ESAIM: PS
April 2006, Vol. 10, p. 216–235
DOI: 10.1051/ps:2006007

ESAIM: Probability and Statistics
www.edpsciences.org/ps

BOOTSTRAPPING THE SHORTH FOR REGRESSION

Cécile Durot[1] and Karelle Thiébot[1, 2]

Abstract. The paper is concerned with the asymptotic distributions of estimators for the length and the centre of the so-called η-shorth interval in a nonparametric regression framework. It is shown that the estimator of the length converges at the $n^{1/2}$-rate to a Gaussian law and that the estimator of the centre converges at the $n^{1/3}$-rate to the location of the maximum of a Brownian motion with parabolic drift. Bootstrap procedures are proposed and shown to be consistent. They are compared with the plug-in method through simulations.

Mathematics Subject Classification. 62E20, 62G05, 62G08, 62G09.

Received March 15, 2005. Revised November 14, 2005.

1. Introduction

This paper is motivated by a practical problem that we explore in the companion paper [3] and that can briefly be described as follows. Each day, the concentration of a given pollutant in the ambient air is measured at a given place at 96 equi-spaced times by Airpl, an organization that maintains a network of air pollution monitoring stations in western France. Visual examinations are daily performed by experts in order to validate the data. One of them consists in checking that the peak of pollution (that is the period of the day when pollution is maximal) occurs at some time which is consistent with respect to a given criterion. Our task is to formalize this visual examination. To do that, we denote by y_i the i-th concentration measure, we suppose that $y_1, \dots, y_{96}$ obey a regression model and we model the peak of pollution by the so-called η-shorth interval for regression with $\eta = 0.25$. Thus our aim is to build statistical tests about the shorth interval for regression. The present paper is concerned with theoretical aspects (we study the distribution of an estimator for the shorth interval in a nonparametric regression model) while the companion paper is concerned with an application of the method to pollution studies.

The usual definition of the shorth interval is based on the empirical distribution function of independent and identically distributed data $X_1, \dots, X_n$: if η denotes a fixed number in $(0, 1)$, then the η-shorth interval of the sample is defined as the shortest interval that contains a fraction η of the sample. The asymptotic properties of the location and length of the η-shorth interval were studied by several authors. Andrews *et al.* [1] first gave a heuristic analysis of the η-shorth estimate of location. Shorack and Wellner [14] then obtained the asymptotic distribution of the estimate. Kim and Pollard [11] finally illustrated the problem where cube root asymptotics arise, through an application of their main theorem on the η-shorth estimator. They established that the centre

Keywords and phrases. Brownian motion with parabolic drift, bootstrap, location of maximum, shorth.

[1] Université Paris-Sud, Bâtiment 425, 91405 Orsay Cedex, France; cecile.durot@math.u-psud.fr

[2] Air Pays de la Loire, 2 rue A. Kastler, BP 30723, 44307 Nantes Cedex 3, France.

Supported by Air Pays De La Loire

© EDP Sciences, SMAI 2006

Article published by EDP Sciences and available at http://www.edpsciences.org/ps or http://dx.doi.org/10.1051/ps:2006007

of the shorth interval converges in law at the rate $n^{1/3}$ to τ, the location of the maximum of a standard two-sided Brownian motion W with parabolic drift:

$$\tau = \underset{t\in\mathbb{R}}{\operatorname{argmax}}\{W(t) - t^2\}. \tag{1.1}$$

Grübel [6] proved that the rate of convergence of the length of the shorth is $n^{1/2}$ and that its limiting distribution is normal. Janaszewska and Nagaev [10] then studied the joint distribution of the shorth height and length.

This paper is concerned with the asymptotic behaviour of empirical estimators for the length and the centre of the η-shorth interval in a nonparametric regression framework, that we define as follows. Let $y_1, \ldots, y_n$ be n observations at time $t_1, \ldots, t_n$ according to the model

$$y_i = f(t_i) + \varepsilon_i,\ 1 \leqslant i \leqslant n. \tag{1.2}$$

Here, $t_i = i/n$ and the ε_i's are independent and identically distributed random variables with zero mean and unknown variance σ^2. The unknown regression function f is assumed to be positive and differentiable on $[0, 1]$. Furthermore, we assume that there exists a unique shortest interval $[\mu_0 - r_0, \mu_0 + r_0] \subset (0, 1)$ that satisfies

$$\int_{\mu_0-r_0}^{\mu_0+r_0} f(s)\mathrm{d}s \geqslant \eta \int_0^1 f(s)\mathrm{d}s \tag{1.3}$$

for some fixed number $\eta \in (0, 1)$ and we refer to this interval as the η-shorth interval. Thus the η-shorth interval can be characterized by either the pair $\{\mu_0, r_0\}$ or the pair $\{\mu_0 - r_0, \mu_0 + r_0\}$. In this paper, we provide consistent estimators μ_n and r_n for μ_0 and r_0 and we describe their asymptotic distributions: it is proved that $n^{1/2}(r_n - r_0)/C_r$ converges to a standard Gaussian law and that $n^{1/3}(\mu_n - \mu_0)/C_\mu$ converges to τ for some explicit quantities C_r and C_μ that only depend on f and σ. These results are similar to those obtained earlier in the i.i.d. case, with different normalizing constants. One can conduct inference about the shorth interval (if n is large enough) by using these convergence results and plug-in estimators for C_r and C_μ. However, C_r and C_μ depend on the unknown regression function and its first derivative f' at the unknown points $\mu_0 - r_0$ and $\mu_0 + r_0$. Since nonparametric estimators of a derivative can converge only slowly to the true derivative (see Stone [15]), we thus suspected that the plug-in method was inappropriate when n is moderate. In particular, we suspected that it was inappropriate for the application we had in view, where we recall that $n = 96$. We thus propose bootstrap procedures as an alternative to the plug-in method. The bootstrap procedures are shown to be consistent and are compared with the plug-in method when $n = 100$ and $\eta = 0.25$ through a simulation study. It can be seen on these simulations that the bootstrap (when properly calibrated) outperforms the plug-in method when one wishes to conduct inference about the parameter μ_0: the bootstrap confidence intervals are shorter than the plug-in ones and their coverage probability is close to the target confidence level $1 - \alpha$. On the other hand, none of the proposed methods is appropriate to conduct inference about r_0 when n is moderate and bootstrap is not better than plug-in in that case. Our conclusion is that for moderate n, one has to consider the pair $\{\mu_0 - r_0, \mu_0 + r_0\}$ but not the pair $\{\mu_0, r_0\}$ in order to conduct inference about the shorth interval. Indeed, it is easy to see that all of the proposed methods apply for estimating $\mu_0 - r_0$ and $\mu_0 + r_0$ and some simulations showed that the results obtained for these parameters are similar to those obtained for μ_0 (so we did not report these results and we only refer to Thiébot [16] for some simulations about these parameters).

The paper is organized as follows. The estimators μ_n and r_n are defined in Section 2. Their asymptotic distributions are given and the bootstrap is shown to be consistent under appropriate assumptions $\mathcal{E}$. In Section 3, we propose several ways to perform the bootstrap so that $\mathcal{E}$ hold. A simulation study is reported in Section 4 and Section 5 is devoted to the proofs.

2. Statement of the main results

Consider the regression model (1.2) where $t_i = i/n$ and the ε_i's are independent and identically distributed random variables with mean zero. Assume that f is positive and that there exists a unique shortest interval

$[\mu_0 - r_0, \mu_0 + r_0] \subset (0,1)$ that satisfies (1.3) for some fixed $\eta \in (0,1)$. Our aim in this section is to provide estimators for μ_0 and r_0 and to study their asymptotic properties as $n \to \infty$. Then we propose a bootstrap method and we give conditions under which it is consistent.

2.1. The estimators

We need to introduce some notation in order to define our estimators. For every $H : [0,1] \to \mathbb{R}$ let

$$r_H = \inf \left\{ r \geqslant 0 : \sup_{\mu} \{H(\mu + r) - H(\mu - r)\} \geqslant \eta H(1) \right\}. \tag{2.1}$$

Thus r_H is well defined if the above set is non-empty. If r_H is well defined and if moreover the function $\mu \mapsto H(\mu + r_H) - H(\mu - r_H)$ achieves its supremum, then let

$$\mu_H = \operatorname*{argmax}_{\mu} \{H(\mu + r_H) - H(\mu - r_H)\}$$

where argmax stands for the infimum of the locations of maximum. It is worth noticing that r_H is well defined whenever $H(0) = 0$ since in that case, the set in (2.1) is non-empty: it contains $1/2$ if $H(1) \geqslant 0$ and 0 otherwise. If furthermore, H is either a continuous or a cadlag step function, then μ_H is also well defined and it maximizes $\mu \mapsto H(\mu + r_H) - H(\mu - r_H)$.

It is easy to see that $(r_0, \mu_0) = (r_F, \mu_F)$, where F is the cumulative function defined by

$$F(t) = \int_0^t f(s) \mathrm{d}s, \ t \in [0,1].$$

Moreover, it is well known that the partial sum process

$$F_n(t) = \frac{1}{n} \sum_{i \leqslant nt} y_i, \ t \in [0,1]$$

is a good estimator for F. We thus consider r_{F_n} and μ_{F_n} (which are well defined) as estimators for r_0 and μ_0 and for notational convenience, we denote them by r_n and μ_n respectively.

2.2. Asymptotic distributions

The asymptotic distributions of the estimators are given in the following theorem, where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution as $n \to \infty$.

Theorem 2.1. *Assume we are given the regression model (1.2) where $t_i = i/n$, $f : [0,1] \to \mathbb{R}$ has a Hölderian first derivative and where $\varepsilon_1, \ldots, \varepsilon_n$ are centered i.i.d. random variables with $\mathbb{E}|\varepsilon_1|^p < \infty$ for some $p > 3$. Assume $\inf_t f(t) > 0$, $\sigma^2 = \mathbb{E}(\varepsilon_i^2) > 0$ and there exists a unique shortest interval $[\mu_0 - r_0, \mu_0 + r_0] \subset (0,1)$ that satisfies (1.3) for some fixed $\eta \in (0,1)$. If $f'(\mu_0 - r_0) > f'(\mu_0 + r_0)$ then*

$$n^{1/3}(\mu_n - \mu_0) \xrightarrow{\mathcal{D}} \frac{2\sigma^{2/3}}{(f'(\mu_0 - r_0) - f'(\mu_0 + r_0))^{2/3}} \, \tau,$$

where τ is defined by (1.1). Moreover,

$$n^{1/2}(r_n - r_0) \xrightarrow{\mathcal{D}} \frac{\sigma}{2f(\mu_0 + r_0)} \mathcal{N}\left(0, \eta^2 + 2r_0(1 - 2\eta)\right).$$

The distribution of τ has been precisely described by Groeneboom [5] and tabulated by Narayanan and Sager [12]. The limiting distributions of μ_n and r_n are thus known up to a finite number of parameters so one can plug-in estimators to get a pivotal asymptotic law. To be more precise, let us denote by $\hat{\sigma}_n^2$ a consistent estimator for σ^2 (see *e.g.* Hall *et al.* [9]) and by $\hat{f}_{n,h}$ a differentiable estimator for f with smoothing parameter h. If $\hat{f}_{n,h}$ is properly chosen then its first derivative $\hat{f}'_{n,h}$ is a reasonable estimator for f', see Yatracos [17]. So let

$$\widehat{C}_{\mu,h_\mu} = \frac{2\hat{\sigma}_n^{2/3}}{(\hat{f}'_{n,h_\mu}(\mu_n - r_n) - \hat{f}'_{n,h_\mu}(\mu_n + r_n))^{2/3}} \quad \text{and} \quad \widehat{C}_{r,h_r} = \frac{\hat{\sigma}_n\sqrt{\eta^2 + 2r_n(1-2\eta)}}{2\hat{f}_{n,h_r}(\mu_n + r_n)}. \tag{2.2}$$

Then for properly chosen h_μ and h_r,

$$n^{1/3}(\mu_n - \mu_0)/\widehat{C}_{\mu,h_\mu} \xrightarrow{\mathcal{D}} \tau \quad \text{and} \quad n^{1/2}(r_n - r_0)/\widehat{C}_{r,h_r} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

One can then use these pivotal statistics to build statistical tests or confidence intervals with prescribed asymptotic level. For instance, fix $\alpha \in (0,1)$ and let q be the quantile of order $(1-\alpha/2)$ of τ. The distribution of τ is symmetrical about zero so

$$\left[\mu_n - n^{-1/3} q\, \widehat{C}_{\mu,h_\mu}, \mu_n + n^{-1/3} q\, \widehat{C}_{\mu,h_\mu}\right] \tag{2.3}$$

is a confidence interval for μ_0 with asymptotic level α.

2.3. **Consistency of the bootstrap**

An alternative to the plug-in method described above lies in bootstrap. As is customary, we denote by $\mathbb{P}^*$ the conditional probability given $(y_1, \dots, y_n)$ and by $\mathbb{E}^*$ the associated expectation. Let $\varepsilon_1^*, \dots, \varepsilon_n^*$ satisfy the following conditions denoted by $\mathcal{E}$ (examples of construction will be given below).

$\mathcal{E}$: Conditionally on $(y_1, \dots, y_n)$, $\varepsilon_1^*, \dots, \varepsilon_n^*$ are i.i.d. random variables with mean zero and finite variance σ_n^{*2}. Moreover, σ_n^{*2} stochastically converges to σ^2 as $n \to \infty$ and there exists some constant $C > 0$ that does not depend on n such that

$$\lim_{n\to\infty} \mathbb{P}\left(\mathbb{E}^*|\varepsilon_1^*|^p < C\right) = 1. \tag{2.4}$$

In order to build our bootstrap estimators we need to smooth the function F_n. Thus we consider the smoothed version of F_n given by

$$G_n(t) = \frac{1}{h_n}\int_{\mathbb{R}} F_n(x) K\left(\frac{t-x}{h_n}\right)\, \mathrm{d}x,\ t \in [0,1],$$

where K is a kernel function, $h_n > 0$ and where we set $F_n(t) = F_n(0) = 0$ for every $t \leqslant 0$ and $F_n(t) = F_n(1)$ for every $t \geqslant 1$. Denoting by g_n the first derivative of G_n, we then define the bootstrap observations as

$$y_i^* = g_n(t_i) + \varepsilon_i^*,\ i = 1, \dots, n.$$

Finally, we consider the bootstrap version of F_n given by

$$G_n^*(t) = \frac{1}{n}\sum_{i\leqslant nt} y_i^*,\ t \in [0,1]$$

and we define the bootstrap estimators as $r_{G_n^*}$ and $\mu_{G_n^*}$ (note that they are indeed well defined). It is proved in Theorem 2.2 below that the conditional asymptotic distributions of the bootstrap estimators (given $y_1, \dots, y_n$) are identical to the asymptotic distributions of the first estimators μ_n and r_n, provided K and h_n are well chosen. To be more specific, we assume in the sequel the following assumptions $\mathcal{K}$.

$\mathcal{K}$: * K is a symmetric probability density that vanishes outside $[-1,1]$;
* K is twice differentiable on $(-1,1)$ with a bounded second derivative;
* $\int K'' = 0$, $\int_{\mathbb{R}} x^2 K''(x)\mathrm{d}x = 2$ and $\int_{\mathbb{R}} xK'(x)\mathrm{d}x = -1$;
* $h_n > 0$, $h_n^{-1} = o(n^{\alpha})$ for some $\alpha < 1/3$ and $h_n = o(n^{-1/6}/\log(n))$.

An example of kernel that satisfies the above conditions is given by the quartic function

$$K(x) = \frac{15}{16}\left(1-x^2\right)^2 \mathbb{1}_{[-1,1]}(x). \tag{2.5}$$

Hereafter, $\xrightarrow{\mathcal{D}^*}$ denotes conditional convergence in distribution given $(y_1,\dots,y_n)$.

Theorem 2.2. *Let $\varepsilon_1^*,\dots,\varepsilon_n^*$ satisfy $\mathcal{E}$ and let K and h_n satisfy $\mathcal{K}$. Under the assumptions of Theorem 2.1, r_{G_n} and μ_{G_n} are well defined with probability that tends to one. Moreover,*

$$n^{1/3}(\mu_{G_n^*} - \mu_{G_n}) \xrightarrow{\mathcal{D}^*} \frac{2\sigma^{2/3}}{(f'(\mu_0-r_0)-f'(\mu_0+r_0))^{2/3}}\,\tau \text{ in probability}$$

and

$$n^{1/2}(r_{G_n^*} - r_{G_n}) \xrightarrow{\mathcal{D}^*} \frac{\sigma}{2f(\mu_0+r_0)}\mathcal{N}(0,\eta^2+2r_0(1-2\eta)) \text{ in probability.}$$

The limiting distributions are continuous so one can build statistical tests or confidence intervals with prescribed asymptotic level using bootstrap. For instance, let $q_{1,n,B}^*$ and $q_{2,n,B}^*$ be the quantiles of order $(\alpha/2)$ and $(1-\alpha/2)$ of B independent copies of $\mu_{G_n^*} - \mu_{G_n}$ conditionally on $(y_1,\dots,y_n)$. Under the assumptions of Theorem 2.2 we have

$$\lim_{n\to\infty}\lim_{B\to\infty}\mathbb{P}(q_{1,n,B}^* \leqslant \mu_n - \mu_0 \leqslant q_{2,n,B}^*) = 1-\alpha.$$

Thus

$$[\mu_n - q_{2,n,B}^*, \mu_n - q_{1,n,B}^*] \tag{2.6}$$

is a confidence interval for μ_0 with asymptotic level α.

To conclude this section, let us notice that the bootstrap procedure we propose here involves a smoothing parameter h_n although the first estimators μ_n and r_n are entirely data-driven. It would be best if the bootstrap procedure were also data-driven so one may wonder whether it is necessary or not to introduce this smoothing parameter. The answer is positive. Indeed if no smoothing parameter were involved, we shall consider the bootstrap partial sum process $F_n^*(t) = \sum_{i\leqslant nt}(y_i+\varepsilon_i^*)/n$, $t\in[0,1]$, where $\varepsilon_1^*,\dots,\varepsilon_n^*$ satisfy $\mathcal{E}$. It is proved in Section 5 that under the assumptions of Theorem 2.1,

$$n^{1/2}(r_{F_n^*} - r_n) \xrightarrow{\mathcal{D}^*} \frac{\sigma}{2f(\mu_0+r_0)}\mathcal{N}(0,\eta^2+2r_0(1-2\eta)) \text{ in probability} \tag{2.7}$$

so this data-driven bootstrap procedure is consistent for estimating r_0. However, it is proved in Section 5 that

$$n^{1/3}(\mu_{F_n^*} - \mu_n) \xrightarrow{\mathcal{D}^*} \frac{2\sigma^{2/3}}{(f'(\mu_0-r_0)-f'(\mu_0+r_0))^{2/3}}\,Z \text{ in probability,} \tag{2.8}$$

where Z does not have the same distribution as τ. Thus the data-driven bootstrap is not consistent for estimating μ_0 and one indeed has to smooth for estimating this parameter (and also, for estimating the boundaries μ_0+r_0 or μ_0-r_0 of the shorth interval).

3. How to bootstrap

In this section, we propose several possible constructions of the bootstrap residuals so that the assumptions $\mathcal{E}$ hold and we discuss studentized bootstrap.

3.1. Construction of the bootstrap residuals

Parametric bootstrap. The simplest construction corresponds to the parametric bootstrap and is relevant in situations where the residuals $\varepsilon_1, \ldots, \varepsilon_n$ are known to be (close to) Gaussian. It consists in generating (conditionally on $y_1, \ldots, y_n$) $\varepsilon_1^*, \ldots, \varepsilon_n^*$ as i.i.d. Gaussian variables with mean zero and variance $\hat{\sigma}_n^2$, where $\hat{\sigma}_n^2$ is a consistent estimator of σ^2 (see *e.g.* Hall *et al.* [9] for examples of such estimators).

Naive bootstrap. In the case when nothing is known about the distribution of the residuals, one has to perform a nonparametric bootstrap. To do so, let us set

$$\tilde{\varepsilon}_i = \hat{\varepsilon}_i - \frac{1}{n}\sum_{j=1}^{n} \hat{\varepsilon}_j, \ i = 1, \ldots, n$$

where $\hat{\varepsilon}_j = y_j - g_n(t_j)$ and where we recall that $g_n = G_n'$ is an estimator for f. The naive bootstrap consists in generating (conditionally on $y_1, \ldots, y_n$) $\varepsilon_1^*, \ldots, \varepsilon_n^*$ as a random sample of size n from the distribution that puts mass $1/n$ at each point $\tilde{\varepsilon}_1, \ldots, \tilde{\varepsilon}_n$.

Smoothed bootstrap. With the naive bootstrap, the bootstrap residuals are generated according to a discrete distribution P_n, the empirical distribution of the centered residuals $\tilde{\varepsilon}_1, \ldots, \tilde{\varepsilon}_n$ defined above. It is known that in some situations, generating the bootstrap residuals according to a smoothed version of P_n instead of P_n itself may improve rates of convergence, see *e.g.* Hall *et al.* [8], De Angelis *et al.* [2], Falk and Reiss [4]. So we also consider here a smoothed bootstrap that we define now. Let ϕ be a probability density and let $h_n' > 0$. The smoothed bootstrap consists in generating (conditionally on $y_1, \ldots, y_n$) $\varepsilon_1^*, \ldots, \varepsilon_n^*$ as a random sample of size n from the distribution $\phi_{h_n'} * P_n$, where $\phi_h(t) = \phi(t/h)/h$. As noticed by Falk and Reiss [4], this remains to generate independent random variables $X_1, \ldots, X_n, V_1, \ldots, V_n$ where the X_i's have common distribution P_n and the V_i's have common density function ϕ and to take $\varepsilon_i^* = X_i + h_n' V_i$.

It is stated in the following proposition that the three bootstrap constructions proposed here satisfy $\mathcal{E}$, so it follows from Theorem 2.2 that the three methods are consistent. Note that we chose to construct the nonparametric bootstrap residuals using the estimator g_n for simplicity but that other choices of estimators are allowed.

Proposition 3.1. *Assume the assumptions of Theorem 2.1 and let K and h_n satisfy $\mathcal{K}$. Assume that $\varepsilon_1^*, \ldots, \varepsilon_n^*$ are generated by either the parametric bootstrap or the naive bootstrap or the smoothed bootstrap with some bandwidth h_n' that tends to zero as $n \to \infty$ and some ϕ that satisfies*

$$\int_{\mathbb{R}} t\phi(t)\mathrm{d}t = 0, \quad \int_{\mathbb{R}} t^2\phi(t)\mathrm{d}t = 1 \quad \textit{and} \quad \int_{\mathbb{R}} |t|^p \phi(t)\mathrm{d}t < \infty.$$

Then $\mathcal{E}$ holds.

3.2. Studentized bootstrap

The studentized bootstrap is known to outperform the ordinary bootstrap in regular models (if one observes i.i.d. random variables $X_1, \ldots, X_n$ and one estimates a smooth function of $\mathbb{E}(X_i)$, see Hall [7]) so it can be interesting to use studentized bootstrap instead of ordinary bootstrap in our framework. In the case when the parameter of interest is μ_0, it consists in approximating the distribution of $(\mu_n - \mu_0)/\widehat{C}_{\mu,h_\mu}$ by that of

$(\mu_{G_n^*} - \mu_{G_n})/C^*_{\mu,h_\mu}$, where $\widehat{C}_{\mu,h_\mu}$ is defined as in (2.2) and where C^*_{μ,h_μ} is computed in the same manner as $\widehat{C}_{\mu,h_\mu}$ but with the observations $y_1, \dots, y_n$ replaced by the bootstrap observations $y_1^*, \dots, y_n^*$. For instance, the confidence interval for μ_0 obtained with the studentized bootstrap is

$$[\mu_n - p^*_{2,n,B}, \mu_n - p^*_{1,n,B}], \tag{3.9}$$

where $p^*_{1,n,B}$ and $p^*_{2,n,B}$ are the quantiles of order $(\alpha/2)$ and $(1-\alpha/2)$ of B independent copies of

$$\widehat{C}_{\mu,h_\mu}(\mu_{G_n^*} - \mu_{G_n})/C^*_{\mu,h_\mu} \tag{3.10}$$

conditionally on $(y_1, \dots, y_n)$. It has asymptotic level α if $\widehat{C}_{\mu,h_\mu}$ is $\mathbb{P}$-consistent and C^*_{μ,h_μ} is $\mathbb{P}^*$-consistent. One can define the studentized bootstrap in the same way in the case when the parameter of interest is r_0.

4. Simulations

We compared the performances of the proposed methods through some simulations. The methods can be compared with each other only if they are properly calibrated (if the involved smoothing parameter is not properly calibrated then the method may behave poorly). Thus each smoothing parameter is calibrated here so as to be optimal in some sense, see Section 4.2.

4.1. **The simulation experiment**

The regression functions we considered are the following:

$$\begin{aligned}
f_1(t) &= \exp(-8(t-0.5)^2) \\
f_2(t) &= \exp(-8(t-0.3)^2) \\
f_3(t) &= 0.05 + \exp(-16(t-0.5)^2) \\
f_4(t) &= \exp(-40(t-0.3)^2) + 1.5\exp(-10(t-0.7)^2) \\
f_5(t) &= \exp(-40(t-0.2)^2) + 1.5\exp(-10(t-0.7)^2).
\end{aligned}$$

For each regression function, we computed the parameters μ_0 and r_0 using a discretization of $[0,1]$ into 2000 equi-spaced points. We fixed $n = 100$, $\sigma = 0.1\,F(1)$ so that the signal/noise ratio remains stable and we generated $\varepsilon_1, \dots, \varepsilon_n$ as independent centered Gaussian variables with variance σ^2. Throughout the simulations, K is the quartic function defined by (2.5), $\hat{\sigma}_n^2$ is the optimal 5th-order sequence estimator defined by Hall *et al.* [9] and the estimator $\hat{f}_{n,h}$ in (2.2) is equal to g_n with bandwidth h. Moreover for the smoothed bootstrap, we chose ϕ to be the density function of a standard Gaussian law. For a given regression function and given smoothing parameters h_μ, h_n and h'_n, we built a confidence interval for μ_0 with asymptotic level $\alpha \in \{0.05, 0.1\}$ using each of the proposed methods (thus we built the confidence intervals given in (2.3), (2.6) and (3.9) for the three bootstrap constructions). The bootstrap resampling was performed $B = 250$ times for each sample. Then we estimated the probability that μ_0 lies in the confidence interval (and also the probability that μ_0 is below, resp. above, the confidence interval) using 3000 replications. We also used these 3000 replications to estimate the mean length of the confidence interval. We did the same work for the parameter r_0.

4.2. **Calibration of the smoothing parameters**

Since the regression function f is known (which of course is not the case in practice), we can calibrate the smoothing parameters so as to minimize a given risk. We first calibrated the smoothing parameters h_μ and h_r involved in the plug-in and in the studentized bootstrap procedures: we chose h_μ and h_r that approximately minimize

$$\mathbb{E}\left(\widehat{C}_{\mu,h} - C_\mu\right)^2 \quad \text{and} \quad \mathbb{E}\left(\widehat{C}_{r,h} - C_r\right)^2$$

TABLE 1. Smoothing parameters h_μ and h_r.

	f_1	f_2	f_3	f_4	f_5
h_μ	0.16	0.14	0.07	0.09	0.18
h_r	0.03	0.03	0.02	0.04	0.04

respectively over h, where C_μ and C_r are the constants that appear in the asymptotic distributions of μ_n and r_n. Precisely, we set $\mathcal{H} = \{0.01, 0.02, \ldots, 0.3\}$, we estimated the above expectations using 5000 simulations for every $h \in \mathcal{H}$ and then chose h_μ and h_r that minimize these estimated expectations. The values for h_μ and h_r we obtained are given in Table 1.

The calibration of the bootstrap smoothing parameter h_n for a given regression function was performed as follows. For notational convenience we set $\mu_{G_n^*} = \mu_{F_n^*}$ and $\mu_{G_n} = \mu_n$ when $h_n = 0$ so in that case, $\mu_{G_n^*}$ and μ_{G_n} are the estimator and parameter obtained with the data-driven bootstrap procedure described at the end of Section 2.3. We set $D = 50$ and $\alpha_i = i/D$ for every $i = 1, \ldots, D$. We fixed $S = 2000$ and for every $s = 1, \ldots, S$ we simulated observations, we computed the estimator $\mu_n(s)$ based on these observations and we performed the bootstrap step with a given bootstrap smoothing parameter h. For every $h \in \mathcal{H} \cup \{0\}$, let $q_i(s, h)$ be the quantile of order α_i of the 250 replications of $\mu_{G_n^*} - \mu_{G_n}$ thus obtained. In order to calibrate the ordinary parametric bootstrap for estimating μ_0, we chose h_n that minimizes

$$\sum_{i=1}^{D} \left(\frac{1}{S} \sum_{s=1}^{S} \mathbb{I}_{\mu_n(s) - \mu_0 \leqslant q_i(s,h)} - \alpha_i \right)^2$$

over $\mathcal{H} \cup \{0\}$. The studentized parametric bootstrap was calibrated in the same way with $q_i(s, h)$ defined as the quantile of order α_i of the 250 replications of the variable given in (3.10). The values of h_n we obtained are given in Table 2. We proceeded in the same way for calibrating h_n for the ordinary and studentized naive bootstraps except that we used $S = 3000$ simulations in that case. The values of h_n we obtained are given in Table 3. Finally for the smoothed bootstrap, we fixed h_n as in Table 3, $S = 2000$ and we chose h'_n that minimizes the above criterion over

$$h \in \{0.001, 0.002, \ldots, 0.01\} \cup \{0.01, 0.02, \ldots, 0.1\}.$$

The values of h'_n we obtained are given in Table 4. We did the same work for the parameter r_0.

4.3. Results for μ_0

In Tables 2–4, we give in column $\in$ the estimated probability that μ_0 belongs to the confidence interval obtained with a given method and in column $<$ (resp. $>$) the estimated probability that μ_0 is below (resp. above) this confidence interval. The values we used for h_n and h'_n are also given in these tables. Note first that the value $h_n = 0$ has never been chosen so the data-driven bootstrap is not recommended (this is consistent with our remark at the end of Sect. 2.3). Now compare the coverage probabilities given in columns $\in$. All of the proposed methods provide satisfactory results since the coverage probabilities are rather close to the target probability $1 - \alpha$. However, the plug-in method is the less efficient one: the coverage probabilities obtained with this method are always larger than $1 - \alpha$ so the confidence intervals are too long. The coverage probabilities are generally less close to $1 - \alpha$ with plug-in than with bootstrap. The different bootstraps perform quite similarly to each other but the smoothed bootstrap tends to be better than the others. In most cases, the studentized bootstrap is slightly better than the ordinary one. However surprisingly, the studentized bootstrap does not perform well when the regression function is f_4. It is also interesting to compare the probabilities given in

TABLE 2. Plug-in and parametric bootstrap: estimated probabilities that μ_0 belongs to the confidence interval with asymptotic level 0.05 (roman) or 0.1 (italic), computed with the smoothing parameter h_n.

	plug-in			ordinary bootstrap				studentized bootstrap			
	$<$	$\in$	$>$	$<$	$\in$	$>$	h_n	$<$	$\in$	$>$	h_n
f_1	0.006	0.980	0.014	0.022	0.962	0.015	0.09	0.025	0.955	0.020	0.15
	0.017	*0.943*	*0.040*	*0.043*	*0.919*	*0.038*		*0.046*	*0.911*	*0.042*	
f_2	0.011	0.969	0.020	0.028	0.953	0.019	0.08	0.024	0.951	0.026	0.10
	0.021	*0.923*	*0.056*	*0.048*	*0.913*	*0.040*		*0.045*	*0.900*	*0.055*	
f_3	0.007	0.983	0.010	0.013	0.976	0.011	0.07	0.028	0.963	0.009	0.08
	0.022	*0.948*	*0.030*	*0.031*	*0.943*	*0.026*		*0.044*	*0.943*	*0.013*	
f_4	0.007	0.974	0.019	0.025	0.959	0.016	0.07	0.019	0.907	0.074	0.08
	0.016	*0.936*	*0.048*	*0.045*	*0.924*	*0.031*		*0.032*	*0.873*	*0.095*	
f_5	0.001	0.986	0.012	0.022	0.956	0.021	0.08	0.023	0.945	0.032	0.11
	0.004	*0.962*	*0.034*	*0.040*	*0.921*	*0.039*		*0.041*	*0.906*	*0.053*	

TABLE 3. Naive bootstrap: estimated probabilities that μ_0 belongs to the confidence interval with asymptotic level 0.05 (roman) or 0.1 (italic), computed with the smoothing parameter h_n.

	ordinary bootstrap				studentized bootstrap			
	$<$	$\in$	$>$	h_n	$<$	$\in$	$>$	h_n
f_1	0.035	0.935	0.030	0.15	0.037	0.935	0.028	0.15
	0.056	*0.891*	*0.052*		*0.057*	*0.890*	*0.053*	
f_2	0.018	0.963	0.019	0.13	0.022	0.955	0.023	0.14
	0.038	*0.925*	*0.037*		*0.043*	*0.906*	*0.051*	
f_3	0.024	0.958	0.018	0.14	0.050	0.934	0.016	0.08
	0.049	*0.914*	*0.037*		*0.086*	*0.880*	*0.035*	
f_4	0.031	0.945	0.024	0.10	0.022	0.925	0.053	0.11
	0.053	*0.903*	*0.043*		*0.035*	*0.890*	*0.075*	
f_5	0.029	0.947	0.025	0.11	0.022	0.945	0.033	0.16
	0.054	*0.897*	*0.049*		*0.038*	*0.909*	*0.053*	

columns $<$ and $>$: it is expected that they are close to each other and close to $(1-\alpha/2)$. The plug-in method is the worst one with respect to this criterion: μ_0 is often above the plug-in confidence interval.

Finally, we give in Table 5 the estimated length of the confidence intervals. The plug-in method is the less efficient one with respect to this criterion. For each of the three bootstrap constructions, the studentized bootstrap is better than the ordinary one. Moreover, the naive and the smoothed bootstrap are very similar to each other and are better than the parametric bootstrap.

4.4. Results for r_0

Some of the results obtained for r_0 are given in Table 6. We did not reported the other results since they are not better than those given in Table 6. We conclude that asymptotics can not be used for the parameter r_0 when n is moderate: none of the proposed methods provides satisfactory results when $n = 100$.

TABLE 4. Smoothed bootstrap: estimated probabilities that μ_0 belongs to the confidence interval with asymptotic level 0.05 (roman) or 0.1 (italic), computed with the smoothing parameters h_n and h'_n.

	ordinary bootstrap					studentized bootstrap				
	$<$	$\in$	$>$	h_n	h'_n	$<$	$\in$	$>$	h_n	h'_n
f_1	0.027	0.949	0.024	0.15	0.02	0.030	0.948	0.022	0.15	0.03
	0.051	*0.903*	*0.046*			*0.051*	*0.899*	*0.049*		
f_2	0.021	0.963	0.016	0.13	0.001	0.024	0.956	0.020	0.14	0.009
	0.037	*0.923*	*0.040*			*0.043*	*0.905*	*0.051*		
f_3	0.026	0.955	0.019	0.14	0.002	0.041	0.949	0.010	0.08	0.03
	0.049	*0.915*	*0.036*			*0.065*	*0.913*	*0.021*		
f_4	0.031	0.945	0.024	0.10	0.006	0.021	0.927	0.052	0.11	0.002
	0.054	*0.904*	*0.042*			*0.036*	*0.889*	*0.075*		
f_5	0.031	0.941	0.027	0.11	0.008	0.021	0.949	0.030	0.16	0.001
	0.054	*0.899*	*0.047*			*0.038*	*0.910*	*0.052*		

TABLE 5. Estimated mean length of the confidence intervals for μ_0 with asymptotic level 0.05 (roman) or 0.1 (italic).

	Plug-in	parametric bootstrap		naive bootstrap		smoothed bootstrap	
		ordinary	studentized	ordinary	studentized	ordinary	studentized
f_1	0.089	0.085	0.079	0.075	0.075	0.078	0.077
	0.076	*0.073*	*0.067*	*0.064*	*0.063*	*0.066*	*0.065*
f_2	0.086	0.086	0.080	0.082	0.080	0.082	0.080
	0.074	*0.073*	*0.067*	*0.070*	*0.067*	*0.070*	*0.067*
f_3	0.054	0.060	0.061	0.052	0.045	0.052	0.051
	0.046	*0.052*	*0.049*	*0.044*	*0.037*	*0.044*	*0.041*
f_4	0.091	0.088	0.077	0.077	0.073	0.077	0.073
	0.077	*0.074*	*0.063*	*0.065*	*0.060*	*0.065*	*0.060*
f_5	0.082	0.076	0.068	0.068	0.068	0.068	0.068
	0.070	*0.065*	*0.058*	*0.058*	*0.058*	*0.058*	*0.058*

5. PROOFS

5.1. Some preliminary results

Lemma 5.1 below is useful to prove that F_n and G_n^* are uniformly close to a Brownian motion with smooth deterministic drift. The approximation result concerning G_n^* is stated in Lemma 5.2. Finally, Lemma 5.3 is an analytic tool that is useful to prove Theorems 2.1 and 2.2.

Lemma 5.1. *Let $\varepsilon_1, \dots, \varepsilon_n$ be i.i.d. variables with mean zero and variance 1. Assume $\mathbb{E}|\varepsilon_1|^p < C$ for some $C > 0$ and $p \geqslant 2$. If the ε_i's are defined on some rich enough probability space then there exist some standard*

Table 6. Plug-in and parametric bootstrap: estimated probabilities that r_0 belongs to the confidence interval with asymptotic level 0.05 (roman) or 0.1 (italic), computed with the smoothing parameter h_n.

	plug-in			ordinary bootstrap				studentized bootstrap			
	$<$	$\in$	$>$	$<$	$\in$	$>$	h_n	$<$	$\in$	$>$	h_n
f_1	0.683	0.317	0.000	0.400	0.299	0.301	0.05	0.231	0.754	0.014	0.22
	0.683	*0.317*	*0.000*	*0.403*	*0.279*	*0.318*		*0.441*	*0.507*	*0.052*	
f_2	0.239	0.761	0.000	0.259	0.476	0.265	0.10	0.259	0.407	0.335	0.10
	0.239	*0.761*	*0.000*	*0.259*	*0.443*	*0.298*		*0.259*	*0.348*	*0.393*	
f_3	0.195	0.805	0.000	0.192	0.751	0.058	0.10	0.189	0.807	0.004	0.12
	0.195	*0.805*	*0.000*	*0.192*	*0.751*	*0.058*		*0.190*	*0.804*	*0.006*	
f_4	0.176	0.821	0.003	0.183	0.769	0.048	0.14	0.188	0.782	0.030	0.12
	0.176	*0.821*	*0.003*	*0.190*	*0.673*	*0.137*		*0.192*	*0.670*	*0.138*	
f_5	0.090	0.899	0.010	0.098	0.891	0.011	0.09	0.099	0.890	0.011	0.05
	0.090	*0.898*	*0.012*	*0.101*	*0.883*	*0.016*		*0.099*	*0.886*	*0.015*	

Brownian motion W_n and some positive number C_p that only depends on p and C such that for all $x > 0$

$$\mathbb{P}\left(\sup_{t\in[0,1]}\left|\frac{1}{n}\sum_{i\leqslant nt}\varepsilon_i - \frac{1}{\sqrt{n}}W_n(t)\right| > x\right) \leqslant C_p n^{1-p}x^{-p}.$$

Proof. By Sakhanenko's [13] construction there exist some standard Brownian motion B_n and some $A_p > 0$ that only depends on p such that

$$\mathbb{E}\left(\sup_{1\leqslant k\leqslant n}\left|\sum_{i=1}^{k}\varepsilon_i - B_n(k)\right|^p\right) \leqslant A_p n\mathbb{E}|\varepsilon_i|^p,$$

provided the ε_i's are defined on some rich enough probability space. For every $u \geqslant 0$, let $[u]$ denote the integer part of u. By exponential inequality we have for every $x > 0$

$$\mathbb{P}\left(\sup_{t\in[0,1]}|B_n(nt) - B_n([nt])| > nx\right) \leqslant 2n\exp\left(-\frac{n^2x^2}{2}\right).$$

There thus exists some $C_p > 0$ that does not depend on n such that for every $x > 0$

$$\begin{aligned}\mathbb{P}\left(\sup_{t\in[0,1]}\left|\sum_{i\leqslant nt}\varepsilon_i - B_n(nt)\right| > 2nx\right) &\leqslant 2n\exp\left(-\frac{n^2x^2}{2}\right) + A_p n^{1-p}\mathbb{E}|\varepsilon_i|^p x^{-p}\\ &\leqslant C_p n^{1-p}x^{-p}.\end{aligned}$$

Setting $W_n(t) = B_n(nt)/\sqrt{n}$ yields Lemma 5.1. □

Lemma 5.2. *Under the assumptions of Theorem 2.2,*

$$\sup_{t\in[h_n,1-h_n]}|g_n(t) - f(t)| = o_{\mathbb{P}}(1) \quad \textit{and} \quad \sup_{t\in[h_n,1-h_n]}|g_n'(t) - f'(t)| = o_{\mathbb{P}}(1). \tag{5.1}$$

Moreover, there exists some $\mathbb{P}^*$-Brownian motion W_n^* such that

$$n^{2/3} \sup_{t\in[0,1]} \left| G_n^*(t) - G_n(t) - \sigma_n^* \frac{W_n^*(t)}{\sqrt{n}} \right| = o_{\mathbb{P}}(1). \tag{5.2}$$

Proof. We assume without loss of generality that $y_1, \ldots, y_n$ are defined on some rich enough probability space. Both f and f' are assumed to be bounded so

$$\sup_{t\in[0,1]} \left| \frac{1}{n} \sum_{i\leqslant nt} f(t_i) - F(t) \right| = O\left(\frac{1}{n}\right).$$

By Lemma 5.1 there thus exists some standard Brownian motion W_n such that

$$n^{2/3} \sup_{t\in[0,1]} \left| F_n(t) - F(t) - \frac{\sigma}{\sqrt{n}} W_n(t) \right| = o_{\mathbb{P}}(1). \tag{5.3}$$

Let H denote either K' or K''. Then $\int H = 0$, H is bounded and vanishes outside $[-1,1]$. There thus exists some $c > 0$ such that

$$\sup_{t\in[0,1]} \left| \int_{\mathbb{R}} W_n(t - xh_n) H(x) \mathrm{d}x \right| \leqslant c \sup_{t\in[0,1],\ |t-u|\leqslant h_n} |W_n(t) - W_n(u)|,$$

where the latter term is of order of magnitude $O_{\mathbb{P}}((h_n \log(1/h_n))^{1/2})$. It is assumed that $h_n^{-1} = o\,(n^{\alpha})$ for some $\alpha < 1/3$ so

$$\frac{1}{\sqrt{n} h_n^2} \sup_{t\in[0,1]} \left| \int_{\mathbb{R}} W_n(t - xh_n) H(x) \mathrm{d}x \right| = o_{\mathbb{P}}(1).$$

By (5.3) we thus have

$$\sup_{t\in[h_n,1-h_n]} \left| g_n(t) - \frac{1}{h_n} \int_{-1}^{1} F(t - xh_n) K'(x) \mathrm{d}x \right| = o_{\mathbb{P}}(1)$$

and

$$\sup_{t\in[h_n,1-h_n]} \left| g_n'(t) - \frac{1}{h_n^2} \int_{-1}^{1} F(t - xh_n) K''(x) \mathrm{d}x \right| = o_{\mathbb{P}}(1).$$

It is assumed that f' is bounded so $F(t - xh_n) = F(t) - xh_n f(t) + O(h_n^2)$ where $O(h_n^2)$ is uniform in t and $x \in [-1,1]$. We obtain

$$\sup_{t\in[h_n,1-h_n]} |g_n(t) - f(t)| = o_{\mathbb{P}}(1)$$

by using $\int K' = 0$ and $\int_{\mathbb{R}} xK'(x) = -1$. Since f' is Hölderian we also have

$$F(t - xh_n) = F(t) - xh_n f(t) + x^2 h_n^2 f'(t)/2 + o(h_n^2),$$

where $o(h_n^2)$ is uniform in t and $x \in [-1,1]$. But K'' is symmetric about zero so $\int_{\mathbb{R}} xK''(x)dx = 0$ and we get

$$\sup_{t\in[h_n,1-h_n]} |g_n'(t) - f'(t)| = o_{\mathbb{P}}(1),$$

which completes the proof of (5.1).

For all $t \leqslant 0$, $F_n(t) = 0$ so

$$\sup_{t\in[0,h_n]} |g_n(t)| = \sup_{t\in[0,h_n]} \frac{1}{h_n} \left| \int_{-1}^{t/h_n} F_n(t - xh_n) K'(x) \, \mathrm{d}x \right|,$$

which is stochastically bounded since $F(t - xh_n) \leqslant |t - xh_n| \sup_s |f(s)|$. One can prove in the same way that $\sup_{t\in[1-h_n,1]} |g_n(t)|$ is stochastically bounded so

$$\sup_{t\in[0,1]} |g_n(t)| = O_{\mathbb{P}}(1). \tag{5.4}$$

We also obtain in the same way that $\sup_{t\in[0,1]} |g'_n(t)| = O_{\mathbb{P}}(h_n^{-1})$ and therefore,

$$\sup_{t\in[0,1]} \left| \frac{1}{n} \sum_{i\leqslant nt} g_n(t_i) - G_n(t) \right| = o_{\mathbb{P}} \left(n^{-2/3} \right). \tag{5.5}$$

But under the assumptions of Theorem 2.2, there exists some $C > 0$ such that

$$\lim_{n\to\infty} \mathbb{P} \left(\mathbb{E}^* |\varepsilon_1^* / \sigma_n^*|^p < C \right) = 1.$$

Lemma 5.2 then follows from Lemma 5.1 (where ε_i stands for $\varepsilon_i^*/\sigma_n^*$) and from (5.5). □

Lemma 5.3. *Let $H : [0,1] \to \mathbb{R}$. If r_H is well defined then under the assumptions of Theorem 2.1,*

$$|r_H - r_0| \leqslant C \sup_{t\in[0,1]} |H(t) - F(t)|$$

for some $C > 0$ that only depends on f.

Proof. Let c be some positive real number with $c < 2 \inf_t f(t)$ and let Δ and δ be defined by $\Delta = \sup_t |H(t) - F(t)|$ and $\delta = 3\Delta/c$ respectively. It follows from Taylor's expansion that for all positive γ,

$$\sup_\mu \{ F(\mu + r_0 - \gamma) - F(\mu + r_0) - F(\mu - r_0 + \gamma) + F(\mu - r_0) \} < -c\gamma.$$

By definition of r_0 and μ_0, $F(\mu+r_0) - F(\mu-r_0) \leqslant F(\mu_0+r_0) - F(\mu_0-r_0)$ for any μ and $F(\mu_0+r_0) - F(\mu_0-r_0) = \eta F(1)$ and therefore

$$\sup_\mu \{ F(\mu + r_0 - \gamma) - F(\mu - r_0 + \gamma) \} < \eta F(1) - c\gamma.$$

We thus have for all positive γ

$$\sup_\mu \{ H(\mu + r_0 - \gamma) - H(\mu - r_0 + \gamma) \} < \eta H(1) - c\gamma + 3\Delta,$$

which implies

$$\sup_\mu \{ H(\mu + r_0 - \gamma) - H(\mu - r_0 + \gamma) \} < \eta H(1)$$

for all $\gamma \geqslant \delta$. So $r_0 - \delta \leqslant r_H$. It follows from Taylor's expansion that

$$F(\mu_0 + r_0 + \delta) - F(\mu_0 + r_0) - F(\mu_0 - r_0 - \delta) + F(\mu_0 - r_0) > c\delta,$$

and therefore

$$H(\mu_0 + r_0 + \delta) - H(\mu_0 - r_0 - \delta) > \eta H(1) + c\delta - 3\Delta.$$

So $r_0 + \delta \geqslant r_H$. We thus have $|r_H - r_0| \leqslant \delta$, which proves the lemma. □

5.2. **Proof of Theorem 2.1**

We assume without loss of generality that $\sigma = 1$ and that $\varepsilon_1, \dots, \varepsilon_n$ are defined on some rich enough probability space so that Lemma 5.1 applies. Then there exists some standard Brownian motion W_n such that

$$\sup_{t\in[0,1]} \left| F_n(t) - F(t) - \frac{1}{\sqrt{n}} W_n(t) \right| = o_{\mathbb{P}}\left(n^{-2/3}\right). \tag{5.6}$$

Therefore, $\sup_t |F_n(t) - F(t)| = O_{\mathbb{P}}(n^{-1/2})$ so by Lemma 5.3 (where H stands for F_n)

$$r_n - r_0 = O_{\mathbb{P}}\left(n^{-1/2}\right). \tag{5.7}$$

For every $a > 0$, there exists some $\varepsilon > 0$ such that

$$\sup_{|x|\geqslant a} \{F(\mu_0 + r_0 + x) - F(\mu_0 + r_0) - F(\mu_0 - r_0 + x) + F(\mu_0 - r_0)\} < -\varepsilon, \tag{5.8}$$

since the supremum in the latter inequality is achieved and since μ_0 and r_0 are uniquely defined. By (5.7) it follows that for every $a > 0$, there exists some $\varepsilon > 0$ such that

$$\sup_{|x|\geqslant a} \{F(\mu_0 + r_n + x) - F(\mu_0 + r_n) - F(\mu_0 - r_n + x) + F(\mu_0 - r_n)\} < -\varepsilon + o_{\mathbb{P}}(1).$$

By assumption μ_0 maximizes $\mu \mapsto F(\mu+r_0)-F(\mu-r_0)$ and the supremum is achieved in the open set $(r_0, 1-r_0)$ so $f(\mu_0 + r_0) = f(\mu_0 - r_0)$. We recall furthermore that f' is Hölderian and that

$$f'(\mu_0 + r_0) < f'(\mu_0 - r_0).$$

By using again Taylor's expansion and (5.7), we obtain that there exist some $a > 0$ and $b > 0$ such that for any $C > 0$ and any t with $C \leqslant |t| \leqslant an^{1/3}$,

$$F\left(\mu_0 + r_n + tn^{-1/3}\right) - F(\mu_0 + r_n) - F\left(\mu_0 - r_n + tn^{-1/3}\right) + F(\mu_0 - r_n)$$
$$< -bt^2 n^{-2/3} + t^2 n^{-2/3} o_{\mathbb{P}}(1).$$

Here, $o_{\mathbb{P}}(1)$ is uniform in t, where $C \leqslant |t| \leqslant an^{1/3}$. So there exists some $b > 0$ such that for any $C > 0$ and any t with $|t| \geqslant C$,

$$F\left(\mu_0 + r_n + tn^{-1/3}\right) - F(\mu_0 + r_n) - F\left(\mu_0 - r_n + tn^{-1/3}\right) + F(\mu_0 - r_n)$$
$$< -bt^2 n^{-2/3} + t^2 n^{-2/3} o_{\mathbb{P}}(1), \tag{5.9}$$

where $o_{\mathbb{P}}(1)$ is uniform in t, $|t| \geqslant C$. Let

$$I_n = \left[n^{1/3}(r_n - \mu_0), n^{1/3}(1 - r_n - \mu_0)\right]$$

and let M_n be the process defined for $t \in I_n$ by

$$M_n(t) = n^{2/3}\left\{F_n\left(\mu_0 + r_n + tn^{-1/3}\right) - F_n\left(\mu_0 - r_n + tn^{-1/3}\right)\right\}.$$

Then $n^{1/3}(\mu_n - \mu_0)$ is the location of the maximum of M_n so for every $C > 0$,

$$\mathbb{P}\left(n^{1/3}|\mu_n - \mu_0| > C\right) \leqslant \mathbb{P}\left(\sup_{|t|>C,\ t\in I_n}\{M_n(t)\} \geqslant M_n(0)\right).$$

By scaling and time homogeneity properties of Brownian motion, it follows from (5.7) that the process

$$\left\{n^{1/6}\left(W_n\left(\mu_0 + r_n + tn^{-1/3}\right) - W_n\left(\mu_0 + r_n\right)\right),\ t \in I_n\right\}$$

is asymptotically identical in law to a Brownian motion restricted to I_n. It thus follows from (5.6), (5.7) and (5.9) that there exists some $b > 0$ such that for every $C > 0$

$$\mathbb{P}\left(n^{1/3}|\mu_n - \mu_0| > C\right) \leqslant 2\mathbb{P}\left(\sup_{|t|\geqslant C}\left\{W_n(t) - bt^2\right\} \geqslant 0\right) + o(1). \tag{5.10}$$

By using time inversion and scaling properties of Brownian motion, one can easily prove that for all positive numbers b and C,

$$\mathbb{P}\left(\sup_{|t|\geqslant C}\left\{W_n(t) - bt^2\right\} \geqslant 0\right) \leqslant 2\exp(-b^2C^3/2). \tag{5.11}$$

Therefore, the right hand term of (5.10) converges to zero as n and C go to infinity, which proves that

$$\mu_n - \mu_0 = O_{\mathbb{P}}(n^{-1/3}).$$

We now derive the asymptotic distributions of μ_n and r_n. Let M be the process defined for $|t| \leqslant \log n$ by

$$M(t) = -t^2\frac{f'(\mu_0 - r_0) - f'(\mu_0 + r_0)}{2} + \sqrt{2}W_1(t),$$

where the process $W_1(t)$ is defined for $t \in [-\log n, \log n]$ by

$$\frac{n^{1/6}}{\sqrt{2}}\left(W_n\left(\mu_0 + r_0 + tn^{-1/3}\right) - W_n\left(\mu_0 + r_0\right) - W_n\left(\mu_0 - r_0 + tn^{-1/3}\right) + W_n\left(\mu_0 - r_0\right)\right). \tag{5.12}$$

In the sequel, we assume n large enough so that W_1 is a standard Brownian motion restricted to $[-\log n, \log n]$. Let T_n and τ_n be defined by

$$T_n = \operatorname*{argmax}_{|t|\leqslant\log n}\{M_n(t) - M_n(0)\} \quad \text{and} \quad \tau_n = \operatorname*{argmax}_{|t|\leqslant\log n}\{M(t)\}.$$

Note that according to scaling property of Brownian motion, the distribution of the location of the maximum of $\{W_1(t) - c_0^{3/2}t^2,\ t \in \mathbb{R}\}$ is identical to that of $c_0^{-1}\tau$, where

$$c_0 = \frac{1}{2}\left(f'(\mu_0 - r_0) - f'(\mu_0 + r_0)\right)^{2/3}. \tag{5.13}$$

Similar to (5.11), the probability that τ_n differs from the location of the maximum of $\{W_1(t) - c_0^{3/2}t^2,\ t \in \mathbb{R}\}$ tends to zero as n goes to infinity. Therefore, τ_n converges in distribution to $c_0^{-1}\tau$ as n goes to infinity. Since $f(\mu_0 + r_0) = f(\mu_0 - r_0)$, it follows from (5.6) that

$$\sup_{|t|\leqslant\log n}|M_n(t) - M_n(0) - M(t)|$$

converges in probability to zero as n goes to infinity so $T_n - \tau_n$ converges to zero in probability. Finally, $n^{1/3}(\mu_n - \mu_0)$ differs from T_n if and only if $n^{1/3}|\mu_n - \mu_0| > \log n$. But $\mu_n - \mu_0 = O_{\mathbb{P}}(n^{-1/3})$ so we get

$$n^{1/3}(\mu_n - \mu_0) - T_n \xrightarrow{\mathcal{D}} 0 \quad \text{as } n \to \infty$$

and we obtain

$$n^{1/3}(\mu_n - \mu_0) \xrightarrow{\mathcal{D}} c_0^{-1}\tau \quad \text{as } n \to \infty.$$

Let U_n be the random variable defined by

$$U_n = \inf\left\{t,\ |t| \leqslant \log n,\ F_n\left(\mu_n + r_0 + tn^{-1/2}\right) - F_n\left(\mu_n - r_0 - tn^{-1/2}\right) \geqslant \eta F_n(1)\right\}. \tag{5.14}$$

We have $f(\mu_0 - r_0) = f(\mu_0 + r_0)$ and $\eta F(1) = F(\mu_0 + r_0) - F(\mu_0 - r_0)$. Since $\mu_n = \mu_0 + O_{\mathbb{P}}(n^{-1/3})$, we thus obtain by using (5.6), Taylor's expansion and standard properties of Brownian motion that U_n is the smallest t such that $|t| \leqslant \log n$ and

$$2tf(\mu_0 + r_0) + W_n(\mu_0 + r_0) - W_n(\mu_0 - r_0) \geqslant \eta W_n(1) + o_{\mathbb{P}}(1).$$

So U_n converges in probability as n goes to infinity towards

$$U = \frac{1}{2f(\mu_0 + r_0)}\left(\eta W(1) - W(\mu_0 + r_0) + W(\mu_0 - r_0)\right)$$

where W is a standard Brownian motion. By definition of r_n and μ_n we have

$$r_n = \inf\{r \geqslant 0,\ F_n(\mu_n + r) - F_n(\mu_n - r) \geqslant \eta F_n(1)\}$$

and therefore,

$$\sqrt{n}(r_n - r_0) = \inf\left\{t,\ F_n\left(\mu_n + r_0 + tn^{-1/2}\right) - F_n\left(\mu_n - r_0 - tn^{-1/2}\right) \geqslant \eta F_n(1)\right\}.$$

It thus follows from (5.7) that $n^{1/2}(r_n - r_0) - U_n$ converges in probability to zero as n goes to infinity. So $n^{1/2}(r_n - r_0)$ converges in distribution to U as n goes to infinity. For any real numbers s and t, $\mathrm{cov}(W(s), W(t)) = \min(s,t)$, so $2f(\mu_0 + r_0)U$ is a centered Gaussian variable with variance $\eta^2 + 2r_0(1 - 2\eta)$, which completes the proof of Theorem 2.1.

5.3. **Proof of Theorem 2.2**

We assume without loss of generality $\sigma = 1$ and there exists some $\mathbb{P}^*$-Brownian motion W_n^* such that (5.2) holds. We assume moreover that there exists some $\mathbb{P}$-Brownian motion W_n with (5.6). Expanding F proves that G_n converges in probability to F in the supremum-distance sense, which ensures that r_{G_n} and μ_{G_n} are well defined with probability that tends to one. Moreover, $r_{G_n} - r_0 = o_{\mathbb{P}}(1)$ by Lemma 5.3. Let C be some positive number. By definition, we can have $|\mu_{G_n} - \mu_0| > C$ only if

$$\sup_{|t|>C}\{G_n(\mu_0 + t + r_{G_n}) - G_n(\mu_0 + t - r_{G_n})\} \geqslant G_n(\mu_0 + r_{G_n}) - G_n(\mu_0 - r_{G_n}).$$

The latter inequality implies

$$\sup_{|t|>C}\{F(\mu_0 + t + r_0) - F(\mu_0 + t - r_0)\} \geqslant F(\mu_0 + r_0) - F(\mu_0 - r_0) + o_{\mathbb{P}}(1).$$

It thus follows from (5.8) that $\mu_{G_n} - \mu_0 = o_{\mathbb{P}}(1)$. By (5.2), $G_n^* - F$ converges in probability to zero in the supremum-distance sense, so we obtain in the same way as above $\mu_{G_n^*} - \mu_0 = o_{\mathbb{P}}(1)$ and $r_{G_n^*} - r_0 = o_{\mathbb{P}}(1)$. The kernel K is assumed to be symmetric about zero so $\int xK(x)\,\mathrm{d}x = 0$. Moreover, $h_n = o(n^{-1/6}/\log n)$ so

$$\sup_{t\in[h_n,1-h_n]} |G_n(t) - F(t)| = o_{\mathbb{P}}\left(n^{-1/3}/(\log n)^2\right).$$

Lemma 5.3 then yields $r_{G_n} - r_0 = o_{\mathbb{P}}\left(n^{-1/3}/(\log n)^2\right)$ and $r_{G_n^*} - r_0 = o_{\mathbb{P}}(n^{-1/3}/(\log n)^2)$.

It is assumed that $[\mu_0 - r_0, \mu_0 + r_0] \subset (0,1)$ so we have with probability that tends to one $[\mu_{G_n} - r_{G_n}, \mu_{G_n} + r_{G_n}] \subset (0,1)$ and $g_n(\mu_{G_n} + r_{G_n}) = g_n(\mu_{G_n} - r_{G_n})$ (since μ_{G_n} maximizes $\mu \mapsto G_n(\mu + r_{G_n}) - G_n(\mu - r_{G_n})$ in an open set). We thus assume in the sequel

$$g_n(\mu_{G_n} + r_{G_n}) = g_n(\mu_{G_n} - r_{G_n}).$$

Let

$$I_n^* = \left[n^{1/3}(h_n + r_{G_n^*} - \mu_{G_n}), n^{1/3}(1 - h_n - r_{G_n^*} - \mu_{G_n})\right].$$

By Taylor's expansion and (5.1) there exist some positive a and b such that for all $C > 0$ and all $t \in I_n^*$ with $C < |t| < an^{1/3}$,

$$G_n\left(\mu_{G_n} + r_{G_n^*} + tn^{-1/3}\right) - G_n(\mu_{G_n} + r_{G_n^*}) - G_n\left(\mu_{G_n} - r_{G_n^*} + tn^{-1/3}\right)$$
$$+G_n(\mu_{G_n} - r_{G_n^*}) < -bt^2n^{-2/3} + t^2n^{-2/3}o_{\mathbb{P}}(1),$$

where $o_{\mathbb{P}}(1)$ is uniform in t. Since G_n converges to F in the supremum-distance sense, it follows from (5.8) and the previous inequality that there exists some positive b such that for all $C > 0$ and all $t \in I_n^*$ with $|t| > C$,

$$G_n\left(\mu_{G_n} + r_{G_n^*} + tn^{-1/3}\right) - G_n\left(\mu_{G_n} + r_{G_n^*}\right) - G_n\left(\mu_{G_n} - r_{G_n^*} + tn^{-1/3}\right)$$
$$+G_n(\mu_{G_n} - r_{G_n^*}) < -bt^2n^{-2/3} + t^2n^{-2/3}o_{\mathbb{P}}(1), \tag{5.15}$$

where $o_{\mathbb{P}}(1)$ is uniform in t. Let M_n^* be the process defined for $t \in I_n^*$ by

$$M_n^*(t) = n^{2/3}\left\{G_n^*\left(\mu_{G_n} + r_{G_n^*} + tn^{-1/3}\right) - G_n^*\left(\mu_{G_n} - r_{G_n^*} + tn^{-1/3}\right)\right\}.$$

With probability that tends to one, $[r_{G_n^*} - \mu_{G_n^*}, r_{G_n^*} + \mu_{G_n^*}] \subset [h_n, 1-h_n]$ so we may assume that $n^{1/3}(\mu_{G_n^*} - \mu_{G_n})$ is the location of the maximum of M_n^*. Therefore, for any $C > 0$,

$$\mathbb{P}^*\left(n^{1/3}|\mu_{G_n^*} - \mu_{G_n}| > C\right) \leqslant \mathbb{P}^*\left(\sup_{|t|>C,\ t\in I_n^*}\{M_n^*(t)\} \geqslant M_n^*(0)\right).$$

With probability that tends to one, $n^{1/3}(\log n)^2(r_{G_n^*} - r_{G_n}) \leqslant 1$ so it follows from scaling and time-homogeneity properties of Brownian motion that the process

$$\left\{n^{1/6}\left(W_n^*\left(\mu_{G_n} + r_{G_n^*} + tn^{-1/3}\right) - W_n^*\left(\mu_{G_n} + r_{G_n^*}\right)\right),\ t \in I_n^*\right\}$$

is asymptotically distributed under $\mathbb{P}^*$ as a restricted Brownian motion. By using (5.15) and the same arguments as in the proof of Theorem 2.1 we thus obtain

$$n^{1/3}(\mu_{G_n^*} - \mu_{G_n}) = O_{\mathbb{P}^*}(1), \tag{5.16}$$

so $n^{1/3}(\mu_{G_n^*} - \mu_{G_n})$ has the same asymptotic distribution as the location of the maximum of $\{M_n^*(t) - M_n^*(0),\ t \in [-\log n, \log n]\}$ and we obtain

$$n^{1/3}(\mu_{G_n^*} - \mu_{G_n}) \xrightarrow{\mathcal{D}^*} {c_0}^{-1}\tau \quad \textit{in probability}$$

as in the proof of Theorem 2.1. Let U_n^* be the random variable defined as the smallest t with $|t| \leqslant n^{1/6}$ and

$$G_n^*\left(\mu_{G_n^*} + r_{G_n} + tn^{-1/2}\right) - G_n^*\left(\mu_{G_n^*} - r_{G_n} - tn^{-1/2}\right) \geqslant \eta G_n^*(1).$$

We have $g_n(\mu_{G_n} + r_{G_n}) = g_n(\mu_{G_n} - r_{G_n})$ and

$$\eta G_n(1) = G_n(\mu_{G_n} + r_{G_n}) - G_n(\mu_{G_n} - r_{G_n})$$

with probability that tends to one. Moreover it follows from Lemma 5.2 that $\sup_{t\in[h_n,1-h_n]} |g_n'(t)|$ is stochastically bounded. By (5.2) and (5.16) , U_n^* is thus the smallest t such that $|t| \leqslant n^{1/6}$ and

$$2tg_n(\mu_{G_n} + r_{G_n}) + \sigma_n^* W_n^*(\mu_{G_n} + r_{G_n}) - \sigma_n^* W_n^*(\mu_{G_n} - r_{G_n}) \geqslant \eta\sigma_n^* W_n^*(1) + o_{\mathbb{P}^*}(1).$$

By Lemma 5.2, $g_n(\mu_{G_n} + r_{G_n})$ approaches $f(\mu_{G_n} + r_{G_n})$ as n goes to infinity. Moreover, μ_{G_n} and r_{G_n} converge in probability to μ_0 and r_0 respectively and σ_n^* converges in probability to σ, so we obtain the asymptotic distribution of $\sqrt{n}(r_{G_n^*} - r_{G_n})$ by using the same arguments as in the proof of Theorem 2.1. □

5.4. Proof of Proposition 3.1

Note first that the bootstrap residuals are conditionally i.i.d. and that $\mathbb{E}^*(\varepsilon_i^*) = 0$ with the three bootstrap constructions.

If the bootstrap residuals are generated with the parametric bootstrap then their common variance under $\mathbb{P}^*$ is $\sigma_n^{*2} = \hat{\sigma}_n^2$, which stochastically converges to σ^2. Moreover, $\varepsilon_i^*/\hat{\sigma}_n$ is well defined and standard Gaussian with probability that tends to one so (2.4) holds for some $C > 0$.

If the bootstrap residuals are generated with the naive bootstrap then $\sigma_n^{*2} = \tilde{\sigma}_n^2$, where

$$\tilde{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^n \tilde{\varepsilon}_i^2 = \frac{1}{n}\sum_{i=1}^n \hat{\varepsilon}_i^2 - \left(\frac{1}{n}\sum_{i=1}^n \hat{\varepsilon}_i\right)^2.$$

It follows from (5.1), (5.4) and law of large numbers that $\tilde{\sigma}_n^2$ stochastically converges to σ^2. Moreover

$$\mathbb{E}^*|\varepsilon_i^*|^p = \frac{1}{n}\sum_{i=1}^n \left|\hat{\varepsilon}_i - \frac{1}{n}\sum_{i=1}^n \hat{\varepsilon}_i\right|^p \leqslant 4^{p-1}\left(\frac{1}{n}\sum_{i=1}^n (|\varepsilon_i|^p + |g_n(t_i) - f(t_i)|^p)\right) + 2^{p-1}\left|\frac{1}{n}\sum_{i=1}^n \hat{\varepsilon}_i\right|^p$$

so we obtain with the same arguments that (2.4) holds for some $C > 0$.

If the bootstrap residuals are generated with the smoothed bootstrap then $\sigma_n^{*2} = \tilde{\sigma}_n^2 + h_n'^2$ stochastically converges to σ^2 provided $h_n' \to 0$. Moreover, denoting by $V_1, \dots, V_n$ independent variables with common density ϕ we get

$$\mathbb{E}^*|\varepsilon_i^*|^p = \frac{1}{n}\sum_{i=1}^n |\tilde{\varepsilon}_i + h_n' V_i|^p \leqslant 2^{p-1}\left(\frac{1}{n}\sum_{i=1}^n |\tilde{\varepsilon}_i|^p + \frac{{h_n'}^p}{n}\sum_{i=1}^n |V_i|^p\right)$$

so (2.4) holds for some $C > 0$.

5.5. **Proof of (2.7) and (2.8)**

Arguments involved here are close to those involved in the proof of Theorem 2.1 so we do not explain them in full details. Once again, we assume $\sigma = 1$ and (5.3) holds for some $\mathbb{P}$-Brownian motion W_n. By Sakhanenko's construction we may assume furthermore that there exists some $\mathbb{P}^*$-Brownian motion W_n^* such that

$$n^{2/3} \sup_{t\in[0,1]} \left| F_n^*(t) - F_n(t) - \sigma_n^* \frac{W_n^*(t)}{\sqrt{n}} \right| = o_{\mathbb{P}}(1). \tag{5.17}$$

For every $a > 0$ there exists some $\varepsilon > 0$ with (5.8). Moreover, by Theorem 2.1 and Lemma 5.3 we have

$$n^{1/3}(\mu_n - \mu_0) = O_{\mathbb{P}}(1) \quad \text{and} \quad \sqrt{n}(r_{F_n^*} - r_0) = O_{\mathbb{P}}(1).$$

So for every $a > 0$ there exists some $\varepsilon > 0$ such that

$$\sup_{|x|\geqslant a} \left\{ F(\mu_n + r_{F_n^*} + x) - F(\mu_n + r_{F_n^*}) - F(\mu_n - r_{F_n^*} + x) + F(\mu_n - r_{F_n^*}) \right\} < -\varepsilon + o_{\mathbb{P}}(1).$$

Fix $\alpha \in (0, 1/3)$. It follows from the latter inequality, (5.3) and Taylor's expansion that there exists some $b > 0$ such that

$$\begin{aligned} F_n(\mu_n + r_{F_n^*} + tn^{-1/3}) - F_n(\mu_n + r_{F_n^*}) - F_n(\mu_n - r_{F_n^*} + tn^{-1/3}) + F_n(\mu_n - r_{F_n^*}) \\ \leqslant -bn^{-2/3}t^2 + t^2 n^{-2/3} o_{\mathbb{P}}(1) \end{aligned}$$

for all t with $|t| \geqslant n^{\alpha}$. Let $\beta \in (\alpha, 1/3)$, let M_n^* be the process defined by

$$M_n^*(t) = \left\{ n^{2/3} \left(F_n^* \left(\mu_n + r_{F_n^*} + tn^{-1/3} \right) - F_n^* \left(\mu_n - r_{F_n^*} + tn^{-1/3} \right) \right) \right\}$$

and let T_n^* be the location of the maximum of $\{M_n^*(t),\ |t| \leqslant n^{\beta}\}$. Then, $n^{1/3}(\mu_{F_n^*} - \mu_n)$ is the location of the maximum of M_n^* so we have $n^{1/3}(\mu_{F_n^*} - \mu_n) = T_n^*$ with probability that tends to one. By expanding F in both neighbourhoods of $\mu_0 + r_0$ and $\mu_0 - r_0$ and by setting β small enough, we obtain that T_n^* has the same distribution as the location of the maximum of

$$\left\{ -c_0^{3/2}(n^{1/3}(\mu_n - \mu_0) + t)^2 + W_1(n^{1/3}(\mu_n - \mu_0) + t) + \sigma_n^* W_n^*(t) + R_n(t),\ |t| \leqslant n^{\beta} \right\},$$

where R_n is negligeable and where c_0 and W_1 are given by (5.13) and (5.12) respectively. But we can approximate $n^{1/3}c_0(\mu_n - \mu_0)$ with the location of the maximum of

$$\left\{ -t^2 + W_2(t),\ |t| \leqslant n^{\beta} \right\},$$

where $W_2(t) = c_0^{1/2} W_1(c_0^{-1}t)$. Denoting by T_n this location of maximum, we get that $c_0 T_n^*$ has the same distribution as the location of the maximum of

$$\{ -(T_n + t)^2 + W_2(T_n + t) + \sigma_n^* W_n^*(t) + R_n'(t),\ |t| \leqslant c_0 n^{\beta} \}$$

where R_n' is negligeable. There thus exist a Brownian path W and a $\mathbb{P}^*$-Brownian motion W^* such that the asymptotic conditional distribution of $n^{1/3}c_0(\mu_{F_n^*} - \mu_n)$ is that of the location of the maximum of

$$\left\{ -(T + t)^2 + W(T + t) + W^*(t),\ t \in \mathbb{R} \right\},$$

where T is the location of the maximum of $\{-t^2+W(t),\ t\in\mathbb{R}\}$. This has not the same distribution as τ, which proves (2.8). To prove (2.7), consider the random variable U_n defined by (5.14) and let

$$U_n^* = \inf\left\{t,\ |t|\leqslant \log n,\ F_n^*\left(\mu_{F_n^*}+r_n+tn^{-1/2}\right)-F_n^*\left(\mu_{F_n^*}-r_n-tn^{-1/2}\right)\geqslant \eta F_n^*(1)\right\}.$$

Then $\sqrt{n}(r_n-r_0)=U_n+o_{\mathbb{P}^*}(1)$ so

$$2n^{1/2}(r_n-r_0)f(\mu_0+r_0)=\eta W_n(1)-W_n(\mu_0+r_0)+W_n(\mu_0-r_0)+o_{\mathbb{P}^*}(1).$$

So we obtain from (5.17), Taylor's expansion and standard properties of Brownian motion that U_n^* is the smallest t such that $|t|\leqslant \log n$ and

$$2tf(\mu_0+r_0)+\sigma_n^*W^*\left(\mu_0+r_0\right)-\sigma_n^*W_n^*\left(\mu_0-r_0\right)\geqslant \eta\sigma_n^*W_n^*(1)+o_{\mathbb{P}^*}(1).$$

The result then follows from the fact that $n^{1/2}(r_{F_n^*}-r_n)-U_n^*$ converges in probability to zero as n goes to infinity. □

References

[1] D.F. Andrews, P.J. Bickel, F.R. Hampel, P.J. Huber, W.H. Rogers and J.W. Tukey, *Robust estimates of location. Survey and advances.* Princeton Univ. Press, Princeton, N.J. (1972).

[2] D. De Angelis, P. Hall and G.A. Young, Analytical and bootstrap approximations to estimator distributions in l^1 regression. *J. Am. Stat. Assoc.* **88** (1993) 1310–1316.

[3] C. Durot and K. Thiébot, Detecting atypical data in air pollution studies by using shorth intervals for regression. *ESAIM: PS* **9** (2005) 230–240.

[4] M. Falk and R.-D. Reiss, Weak convergence of smoothed and nonsmoothed bootstrap quantiles estimates. *Ann. Probab.* **17** (1989) 362–371.

[5] P. Groeneboom, Brownian motion with a parabolic drift and Airy functions *Probab. Th. Rel. Fields* **81** (1989) 79–109.

[6] R. Grübel, The length of the shorth. *Ann. Statist.* **16** (1988) 619–628.

[7] P. Hall, Theoretical comparison of bootstrap confidence intervals. *Ann. Statist.* **16** (1988) 927–953.

[8] P. Hall, T.J. DiCiccio and J.P. Romano, On smoothing and the bootstrap. *Ann. Statist.* **17** (1989) 692–704.

[9] P. Hall, J.W. Kay and D.M. Titterington, Asymptotically optimal difference-based estimation of variance in nonparametric regression. *Biometrika* **77** (1990) 521–528.

[10] E. Janaszewska and A.V. Nagaev, On the joint distribution of the shorth height and length. *Math. Meth. Statist.* **7** (1998) 210–229.

[11] J. Kim and D. Pollard, Cube root asymptotics. *Ann. Statist.* **18** (1990) 191–219.

[12] A. Narayanan and T.W. Sager, Table for the asymptotic distribution of univariate mode estimators. *J. Stat. Comput. Simul.* **33** (1989) 37–51.

[13] A.I. Sakhanenko, Estimates in the invariance principle. *Predel'nye Teoremy Teorii Veroyatnostej, Tr. Inst. Mat.* **5** (1985) 27–44.

[14] G.R. Shorack and J.A. Wellner, *Empirical processes with applications to statistics.* New York, Wiley (1986).

[15] C.J. Stone, Optimal uniform rate of convergence for nonparametric estimators of a density function and its derivatives. *Recent Advances in Statistics*, Academic Press, New York (1983) 293–406.

[16] K. Thiébot, *Analyses statistiques et validation de données de pollution atmosphérique.* Ph.D. thesis, Université Paris-Sud Orsay, France (2001).

[17] Y.G. Yatracos, On the estimation of the derivatives of a function with the derivatives of an estimate. *J. Multivariate Anal.* **28** (1989) 172–175.