HAL
open science

Approximation for general bootstrap of empirical processes with an application to kernel-type density estimation

Salim Bouzebda, Omar El-Dakkak

To cite this version:

Salim Bouzebda, Omar El-Dakkak. Approximation for general bootstrap of empirical processes with an application to kernel-type density estimation. 2009. hal-00371340

HAL Id: hal-00371340

https://hal.science/hal-00371340

Preprint submitted on 27 Mar 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Approximations for general bootstrap of empirical processes with an application to kernel-type density estimation

Salim BOUZEBDA* and Omar EL-DAKKAK†
Laboratoire de Statistique Théorique et Appliquée (L.S.T.A.)
Université Paris VI

Abstract

The purpose of this note is to provide an approximation for the generalized bootstrapped empirical process achieving the rate in Komlós *et al.* (1975). The proof is based on much the same arguments used in Horváth *et al.* (2000). As a consequence, we establish an approximation of the bootstrapped kernel-type density estimator.

Key words: General bootstrap; Brownian bridge; Best approximation; kernel density estimator
AMS Classifications: 62G30 ; 60F17.

1 Introduction and Main Results

Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed [i.i.d.] random variables with common distribution function $F(t) = P(X_1 \leq t)$. The empirical distribution function of $X_1, \ldots, X_n$ is

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{X_i \leq t\}, \quad -\infty < t < \infty, \tag{1}$$

where $\mathbb{1}\{A\}$ stands for the indicator function of the event A. Given the sample $X_1, \ldots, X_n$, let $X_1^*, \ldots, X_m^*$, be conditionally independent random variables with common distribution function F_n. Let

$$F_{m,n}(t) = \frac{1}{m}\sum_{i=1}^{m} \mathbb{1}\{X_i^* \leq t\}, \quad -\infty < t < \infty, \tag{2}$$

denote the *classical* Efron (or multinomial) bootstrap (see, e.g. Efron (1979) and Efron and Tibshirani (1993) for more details). Define the *bootstrapped empirical process*, $\widehat{\alpha}_{m,n}$, by

$$\alpha_{m,n}(t) := \sqrt{n}(F_{m,n}(t) - F_n(t)), \quad -\infty < t < \infty. \tag{3}$$

Among many other things, Bickel and Freedman (1981) established weak convergence of the process in (3), which enabled them to deduce the asymptotic validity of the bootstrap method in forming confidence bounds for $F(\cdot)$. Shorack (1982) gave a simple proof of weak convergence of the process in (3) [see also Shorack and Wellner (1986), Section 23.1]. The Bickel and Freedman result for $\alpha_{m,n}$ has been subsequently generalized for empirical processes based on observations in $\mathbb{R}^d$, $d > 1$ as well as in very general sample spaces and for various set and function-indexed random objects [see, for example Beran (1984), Beran and Millar (1986), Beran *et al.* (1987), Gaenssler (1992), Lohse (1987)]. This line of research found its "final results" in the work of Giné and Zinn (1989, 1990) and Csörgő and Mason (1989).
By now, the bootstrap is a widely used tool and, therefore, the properties of $\alpha_{m,n}(t)$ are of great interest in applied as well as in theoretical statistics. In fact, several procedures can actually be described in terms of the empirical process $\alpha_n(t)$, the limit distributions being functionals of $B(F(t))$, where B is a Brownian bridge.

*e-mail: salim.bouzebda@upmc.fr
†e-mail: omar.eldakkak@gmail.com

The fact that the limits may depend on the unknown distribution $F(t)$ makes it important that good approximations of these limiting distributions be found and that is where the bootstrap proved to be a very effective tool. There is a huge literature on the application of the bootstrap methodology to nonparametric kernel density and regression estimation, among other statistical procedures, and it is not the purpose of this paper to survey this extensive literature. This being said, it is worthwhile mentioning that the bootstrap as per Efron's original formulation (see Efron (1979)) presents some drawbacks. Namely, some observations may be used more than once while others are not sampled at all. To overcome this difficulty, a more general formulation of the bootstrap has been devised: the *weighted* (or *smooth*) bootstrap, which has also been shown to be computationally more efficient in several applications. For a survey of further results on weighted bootstrap the reader is referred to Barbe and Bertail (1995). Exactly as for Efron's bootstrap, the question of rates of convergence is an important one (both in probability and in statistics) and has occupied a great number of authors (see Csörgő and Révész (1981), Komlós *et al.* (1975) Horváth *et al.* (2000) and the references therein).

In this note, we will consider a version of the Mason-Newton bootstrap (see Mason and Newton (1992), and the references therein). As will be clear, this approach to bootstrap is very general and allows for a great deal of flexibility in applications. Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. random variables defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We extend $(\Omega, \mathcal{A}, \mathbb{P})$ to obtain a probability space $(\Omega^{(\pi)}, \mathcal{A}^{(\pi)}, P)$. The latter will carry the independent sequences $(X_n)_{n\geq 1}$ and $(Z_n)_{n\geq 1}$ (defined below) and will be considered rich enough as to allow the definition of another sequence (B_n^*) of Brownian bridges, independent of all the preceding sequences. The possibility of such an extension is discussed in detail in literature; the reader is referred, e.g., to Csörgő and Révész (1981), Komlós *et al.* (1975) and Berkes and Philipp (1977). In the sequel, whenever an almost sure property is stated, it will be tacitly assumed that it holds with respect the the p.m. P defined on the extended space.

Define a sequence $(Z_n)_{n\geq 1}$ of i.i.d. replicæ of a *strictly positive* random variable Z with distribution function $G(\cdot)$, independent of the X_n's. In the sequel, the following assumptions on the Z_n's will prevail:

(A1) $\mathbb{E}(Z) = 1$; $\quad \mathbb{E}(Z^2) = 2$ (or, equivalently, $\mathbf{Var}(Z) = 1$).

(A2) There exists an $\varepsilon > 0$, such that
$$\mathbb{E}(e^{tZ}) < \infty \text{ for all } |t| \leq \varepsilon.$$

For all $n \geq 1$, let $T_n = Z_1 + \cdots + Z_n$ and define the random weights,

$$\mathscr{W}_{i;n} := \frac{Z_i}{T_n}, \qquad i = 1, \ldots, n. \tag{4}$$

The quantity

$$F_n^*(t) = \sum_{i=1}^{n} \mathscr{W}_{i;n} \mathbb{1}\{X_i \leq t\}, \text{ for } -\infty < t < \infty. \tag{5}$$

will be called *generalized (or weighted) bootstrapped empirical distribution function*. Analogously, recalling the empirical process based on $X_1, \ldots, X_n$,

$$\alpha_n(t) = n^{1/2}(F_n(t) - F(t)), \ -\infty < t < \infty, \tag{6}$$

define the corresponding *generalized (or weighted) bootstrapped empirical process* by

$$\alpha_n^*(t) = n^{1/2}(F_n^*(t) - F_n(t)), \ -\infty < t < \infty. \tag{7}$$

The system of weights defined in (4) appears in Mason and Newton (1992), p.1617 where it is shown that it satisfies assumptions ($\mathscr{W}_I$), ($\mathscr{W}_{II}$) and ($\mathscr{W}_{III}$) on p.1612 of the same reference, so that all the results therein hold for the objects to be treated in this note. In particular, weak convergence for the process α_n^* to a Brownian bridge is proved. For more results concerning this version of the the weighted boostrapped empirical process, we refer the reader to Deheuvels and Derzko (2008). Note that, as a special case of the system of weights we are considering, one can obtain the one used for Bayesian bootstrap (see Rubin (1981)).

In what follows, we obtain a KMT rate of convergence for this process in sup norm. More precisely, we consider deviations between the generalized bootstrapped empirical process $\{\alpha_n^*(t) : t \in \mathbb{R}\}$ and a sequence of approximating Brownian bridges $\{B_n^*(F(t)) : t \in \mathbb{R}\}$ on $\mathbb{R}$. Our main result goes as follows.

Theorem 1 *Let assumptions* (A1) *and* (A2) *hold. Then, it is possible to define a sequence of Brownian bridges* $\{B_n^*(y) : 0 \le y \le 1\}$ *such that, for all* $\varepsilon, \eta > 0$*, there exists* $N = N(\varepsilon, \eta)$*, such that, for all* $n \ge N$ *and all* $x > 0$,

$$P\left(\sup_{-\infty<t<\infty} |\alpha_n^*(t) - B_n^*(F(t))| > 3n^{-1/2}(K_1 \log n + x)\right) \le K_2 \exp\left(-\frac{K_3 x}{(1+\varepsilon)^2}\right) + \eta, \tag{8}$$

where K_1*,* K_2 *and* K_3 *are positive universal constants.*

The proof of Theorem 1 is given in Section 3.

Remark 1 Theorem 1 implies the following approximation of the weighted bootstrap:

$$\sup_{-\infty<t<\infty} |\alpha_n^*(t) - B_n^*(F(t))| = O_P\left(\frac{\log n}{n^{1/2}}\right). \tag{9}$$

Remark 2 Theorem 1 turns out be useful in obtaining confidence bands for the distribution function of the sample data. We formalize this idea as follows: for $0 < \alpha < 1$, one has

$$\lim_{n\to\infty} P\left(\sup_{-\infty<t<\infty} \sqrt{n}|F_n(t) - F(t)| \le c(\alpha)\right) = P\left(\sup_{-\infty<t<\infty} |B(F(t))| \le c(\alpha),\right). \tag{10}$$

Note that for each fixed t, $B(F(t))$ is a zero-mean Gaussian random variable with covariance structure

$$E(B(F(t))B(F(s))) = F(t \wedge s) - F(t)F(s)$$

where $t \wedge s := \min(t, s)$. In practice, $c(\alpha)$ can, of course, not be computed since the covariance structure of $B(F(t))$ depends on the unknown cdf F. Instead, suppose $(Z_1^{(1)}, \ldots, Z_n^{(1)}), \ldots, (Z_1^{(N)}, \ldots, Z_n^{(N)})$ are N independent vectors of i.i.d. copies of Z, sampled independently of the X_i's. Define the random variables

$$\psi^j := \sup_{-\infty<t<\infty} \left|\alpha_{n,j}^*(t)\right|, \qquad j = 1, \ldots, N, \tag{11}$$

where $\alpha_{n,j}^*$ denotes the generalized bootstrapped empirical process constructed with the sample $(Z_1^{(j)}, \ldots, Z_n^{(j)})$, $j = 1, \ldots, N$. Theorem 1 accounts for the use of the smallest $z > 0$ such that

$$\frac{1}{N}\sum_{i=1}^{N} \mathbb{1}\left\{\psi^j \le z\right\} \ge 1 - \alpha.$$

as an estimator of $c(\alpha)$.

A direct consequence of Theorem 1 and Theorem 1.5 in Horváth *et al.* (2000) is the following approximation for $\alpha_n^*(\cdot)$ based on a Kiefer process

Theorem 2 *There is a Kiefer process* $\{K(t; x); 0 \le t \le 1; 0 \le x \le \infty\}$ *such that*

$$\max_{1\le k\le n} \sup_{-\infty<t<\infty} \left|\sum_{i=1}^{k} (\mathscr{W}_{i;n} - 1/n)\mathbb{1}\{X_i \le t\} - K(F(t), k)\right| = O_P(n^{1/4}(\log n)^{1/2}). \tag{12}$$

2 An application to kernel density estimation

Let $X_1, \ldots, X_n$ be independent random replicæ of a random variable $X \in \mathbb{R}$ with distribution function $F(\cdot)$. We assume that the distribution function $F(\cdot)$ has a density $f(\cdot)$ (with respect to the Lebesgue measure in $\mathbb{R}$). First of all, we introduce a kernel density estimator of $f(\cdot)$. To this end, let $K(\cdot)$ be a measurable function fulfilling the following conditions

(K1) $K(\cdot)$ is of bounded variation and compactly supported on $\mathbb{R}$;

(K2) $K \geq 0$ and $\int K(u)du = 1$.

Now, define the Akaike-Parzen-Rosenblatt kernel density estimator of $f(\cdot)$ (see Akaike (1954), Parzen (1962) and Rosenblatt (1956)) as follows: for all $x \in \mathbb{R}$, estimate $f(x)$ by

$$f_{n,h_n}(x) = \frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right), \tag{13}$$

where $\{h_n : n \geq 1\}$ is a sequence of positive constants satisfying the conditions

$$h_n \downarrow 0 \quad \text{and} \quad nh_n \uparrow \infty, \quad \text{as} \quad n \to \infty.$$

Secondly, we define the bootstrapped version of $f_{n,h_n}(\cdot)$, by setting for all $h_n > 0$ and $x \in \mathbb{R}$,

$$f^*_{n,h_n}(x) = \frac{1}{h_n} \sum_{i=1}^{n} \mathscr{W}_{i;n} K\left(\frac{x - X_i}{h_n}\right), \tag{14}$$

where $\mathscr{W}_{i;n}$ is defined in (4). We will provide an approximation rate for the following process

$$\gamma^*_n(x) = \sqrt{nh_n^2}\left(f^*_{n,h_n}(x) - f_{n,h_n}(x)\right), \quad -\infty < x < \infty. \tag{15}$$

The following theorem, proved in the next Section, shows that a single bootstrap suffices to obtain the desired approximation for non-parametric kernel-type density estimators.

Theorem 3 *Let conditions* (A1), (A2), (K1) *and* (K2) *prevail. Then we can define Brownian bridges* $\{B^*_n(y) : 0 \leq y \leq 1\}$ *such that almost surely along* $X_1, X_2, \ldots,$ *as* n *tends to infinity, we have*

$$\sup_{-\infty<x<\infty} \left|\gamma^*_n(x) - \int K\left(\frac{x-s}{h_n}\right) \mathrm{d}B^*_n(F(s))\right| = O_P\left(\frac{\log n}{\sqrt{n}}\right). \tag{16}$$

If, moreover, we suppose boundedness of the unknown density, f, i.e. if we suppose the existence of $M > 0$ *such that* $\sup_{-\infty<x<\infty} f(x) \leq M$, *then, almost surely along* $X_1, X_2, \ldots,$ *as* n *tends to infinity,*

$$\sup_{-\infty<x<\infty} \left|\gamma^*_n(x) - B^*_n(F(x)) \int K(t)\mathrm{d}t\right| = O_P\left(\frac{\log n}{\sqrt{n}} + h_n\sqrt{\log h_n^{-1}}\right). \tag{17}$$

Remark 3. Under appropriate conditions, and using the same arguments rehearsed in the proof of Theorem 3, it is possible to obtain an approximation of a smoothed version of F^*_n.

3 Proofs

Proof of Theorem 1. In the sequel, we will write $\|\cdot\|$ to indicate $\sup_{-\infty<t<+\infty}|\cdot|$. We have that

$$\|\alpha^*_n(t) - B^*_n(F(t))\| = \|\sqrt{n}(F^*_n(t) - F_n(t)) - B^*_n(F(t))\|.$$

Now, it is easily seen that

$$\sqrt{n}(F^*_n(t) - F_n(t)) = \left(\frac{n}{T_n}\right)\left[\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n} Z_i \mathbb{1}\{X_i \leq t\} - F(t)T_n + (F(t) - F_n(t))T_n\right)\right], \tag{18}$$

so that

$$\|\alpha^*_n(t) - B^*_n(F(t))\| \leq S_1(n) + S_2(n) + S_3(n),$$

where

$$S_1(n) := \left(\frac{n}{T_n}\right)\left\|\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n} Z_i \mathbb{1}\{X_i \leq t\} - T_n F(t)\right) - B^*_n(F(t))\right\|, \tag{19}$$

where

$$S_2(n) := \left(\frac{n}{T_n}\right) \left\| \frac{T_n}{\sqrt{n}}(F(t) - F_n(t)) \right\|, \tag{20}$$

and where

$$S_3(n) := \left| \frac{n}{T_n} - 1 \right| \|B_n^*(F(t))\|. \tag{21}$$

We start by dealing with the term $S_3(n)$. We will treat the cases $x > Cn$ and $x \leq Cn$ (C being a strictly positive constant) separately. Fix $x > Cn$ arbitrarily. Union bound gives for all n,

$$P\left(S_3(n) \geq n^{-1/2}(x + c\log n)\right) \leq P\left(S_4(n) \geq \frac{x}{2\sqrt{n}}\right) + P\left(\|B_n^*(F(t))\| \geq \frac{x}{2\sqrt{n}}\right),$$

where

$$S_4(n) := \left(\frac{n}{T_n}\right) \|B_n^*(F(t))\|. \tag{22}$$

Now, it is known that, for all $n \geq 1$ and all $x > n \geq 1$, there exists a positive constant c_1, such that

$$P\left(\|B_n^*(F(t))\| \geq \frac{x}{2\sqrt{n}}\right) \leq c_1 \exp\left(-\frac{x^2}{4n}\right) \leq \exp\left(-\frac{x}{4}\right). \tag{23}$$

On the other hand, since strong law of large numbers gives

$$\left| \frac{n}{T_n} - 1 \right| \stackrel{a.s.}{\to} 0,$$

for all $\varepsilon, \eta > 0$, there exists $N_1 = N_1(\varepsilon, \eta)$, such that, for all $n \geq N_1$,

$$P\left(\left| \frac{n}{T_n} - 1 \right| \in (0, \varepsilon)\right) \geq 1 - \eta. \tag{24}$$

Consequently, denoting the law of $\frac{n}{T_n}$ by $\mathcal{L}_{\frac{n}{T_n}}$, independence of the Z_n's from the B_n's gives

$$\begin{aligned}
P\left(S_4(n) \geq \frac{x}{2\sqrt{n}}\right) &= P\left(S_4(n) \geq \frac{x}{2\sqrt{n}}, \left| \frac{n}{T_n} - 1 \right| \in (0, \varepsilon)\right) \\
&\quad + P\left(S_4(n) \geq \frac{x}{2\sqrt{n}}, \left| \frac{n}{T_n} - 1 \right| \notin (0, \varepsilon)\right) \\
&\leq P\left(\frac{n}{T_n}\|B_n^*(F(t))\| \geq \frac{x}{2\sqrt{n}} \mid \left| \frac{n}{T_n} - 1 \right| \in (0, \varepsilon)\right) \\
&\quad + P\left(\left| \frac{n}{T_n} - 1 \right| \notin (0, \varepsilon)\right) \\
&\leq \int_{1-\varepsilon}^{1+\varepsilon} P\left(\|B_n^*(F(t))\| > \frac{x}{2\sqrt{ny^2}} \mid \frac{n}{T_n} = y\right) \mathcal{L}_{\frac{n}{T_n}}(\mathrm{d}y) + \eta \\
&\leq P\left(\|B_n^*(F(t))\| > \frac{x}{2\sqrt{n(1+\varepsilon)^2}}\right) + \eta \\
&\leq c_1 \exp\left(-\frac{x}{4(1+\varepsilon)^2}\right) + \eta,
\end{aligned} \tag{25}$$

where, in the last inequality, we have used (23). Combining (23) and (25), we have that, for all $\varepsilon, \eta > 0$, there exists $N_1 = N_1(\varepsilon, \eta)$, such that, for all $n \geq N_1$,

$$P\left(S_3(n) \geq n^{-1/2}(x + c\log n)\right) \leq (1 + c_1) \exp\left(-\frac{x}{4(1+\varepsilon)^2}\right) + \eta. \tag{26}$$

Now we turn to the case $0 < x \leq Cn$. Again, by the union bound,

$$P\left(S_3(n) \geq n^{-1/2}(x + c\log n)\right) \leq P\left(\left|\frac{n}{T_n} - 1\right| > \sqrt{\frac{x}{n}}\right) + P\left(\|B_n^*(F(t))\| > \sqrt{x}\right). \tag{27}$$

Again by (23), we have that for all n,

$$P\left(\|B_n^*(F(t))\| > \sqrt{x}\right) \leq c_1 \exp(-x/2). \tag{28}$$

On the other hand, by (24), for all $\varepsilon, \eta > 0$, there exists $N_1 = N_1(\varepsilon, \eta)$ such that for all $n \geq N_1$,

$$\begin{aligned} P\left(\left|\frac{n}{T_n} - 1\right| > \sqrt{\frac{x}{n}}\right) &= P\left(\left|\frac{n}{T_n} - 1\right| > \sqrt{\frac{x}{n}}, \left|\frac{n}{T_n} - 1\right| \in (0, \varepsilon)\right) \\ &\quad + P\left(\left|\frac{n}{T_n} - 1\right| > \sqrt{\frac{x}{n}}, \left|\frac{n}{T_n} - 1\right| \notin (0, \varepsilon)\right) \\ &\leq P\left(\left(\frac{n}{T_n}\right)\left|\frac{T_n}{n} - 1\right| > \sqrt{\frac{x}{n}}, \left|\frac{n}{T_n} - 1\right| \in (0, \varepsilon)\right) + \eta \\ &\leq P\left(\left|\frac{T_n}{n} - 1\right| > \sqrt{\frac{x}{n(1+\varepsilon)^2}}\right) + \eta \end{aligned} \tag{29}$$

Use Theorem 2.6 in Petrov (1995) to find constants c_2 and c_3 such that

$$P\left(\left|\frac{T_n}{n} - 1\right| > \sqrt{\frac{x}{n(1+\varepsilon)^2}}\right) \leq c_2 \exp\left(\frac{-c_3 x}{(1+\varepsilon)^2}\right). \tag{30}$$

Combining (28), (29) and (30), and plugging in (27), we deduce the existence of positive universal constants c_4 and c_5 such that

$$P\left(S_3(n) \geq n^{-1/2}(x + c\log n)\right) \leq c_4 \exp\left(\frac{-c_5 x}{(1+\varepsilon)^2}\right) + \eta, \tag{31}$$

so that one concludes, from (26) and (31), that for all $\varepsilon, \eta > 0$, there exists $N = N(\varepsilon, \eta)$, such that, for all $n \geq N$, and all $x > 0$

$$P\left(S_3(n) \geq n^{-1/2}(x + c\log n)\right) \leq c_6 \exp\left(\frac{-c_7 x}{(1+\varepsilon)^2}\right) + \eta, \tag{32}$$

for some universal constants c_6 and c_7.
The proof is concluded once we show the existence of universal positive constants c_8, c_9, c_{10} and c_{11} such that, for all $\varepsilon, \eta > 0$, there exists $N_2 = N_2(\varepsilon, \eta)$, and $N_3 = N_3(\varepsilon, \eta)$ such that, for all $n \geq N_2$, and all $x > 0$

$$P\left(S_1(n) \geq n^{-1/2}(x + c\log n)\right) \leq c_8 \exp\left(\frac{-c_9 x}{(1+\varepsilon)^2}\right) + \eta, \tag{33}$$

and for all $n \geq N_3$ and all $x > 0$,

$$P\left(S_2(n) \geq n^{-1/2}(x + c\log n)\right) \leq c_{10} \exp\left(\frac{-c_{11} x}{(1+\varepsilon)^2}\right) + \eta. \tag{34}$$

Since

$$S_1(n) = \left(\frac{n}{T_n}\right)\left\|\frac{1}{\sqrt{n}}\left(\sum_{i=1}^n Z_i \mathbb{1}\{X_i \leq t\} - T_n F(t)\right) - B_n^*(F(t))\right\|,$$

formula (3.7) in Horváth *et al.* (2000) combined with arguments similar to those used for the term $S_3(n)$ imply (33). As for (34), formula (3.5) in Horváth *et al.* (2000) together with the by now usual ε, η argument conclude the proof. □

Proof of Theorem 3. We start by proving (16). We have for $x \in \mathbb{R}$

$$\begin{aligned}\sqrt{nh_n^2}\left(\widehat{f}^*_{n,h_n}(x) - \widehat{f}_{n,h_n}(x)\right) &= \int K\left((x-s)/h_n\right) \mathrm{d}\{n^{1/2}(F^*_n(s) - F_n(s))\} \\ &= \int K\left((x-s)/h_n\right) \mathrm{d}\alpha^*_n(s).\end{aligned}$$

Integration by parts implies that

$$\int K\left(\frac{x-s}{h_n}\right) \mathrm{d}\alpha^*_n(s) = -\int \alpha^*_n(x-th_n)\mathrm{d}K(t), \tag{35}$$

and

$$\int K\left(\frac{x-s}{h_n}\right) \mathrm{d}B^*_n(F(s)) = -\int B^*_n(F(x-th_n))\mathrm{d}K(t). \tag{36}$$

Now, Theorem 1 together with condition (K1) give

$$\begin{aligned}&\sup_{-\infty<x<\infty}\left|\int \alpha^*_n(x-th_n)\mathrm{d}K\left(t\right) - \int B^*_n(F(x-th_n))\mathrm{d}K\left(t\right)\right| \\ &\qquad \leq \sup_{-\infty<u<\infty}|\alpha^*_n(u) - B^*_n(F(u))| \int \mathrm{d}|K\left(t\right)| = O_P\left(\frac{\log n}{\sqrt{n}}\right),\end{aligned} \tag{37}$$

thus proving (16).
Once (16) is at hand, to prove (17), it suffices to bound

$$\left|\int B^*_n(F(x-th_n))\mathrm{d}K\left(t\right) - B^*_n(F(x))\right| \leq \int |B^*_n(F(x-th_n)) - B^*_n(F(x))| \,\mathrm{d}K\left(t\right), \tag{38}$$

in probability. By condition (K1), and provided the unknown density f is bounded (by a strictly positive constant, say M), for n large enough,

$$|B^*_n(F(x-th_n)) - B^*_n(F(x))| \leq \sup_{|u-v|\leq\delta_n}|B^*_n(u) - B^*_n(v)| \tag{39}$$

where $\delta_n = Mh_n$. Now, it is always possible to define a Brownian Bridge, $\{B^*(y) : 0 \leq y \leq 1\}$, on the same probability space carrying the sequence of Brownian Bridges $\{B^*_n(y) : 0 \leq y \leq 1\}_{n\geq1}$, such that for all n, and all $\varepsilon > 0$

$$\begin{aligned}&P\left(\{2\delta_n \log\delta_n^{-1}\}^{-1/2} \sup_{|u-v|<h}\sup_{h\in[0,\delta_n]}|B^*_n(u) - B^*_n(v)| > 1+\varepsilon\right) \\ &= P\left(\{2\delta_n \log\delta_n^{-1}\}^{-1/2} \sup_{|u-v|<h}\sup_{h\in[0,\delta_n]}|B^*(u) - B^*(v)| > 1+\varepsilon\right).\end{aligned}$$

Since $\delta_n \to 0$, by Theorem 1.4.1 in Csörgő and Révész (1981), we have with probability one

$$\lim_{n\to\infty}\{2\delta_n \log\delta_n^{-1}\}^{-1/2} \sup_{|u-v|<h}\sup_{h\in[0,\delta_n]}|B^*(u) - B^*(v)| = 1. \tag{40}$$

Thus, as $n \to \infty$,

$$P\left(\{2\delta_n \log\delta_n^{-1}\}^{-1/2} \sup_{|u-v|<h}\sup_{h\in[0,\delta_n]}|B^*_n(u) - B^*_n(v)| > 1+\varepsilon\right) \to 0,$$

giving

$$\sup_{|u-v|\leq h}\sup_{h\in[0,\delta_n]}|B^*_n(u) - B^*_n(v)| = O_P\left(\sqrt{2\delta_n \log\delta_n^{-1}}\right). \tag{41}$$

Put (35), (36), (38), (39) and (41) together to obtain

$$\sup_{-\infty<x<\infty}\left|\gamma^*_n(x) - B^*_n(F(x))\int \mathrm{d}K(t)\right| = O_P\left(\frac{\log n}{\sqrt{n}} + h_n\sqrt{\log h_n^{-1}}\right),$$

thus completing the proof of Theorem. □

Acknowledgement The authors are grateful to Professors Paul Deheuvels and Giovanni Peccati for a number of very useful discussions.

References

Akaike, H. (1954). An approximation to the density function. *Ann. Inst. Statist. Math., Tokyo*, **6**, 127–132.

Barbe, P. and Bertail, P. (1995). *The weighted bootstrap*, volume 98 of *Lecture Notes in Statistics*. Springer-Verlag, New York.

Beran, R. (1984). Bootstrap methods in statistics. *Jahresber. Deutsch. Math.-Verein.*, **86**(1), 14–30.

Beran, R. and Millar, P. W. (1986). Confidence sets for a multivariate distribution. *Ann. Statist.*, **14**(2), 431–443.

Beran, R. J., Le Cam, L., and Millar, P. W. (1987). Convergence of stochastic empirical measures. *J. Multivariate Anal.*, **23**(1), 159–168.

Berkes, I. and Philipp, W. (1977). An almost sure invariance principle for the empirical distribution function of mixing random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **41**(2), 115–137.

Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.*, **9**(6), 1196–1217.

Csörgő, M. and Révész, P. (1981). *Strong approximations in probability and statistics*. Probability and Mathematical Statistics. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York.

Csörgő, S. and Mason, D. M. (1989). Bootstrapping empirical functions. *Ann. Statist.*, **17**(4), 1447–1471.

Deheuvels, P. and Derzko, G. (2008). Asymptotic certainty bands for kernel density estimators based upon a bootstrap resampling scheme. *Statistical Models and Methods for Biomedical and Technical Systems*, **III**, 171–186.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.*, **7**(1), 1–26.

Efron, B. and Tibshirani, R. J. (1993). *An introduction to the bootstrap*, volume 57 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, New York.

Gaenssler, P. (1992). Confidence bands for probability distributions on Vapnik-Chervonenkis classes of sets in arbitrary sample spaces using the bootstrap. In *Bootstrapping and related techniques (Trier, 1990)*, volume 376 of *Lecture Notes in Econom. and Math. Systems*, pages 57–61. Springer, Berlin.

Giné, E. and Zinn, J. (1989). Necessary conditions for the bootstrap of the mean. *Ann. Statist.*, **17**(2), 684–691.

Giné, E. and Zinn, J. (1990). Bootstrapping general empirical measures. *Ann. Probab.*, **18**(2), 851–869.

Horváth, L., Kokoszka, P., and Steinebach, J. (2000). Approximations for weighted bootstrap processes with an application. *Statist. Probab. Lett.*, **48**(1), 59–70.

Komlós, J., Major, P., and Tusnády, G. (1975). An approximation of partial sums of independent RV's and the sample DF. I. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **32**, 111–131.

Lohse, K. (1987). Consistency of the bootstrap. *Statist. Decisions*, **5**(3-4), 353–366.

Mason, D. M. and Newton, M. A. (1992). A rank statistics approach to the consistency of a general bootstrap. *Ann. Statist.*, **20**(3), 1611–1624.

Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.*, **33**, 1065–1076.

Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.*, **27**, 832–837.

Rubin, D. B. (1981). The Bayesian bootstrap. *Ann. Statist.*, **9**(1), 130–134.

Shorack, G. R. (1982). Bootstrapping robust regression. *Comm. Statist. A—Theory Methods*, **11**(9), 961–972.

Shorack, G. R. and Wellner, J. A. (1986). *Empirical processes with applications to statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York.