Estimation of boundary and discontinuity points in deconvolution problems

A. Delaigle[1,*] and I. Gijbels[2,**]

[1]Department of Mathematics, University of California, San Diego, CA 92122 USA

[2]Universitair Centrum voor Statistiek, Katholieke Universiteit Leuven, Belgium.

Abstract

We consider estimation of the boundary of the support of a density function when only a contaminated sample from the density is available. This estimation problem is a necessary step when estimating a density with unknown 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. 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 limiting distribution of the proposed estimator as well as approximate expressions for its mean squared error, for various types of error densities, and deduce rates of convergence of the estimator. The finite sample performance of the procedure is investigated via simulation.

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

* Supported in part by a Belgian American and Educational Foundation Post-Doctoral Fellowship.

** 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 Government (Belgian Science Policy) is also gratefully acknowledged.

1 Introduction

In this paper, we consider kernel estimation of boundary points and discontinuity points of a density from a contaminated sample of that density, i.e. from a sample that contains measurement errors. The contamination problem, often referred to as a deconvolution problem, has applications in many different fields such as chemistry or public health. In this context a so-called deconvolution kernel estimator of the density has been proposed in the literature. See, for example, Carroll and Hall (1988) and Stefanski and Carroll (1990). This deconvolution kernel density estimator, however, is not consistent at a discontinuity point or at a finite left/right endpoint of the density to estimate, and has to be modified by taking these points into account. See for example Zhang and Karunamuni (2000) for the modifications to apply in the case of boundary points. It is necessary to provide good estimators of these boundary points or, more generally, discontinuity points when they are unknown.

Boundary estimation also arises when investigating efficiencies of firms like banks, or public services. These investigations involve estimation of quantities such as the maximum level of output that can be produced for a given level of input, which is often referred to as an economic frontier estimation problem, but can be seen as a problem of estimation of the boundary of a density. Many different methods have been proposed to estimate a frontier in the case where the observations do not contain any measurement error, but these methods do generally not provide consistent estimators of the more realistic stochastic frontiers, i.e. of frontiers (or boundary points) to be estimated from data that are contaminated by noise.

Boundary estimation from contaminated data has been studied by Kneip and Simar (1996), Neumann (1997b) or Hall and Simar (2002), for example. These papers, however, either focus on very specific contexts, or propose methods that are difficult to implement in practice. Our goal is to provide a method that works in more general contexts, and to provide a way to implement the method in practice. The idea 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 the proposed 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.

This paper is organized as follows. In Section 2 we present the problem of boundary point estimation and introduce the estimation procedure. In Section 3 we establish the asymptotic distribution for the estimator and deduce approximate expressions for its bias and variance. In Section 4, the finite sample performance of the procedure is illustrated on simulated examples. The proofs of the results are given in Section 5.

2 The estimation method

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 $Y_i = X_i + Z_i, i = 1, 2, \ldots, n$, 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

$$\widehat{f}_X(x; h) = \frac{1}{nh} \sum_{j=1}^{n} K_Z\Big(\frac{x - Y_j}{h}\Big), \tag{2.1}$$

where $K_Z(u) = (2\pi)^{-1} \int e^{-itu} \varphi_K(t)/\varphi_Z(t/h)\, dt$, with φ_L the Fourier transform (resp. characteristic function) of a function (resp. random variable) L. See Carroll and Hall (1988) and Stefanski and Carroll (1990) for an introduction to this estimator. Throughout this paper, we assume that f_Z is real and symmetric and that, for all $t \in I\!R$, $\varphi_Z(t) \neq 0$. In order to guarantee that the integral in (2.1) exists, we choose the kernel K to be a real, continuous and symmetric function, such that φ_K has a compact support $[-B_K, B_K]$, with $0 < B_K < \infty$. Note that, under our assumptions, K_Z is real and symmetric and $\|\varphi_K\|_\infty < \infty$.

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, which is of particular interest when estimating an economic frontier or as a first step to kernel estimation of f_X. 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. Hence, we need a more elaborated procedure.

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 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 Müller (1992), Wu and Chu (1993), Gijbels, Hall and Kneip (1999), Goderniaux (2001) and Gijbels and Goderniaux (2004), among others, for similar methods in the regression context.

We propose a diagnostic function based on derivative estimation. For a density f_X with a single boundary point τ, we define the estimator of τ by

$$\widehat{\tau} = \text{argmax}_x \, |\widehat{J}(x)|, \tag{2.2}$$

where the diagnostic function $\widehat{J}(x)$ is proportional to the derivative of the deconvolution kernel density estimator of f_X: $\widehat{J}(x) = \frac{1}{nh} \sum_{i=1}^{n} K_Z'\left(\frac{x-Y_i}{h}\right)$. Unlike f_X, the kernel estimate $\widehat{f}_X$ is a smooth function, even in τ. 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). In the next section, we prove that the method is consistent, and provide asymptotic distribution of the estimator.

3 Asymptotic distribution of the estimator

Consider a density f_X with a single boundary or discontinuity point τ. In this section we show that the estimator for τ introduced in the previous section is a consistent estimator, and establish its asymptotic law. The basic ideas of proof use techniques and conditions somewhat similar to those used for proving the consistency of a discontinuity point estimator in the error free case. See for example Müller (1992), Chu and Cheng (1996) and Couallier (1999). In particular, we assume that τ lies in a compact interval $[A, B]$, and thus our estimator is defined as $\widehat{\tau} = \operatorname{argmax}_{x \in [A,B]} |\widehat{J}(x)|$. We partition the interval $[A, B]$ in n^{1+q} intervals of equal size, and define E_n as the set of endpoints of the partition. More precisely, let $E_n = \{z_0, \ldots, z_{n^{1+q}}\}$, where $z_0 = A < z_1 < \ldots < z_{n^{1+q}} = B$, and $z_{j+1} - z_j = (B - A)/n^{1+q}$ for $j = 0, \ldots, n^{1+q} - 1$.

Note that, unlike the error free case, where the kernel K is usually a positive, bounded and compactly supported function, the pseudo kernel K_Z of the error case is supported on the whole real line, not positive everywhere and is asymptotically unbounded. Hence, despite some similarities in the main ideas, the error case is much more difficult to deal with.

In particular, the asymptotic properties of our estimator depend strongly on the error distribution, since the latter dictates the behaviour of K_Z. As in Fan (1991c) we consider two types of error distributions, the ordinary smooth distributions and the supersmooth distributions.

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

(i) ordinary smooth of order β if its characteristic function $\varphi_Z(t)$ satisfies: $d_0|t|^{-\beta} \leq |\varphi_Z(t)| \leq d_1|t|^{-\beta}$ as $t \to \infty$, for some positive constants d_0, d_1 and β.

(ii) supersmooth of order β if its characteristic function $\varphi_Z(t)$ satisfies: $d_0|t|^{\beta_0} \exp(-|t|^{\beta}/\gamma) \leq |\varphi_Z(t)| \leq d_1|t|^{\beta_1} \exp(-|t|^{\beta}/\gamma)$ as $t \to \infty$, for some positive constants d_0, d_1, β, γ and

constants β_0 and β_1;

We will see that for supersmooth error densities (e.g. normal and Cauchy densities) the rate of convergence of the estimator is logarithmic. 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)).

In what follows, we treat the ordinary smooth and the supersmooth error cases separately. We first define some useful notations. For any set $D \subset I\!R$ and positive integer m, let $\mathcal{C}_m(D)$ denote the set of functions m times continuously differentiable on D and define $\mathcal{D}_m(D) = \big\{f \in \mathcal{C}_m(D) : \sup_{0\le j\le m}\ \sup_{x\in D} |f^{(j)}(x)| < \infty\big\}$. For a square integrable function f, let also $R(f)$ denote $\int f^2(x)\,dx$. Finally, let d denote the size of the discontinuity of f_X in τ, i.e. $d = f_X(\tau^+) - f_X(\tau^-)$, where, for any function g and point $a \in I\!R$, we use the notation $g(a^+) = \lim_{x \underset{>}{\to} a} g(x)$ and $g(a^-) = \lim_{x \underset{<}{\to} a} g(x)$. Then, we define the function r_X by $r_X = f_X - d\, I_{[\tau,+\infty[}$. Clearly, r_X is continous on $I\!R$, and, in particular in τ, since we have $r_X(\tau^-) = r_X(\tau^+) = f_X(\tau^-)$.

3.1 Ordinary smooth error case

In the case where the error Z is ordinary smooth of order β, we introduce the following assumptions.

Condition A:

- **(A1)** $K \in \mathcal{C}_3(I\!R)$ is a symmetric, kth order kernel ($k \ge 2$ even), such that $\|K\|_\infty = K(0) > \max_{x\neq 0}|K(x)|$, $K''(0) < 0$, and $\int |uK^{(r)}(u)|\,du < \infty$ for $r = 0, 1, 2, 3$;
- **(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)** $K_Z \in \mathcal{C}_4(I\!R)$ is such that $\int |K_Z''(u)|\,du = O(h^{-\beta})$ and $\int |u| \cdot |K_Z''(u)|^2\,du = O(h^{-2\beta})$, and, for $r = 0, 1, \ldots, 4$, $\|K_Z^{(r)}\|_\infty = O(h^{-\beta})$ and $R(K_Z^{(r)}) \sim h^{-2\beta}$;
- **(A5)** $h \to 0$ as $n \to \infty$, such that, for some $0 < \delta < 1/2$ and $p \in I\!N_0 = I\!N \setminus \{0\}$, $nh^{\frac{2+2\delta+\beta}{1+q}} \to \infty$ $\sum_{n=1}^{\infty} n^{1+q-p} h^{-4\delta p - p - 2\beta p} < \infty$, with $q \in I\!N_0$ as at page 4.

Although some of these conditions look rather involved, they are quite common in density deconvolution problems. A discussion on these and similar conditions is provided in Delaigle (2003). There, it is shown that the conditions can be expressed in a rather simple form, but to the extent of less generality of the functions f_Z and K. In particular, Condition (A1) is satisfied by most kernels commonly used in deconvolution problems. See also Fan (1991a,b) and Delaigle and Gijbels (2002,2004a,b).

Under the conditions stated, we prove the asymptotic normality of the estimator. We approximate the first two moments of the estimator by the first two moments of the asymptotic distribution, and deduce an approximation of the Mean Squared Error (MSE) of the estimator $\widehat{\tau}$. We discuss rates of convergence of the estimator to its target point τ. The proofs of the results are deferred to Section 5.

The asymptotic distribution of the estimator is described in the next theorem. This distribution depends on the number of derivatives l of the function r_X. We need to distinguish the case $l = 0$ from the case $l \geq 1$.

Theorem 3.1. *Suppose the error is ordinary smooth of order β. Under Conditions (A1)–(A5), if $r_X \in \mathcal{C}_l(I\!R) \cap \mathcal{D}_3(I\!R \setminus \{\tau\})$ with $l \geq 0$, and $\int |u|^3 |K''(u)|\, du < \infty$. Let $k_2 = 0$ if $l = 0$ and 1 otherwise. Then, for $h = o(n^{-\frac{1}{2\beta+2k_2+5}})$, we have*

$$\sqrt{\frac{n}{h}}\Big[\frac{\tau - \widehat{\tau}}{\sqrt{R(K_Z'')}} - \frac{h^{k_2+2} D_\tau}{dK''(0)\sqrt{R(K_Z'')}}\Big] \xrightarrow{\mathrm{L}} N\Big(0; \frac{B_\tau}{d^2\{K''(0)\}^2}\Big), \tag{3.1}$$

where $D_\tau = \frac{(-1)^{k_2+1}}{(k_2+1)!}[r_X^{(k_2+1)}(\tau^+) + (-1)^{k_2+1}\, r_X^{(k_2+1)}(\tau^-)] \int_{-\infty}^0 u^{k_2+1} K''(u)\, du$ and $B_\tau = [f_Y(\tau^+) + f_Y(\tau^-)]/2$.

An *approximation* of the mean squared error ($\mathcal{A}$MSE) of the estimator $\widehat{\tau}$ of τ can be found by using the moments of the asymptotic distribution.

Corollary 3.1. *Under the conditions of Theorem 3.1, we have,*

$$\mathcal{A}\mathrm{MSE}[\widehat{\tau}] = \frac{h^{2k_2+4} D_\tau^2}{d^2\{K''(0)\}^2} + \frac{hR\big(K_Z''\big) B_\tau}{nd^2\{K''(0)\}^2}. \tag{3.2}$$

When $r_X \in \mathcal{C}_l(I\!\!R)$, with $l > 1$, we see that we obtain the same asymptotic expression whatever the value of l. If $D_\tau = 0$, one has to go one or several steps further in the Taylor expansions used in the proofs, until finding a non-zero leading term.

The above results show that, the larger the discontinuity, the easier the estimation, which is easy to understand intuitively, as a large discontinuity is more likely to produce large derivatives of $\widehat{f}_X$, and thus easily detectable maxima for the diagnostic function $\widehat{J}$.

From Theorem 3.1, we deduce that

$$\tau - \widehat{\tau} = O_{\mathrm{P}}\Big(\sqrt{\frac{hR\big(K_Z''\big)}{n}}\Big) + O(h^{k_2+2}). \tag{3.3}$$

Under the conditions of the theorem, we know that $hR\big(K_Z''\big)$ is of order $h^{1-2\beta}$ (by condition (A4)). To obtain rates of convergence of the estimator, we need to investigate both terms at the right-hand side of expression (3.3) and make 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 (3.2).

If $0 \leq \beta \leq 1/2$ and hence $1 - 2\beta \geq 0$, minimization of (3.3) with respect to h leads to choosing h as small as possible. Under the conditions of the theorem, we can take $h \sim n^{-\frac{1}{2\beta+1}+\eta}$, with $\eta > 0$ which provides a rate of convergence slightly slower than $n^{-\frac{1}{2\beta+1}}$, more precisely, $\tau - \widehat{\tau} = O_{\mathrm{P}}(n^{-\frac{1}{2\beta+1}+\epsilon})$ with $\epsilon > 0$.

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 (3.3) of the same order). If $D_\tau \neq 0$, this bandwidth satisfies $h \sim n^{-\frac{1}{2\beta+2k_2+3}}$. From (3.3), we conclude then that $\tau - \widehat{\tau} = O_{\mathrm{P}}(n^{-\frac{k_2+2}{2\beta+2k_2+3}})$.

3.2 Supersmooth error case

If the error is supersmooth of order β with β_0 and β_1 as in Definition 1, we assume, for simplicity, that the support of φ_K is $[-1,1]$, i.e. $B_K=1$, and we introduce some conditions.

Condition B:

(B1)–(B3): the same as Conditions (A1)–(A3), with the extra condition $\int |t|^{2r-2\beta_0}|\varphi_K(t)|^2\,dt<\infty$ in (B1);

(B4) $K_Z\in\mathcal{C}_4(I\!R)$ is such that, for $r=0,1,\ldots 4$, $\|K_Z^{(r)}\|_\infty=O\big(h^{\beta_0}\exp(h^{-\beta}/\gamma)\big)$, $R(K_Z^{(r)})=O\big(h^{2\beta_0}\exp(2h^{-\beta}/\gamma)\big)$, 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,

$$h^{-1}\int_{-\infty}^{\infty}\{K_Z''(y)\}^2 f_Y(\tau-hy)\,dy\geq c_5 f_Y(\tau^-)h^{c_6}\exp\big(\frac{2}{h^\beta\gamma}-\frac{4\beta b_n}{h^\beta\gamma}\big) \tag{3.4}$$

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

As in the ordinary smooth error case, these conditions are rather common for this (difficult) class of error densities. Fan (1991a) proves that (B5) is satisfied under mild additional assumptions on K and f_Z. See Delaigle (2003) for similar results for Condition (B4).

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 distribution 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. The proofs of the results can be found in Section 5.

Theorem 3.2. *Suppose the error is supersmooth of order* $\beta > 0$ *and* $h = d\cdot(2/\gamma)^{1/\beta}(\ln n)^{-1/\beta}$ *with* $d > 1$*. Under Conditions (B1)–(B5), if* $r_X \in \mathcal{C}_l(I\!R) \cap \mathcal{D}_3(I\!R \setminus \{\tau\})$ *with* $l \geq 0$*, and* $\int |u|^3 |K''(u)|\, du < \infty$*. Let* $k_2 = 0$ *if* $l = 0$ *and* 1 *otherwise. Then*

$$\frac{\tau - \widehat{\tau}}{h\sqrt{\mathrm{Var}(h\widehat{J}'(\tau))}} - \frac{\mathrm{E}[h\widehat{J}'(\tau)]}{dK''(0)\sqrt{\mathrm{Var}(h\widehat{J}'(\tau))}} \xrightarrow{\mathrm{L}} N\Big(0; \frac{1}{d^2\{K''(0)\}^2}\Big), \tag{3.5}$$

with $\mathrm{E}[h\widehat{J}'(\tau)] = h^{k_2+1} D_\tau + O(h^{k_2+2})$*, and where* $\mathrm{Var}(h\widehat{J}'(\tau)) = (nh)^{-1} \int \big\{K''_Z(u)\big\}^2\, f_Y(\tau - hu)\, du + O(n^{-1})$*. Further, if* $\int |u| \cdot |K''_Z(u)|^2\, du = O(R(K''_Z))$*, we can write*

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

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

Corollary 3.2. *Under the conditions of Theorem 3.2, we have*

$$\mathcal{A}\mathrm{MSE}[\widehat{\tau}] = \frac{h^{2k_2+4} D_\tau^2}{d^2\{K''(0)\}^2} + \frac{h^2\,\mathrm{Var}(h\widehat{J}'(\tau))}{d^2\{K''(0)\}^2}. \tag{3.6}$$

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

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

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, we obtain, from (3.3), that $\widehat{\tau} - \tau = O_{\mathrm{P}}\big((\ln n)^{-(k_2+2)/\beta}\big)$, which is a much slower rate of convergence than in the ordinary smooth error case.

4 Simulations

In this section, we illustrate the finite sample performance of the procedure on a few examples. As noted in Section 2, the estimator $\widehat{f}_X$ is expected to have large positive derivative for a left endpoint and large negative derivative for a right endpoint. Hence, in practice, our estimator can be calculated as $\widehat{\tau} = \mathrm{argmax}_x\, \widehat{J}(x)$, (respectively $\widehat{\tau} = \mathrm{argmin}_x\, \widehat{J}(x)$), for a

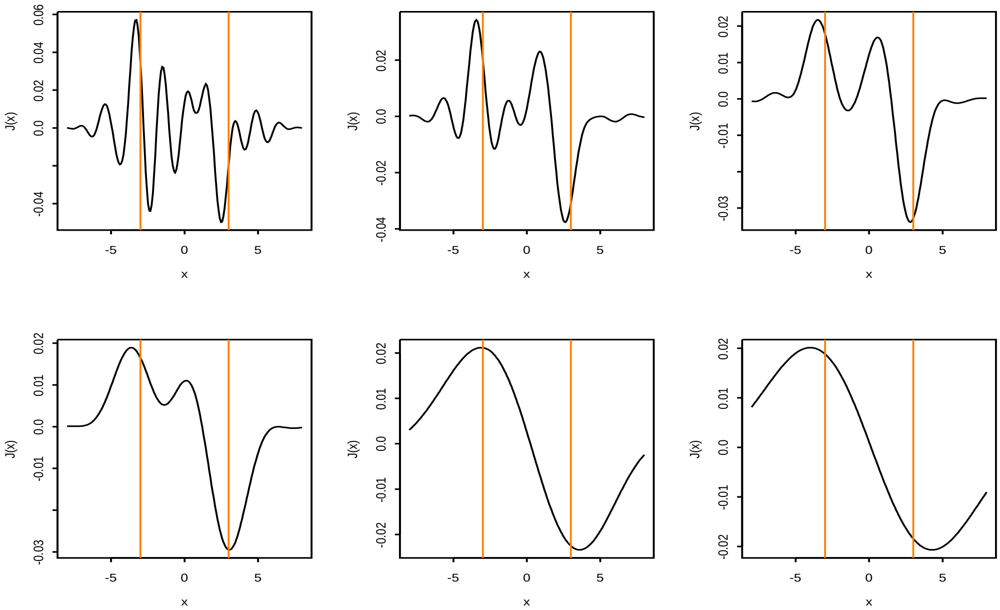

Figure 4.1: *Typical shape of* $\widehat{J}(x)$ *for a sample of size* $n = 100$ *from Density #3 contaminated with a Laplace error, for increasing bandwidths (from the left to the right and from the top to the bottom).*

left (respectively right) boundary point τ. In the case of two boundary points τ_1 and τ_2, we define, using similar ideas, the estimators $\widehat{\tau}_1 = \text{argmax}_x \widehat{J}(x)$ and $\widehat{\tau}_2 = \text{argmin}_x \widehat{J}(x)$.

Note that 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 (2.2).

Typical shape of the diagnostic function $\widehat{J}$ is illustrated in Figure 4.1 for increasing bandwidths (here for a sample of size $n = 100$ from Density #3 below contaminated by a Laplace error), with the actual endpoints (here -3 and 3) 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.

We select the bandwidth by estimating the asymptotic MISE optimal bandwidth for estimating the derivative of a density developed in the case of continuous and differentiable densities, i.e., for a second order kernel K, we estimate the bandwidth that minimizes $\text{AMISE}\left\{\widehat{f}'_X(x;h)\right\} = (2\pi n h^5)^{-1} \int t^4 |\varphi_K(t)|^2 |\varphi_Z(t/h)|^{-2}\, dt + \frac{h^4}{4}\mu_{K,2}^2 \theta_4$, where $\mu_{K,2}$ denotes the second moment of the kernel K, and $\theta_4 = R(f_X^{(4)})$ is unknown. We use a plug-in estima-

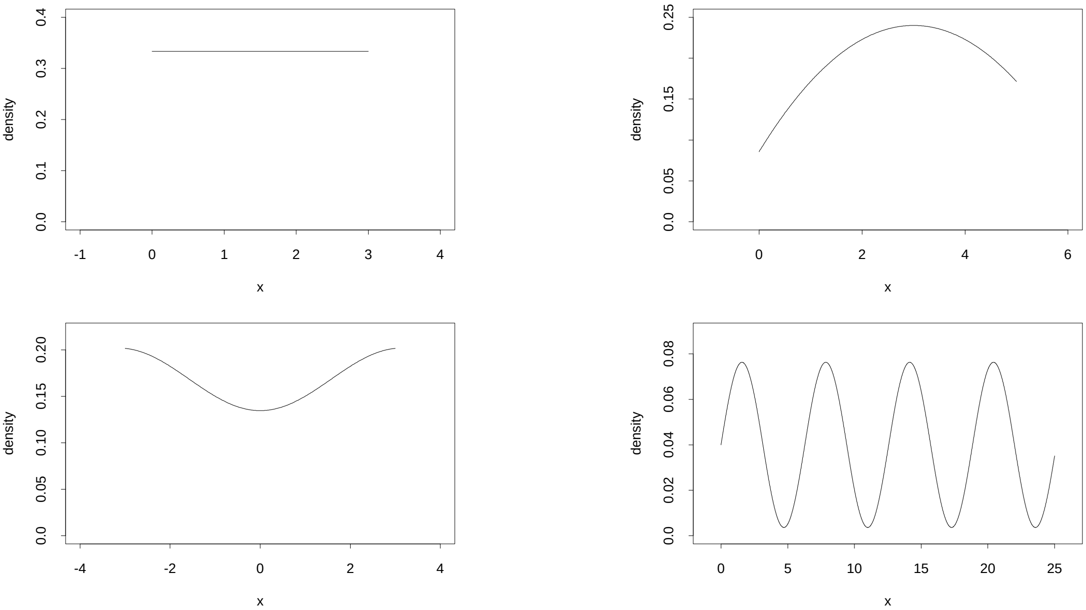

Figure 4.2: *Densities with a bounded support. Density #1 (top left panel), density #2 (top right panel), density #3 (bottom left panel) and density #4 (bottom right panel).*

tion method similar to the plug-in bandwidth selector of Delaigle and Gijbels (2002, 2004b). More precisely, we choose $h = \operatorname{argmin} \widehat{\mathrm{AMISE}}(h)$, with

$$\widehat{\mathrm{AMISE}}(h) = \frac{1}{2\pi n h^5} \int t^4 \frac{|\varphi_K(t)|^2}{|\varphi_Z(t/h)|^2}\, dt + \frac{h^4}{4} \mu_{K,2}^2 \widehat{\theta}_4, \tag{4.7}$$

where $\widehat{\theta}_4$ is the two-stage plug-in estimator of θ_4 proposed by Delaigle and Gijbels (2002). See Delaigle and Gijbels (2002) for detailed implementation of the procedure.

We use the second order kernel K corresponding to $\varphi_K(t) = (1 - t^2)^3 . 1_{[-1,1]}(t)$, commonly used in deconvolution problems. We consider four densities with a compact support, illustrated in Figure 4.2:

1. Density #1 (a uniform density): $f_X(x) = 1/3 \cdot 1_{[0,3]}(x)$;
2. Density #2 (a concave density): $f_X(x) = 3/175(-x^2 + 6x + 5) \cdot 1_{[0,5]}(x)$;
3. Density #3 (a sinus type density): $f_X(x) = (\sin^2(x/2) + 2)/(15 - \sin 3) \cdot 1_{[-3,3]}(x)$;
4. Density #4 (a multimodal density): $f_X(x) = (\sin x + 1.1)/(28.5 - \cos 25) \cdot 1_{[0,25]}(x)$.

For each of the above densities #1—4, we have generated 100 samples of size $n = 100$ and 250, contaminated by a Laplace error with a noise to signal ratio $\operatorname{Var} Z / \operatorname{Var} X = 10\%$.

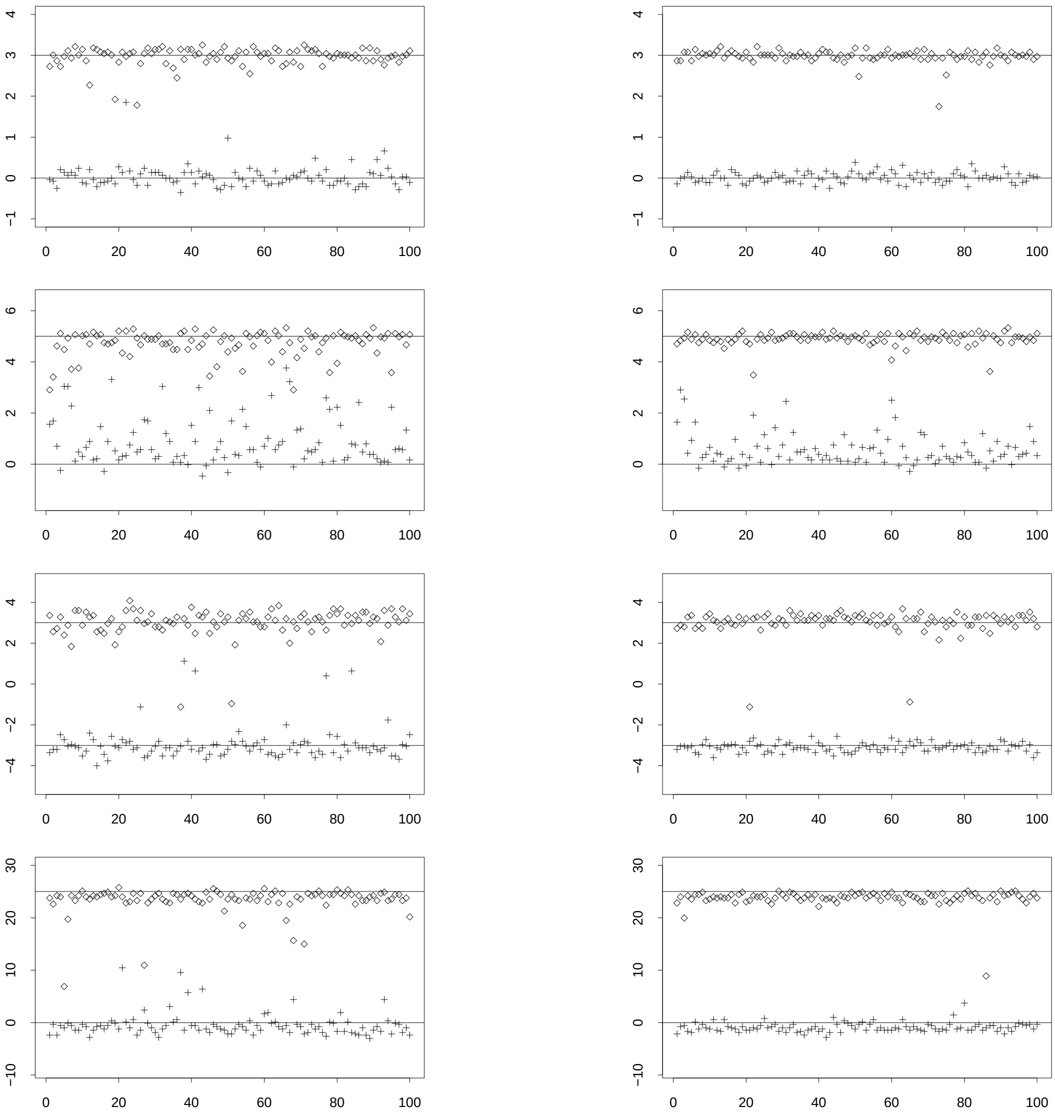

Figure 4.3: *Scatterplots of 100 replicated estimators for samples of size* $n = 100$ *(left panels), or* 250 *(right panels), from density #1 (top panels), density #2 (second line panels), density #3 (third line panels) or density #4 (bottom panels) contaminated by a Laplace error with a noise to signal ratio* $\operatorname{Var} Z / \operatorname{Var} X = 10\%$*. Estimates of the right (respectively left) endpoint are indicated by the character* $\diamond$ *(respectively* $+$*).*

Figure 4.3 shows scatterplots of the 100 replicated estimators of the left and right endpoints, indicated by respectively the characters $+$ and $\diamond$. The true endpoints are indicated by horizontal lines. From that figure, we see that the method performs quite well, even for the more difficult Densities #2 and #4, and the results improve as the sample size increases. As one could have expected, the left endpoint of Density #2 is more difficult to estimate than the right endpoint, because it has a smaller jump size, but yet, we see that the bias decreases as the sample size increases. See Delaigle and Gijbels (2004c) for more detailed results on this type of problem and other more challenging difficulties.

5 Proofs of the results

In this section we prove the results of Sections 3.1 and 3.2. For $u < 0$ and $r_X \in \mathcal{C}_{\ell+1}(I\!R\backslash\{\tau\})$, the ℓth order Taylor expansion of r_X around a boundary point τ may be written as $r_X(\tau+u) = r_X(\tau) + ur'_X(\tau^-) + \ldots + \frac{u^\ell}{\ell!} r_X^{(\ell)}(\tau^-) + R_\ell(\tau)$, where $R_\ell(\tau) = \dfrac{u^{\ell+1}}{(\ell+1)!} r_X^{(\ell+1)}(\tau + \theta u)$, with $0 < \theta < 1$. We obtain a similar expansion for $u > 0$, with τ^+ instead of τ^-.

5.1 Auxilary results for the ordinary smooth case

The following sequence of lemmas lead to the proof of Theorem 3.1 and Corollary 3.1. Throughout, K is a kth order symmetric kernel with $k \geq 2$. The following conditions will be useful.

Condition C: $(\mathbf{C}_1^m)$ $K \in \mathcal{C}_m(I\!R)$ is such that $\lim_{|x|\to\infty} K^{(m-1)}(x) = 0$; $(\mathbf{C}_2^m)$ $K_Z \in \mathcal{C}_m(I\!R)$; $(\mathbf{C}_3^m)$ $\int |u| \cdot |K_Z^{(m)}(u)|^2\, du = O(h^{-2\beta})$; $(\mathbf{C}_4^m)$ $\int |K_Z^{(m)}(u)|\, du = O(h^{-\beta})$; $(\mathbf{C}_5^m)$ $\|K_Z^{(m)}\|_\infty = O(h^{-\beta})$; $(\mathbf{C}_6^m)$ $R(K_Z^{(m)}) \sim h^{-2\beta}$; $(\mathbf{C}_7^m)$ $\int |uK^{(m)}(u)|\, du < \infty$.

The next lemma is a generalization of a result of Stefanski and Carroll (1990). See Delaigle and Gijbels (2002) for a proof.

Lemma 5.1. *Let $r \geq 0$. If $K \in \mathcal{C}_r(I\!R)$, we have* $\mathrm{E}\left[K_Z^{(r)}\left(\frac{x-Y}{h}\right)\right] = \mathrm{E}\left[K^{(r)}\left(\frac{x-X}{h}\right)\right]$.

Lemma 5.2. *Assume Conditions* (C_1^2) *and* (C_2^2), *and* $r_X \in \mathcal{C}_l(I\!R) \cap \mathcal{D}_3(I\!R \setminus \{\tau\})$ *with* $l \geq 0$. *Let* $k_2 = 0$ *if* $l = 0$ *and* 1 *otherwise. Then, if* $K'(0) = 0$ *and* $\int |u|^3 |K''(u)|\, du < \infty$,

$$\mathrm{E}\left[\frac{1}{nh}\sum_{i=1}^n K''_Z\Big(\frac{\tau - Y_i}{h}\Big)\right] = h^{k_2+1} D_\tau + O(h^{k_2+2}). \tag{5.1}$$

Proof. From Lemma 5.1 and the condition $K'(0) = 0$, we can write

$$\mathrm{E}\Big[\frac{1}{nh}\sum_{i=1}^{n} K_Z''\Big(\frac{\tau - Y_i}{h}\Big)\Big] = \int_{-\infty}^{0} K''(u) r_X(\tau - hu)\, du + \int_{0}^{+\infty} K''(u) r_X(\tau - hu)\, du\ .$$

A Taylor expansion of r_X of order 2 around τ^- (resp. τ^+) for $u > 0$ (resp. $u < 0$), combined with the fact that $\int u^j K''(u)\, du = 0$ for $j = 0, 1$, provides the result. $\square$

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*,*

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

uniformly in x*.*

Proof. Follows from Lemma 5.1 and Lipschitz continuity of r_X. $\square$

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*

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

Proof. From Lemma 5.3, a first order Taylor expansion of f_Y around τ^+ or τ^- and the symmetry of $K_Z^{(r)}$, we have

$$\begin{aligned}\mathrm{Var}\left[h^{-1} K_Z^{(r)}\Big(\frac{\tau - Y}{h}\Big)\right] &= h^{-1} \int \left\{K_Z^{(r)}(u)\right\}^2 f_Y(\tau - hu)\, du + O(1) \\ &= \frac{f_Y(\tau^+) + f_Y(\tau^-)}{2h} \int \left\{K_Z^{(r)}(u)\right\}^2 du + R_2 + O(1),\end{aligned}$$

where $|R_2| \leq \sup_{x \in I\!R \setminus \{\tau\}} |f_Y'(x)| \int |u| \cdot |K_Z^{(r)}(u)|^2\, du = O(h^{-2\beta})$. $\square$

The next lemma generalizes a result of Fan (1991a) to the case where the density f_X is not continuous.

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*

$$\frac{\frac{1}{nh}\sum_{i=1}^{n} K_Z^{(r)}\left(\frac{\tau-Y_i}{h}\right) - \mathrm{E}\left[\frac{1}{nh}\sum_{i=1}^{n} K_Z^{(r)}\left(\frac{\tau-Y_i}{h}\right)\right]}{\sqrt{\mathrm{Var}\left[\frac{1}{nh}\sum_{i=1}^{n} K_Z^{(r)}\left(\frac{\tau-Y_i}{h}\right)\right]}} \xrightarrow{\mathrm{L}} N(0;1). \tag{5.2}$$

Proof. Denoting $h^{-1}K_Z^{(r)}\left(\frac{\tau-Y_i}{h}\right)$ by $Y_{n,i}, i = 1, \ldots, n$, where, for all $i \neq j$, $Y_{n,i} \perp\!\!\!\perp Y_{n,j}$, we use the central limit theorem for triangular arrays of random variables. See for example Serfling (1980), page 32. We need to verify the following Lyapounov condition: for some $\eta > 0$,

$$\lim_{n\to\infty} \frac{n\,\mathrm{E}\,|Y_{n,1} - \mathrm{E}(Y_{n,1})|^{2+\eta}}{(n\,\mathrm{Var}(Y_{n,1}))^{(2+\eta)/2}} = 0. \tag{5.3}$$

From Minkowski's inequality, we have $\mathrm{E}\,|Y_{n,1}-\mathrm{E}(Y_{n,1})|^{2+\eta} \leq \left(\left\{\mathrm{E}\,|Y_{n,1}|^{2+\eta}\right\}^{\frac{1}{2+\eta}} + \mathrm{E}\,|Y_{n,1}|\right)^{2+\eta}$. Under the conditions of the lemma, $\mathrm{E}\,|Y_{n,1}| \leq \sup_{x\in I\!R\setminus\{\tau\}} |f_Y(x)| \int |K_Z^{(r)}(u)|\,du = O(h^{-\beta})$, and $h^{1+\eta}\,\mathrm{E}\,|Y_{n,1}|^{2+\eta} \leq ||K_Z^{(r)}||_\infty^{1+\eta} \cdot \sup_{x\in I\!R\setminus\{\tau\}} |f_Y(x)| \int |K_Z^{(r)}(u)|\,du = O(h^{-\beta(2+\eta)})$. We conclude that $h^{1+\eta}\,\mathrm{E}\,|Y_{n,1} - \mathrm{E}(Y_{n,1})|^{2+\eta} = O(h^{-\beta(2+\eta)})$. From the proof of Lemma 5.4, we have $\mathrm{Var}(Y_{n,1}) \sim h^{-2\beta-1}$, and the Lyapounov condition is satisfied. □

Lemma 5.6. *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* $\sup_{x\in I\!R\setminus\{\tau\}} |f_Y(x)| < \infty$ *and if $nh \to \infty$ as $n \to \infty$, we have, for all $p \in I\!N_0$ and for n large enough,*

$$\mathrm{E}[h^r \widehat{J}^{(r)}(x) - h^r\,\mathrm{E}\,\widehat{J}^{(r)}(x)]^{2p} \leq 2n^{-p}h^{-p-2\beta p} \cdot \left\{\frac{2}{\pi}\int |t|^{2r+2+2\beta}|\varphi_K(t)|\,dt \cdot \sup_{x\in I\!R\setminus\{\tau\}} |f_Y(x)|\,(2/d_0)^2\right\}^p.$$

Proof. Let T_j denote $K_Z^{(r+1)}\left(\frac{x-Y_j}{h}\right)$. We have

$$\mathrm{E}[h^r \widehat{J}^{(r)}(x) - h^r\,\mathrm{E}\,\widehat{J}^{(r)}(x)]^{2p} = \frac{1}{(nh)^{2p}}\,\mathrm{E}\sum_{i_1,i_2,\ldots,i_{2p}=1}^{n}\prod_{j=i_1}^{i_{2p}}\left[T_j - \mathrm{E}(T_1)\right]$$

$$= \frac{1}{(nh)^{2p}}\sum_{i=1}^{n}\mathrm{E}\left[T_i - \mathrm{E}(T_1)\right]^{2p} + \frac{1}{(nh)^{2p}}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{l_i=2}^{2p-2}\mathrm{E}\left[T_i - \mathrm{E}(T_1)\right]^{l_i}\cdot\mathrm{E}\left[T_j - \mathrm{E}(T_1)\right]^{2p-l_i}$$

$$+ \ldots + \frac{1}{(nh)^{2p}}\sum_{i_1\neq i_2\neq\ldots\neq i_p}\prod_{j=i_1}^{i_p}\mathrm{E}\left[T_j - \mathrm{E}(T_1)\right]^2, \tag{5.4}$$

since, from $\mathrm{E}\left[T_i - \mathrm{E}(T_i)\right] = 0$, 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. Let C_l^j denote the binomial coefficient. We have, for all $l \geq 2$,

$$\mathrm{E}\left[T_i - \mathrm{E}(T_1)\right]^l = \sum_{j=0}^{l} (-1)^j C_l^j \,\mathrm{E}\left[K_Z^{(r+1)}\Big(\frac{x-Y_i}{h}\Big)\right]^j \left[\mathrm{E}\, K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\right]^{l-j} = O(h^{1-l\beta}),$$

since, by Lemma 5.3, we have $\mathrm{E}\left[K_Z^{(r+1)}\big(\frac{x-Y_1}{h}\big)\right] = O(h)$, and, for all $j \geq 2$, using arguments similar to the proof of Lemma 5.5, we have $\mathrm{E}\left[K_Z^{(r+1)}\big(\frac{x-Y_i}{h}\big)\right]^j = O(h^{1-j\beta})$. Finally we get

$$\mathrm{E}[h^r \widehat{J}^{(r)}(x) - h^r \,\mathrm{E}\,\widehat{J}^{(r)}(x)]^{2p} = O\Big(n^{-p}h^{-p-2\beta p}\Big),$$

since $nh \to \infty$ as $n \to \infty$. From the above calculations, we also see that we can write

$$\begin{aligned}
\mathrm{E}[h^r \widehat{J}^{(r)}(x) - h^r \,\mathrm{E}\,\widehat{J}^{(r)}(x)]^{2p} =& \frac{\prod_{l=0}^{p-1}(n-l)}{(nh)^{2p}} \Big\{ \mathrm{E}\left[T_1 - \mathrm{E}(T_1)\right]^2 \Big\}^p \cdot (1+o(1)) \\
=& \frac{1}{n^p h^{2p}} \Big\{ \mathrm{E}\left[K_Z^{(r+1)}\Big(\frac{x-Y_1}{h}\Big)\right]^2 \Big\}^p \cdot (1+o(1)) \\
\leq& \frac{1}{n^p h^{2p}} \Big\{ 2h \sup_{x \in I\!R \setminus \{\tau\}} |f_Y(x)| \cdot R(K_Z^{(r+1)}) \Big\}^p \cdot (1+o(1)) \\
\leq& n^{-p} h^{-p-2\beta p} \cdot \Big\{ 2c_\beta(K) \sup_{x \in I\!R \setminus \{\tau\}} |f_Y(x)| \cdot (2/d_0)^2 \Big\}^p \cdot (1+o(1)),
\end{aligned}$$

where $c_\beta(K) = \pi^{-1} \int |t|^{2r+2+2\beta} |\varphi_K(t)|^2 \, dt$, and details for obtaining the last inequality can be found in Delaigle and Gijbels (2003). $\square$

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* $\|K^{(r+1)}\|_\infty < \infty$, *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).*

Proof. Let E_n be as defined on page 5. For each x in $[A, B]$ there exists at least one point z in E_n such that $|x - z| \leq (B - A)n^{-(1+q)}$. Denote that point by $z(x)$. For all $\omega \in \Omega$, the sample space, we have

$$\sup_{x \in [A,B]} |h^r \widehat{J}^{(r)}(x) - h^r \,\mathrm{E}\,\widehat{J}^{(r)}(x)| \leq S_{1,n} + S_{2,n} + S_{3,n}, \tag{5.6}$$

where $S_{1,n} \equiv \sup_{x\in[A,B]} |h^r \widehat{J}^{(r)}(x) - h^r \widehat{J}^{(r)}(z(x))|$, $S_{2,n} \equiv \sup_{x\in[A,B]} |h^r \widehat{J}^{(r)}(z(x)) - h^r \operatorname{E} \widehat{J}^{(r)}(z(x))|$ and $S_{3,n} \equiv \sup_{x\in[A,B]} |h^r \operatorname{E} \widehat{J}^{(r)}(z(x)) - h^r \operatorname{E} \widehat{J}^{(r)}(x)|$, and where to simplify the notations, we do not indicate specifically the dependence of the random variables on 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)

$$|h^r \widehat{J}^{(r)}(x) - h^r \widehat{J}^{(r)}(z(x))| \leq h^r |\widehat{J}^{(r+1)}(\xi)| \cdot |x - z(x)| \leq h^{-2} \|K_Z^{(r+2)}\|_\infty \cdot |x - z(x)|$$

where ξ lies between x and $z(x)$ and $\|K_Z^{(r+2)}\|_\infty \leq c_1 h^{-\beta}$, with c_1 a positive constant independent of ω and n. We conclude that

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

For handling the second term in (5.6), note that 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 for all $\epsilon > 0$ and $\ell \geq 1$, we can write

$$\begin{aligned}\sum_{n=1}^{\ell} P\Big(h^{-2\delta} S_{2,n} > \epsilon\Big) &\leq \sum_{n=1}^{\ell} P\Big(h^{-2\delta} \sup_{z\in E_n} |h^r \widehat{J}^{(r)}(z) - h^r \operatorname{E} \widehat{J}^{(r)}(z)| > \epsilon\Big) \\ &\leq \sum_{n=1}^{\ell} \sum_{z\in E_n} \Big[P\Big(h^{-2\delta} |h^r \widehat{J}^{(r)}(z) - h^r \operatorname{E} \widehat{J}^{(r)}(z)| > \epsilon\Big)\Big].\end{aligned}$$

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. Taking the limit as $\ell \to \infty$, we deduce that

$$\sum_{n=1}^{\infty} P\Big(h^{-2\delta} S_{2,n} > \epsilon\Big) \leq M - 1 + c_2 \sum_{n=M}^{\infty} n^{1+q-p} h^{-4\delta p-p-2\beta p} + c_2 \sum_{n=M}^{\infty} n^{-p} h^{-4\delta p-p-2\beta p} < \infty,$$

by Condition (A5). For the third term in (5.6), we have

$$|h^r \operatorname{E} \widehat{J}^{(r)}(z(x)) - h^r \operatorname{E} \widehat{J}^{(r)}(x)| \leq I + II,$$

with $I = |\int K^{(r+1)}(u)[r_X(z(x)-hu) - r_X(x-hu)]\,du| \le (B-A)\,n^{-1-q}L\int |K^{(r+1)}(u)|\,du$ and $II = |\int K^{(r+1)}(u)[dI_{[\tau,+\infty[}(z(x)-hu) - dI_{[\tau,+\infty[}(x-hu)]\,du| \le (B-A)\,n^{-1-q}|d| \cdot \|K^{(r+1)}\|_\infty h^{-1}$, where we used Lipschitz continuity of r_X and the bound $|z(x)-x| \le (B-A)\,n^{-1-q}$. Finally we obtain

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

with c_3 a positive constant independent of n (and of ω). Let ϵ be any positive real number. Since we have shown that, for all $\omega\in\Omega$, $h^{-2\delta}S_{1,n} + h^{-2\delta}S_{3,n} \le c_1\,(B-A)h^{-2-2\delta-\beta}n^{-1-q} + c_3n^{-1-q}h^{-1-2\delta}$, which, under Condition (A5), tends to zero as $n\to\infty$, we have, for n large enough (say $n\ge M$), $h^{-2\delta}S_{1,n} + h^{-2\delta}S_{3,n} \le \epsilon/2$. Thus we can write

$$\sum_{n=1}^{\infty} P\Big(h^{-2\delta}S_{1,n} + h^{-2\delta}S_{2,n} + h^{-2\delta}S_{3,n} > \epsilon\Big) \le (M-1) + \sum_{n=M}^{\infty} P\Big(h^{-2\delta}S_{2,n} > \epsilon/2\Big) < \infty.$$

□

Lemma 5.8. *Suppose that $\widehat{\tau} = \tau + O(h^{1+\eta})$ a.s., with $\eta > 0$. 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 $\|K^{(3)}\|_\infty < \infty$ and $\sup_{x\in I\!R\setminus\{\tau\}}|f_Y(x)| < \infty$. Then, if $nh\to\infty$ as $n\to\infty$, we have*

$$h^2\widehat{J}''(\xi) \xrightarrow{a.s} dK''(0), \tag{5.7}$$

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

Proof. We have

$$|h^2\widehat{J}''(\xi) - dK''(0)| \le |h^2\widehat{J}''(\xi) - h^2\,\mathrm{E}\,\widehat{J}''(\tau)| + |h^2\,\mathrm{E}\,\widehat{J}''(\tau) - dK''(0)| \le T_{1,n} + T_{2,n} + T_{3,n},$$

with $T_{1,n} = \sup_{x\in[A,B]}|h^2\widehat{J}''(x) - h^2\,\mathrm{E}\,\widehat{J}''(x)|$, $T_{2,n} = \sup_{x\in[\tau\wedge\widehat{\tau},\tau\vee\widehat{\tau}]}|h^2\,\mathrm{E}\,\widehat{J}''(x) - h^2\,\mathrm{E}\,\widehat{J}''(\tau)|$, and $T_{3,n} = |\,\mathrm{E}\,h^2\widehat{J}''(\tau) - dK''(0)|$, and where, by Lemmas 5.7 and 5.3, $T_{1,n} \xrightarrow{a.s.} 0$ and $T_{3,n}\to 0$.

Now, by Lemma 5.3 and a Taylor expansion of K'' around 0, we have, for all $x\in[A,B]$, $h^2\,\mathrm{E}\,\widehat{J}''(x) - h^2\,\mathrm{E}\,\widehat{J}''(\tau) = d\frac{x-\tau}{h}K^{(3)}(\theta) + O(h)$, with θ between 0 and $(x-\tau)/h$ and where the remainder term $O(h)$ is uniform in x. In particular, since $[\tau\wedge\widehat{\tau},\tau\vee\widehat{\tau}]\subset[A,B]$ we can write $T_{2,n} \le h^{-1}|d|\cdot\|K^{(3)}\|_\infty\cdot|\widehat{\tau}-\tau| + O(h) = O(h^\eta) + O(h)$, where the last equality holds almost surely. This proves that $T_{2,n} \xrightarrow{a.s} 0$. □

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). Its proof requires Proposition 5.1 below. The proof of Proposition 5.1 follows from standard arguments and is ommitted here. See Delaigle and Gijbels (2003) for more details. See also Couallier (2000). Let $I_n = \{x \in [A, B] : |x - \tau| > h^{1+\delta}\}$, with $\delta > 0$.

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

$$\sum_{n=1}^{\infty} P(\sup_{x \in I_n} |\widehat{J}(x)| \geq |\widehat{J}(\tau)|) < \infty. \tag{5.8}$$

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

$$\widehat{\tau} - \tau = O(h^{1+\delta}) \quad a.s. \ , \tag{5.9}$$

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

Proof. By definition of $\widehat{\tau}$, we have $P\big(\widehat{\tau} \in I_n\big) \leq P\big(\sup_{x \in I_n} |\widehat{J}(x)| \geq |\widehat{J}(\tau)|\big)$, and

$$\sum_{n=1}^{\infty} P(\sup_{x \in I_n} |\widehat{J}(x)| \geq |\widehat{J}(\tau)|) < \infty \Longrightarrow P(A) = 1,$$

by the Borel-Cantelli lemma, if we 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 that $\forall \omega \in A \ : \ \limsup_{n \to \infty} \frac{|\widehat{\tau}_n - \tau|}{h_n^{1+\delta}} \leq 1$, and thus $\widehat{\tau} - \tau = O(h^{1+\delta})$ almost surely. We conclude with Proposition 5.1. $\square$

5.2 Proofs of the main results for the ordinary smooth case

Proof of Theorem 3.1. 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

$$\tau - \widehat{\tau} = \widehat{J}'(\tau)/\widehat{J}''(\xi), \tag{5.10}$$

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

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

where, by Lemma 5.4, $\mathrm{Var}\left[h\widehat{J}'(\tau)\right] = \frac{B_\tau}{nh} R\big(K''_Z\big) + O(n^{-1}h^{-2\beta})$, with $B_\tau = \frac{f_Y(\tau^+)+f_Y(\tau^-)}{2}$, and where $R(K''_Z) \sim h^{-2\beta}$ (see Condition (A4)).

From Lemma 5.2, we know that $\mathrm{E}[h\widehat{J}'(\tau)] = h^{k_2+1}D_\tau + O(h^{k_2+2})$. Hence we have

$$\frac{h\widehat{J}'(\tau) - h^{k_2+1}D_\tau}{\sqrt{\frac{B_\tau}{nh} R\big(K''_Z\big)}} \cdot a_n + b_n \xrightarrow{\mathrm{L}} N(0;1),$$

where $a_n = \frac{\sqrt{\frac{B_\tau}{nh} R(K''_Z)}}{\sqrt{\frac{B_\tau}{nh} R(K''_Z)+O(n^{-1}h^{-2\beta})}} \to 1$ and $b_n = \frac{O(h^{k_2+2})}{\sqrt{\frac{B_\tau}{nh} R(K''_Z)+O(n^{-1}h^{-2\beta})}} \to 0$ $n \to \infty$, since $R(K''_Z) \sim h^{-2\beta}$, and we deduce that

$$\frac{h\widehat{J}'(\tau) - h^{k_2+1}D_\tau}{\sqrt{\frac{B_\tau}{nh} R\big(K''_Z\big)}} \xrightarrow{\mathrm{L}} N(0;1)$$

and the conclusion follows from (5.10) and Lemma 5.8. □

Proof of Corollary 3.1. From Theorem 3.1, an approximate expression for the expectation ($\mathcal{A}$E) may be derived from $\mathcal{A}\mathrm{E}\left[\frac{\tau-\widehat{\tau}}{\sqrt{R\big(K''_Z\big)}} - \frac{h^{k_2+2}D_\tau}{dK''(0)\sqrt{R\big(K''_Z\big)}}\right] = o(\sqrt{h/n})$. Using Condition (A4), we obtain $\mathcal{A}\mathrm{E}\left[\widehat{\tau}\right] = \tau - \frac{h^{k_2+2}D_\tau}{dK''(0)} + o(n^{-1/2}h^{(1-2\beta)/2})$. An approximate expression for the variance ($\mathcal{A}$Var) is given by

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

and thus $\mathcal{A}\mathrm{Var}[\widehat{\tau}] = \frac{hR(K''_Z)B_\tau}{nd^2\{K''(0)\}^2}(1+o(1))$. The conclusion follows immediately. □

5.3 Auxilary results for the supersmooth case

The following sequence of lemmas lead to the proof of Theorem 3.2 and Corollary 3.2. Again, K is a kth order symmetric kernel with $k \geq 2$. The proof of Lemma 5.12 is straightforward, hence it is ommitted. The following conditions will be useful.

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; **($\mathbf{D}_2^m$)** $\|K_Z^{(m)}\|_\infty = O\big(h^{\beta_0}\exp(h^{-\beta}/\gamma)\big)$; **($\mathbf{D}_3^m$)** $R(K_Z^{(m)}) = O\big(h^{2\beta_0}\exp(2h^{-\beta}/\gamma)\big)$.

The next lemma generalizes a result of Fan (1991a) to the case where the density f_X is not continuous.

Lemma 5.9. *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* (3.4) *is satisfied and* $\sup_{x\in I\!R\setminus\{\tau\}} |f_Y(x)| < \infty$, *then statement* (5.2) *in Lemma 5.5 holds.*

Proof. We need to verify the Lyapounov condition (5.3) for $Y_{n,i} = h^{-1}K_Z^{(r)}\big(\frac{\tau - Y_i}{h}\big)$, $i = 1,\ldots,n$. Let $\beta_4 = \min(\beta_0,\beta_3)$. We have $\mathrm{E}\,|Y_{n,1}| \leq \sup_{x\in I\!R\setminus\{\tau\}} |f_Y(x)| \int |K_Z^{(r)}(u)|\,du = O\big(h^{\beta_4}\exp(h^{-\beta}/\gamma)\big)$ and $h^{1+\eta}\,\mathrm{E}\,|Y_{n,1}|^{2+\eta} \leq \|K_Z^{(r)}\|_\infty^{1+\eta} \cdot \sup_{x\in I\!R\setminus\{\tau\}} |f_Y(x)| \int |K_Z^{(r)}(u)|\,du = O\big(h^{(2+\eta)\beta_4}\exp((2+\eta)h^{-\beta}/\gamma)\big)$. We conclude that

$$h^{1+\eta}\,\mathrm{E}\,|Y_{n,1} - \mathrm{E}(Y_{n,1})|^{2+\eta} = O\Big(h^{\beta_4(2+\eta)}\exp((2+\eta)h^{-\beta}/\gamma)\Big). \tag{5.11}$$

For the denominator, note that, from expression (3.4), we have, as in the proof of Lemma 5.4 (using Lemma 5.3)

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

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 have

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

where $c = \big[\frac{2}{c_5 f_Y(\tau^-)}\big]^{(2+\eta)/2}$ and $\beta_5 = (2+\eta)\,\beta_4 - \frac{2+\eta}{2}\,c_6 - \eta - 1$. Under the assumptions of the lemma, we have $\exp\big(2(2+\eta)\beta b h^{-\beta}/\gamma\big) = n^{(2+\eta)\beta d^{-\beta} b}$, 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 (5.11), the Lyapounov condition (5.3) is satisfied. $\square$

Lemma 5.10. *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 $c_{\beta_0}(K) = \int |t|^{2r+2-2\beta_0} |\varphi_K(t)|^2 \, dt$.

Proof. Similar to the proof of Lemma 5.6. In this case, however, for all $j \geq 2$, we find $\mathrm{E}\left[K_Z^{(r+1)}\left(\frac{x-Y_i}{h}\right)\right]^j = O\big(h^{j\beta_0+1} \exp(jh^{-\beta}/\gamma)\big)$ and then get $\mathrm{E}[h^r \widehat{J}^{(r)}(x) - h^r \,\mathrm{E}\, \widehat{J}^{(r)}(x)]^{2p} = O\big(n^{-p} \; h^{-p+2\beta_0 p} \exp(2ph^{-\beta}/\gamma)\big)$, since $nh \to \infty$ as $n \to \infty$. From the above calculations, we also have

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

where details for obtaining the last inequality can be found in Delaigle and Gijbels (2003). $\square$

Lemma 5.11. *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 $\delta > 0$.*

Proof. Similarly as in the proof of Lemma 5.7, we need to bound the sum in (5.6). For the first term, we have, for all $\omega \in \Omega$

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

with $c > 0$ a constant independent of ω and n, and thus $h^{-2\delta} S_{1,n} \leq c_1 n^{(d^{-\beta}/2)-1-q} (\ln n)^{\frac{2+2\delta-\beta_0}{\beta}}$, with $c_1 > 0$ a constant independent of ω and n. For the second term in (5.6), by Chebychev's Inequality and Lemma 5.10, we have that, for any $z \in I\!R$, and n large enough (say $n \geq M$)

$$P\Big(h^{-2\delta} |h^r \widehat{J}^{(r)}(z) - h^r \,\mathrm{E}\, \widehat{J}^{(r)}(z)| > \epsilon\Big) \leq c_2 (\ln n)^{\frac{p+4\delta p-2\beta_0 p}{\beta}} n^{pd^{-\beta}-p},$$

where c_2 is independent of ω and n. We deduce that

$$\sum_{n=1}^{\infty} P\big(h^{-2\delta} S_{2,n} > \epsilon\big) < M + c_2 \sum_{n=M}^{\infty} (\ln n)^{\frac{p+4\delta p-2\beta_0 p}{\beta}} n^{pd^{-\beta}-p+1+q} + c_2 \sum_{n=M}^{\infty} (\ln n)^{\frac{p+4\delta p-2\beta_0 p}{\beta}} n^{pd^{-\beta}-p}$$
$$< \infty,$$

as soon as $pd^{-\beta} - p + 1 + q < -1$, which is equivalent to requiring that $p > \frac{2+q}{1-d^{-\beta}}$, since $d > 1$. For the third term in (5.6), we get $h^{-2\delta} S_{3,n} \leq c_3 n^{-1-q} h^{-1-2\delta}$, with c_3 a positive constant. Since $c_1 n^{(d^{-\beta}/2)-1-q} (\ln n)^{\frac{2+2\delta-\beta_0}{\beta}} + c_3 n^{-1-q} h^{-1-2\delta}$ tends to zero as $n \to \infty$, we conclude as in the proof of Lemma 5.7. $\square$

Lemma 5.12. *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* (5.7) *in Lemma 5.8 holds.*

The proofs of Propositions 5.4 and 5.3 are similar to the ordinary smooth error case.

Proposition 5.3. *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 5.1 holds.*

Proposition 5.4. *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* (5.9) *of Proposition 5.2 holds.*

5.4 Proofs of the main results for the supersmooth case

Proof of Theorem 3.2. As in the proof of Theorem 3.1, we have $\tau - \widehat{\tau} = \widehat{J}'(\tau)/\widehat{J}''(\xi)$, where ξ lies between τ and $\widehat{\tau}$. From Lemma 5.9 for $r = 2$ and Lemma 5.12,

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

which implies

$$\frac{\tau-\widehat{\tau}}{h\sqrt{\operatorname{Var}(h\widehat{J}'(\tau))}}-\frac{\mathrm{E}[h\widehat{J}'(\tau)]}{dK''(0)\sqrt{\operatorname{Var}(h\widehat{J}'(\tau))}} \xrightarrow{\ \mathrm{L}\ } N\Big(0;\frac{1}{d^2\{K''(0)\}^2}\Big), \tag{5.12}$$

where $\mathrm{E}[h\widehat{J}'(\tau)]=h^{k_2+1}D_\tau+O(h^{k_2+2})$. We may further develop the variance term if we note that, as in the proof of Lemma 5.4, we have $\operatorname{Var}(h\widehat{J}'(\tau))=\frac{B_\tau}{nh}R\big(K''_Z\big)+O(n^{-1}R(K''_Z))$. $\square$

Proof of Corollary 3.2. From (5.12), an approximate expression for the variance ($\mathcal{A}$Var) is given by $\mathcal{A}\mathrm{Var}\left[\widehat{\tau}\right]=\frac{h^2\operatorname{Var}(h\widehat{J}'(\tau))}{d^2\{K''(0)\}^2}$. An approximate expression for the expectation ($\mathcal{A}$E) may be derived as follows

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

Hence, we obtain $\mathcal{A}\mathrm{E}[\widehat{\tau}]=\tau-\frac{h^{k_2+2}D_\tau}{dK''(0)}+o\Big(h\sqrt{\operatorname{Var}(h\widehat{J}'(\tau))}\Big)$. Finally an approximate mean squared error may be written as

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

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 the conclusion follows. $\square$

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. (2003). Boundary estimation and estimation of discontinuity points in deconvolution problems, Institut de Statistique, Université catholique de Louvain, *Discussion Paper* # 0320. http://www.stat.ucl.ac.be/ISpub/ISdp.html

Delaigle, A. and Gijbels, I. (2004a). Bootstrap bandwidth selection in kernel density estimation from a contaminated sample. *The Annals of the Institute of Statistical Mathematics*, **56**, 19–47.

Delaigle, A. and Gijbels, I. (2004b). Practical bandwidth selection in deconvolution kernel density estimation. *Computational Statistics and Data Analysis*, **45**, 249–267.

Delaigle, A. and Gijbels, I. (2004c). Practical implementation of deconvolution kernel estimation of boundaries and discontinuity points, Institut de Statistique, Université catholique de Louvain, *Discussion Paper* # 0420.

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. (2004). Bandwidth selection for change point estimation in nonparametric regression. *Technometrics*, **46**, 76–86.

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.

Müller, H-G. (1992). Change-points in noparametric regression analysis. *The Annals of Statistics*, **20**, 737–761.

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.