(2h^{-\beta}/\gamma)\Big\}^p\cdot(1+o(1))\\
&\quad= n^{-p}h^{-p+2\beta_0 p}\exp(2ph^{-\beta}/\gamma)\cdot\Big\{c_{\beta_0}(K)\sup_{x\in I\!R\setminus\{\tau\}}|f_Y(x)|\cdot\frac{8}{\pi d_{0,0}^2}\Big\}^p\cdot(1+o(1)),
\end{aligned}
$$

where $c_{\beta_0}(K) = \int |t|^{2r+2-2\beta_0}|\varphi_K(t)|^2\,dt$. This implies that, for n large enough, we have

$$
\begin{aligned}
&\mathrm{E}[h^r\widehat{J}^{(r)}(x) - h^r\,\mathrm{E}\,\widehat{J}^{(r)}(x)]^{2p}\\
&\qquad\le 2n^{-p}h^{-p+2\beta_0 p}\exp(2ph^{-\beta}/\gamma)\cdot\Big\{c_{\beta_0}(K)\sup_{x\in I\!R\setminus\{\tau\}}|f_Y(x)|\cdot\frac{8}{\pi d_{0,0}^2}\Big\}^p.
\end{aligned}
$$

□

Proof of Lemma 6.3. Similarly as in the proof of Lemma 5.7, we have, for all $\omega \in \Omega$,

$$
\sup_{x\in[A,B]}|h^r\widehat{J}^{(r)}(x) - h^r\,\mathrm{E}\,\widehat{J}^{(r)}(x)| \le T_{1,n} + T_{2,n} + T_{3,n}, \tag{6.2}
$$

where to simplify the notations, we do not indicate the dependence of the random variables in ω. We treat these three terms separately.

For the first term, note that for all $\omega \in \Omega$ and for all $x \in [A, B]$ we have (by the mean-value theorem)

$$|h^r \widehat{J}^{(r)}(x) - h^r \widehat{J}^{(r)}(z(x))| \leq |h^r \widehat{J}^{(r+1)}(\xi)| \cdot |x - z(x)|,$$

where ξ lies between x and $z(x)$. Now we have

$$|h^r \widehat{J}^{(r+1)}(\xi)| = |\frac{1}{nh^2} \sum_{i=1}^{n} K_Z^{(r+2)}\Big(\frac{\xi - Y_i}{h}\Big)| \leq h^{-2} \|K_Z^{(r+2)}\|_\infty = O\Big(h^{\beta_0 - 2} \exp(h^{-\beta}/\gamma)\Big).$$

We conclude that

$$\sup_{x\in[A,B]} |h^r \widehat{J}^{(r)}(x) - h^r \widehat{J}^{(r)}(z(x))| \leq h^{-2} \|K_Z^{(r+2)}\|_\infty |x - z(x)| \leq c n^{-1-q} h^{\beta_0-2} \exp(h^{-\beta}/\gamma),$$

with c a positive constant independent of ω and n, and

$$\begin{aligned} h^{-2\delta} T_{1,n} =& h^{-2\delta} \sup_{x\in[A,B]} |h^r \widehat{J}^{(r)}(x) - h^r \widehat{J}^{(r)}(z(x))| \\ \leq& c \cdot n^{-1-q} h^{\beta_0-2-2\delta} \exp(h^{-\beta}/\gamma) \\ =& c \cdot n^{-1-q} h^{\beta_0-2-2\delta} n^{d^{-\beta}/2} \\ =& c \cdot (d \cdot (2/\gamma)^{1/\beta})^{\beta_0-2-2\delta} n^{(d^{-\beta}/2)-1-q} (\ln n)^{\frac{2+2\delta-\beta_0}{\beta}} \\ =& c_1 n^{(d^{-\beta}/2)-1-q} (\ln n)^{\frac{2+2\delta-\beta_0}{\beta}}, \end{aligned}$$

with c_1 a positive constant independent of ω and n.

For the second term in (6.2), we have, for all $\omega \in \Omega$,

$$\sup_{x\in[A,B]} |h^r \widehat{J}^{(r)}(z(x)) - h^r \operatorname{E} \widehat{J}^{(r)}(z(x))| \leq \sup_{z\in E_n} |h^r \widehat{J}^{(r)}(z) - h^r \operatorname{E} \widehat{J}^{(r)}(z)|.$$

Hence, similarly as in the proof of Lemma 5.7, for all $\epsilon > 0$ and $\ell \geq 1$, we can write

$$\begin{aligned} \sum_{n=1}^{\ell} P\Big(h^{-2\delta} T_{2,n} > \epsilon\Big) \leq& \sum_{n=1}^{\ell} \Big[P\Big(h^{-2\delta} |h^r \widehat{J}^{(r)}(z_0) - h^r \operatorname{E} \widehat{J}^{(r)}(z_0)| > \epsilon\Big) \\ & + \ldots + P\Big(h^{-2\delta} |h^r \widehat{J}^{(r)}(z_{n^{1+q}}) - h^r \operatorname{E} \widehat{J}^{(r)}(z_{n^{1+q}})| > \epsilon\Big)\Big]. \end{aligned}$$

Now by Chebychev's Inequality and Lemma 6.2, we have, for any $z \in I\!R$, and n large enough (say $n \geq M$)

$$\begin{aligned} P\Big(h^{-2\delta} |h^r \widehat{J}^{(r)}(z) - h^r \operatorname{E} \widehat{J}^{(r)}(z)| > \epsilon\Big) \leq& \frac{\operatorname{E}[|h^r \widehat{J}^{(r)}(z) - h^r \operatorname{E} \widehat{J}^{(r)}(z)|]^{2p}}{\epsilon^{2p}/h^{-4\delta p}} \\ \leq& c_2\, n^{-p} h^{-p-4\delta p+2\beta_0 p} \exp(2p h^{-\beta}/\gamma) \\ =& c_2' (\ln n)^{\frac{p+4\delta p-2\beta_0 p}{\beta}} n^{p d^{-\beta} - p}, \end{aligned}$$

where c_2 and c_2' are independent of ω and n. Thus we have, for $\ell \geq M$,

$$\begin{aligned}
\sum_{n=1}^{\ell} P\Big(h^{-2\delta} \sup_{z\in E_n} |h^r \widehat{J}^{(r)}(z) - h^r \operatorname{E} \widehat{J}^{(r)}(z)| > \epsilon\Big) \\
\leq &(M-1) + \sum_{n=M}^{\ell} P\Big(h^{-2\delta} \sup_{z\in E_n} |h^r \widehat{J}^{(r)}(z) - h^r \operatorname{E} \widehat{J}^{(r)}(z)| > \epsilon\Big) \\
\leq &(M-1) + c_2' \sum_{n=M}^{\ell} (n^{1+q}+1)(\ln n)^{\frac{p+4\delta p-2\beta_0 p}{\beta}} n^{pd^{-\beta}-p}.
\end{aligned}$$

We deduce that

$$\begin{aligned}
\lim_{\ell\to\infty} \sum_{n=1}^{\ell} P\Big(h^{-2\delta} T_{2,n} > \epsilon\Big) \leq &\lim_{\ell\to\infty} \sum_{n=1}^{\ell} P\Big(h^{-2\delta} \sup_{z\in E_n} |h^r \widehat{J}^{(r)}(z) - h^r \operatorname{E} \widehat{J}^{(r)}(z)| > \epsilon\Big) \\
\leq &(M-1) + c_2' \lim_{\ell\to\infty} \sum_{n=M}^{\ell} (n^{1+q}+1)(\ln n)^{\frac{p+4\delta p-2\beta_0 p}{\beta}} n^{pd^{-\beta}-p} \\
= &(M-1) + c_2' \sum_{n=M}^{\infty} (\ln n)^{\frac{p+4\delta p-2\beta_0 p}{\beta}} n^{pd^{-\beta}-p+1+q} \\
&+ c_2' \sum_{n=M}^{\infty} (\ln n)^{\frac{p+4\delta p-2\beta_0 p}{\beta}} n^{pd^{-\beta}-p} \\
< &\infty,
\end{aligned}$$

as soon as $pd^{-\beta} - p + 1 + q < -1$, which is equivalent to requiring that

$$p[1-d^{-\beta}] > 2+q \iff p > \frac{2+q}{1-d^{-\beta}},$$

since $d > 1$, and thus $d^{-\beta} < 1$.

For the third term in (6.2), we proceed as in the ordinary smooth error case and get

$$h^{-2\delta} T_{3,n} \leq c_3 n^{-1-q} h^{-1-2\delta},$$

with c_3 a positive constant.

Now, let ϵ be any positive real number. Since we have shown that, for all $\omega \in \Omega$,

$$h^{-2\delta} T_{1,n} + h^{-2\delta} T_{3,n} \leq c_1 n^{(d^{-\beta}/2)-1-q} (\ln n)^{\frac{2+2\delta-\beta_0}{\beta}} + c_3 n^{-1-q} h^{-1-2\delta},$$

which tends to zero as $n \to \infty$, we have, for n large enough,

$$h^{-2\delta} T_{1,n} + h^{-2\delta} T_{3,n} \leq \epsilon/2,$$

and we conclude as in the proof of Lemma 5.7.

□

Proof of Lemma 6.4. This proof is similar to the proof of Lemma 5.8, where we replace Lemma 5.7 by Lemma 6.3. □

References

Barry, J. and Diggle, P. (1995). Choosing the smoothing parameter in a Fourier approach to nonparametric deconvolution of a density function. *Journal of Nonparametric Statistics*, **4**, 223–232.

Carroll, R.J. and Hall, P. (1988). Optimal rates of convergence for deconvolving a density. *Journal of the American Statistical Association*, **83**, 1184–1186.

Chu, C.K. and Cheng, P.E. (1996). Estimation of jumps points and jump values of a density function. *Statistica Sinica*, **6**, 79–95.

Couallier, V. (1999). Estimation non paramétrique d'une discontinuité dans une densité. *Comptes Rendus de l'Académie des Sciences, I*, **t.329**, 633–636.

Couallier, V. (2000). Inférence statistique pour des estimateurs de discontinuités dans un cadre non paramétrique. *PhD. thesis*, Laboratoire de Statistique et Probabilités, Université Paul Sabatier, Toulouse III, France.

Delaigle, A. (2003). Kernel estimation in deconvolution problems. *PhD dissertation*. Institut de Statistique, Université catholique de Louvain, Belgium.

Delaigle, A. and Gijbels, I. (2002). Estimation of integrated squared density derivatives from a contaminated sample. *Journal of the Royal Statistical Society, Series B*, **64**, 869–886.

Delaigle, A. and Gijbels, I. (2003a). Bootstrap bandwidth selection in kernel density estimation from a contaminated sample. *The Annals of the Institute of Statistical Mathematics*, to appear.

Delaigle, A. and Gijbels, I. (2003b). Practical implementation of deconvolution kernel estimation of boundaries and discontinuity points. *Manuscript*.

Fan, J. (1991a). Asymptotic normality for deconvolution kernel density estimators. *Sankhya A*, **53**, 97–110.

Fan, J. (1991b). Global behaviour of deconvolution kernel estimates. *Statistica Sinica*, **1**, 541–551.

Fan, J. (1991c). On the optimal rates of convergence for nonparametric deconvolution problems. *The Annals of Statistics*, **19**, 1257–1272.

Gijbels, I. and Goderniaux, A.-C. (2001). Bandwidth selection for change point estimation in nonparametric regression. Institut de Statistique, Université catholique de Louvain, *Discussion Paper # 0024. Under revision.*

Gijbels, I., Hall, P. and Kneip, A. (1999). On the estimation of jump points in smooth curves. *Annals of the Institute of Statistical Mathematics*, **51**, 231–251.

Goderniaux, A-C. (2001). Automatic detection of change-points in nonparametric regression. *PhD. thesis*, Institut de Statistique, Université catholique de Louvain, Belgium.

Hall, P. and Simar, L. (2002). Estimating a changepoint, boundary or frontier in the presence of observation error. *Journal of the American Statistical Association*, **97**, 523–534.

Kneip, A. and Simar, L. (1996). A general framework for frontier estimation with panel data. *Journal of Productivity Analysis*, **7**, 187–212.

Li, T. and Vuong, Q. (1998). Nonparametric estimation of the measurement error model using multiple indicators. *Journal of Multivariate Analysis*, **65**, 139–165.

Neumann, M.H. (1997a). On the effect of estimating the error density in nonparametric deconvolution. *Journal of Nonparametric Statistics*, **7**, 307–330.

Neumann, M.H. (1997b). Optimal change-point estimation in inverse problems. *Scandinavian Journal of Statistics*, **24**, 503–521.

Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

Stefanski, L. and Carroll, R.J. (1990). Deconvoluting kernel density estimators. *Statistics*, **2**, 169–184.

Wu, J.S. and Chu, C.K. (1993). Kernel-type estimators of jump points and jump values of a regression function. *The Annals of Statistics*, **21**, 1545–1566.

Zhang, S. and Karunamuni, R. (2000). Boundary bias correction for nonparametric deconvolution. *Annals of the Institute of Statistical Mathematics*, **52**, 612–629.