The behavior of the NPMLE of a decreasing density near the boundaries of the support

Vladimir N. Kulikov[†]
Hendrik P. Lopuhaä[‡]

01.06.2004

Abstract: We investigate the behavior of the nonparametric maximum likelihood estimator $\hat{f}_n$ for a decreasing density f near the boundaries of the support of f. We establish the limiting distribution of $\hat{f}_n(n^{-\alpha})$, where we need to distinguish between different values of $0 < \alpha < 1$. Similar results are obtained for the upper endpoint of the support, in the case it is finite. This yields consistent estimators for the values of f at the boundaries of the support. The limit distribution of these estimators is established and their performance is compared with the penalized NPMLE of Woodroofe and Sun (1993).

[†]EURANDOM, P.O. Box 513 - 5600 MB Eindhoven, the Netherlands, kulikov@eurandom.nl
[‡]Delft University of Technology, Mekelweg 4 2628 CD Delft, the Netherlands, h.p.lopuhaa@its.tudelft.nl

1 Introduction

In various statistical models, such as density estimation, and estimation of regression curves or hazard rates, monotonicity constraints can arise naturally. For these situations certain isotonic estimators have been in use for considerable time. Often these estimators can be seen as maximum likelihood estimators in a semi-parametric setting. Although conceptually these estimators have great appeal and are easy to formulate, their distributional properties are usually of a very complicated nature.

In the context of density estimation, the non-parametric maximum likelihood estimator $\hat{f}_n$ for a non-increasing density f on $[0, \infty)$ has been discovered by Grenander (1956). It is defined as the left derivative of the least concave majorant (LCM) of the empirical distribution function F_n constructed from a sample $X_1, \ldots, X_n$ from f. Prakasa Rao (1969) obtained the asymptotic pointwise behavior of $\hat{f}_n$. Groeneboom (1985) provided an elegant proof of the same result, which can be formulated as follows. For each $x_0 > 0$,

$$|4f(x_0)f'(x_0)|^{-1/3} n^{1/3} \left\{ \hat{f}_n(x_0) - f(x_0) \right\} \to \operatorname*{argmax}_{t \in \mathbb{R}} \left\{ W(t) - t^2 \right\} \tag{1.1}$$

in distribution, where W denotes standard two-sided Brownian motion originating from zero. The first distributional result for a global measure of deviation for $\hat{f}_n$ was found by Groeneboom (1985), concerning asymptotic normality of the L_1-distance $\|\hat{f}_n - f\|_1$ (see Groeneboom, Hooghiemstra and Lopuhaä (1999) for a rigorous proof).

Apart from estimating a monotone density f on $(0, \infty)$, the estimation of the value of f or its derivatives at zero, is required in various statistical applications. There is a direct connection with renewal processes, where the backward recurrence time in equilibrium has density $f(x) = (1 - G(x))/\mu$, where G and μ are the distribution function and mean of the interarrival times (see Feller (1971)). Clearly, f is decreasing and a natural parameter of interest is $\mu = 1/f(0)$. An interesting application is in the context of natural fecundity of human populations, where one is interested in the time T it takes for a couple from initiating attempts to become pregnant until conception occurs. Keiding et. al. (2002) investigate a current-duration design where data are collected from a cross-sectional sample of couples that are currently attempting to become pregnant. If U is the time to discontinuation without success and V is the time to discontinuation of follow-up, then $X = T \wedge U$ is the waiting time until termination for whatever reason, and $Y = T \wedge U \wedge V$ is the observed experience waiting time. When the initiations happen according to a homogeneous Poisson process, Y is distributed as the backward recurrence time in a renewal process in equilibrium, and the survival function of X is $f(x)/f(0)$, where f is decreasing. Woodroofe and Sun (1993) provide a different application in the context of astronomy. If Y denotes the normalized angular diameter of a galaxy, conditional on that it is being observed, then $1/Y^3$ has a non-increasing density f and the proportion of galaxies that are observed is $1/f(0)$. Another example is from Hampel (1987), who studies the sojourn time of migrating birds. Under certain model assumptions, the expected sojourn time is $-f(0)/f'(0)$, where f is the (convex) decreasing density of the time span between capture and recapture of a bird.

In contrast to (1.1), Woodroofe and Sun (1993) showed that $\hat{f}_n$ is not consistent at zero. They proposed a penalized maximum likelihood estimator $\hat{f}_n^{\mathrm{P}}(0)$ and in Sun and Woodroofe (1996) it was shown that

$$n^{1/3} \left\{ \hat{f}_n^{\mathrm{P}}(0) - f(0) \right\} \to \sup_{t > 0} \frac{W(t) - (c - f(0)f'(0)t^2/2)}{t},$$

where c depends on the penalization. Surprisingly, the inconsistency of $\hat{f}_n$ at zero does not influence the behavior of $\|\hat{f}_n - f\|_1$. Nevertheless, the inconsistency at the boundaries will have an effect if one studies other global measures of deviation, such as the L_k-distance, for k larger than one, or the supremum distance.

In this paper we study the behavior of the Grenander estimator at the boundaries of the support of f. We first consider a non-increasing density f on $[0, \infty)$ and investigate the behavior of

$$n^{\beta}\left\{\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right\}, \tag{1.2}$$

for $c > 0$, where $0 < \alpha < 1$ and $\beta > 0$ are chosen suitably in order to make (1.2) converge in distribution. Our results will imply that when $f'(0) < 0$, then $\hat{f}_n(cn^{-1/3})$ is a consistent estimator for $f(0)$ at rate $n^{1/3}$ with a limiting distribution that is a functional of W. This immediately yields $\hat{f}_n^{\text{S}}(0) = \hat{f}_n(n^{-1/3})$ as a simple estimator for $f(0)$. A more adaptive alternative would be to find the value of c that minimizes the asymptotic mean squared error. This turns out to depend on f and then has to be estimated. The resulting estimator $\hat{f}_n^{\text{A}}(0) = \hat{f}_n(\hat{c}n^{-1/3})$ will be compared with the penalized maximum likelihood estimator from Sun and Woodroofe (1996). We will also consider the case where $f'(0) = 0$ and $f''(0) < 0$, which requires different values for c and α. For non-increasing f with compact support, say $[0, 1]$, we also investigate the behavior near one. Similarly, this leads to a consistent estimator for $f(1)$. Moreover, the results on the behavior of $\hat{f}_n$ at the boundaries of $[0, 1]$ allows an adequate treatment of the L_k-distance between $\hat{f}_n$ and f. It turns out that for $k > 2.5$, the inconsistency of $\hat{f}_n$ starts to affect the behavior of $\|\hat{f}_n - f\|_k$ (see Kulikov and Lopuhaä (2004a)).

In Section 2 we give a brief outline of our approach for studying differences such as (1.2), and state some preliminary results for the argmax functional. Section 3 is devoted to the behavior of $\hat{f}_n$ near zero. Section 4 deals with the behavior of $\hat{f}_n$ near the boundary at the other end of the support for a density f on $[0, 1]$. In Section 5 we propose two estimators $\hat{f}_n^{\text{S}}(0)$ and $\hat{f}_n^{\text{A}}(0)$ based on the presented theory, and compare these with the penalized maximum likelihood estimator from Sun and Woodroofe (1996).

2 Preliminaries

Instead of studying the process $\{\hat{f}_n(t) : t \in [0, 1]\}$ itself, we will use the more tractable inverse process $\{U_n(a) : a \geq 0\}$, where $U_n(a)$ is defined as the last time that the process $F_n(t) - at$ attains its maximum:

$$U_n(a) = \underset{t \in [0,\infty)}{\text{argmax}} \left\{F_n(t) - at\right\}.$$

Its relation with $\hat{f}_n$ is as follows: with probability one

$$\hat{f}_n(x) \leq a \Leftrightarrow U_n(a) \leq x. \tag{2.1}$$

Let us first describe the line of reasoning used to prove convergence in distribution of (1.2). We illustrate things for the case $c = 1$, $0 < \alpha < 1/3$, and $f'(0) < 0$. It turns out that in this case the proper choice for β is $1/3$. Hence, we will consider events of the following type

$$n^{1/3}\left\{\hat{f}_n(n^{-\alpha}) - f(n^{-\alpha})\right\} \leq x.$$

According to relation (2.1), this event is equivalent with

$$U_n\left(f(n^{-\alpha})+xn^{-1/3}\right)-n^{-\alpha}\leq 0.$$

The left hand side is the argmax of the process

$$Z_n(t)=F_n(t+n^{-\alpha})-f(n^{-\alpha})t-xtn^{-1/3}.$$

With suitable scaling, the process Z_n converges in distribution to some Gaussian process Z. The next step is to use an argmax version of the continuous mapping theorem from Kim and Pollard (1990). The version that suffices for our purposes, is stated below for further reference.

Theorem 2.1 *Let $\{Z(t) : t \in \mathbb{R}\}$ be a continuous random process satisfying*

(i) Z has a unique maximum with probability one,

(ii) $Z(t) \to -\infty$, as $|t| \to \infty$, with probability one.

Let $\{Z_n(t) : t \in \mathbb{R}\}$ be a sequence of random processes satisfying

(iii) $\operatorname{argmax}_{t\in\mathbb{R}} Z_n(t) = O_p(1)$, *as $n \to \infty$.*

If Z_n converges in distribution to Z, in the topology of uniform convergence on compacta, then $\operatorname{argmax}_{t\in\mathbb{R}} Z_n(t)$ *converges in distribution to* $\operatorname{argmax}_{t\in\mathbb{R}} Z(t)$.

Application of this theorem yields that $U_n\left(f(n^{-\alpha})+xn^{-1/3}\right)$, properly scaled, convergences in distribution to the argmax of a Gaussian process. Convergence of (1.2) then follows from another application of (2.1).

The main difficulty in verifying the conditions of Theorem 2.1, is showing that (iii) holds. It requires careful handling of all small order terms in the expansion of the process. In the process of proving condition (iii) we will frequently use the following lemma, which enables us to suitably bound the argmax from above.

Lemma 2.1 *Let f and g be continuous functions on $K \subset \mathbb{R}$.*

(i) Suppose that g is non-increasing. Then $\operatorname{argmax}_{x\in K}\{f(x)+g(x)\} \leq \operatorname{argmax}_{x\in K} f(x)$.

(ii) Let $C > 0$ and suppose that for all $s, t \in K$, such that $t \geq C+s$, we have that $g(t) \leq g(s)$. Then $\operatorname{argmax}_{x\in K}\{f(x)+g(x)\} \leq C + \operatorname{argmax}_{x\in K} f(x)$.

In studying processes like Z_n we will use a Brownian approximation similar to the one used in in Groeneboom, Hooghiemstra and Lopuhaä (1999). Let E_n denote the empirical process $\sqrt{n}(F_n - F)$. For $n \geq 1$, let B_n be versions of the Brownian bridge constructed on the same probability space as the uniform empirical process $E_n \circ F^{-1}$ via the Hungarian embedding, where

$$\sup_{t\in[0,1]} |E_n(t) - B_n(F(t))| = O_p(n^{-1/2}\log n) \tag{2.2}$$

(see Komlos, Major and Tusnady (1975)). Define versions W_n of Brownian motion by

$$W_n(t) = B_n(t) + \xi_n t, \quad t \in [0,1],$$

where ξ_n is standard normal random variable independent of B_n. This means that we can represent B_n by the pathwise equality $B_n(t) = W_n(t) - tW_n(1)$.

We will often use a Brownian scaling argument in connection with argmax functionals. Note that $\text{argmax}_t\{Z(t)\}$ does not change by multiplying Z with a constant, and that the process $W(bt)$ has the same distribution as the process $b^{1/2}W(t)$. This implies that

$$\begin{aligned} a \operatorname*{argmax}_{t\in I} \left\{W(bt) - ct^k\right\} &= \operatorname*{argmax}_{t\in aI} \left\{W(ba^{-1}t) - ca^{-k}t^k\right\} \\ &\stackrel{d}{=} \operatorname*{argmax}_{t\in aI} \left\{b^{1/2}a^{-1/2}W(t) - ca^{-k}t^k\right\} \\ &= \operatorname*{argmax}_{t\in aI} \left\{W(t) - cb^{-1/2}a^{-k+1/2}t^k\right\}, \end{aligned} \tag{2.3}$$

for $I \subset \mathbb{R}$ and constants $a, b > 0$ and $c \in \mathbb{R}$.

3 Behavior near zero

We first consider the case that f is a non-increasing density on $[0, \infty)$ satisfying

(C1) $0 < f(0) = \lim_{x\downarrow 0} f(x) < \infty$.

(C2) For some $k \geq 1$,
$$0 < |f^{(k)}(0)| \leq \sup_{s\geq 0} |f^{(k)}(s)| < \infty,$$
with $f^{(k)}(0) = \lim_{x\downarrow 0} f^{(k)}(x)$, and $f^{(i)}(0) = 0$ for $1 \leq i \leq k-1$.

Under these conditions we determine the behavior of the Grenander estimator near zero. With the proper normalizing constants the limit distribution of $n^{-\beta}\big(\hat{f}_n(n^{-\alpha}) - f(n^{-\alpha})\big)$ is independent of f. Define $D[Z(t)](a)$ as the right derivative of the LCM on $\mathbb{R}$ of the process $Z(t)$ at the point $t = a$, and define D_R similarly, where the LCM is restricted to the set $t \geq 0$.

Theorem 3.1 *Suppose f satisfies conditions (C1)-(C2) and let $c > 0$. Then*

(i) for $1/(2k+1) < \alpha < 1$ and $A_1 = (c/f(0))^{1/2}$, the sequence
$$A_1 n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right)$$
converges in distribution to $D_R\,[W(t)]\,(1)$, as $n \to \infty$.

(ii) for $A_{2k} = \sqrt{B_{2k}/f(0)}$ and $B_{2k} = \left(f(0)^{1/2}|f^{(k)}(0)|^{-1}(k+1)!\right)^{2/(2k+1)}$, the sequence
$$A_{2k}\left\{n^{k/(2k+1)}\left(\hat{f}_n\big(cB_{2k}n^{-1/(2k+1)}\big) - f\big(cB_{2k}n^{-1/(2k+1)}\big)\right) - \frac{|f^{(k)}(0)|(cB_{2k})^k}{k!}\right\}$$
converges in distribution to $D_R\left[W(t) - t^{k+1}\right](c)$, as $n \to \infty$.

(iii) for $0 < \alpha < 1/(2k+1)$ and $A_{3k} = (2(k-1)!)^{1/3}\,|f(0)f^{(k)}(0)c^{k-1}|^{-1/3}$, the sequence
$$A_{3k}n^{1/3+\alpha(k-1)/3}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right)$$
converges in distribution to $D\left[W(t) - t^2\right](0)$, as $n \to \infty$.

Remark 3.1 *In order to formulate the limiting distributions in Theorem 3.1 in a similar way, they have been expressed in terms of slopes of least concave majorants. However, note that similar to switching relation (2.1), one finds that*

$$\begin{aligned} D_R\left[W(t)\right](1) &\stackrel{d}{=} \sqrt{\underset{t\in[0,\infty)}{\operatorname{argmax}}\{W(t)-t\}} \\ D\left[W(t)-t^2\right](0) &\stackrel{d}{=} 2\underset{t\in\mathbb{R}}{\operatorname{argmax}}\left\{W(t)-t^2\right\}. \end{aligned}$$

In studying the behavior of (1.2), we follow the line of reasoning described in Section 2. We start by establishing convergence in distribution of the relevant processes. It turns out that we have to distinguish between three cases concerning the rate at which $n^{-\alpha}$ tends to zero.

Lemma 3.1 *Suppose f satisfies (C1)-(C2) and let W denote standard two-sided Brownian motion on $\mathbb{R}$. For $1/(2k+1) \le \alpha < 1$, $t \ge 0$ and $x \in \mathbb{R}$, define*

$$Z_{n1}(x,t) = n^{(1+\alpha)/2}\left(F_n(tn^{-\alpha}) - f(0)tn^{-\alpha}\right) - xt.$$

- *(i) For $1/(2k+1) < \alpha < 1$, the process $\{Z_{n1}(x,t) : t \in [0,\infty)\}$ converges in distribution, in the uniform topology on compacta, to the process $\{W(f(0)t) - xt : t \in [0,\infty)\}$.*
- *(ii) For $\alpha = 1/(2k+1)$, the process $\{Z_{n1}(x,t) : t \in [0,\infty)\}$ converges in distribution, in the uniform topology on compacta, to $\{W(f(0)t) - xt - |f^{(k)}(0)|t^{k+1}/(k+1)! : t \in [0,\infty)\}$.*
- *(iii) For $0 < \alpha < 1/(2k+1)$, $b = (1-2\alpha(k-1))/3$, $t \ge -cn^{b-\alpha}$ and $x \in \mathbb{R}$, define*

$$Z_{n2}(x,t) = n^{(b+1)/2}\left(F_n(cn^{-\alpha}+tn^{-b}) - F_n(cn^{-\alpha}) - f(cn^{-\alpha})tn^{-b}\right) - xt.$$

Then the process $\{Z_{n2}(x,t) : t \in [-cn^{b-\alpha},\infty)\}$ converges in distribution, in the uniform topology on compacta, to the process $\{W(f(0)t) - xt - c^{k-1}|f^{(k)}(0)|t^2/(2(k-1)!) : t \in \mathbb{R}\}$.

The next step is to use Theorem 2.1. The major difficulty is to verify condition (iii) of this theorem. The following lemma ensures that this condition is satisfied.

Lemma 3.2 *Let f satisfy (C1)-(C2) and let Z_{n1}, Z_{n2}, and b be defined as in Lemma 3.1.*

- *(i) For $1/(2k+1) < \alpha < 1$ and $x > 0$, $\operatorname{argmax}_{t\in[0,\infty)} Z_{n1}(x,t) = \mathcal{O}_p(1)$.*
- *(ii) For $\alpha = 1/(2k+1)$ and $x \in \mathbb{R}$, $\operatorname{argmax}_{t\in[0,\infty)} Z_{n1}(x,t) = \mathcal{O}_p(1)$.*
- *(iii) For $0 < \alpha < 1/(2k+1)$ and $x \in \mathbb{R}$, $\operatorname{argmax}_{t\in[-cn^{b-\alpha},\infty)} Z_{n2}(x,t) = \mathcal{O}_p(1)$.*

With Lemmas 3.1 and 3.2 at hand, the proof of Theorem 3.1 consists of using switching relation (2.1) and an application of Theorem 2.1.

Proof of Theorem 3.1: (i) First note that by condition (C2),

$$n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right) = n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(0)\right) + \mathcal{O}(n^{(1-(2k+1)\alpha)/2}),$$

where $(1-(2k+1)\alpha)/2 < 0$. For $x > 0$, according to (2.1),

$$P\left\{n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(0)\right) \le x\right\} = P\left\{n^{\alpha}U_n(f(0)+xn^{-(1-\alpha)/2}) \le c\right\}. \tag{3.1}$$

If Z_{n1} is the process defined in Lemma 3.2(i), then

$$0 \leq n^{\alpha} U_n(f(0) + xn^{-(1-\alpha)/2}) = \underset{t\in[0,\infty)}{\operatorname{argmax}} Z_{n1}(x,t) = \mathcal{O}_p(1), \tag{3.2}$$

where, according to Lemma 3.1, the process $\{Z_{n1}(x,t) : t \in [0,\infty)\}$ converges in distribution to the process $\{W(f(0)t) - xt : t \in [0,\infty)\}$. To apply Theorem 2.1, we have to extend the above processes to the whole real line. Therefore define

$$\tilde{Z}_{n1}(t) = \begin{cases} Z_{n1}(x,t) & ,t \geq 0, \\ t & ,t \leq 0. \end{cases}$$

Then, for x fixed, $\tilde{Z}_{n1}$ converges in distribution to the process Z_1, where

$$Z_1(t) = \begin{cases} W(f(0)t) - xt & ,t \geq 0, \\ t & ,t \leq 0. \end{cases}$$

Moreover, since $Z_{n1}(x,0) = 0$, together with (3.2), it follows that

$$\underset{t\in\mathbb{R}}{\operatorname{argmax}} \tilde{Z}_{n1}(t) = \underset{t\in[0,\infty)}{\operatorname{argmax}} \tilde{Z}_{n1}(t) = n^{\alpha} U_n(f(0) + xtn^{-(1-\alpha)/2}) = \mathcal{O}_p(1).$$

The process Z_1 is continuous and since $\operatorname{Var}(Z_1(s) - Z_1(t)) \neq 0$, for $s,t > 0$ with $s \neq t$, it follows from Lemma 2.6 in Kim and Pollard (1990) that Z_1 has a unique maximum with probability one. By an application of the law of iterated logarithm for Brownian motion:

$$P\left\{\limsup_{|u|\to\infty} \frac{W(u)}{\sqrt{2|u|\log\log|u|}} = 1\right\} = 1, \tag{3.3}$$

it can be seen that $Z_1(t) \to -\infty$, as $|t| \to \infty$. Theorem 2.1 now yields that $\operatorname{argmax}_{t\in\mathbb{R}} \tilde{Z}_{n1}(t)$ converges in distribution to $\operatorname{argmax}_{t\in\mathbb{R}} Z_1(t) = \operatorname{argmax}_{t\geq 0} \{W(f(0)t) - xt\}$. Using (3.1), together with (2.3) this implies that

$$\begin{aligned} P\left\{n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(0)\right) \leq x\right\} &= P\left\{\underset{t\in\mathbb{R}}{\operatorname{argmax}} \tilde{Z}_{n1}(t) \leq c\right\} \\ &\to P\left\{\underset{t\geq 0}{\operatorname{argmax}} \{W(f(0)t) - xt\} \leq c\right\} \\ &= P\left\{\underset{t\geq 0}{\operatorname{argmax}} \left\{W(t) - \frac{xc^{1/2}t}{f(0)^{1/2}}\right\} \leq 1\right\}. \end{aligned}$$

Similar to switching relation (2.1), the right hand side equals $P\{(f(0)/c)^{1/2} D_R[W(t)](1) \leq x\}$, so that it remains to show that $P\{n^{(1-\alpha)/2}(\hat{f}_n(cn^{-\alpha}) - f(0)) \leq 0\} \to 0$. But this is evident, as for any $\epsilon > 0$, using (2.3) once more,

$$\begin{aligned} P\left\{n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(0)\right) \leq 0\right\} &\leq P\left\{n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(0)\right) \leq \epsilon\right\} \\ &\to P\left\{\underset{t\geq 0}{\operatorname{argmax}} \left\{W(t) - \frac{\epsilon t}{\sqrt{f(0)}}\right\} \leq c\right\} \\ &= P\left\{\underset{t\geq 0}{\operatorname{argmax}} \{W(t) - t\} \leq \frac{c\epsilon^2}{f(0)}\right\}. \end{aligned}$$

When $\epsilon \downarrow 0$, the right hand side tends to zero, which can be seen from

$$P\left\{\limsup_{t\downarrow 0}\frac{W(t)}{\sqrt{2t\log\log(1/t)}}=1\right\}=1.$$

This proves (i).

(ii) First note that by (C2) and the fact that $f^{(k)}(0)<0$ (see the proof of Lemma 3.1),

$$\begin{aligned}&n^{k/(2k+1)}\left(\hat{f}_n(cB_{2k}n^{-1/(2k+1)})-f(cB_{2k}n^{-1/(2k+1)})\right)-|f^{(k)}(0)|\frac{(cB_{2k})^k}{k!}\\&=n^{k/(2k+1)}\left(\hat{f}_n(cB_{2k}n^{-1/(2k+1)})-f(0)\right)+o(1),\end{aligned}$$

and that according to (2.1), $P\left\{n^{k/(2k+1)}\left(\hat{f}_n(cB_{2k}n^{-1/(2k+1)})-f(0)\right)\le x\right\}$ is equal to

$$P\left\{B_{2k}^{-1}n^{1/(2k+1)}U_n\left(f(0)+xn^{-k/(2k+1)}\right)\le c\right\}.$$

With Z_{n1} being the process defined in Lemma 3.1 with $\alpha=1/(2k+1)$, we get

$$B_{2k}^{-1}n^{1/(2k+1)}U_n\left(f(0)+xn^{-k/(2k+1)}\right)=\underset{t\in[0,\infty)}{\operatorname{argmax}}\left\{Z_{n1}(x,B_{2k}t)\right\}=\mathcal{O}_p(1).$$

Again, we first extend the above process to the whole real line:

$$\tilde{Z}_{n1}(t)=\begin{cases}Z_{n1}(x,B_{2k}t) & ,t\ge 0,\\ t & ,t\le 0.\end{cases}$$

Then, according to Lemma 3.1, $\tilde{Z}_{n1}$ converges in distribution to the process

$$Z_2(t)=\begin{cases}W(f(0)B_{2k}t)-B_{2k}xt-|f^{(k)}(0)|B_{2k}^{k+1}t^{k+1}/(k+1)! & ,t\ge 0,\\ t & ,t\le 0.\end{cases}$$

Similar to the proof of (i), it follows from Theorem 2.1 that $\operatorname{argmax}_t \tilde{Z}_{n1}(t)$ converges in distribution to $\operatorname{argmax}_t Z_2(t)$. This implies that

$$\begin{aligned}&P\left\{A_{2k}n^{k/(2k+1)}\left(\hat{f}_n(cB_{2k}n^{-1/(2k+1)})-f(0)\right)\le x\right\}\\&\to P\left\{\underset{t\ge 0}{\operatorname{argmax}}\left\{W(f(0)B_{2k}t)-A_{2k}^{-1}B_{2k}xt-|f^{(k)}(0)|B_{2k}^{k+1}t^{k+1}/(k+1)!\right\}\le c\right\}\\&=P\left\{\underset{t\ge 0}{\operatorname{argmax}}\left\{W(t)-xt-t^{k+1}\right\}\le c\right\}\\&=P\left\{D_R\left[W(t)-t^{k+1}\right](c)\le x\right\},\end{aligned}$$

by means of Brownian scaling similar to (2.3), and a switching relation similar to (2.1).

(iii) According to (2.1), we have

$$\begin{aligned}&P\left\{n^{(1-b)/2}\left(\hat{f}_n(cn^{-\alpha})-f(cn^{-\alpha})\right)\le x\right\}\\&=P\left\{n^b\left(U_n(f(cn^{-\alpha})+xn^{-(1-b)/2})-cn^{-\alpha}\right)\le 0\right\},\end{aligned}\tag{3.4}$$

and with Z_{n2} as defined in Lemma 3.2(iii), we get

$$n^b \left(U_n(f(cn^{-\alpha}) + xn^{-(1-b)/2}) - cn^{-\alpha} \right) = \underset{t \in [-cn^{b-\alpha}, \infty)}{\text{argmax}} \; Z_{n2}(x,t) = \mathcal{O}_p(1).$$

As in the proof of (i) and (ii), we extend the above process to the whole real line:

$$\tilde{Z}_{n2}(t) = \begin{cases} Z_{n2}(x,t) & , t \geq -cn^{b-\alpha}, \\ Z_{n2}(x, -cn^{b-\alpha}) + (t + cn^{b-\alpha}) & , t < -cn^{b-\alpha}. \end{cases}$$

Then, by Lemma 3.1, Z_{n2} converges in distribution to the process Z_3, where

$$Z_3(t) = W(f(0)t) - xt - \frac{|f^{(k)}(0)|c^{k-1}}{2(k-1)!} t^2, \quad t \in \mathbb{R}.$$

Similar to the proof of (i) and (ii), it follows from Theorem 2.1 that $\text{argmax}_t \, Z_{n2}(t)$ converges in distribution to $\text{argmax}_t \, Z_3(t)$. Together with (3.4), this implies that

$$\begin{aligned} & P\left\{ n^{(1-b)/2} A_{3k} \left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha}) \right) \leq x \right\} \\ & \to P\left\{ \underset{t \in \mathbb{R}}{\text{argmax}} \left\{ W(f(0)t) - A_{3k}^{-1} xt - \frac{|f^{(k)}(0)|c^{k-1}}{2(k-1)!} t^2 \right\} \leq 0 \right\} \\ & = P\left\{ \underset{t \in \mathbb{R}}{\text{argmax}} \left\{ W(t) - xt - t^2 \right\} \leq 0 \right\} \\ & = P\left\{ D\left[W(t) - t^2 \right](0) \leq x \right\}, \end{aligned}$$

again using Brownian scaling similar to (2.3), and a switching relation similar to (2.1). ■

4 Behavior near the end of the support

Suppose that f has compact support and, without loss of generality, assume this to be the interval $[0,1]$. In this section we investigate the behavior of $\hat{f}_n$ near one. Although there seems to be no simple symmetry argument to derive the behavior near one from the results in Section 3, the arguments to obtain the behavior of

$$n^\beta \left\{ f(1 - n^{-\alpha})) - \hat{f}_n(1 - n^{-\alpha}) \right\},$$

are similar to the ones used in studying (1.2). If $f(1) > 0$, then $\hat{f}_n(1)$ will always under estimate $f(1)$, since by definition $\hat{f}_n(1) = 0$. Nevertheless, the behavior near the end of the support is similar to the behavior near zero.

For this reason, we only provide the statement of a theorem for the end of the support, which is analogous to Theorem 3.1. Motivations for studying the behavior near the end of the support are not so strong as for the behavior near zero. However, the behavior near one is required for establishing the asymptotic normality of the L_k-distance between $\hat{f}_n$ and f. Similar to (C1) and (C2) we will assume that

(C3) $0 < f(1) = \lim_{x \uparrow 1} f(x) < \infty$.

(C4) For some $k \geq 1$

$$0 < |f^{(k)}(1)| \leq \sup_{0 \leq s \leq 1} |f^{(k)}(s)| < \infty,$$

with $f^{(k)}(1) = \lim_{x \uparrow 1} f^{(k)}(x)$, and $f^{(i)}(1) = 0$ for $1 \leq i \leq k-1$.

We then have the following theorem.

Theorem 4.1 *Suppose f satisfies conditions (C3)-(C4) and $c > 0$. Then*

(i) for $1/(2k+1) < \alpha < 1$ and $\tilde{A}_1 = (c/f(1))^{1/2}$, the sequence

$$\tilde{A}_1 n^{(1-\alpha)/2} \left(f(1 - cn^{-\alpha}) - \hat{f}_n(1 - cn^{-\alpha}) \right)$$

converges in distribution to $D_R \left[W(t) \right](1)$, as $n \to \infty$.

(ii) for $\tilde{A}_{2k} = \sqrt{\tilde{B}_{2k}/f(1)}$ and $\tilde{B}_{2k} = (f(1)^{1/2}|f^{(k)}(1)|^{-1}((k+1)!))^{2/(2k+1)}$, the sequence

$$\tilde{A}_{2k} \left\{ n^{k/(2k+1)} \left(f(1 - c\tilde{B}_{2k}n^{-1/(2k+1)}) - \hat{f}_n(1 - c\tilde{B}_{2k}n^{-1/(2k+1)}) \right) - \frac{|f^{(k)}(1)|(c\tilde{B}_{2k})^k}{k!} \right\}$$

converges in distribution to $D_R \left[W(t) - t^{k+1} \right](c)$, as $n \to \infty$.

(iii) for $0 < \alpha < 1/(2k+1)$ and $\tilde{A}_{3k} = ((k-1)!)^{1/3}|4f(1)f^{(k)}(1)c^{k-1}|^{-1/3}$, the sequence

$$\tilde{A}_{3k} n^{1/3 + \alpha(k-1)/3} \left(f(1 - cn^{-\alpha}) - \hat{f}_n(1 - cn^{-\alpha}) \right)$$

converges in distribution to $D \left[W(t) - t^2 \right](0)$, as $n \to \infty$.

The proof of this theorem follows from convergence in distribution of the relevant processes, as described in the next lemma.

Lemma 4.1 *Suppose f satisfies (C3)-(C4) and let W denote standard two-sided Brownian motion on $\mathbb{R}$. For $1/(2k+1) \leq \alpha < 1$ and $t, x \in \mathbb{R}$, define $Y_{n1}(x,t)$ by*

$$Y_{n1}(x,t) = n^{(1+\alpha)/2} \left(F_n(1 - tn^{-\alpha}) - F_n(1) + f(1)tn^{-\alpha} \right) - xt.$$

(i) For $1/(2k+1) < \alpha < 1$, the process $\{Y_{n1}(x,t) : t \in [0, n^\alpha]\}$ converges in distribution, in the uniform topology on compacta, to the process $\{W(f(1)t) - xt : t \in [0, \infty)\}$.

(ii) For $\alpha = 1/(2k+1)$, the process $\{Y_{n1}(x,t) : t \in [0, n^\alpha]\}$ converges in distribution, in the uniform topology on compacta, to $\{W(f(1)t) - xt - |f^{(k)}(1)|t^{k+1}/(k+1)! : t \in [0, \infty)\}$.

(iii) For $0 < \alpha < 1/(2k+1)$ and $b = (1 - 2\alpha(k-1))/3$, define $Y_{n2}(x,t)$ by

$$Y_{n2}(x,t) = n^{(1+b)/2} \left(F_n(1 - cn^{-\alpha} - tn^{-b}) - F_n(1 - cn^{-\alpha}) + f(1 - cn^{-\alpha})tn^{-b} \right) - xt.$$

Then the process $\{Y_{n2}(x,t) : t \in [-cn^{b-\alpha}, n^b(1 - cn^{-\alpha})]\}$ converges in distribution, in the uniform topology on compacta, to $\{W(f(1)t) - xt - c^{k-1}|f^{(k)}(1)|t^2/(2(k-1)!) : t \in \mathbb{R}\}$.

Next, we apply Theorem 2.1. Again the major difficulty is establishing condition (iii) of this theorem. This is ensured by the following lemma.

Lemma 4.2 *Let f satisfies (C3)-(C4) and let Y_{n1}, Y_{n2}, and b be defined as in Lemma 4.1.*

(i) For $1/(2k+1) < \alpha < 1$ and $x > 0$, $\mathrm{argmax}_{t \in [0, n^\alpha]} Y_{n1}(x,t) = O_p(1)$.

(ii) For $\alpha = 1/(2k+1)$ and $x \in \mathbb{R}$, $\mathrm{argmax}_{t \in [0, n^\alpha]} Y_{n1}(x,t) = O_p(1)$.

(iii) For $0 < \alpha < 1/(2k+1)$, $\mathrm{argmax}_{t \in [-cn^{b-\alpha}, n^b(1 - cn^{-\alpha})]} Y_{n2}(x,t) = O_p(1)$.

5 A comparison with the penalized NPMLE

Consider a decreasing density f on $[0,\infty)$. We first consider the case where $f'(0)<0$. As pointed out in Woodroofe and Sun (1993), the NPMLE $\hat f_n$ for f is not consistent at zero. They proposed a penalized NPMLE $\hat f_n^{\mathrm{P}}(\alpha,0)$, and in Sun and Woodroofe (1996) it was shown that

$$n^{1/3}\left\{\hat f_n^{\mathrm{P}}(\alpha_n,0)-f(0)\right\}\to\sup_{t>0}\frac{W(t)-(c-\frac{1}{2}f(0)f'(0)t^2)}{t},$$

where c is related to the smoothing parameter $\alpha_n=cn^{-2/3}$. Sun and Woodroofe (1996) also provide (to some extent) an adaptive choice for c that leads to an estimate $\hat\alpha_n$ of the smoothing parameter, and report some results of a simulation experiment for $\hat f_n^{\mathrm{P}}(\hat\alpha_n,0)$.

We propose two consistent estimators of $f(0)$ both converging at rate $n^{1/3}$. A simple estimator is $\hat f_n^{\mathrm{S}}(0)=\hat f_n(n^{-1/3})$. This estimator is straightforward and does not have any additional smoothing parameters. According to Theorem 3.1(ii), $\hat f_n^{\mathrm{S}}(0)$ is a consistent estimator for $f(0)$, converging at rate $n^{1/3}$. It has a limiting distribution that is a functional of W:

$$A_{21}n^{1/3}\left\{\hat f_n^{\mathrm{S}}(0)-f(0)\right\}\to D_R\left[W(t)-t^2\right](1/B_{21}),$$

where A_{21} and B_{21} are defined in Theorem 3.1(ii). In order to reduce the mean squared error, we also propose an adaptive estimator $\hat f_n^{\mathrm{A}}(0)=\hat f_n(c_1^*\hat B_{21}n^{-1/3})$ for $f(0)$. Here c_k^* is the value that minimizes $E(D_R[W(t)-t^{k+1}](c))^2$, and $\hat B_{21}$ is an estimate for the constant $B_{21}=4^{1/3}f(0)^{1/3}|f'(0)|^{-2/3}$ in Theorem 3.1(ii). Computer simulations show that $c_k^*\approx 0.345$ for both $k=1$ and $k=2$. For $\hat B_{21}$ we take $\hat B_{21}=4^{1/3}\hat f_n^{\mathrm{S}}(0)^{1/3}|\tilde f_n'(0)|^{-2/3}$, where $\tilde f_n'(0)=\min(n^{1/6}(\hat f_n(n^{-1/6})-\hat f_n(n^{-1/3})),-n^{-1/3})$ is an estimate for $f'(0)$. As we have seen above, $\hat f_n^{\mathrm{S}}(0)$ is consistent for $f(0)$, and according to Theorem 3.1, $\tilde f_n'(0)$ is consistent for $f'(0)$. When f is twice continuously differentiable, it converges at rate $n^{1/6}$. Therefore $\hat B_{21}$ is consistent for B_{21} and $\hat f_n^{\mathrm{A}}(0)$ is a consistent estimator of $f(0)$, converging with rate $n^{1/3}$. It has the following limit behavior:

$$A_{21}n^{1/3}\left\{\hat f_n^{\mathrm{A}}(0)-f(0)\right\}\to D_R\left[W(t)-t^2\right](c_1^*),$$

where A_{21} is defined in Theorem 3.1(ii).

We simulated 10000 samples of sizes $n=50,100,200$, and 10000 from a standard exponential distribution with mean one. For each sample, the values of $n^{1/3}\{\hat f_n^{\mathrm{S}}(0)-f(0)\}$, $n^{1/3}\{\hat f_n^{\mathrm{A}}(0)-f(0)\}$ and $n^{1/3}\{\hat f_n^{\mathrm{P}}(\hat\alpha_n,0)-f(0)\}$ were computed. The value of $\hat\alpha_n$ was computed as proposed in Sun and Woodroofe (1996), $\hat\alpha_n=0.649\cdot\hat\beta_n^{-1/3}n^{-2/3}$, where

$$\hat\beta_n=\max\left\{\hat f_n^{\mathrm{P}}(\alpha_0,0)\frac{\hat f_n^{\mathrm{P}}(\alpha_0,0)-\hat f_n^{\mathrm{P}}(\alpha_0,x_m)}{2x_m},n^{-q}\right\},$$

is an estimate for $\beta=-f(0)f'(0)/2$. Here, x_m denotes the second point of jump of the penalized NPMLE $\hat f_n^{\mathrm{P}}(\alpha_0,\cdot)$ computed with smoothing parameter α_0. The parameter $\alpha_0=c_0n^{-2/3}$, and q should be taken between 0 and 0.5. However, Sun and Woodroofe (1996) do not specify how to choose q and c_0 in general. We took $q=1/3$, and for α_0 the values as listed in their Table 2: $\alpha_0=0.0516, 0.0325$ and 0.0205 for sample sizes $n=50,100$ and 200. For sample size $n=10000$ we took the theoretical optimal value $\alpha_0=0.649\,\beta^{-1/3}n^{-2/3}$, with

n	$n^{1/3}\{\hat f_n^{\mathrm{S}}(0)-f(0)\}$			$n^{1/3}\{\hat f_n^{\mathrm{A}}(0)-f(0)\}$			$n^{1/3}\{\hat f_n^{\mathrm{P}}(\hat\alpha_n,0)-f(0)\}$		
	Mean	Variance	MSE	Mean	Variance	MSE	Mean	Variance	MSE
50	−0.847	0.439	1.157	−0.738	0.934	1.478	−0.072	1.296	1.301
100	−0.853	0.484	1.211	−0.777	0.742	1.345	−0.079	1.530	1.537
200	−0.868	0.536	1.289	−0.793	0.807	1.436	−0.075	1.732	1.738
10000	−0.917	0.700	1.541	−0.643	1.045	1.458	−0.195	1.913	1.951

Table 1: Simulated mean, variances and mean squared error for the three estimators at the standard exponential distribution.

$\beta = 0.5$. It is worth noticing that Sun and Woodroofe (1996) do not optimize the MSE, but $n^{1/3}E|\hat f_n^{\mathrm{P}}(\hat\alpha_n,0)-f(0)|$. Nevertheless, computer simulations show that the α_n minimizing the MSE is approximately the same and that $n^{2/3}E|\hat f_n^{\mathrm{P}}(\alpha,0)-f(0)|^2$ is a very flat function in a neighborhood of α_n. A similar property holds for the value c_k^* minimizing the AMSE of our estimator.

In Table 1 we listed simulated values for the mean, variance and mean squared error of the three estimators. The penalized NPMLE is less biased, but has a larger variance. Estimator $\hat f_n^{\mathrm{A}}(0)$ performs better in the sense of mean squared error, approaching the best theoretically expected performance. It is also remarkable how good it mimics its limiting distribution for already small samples. Estimator $\hat f_n^{\mathrm{S}}(0)$ performs a little worse than $\hat f_n^{\mathrm{A}}(0)$, having the largest bias, but the smallest variance.

If $k=2$ in condition (C2), it is possible to estimate $f(0)$ at a rate faster than $n^{1/3}$. If it is known in advance that $k=2$, we can produce two consistent estimators of $f(0)$ converging at rate $n^{2/5}$. Similar to the previous case, a simple estimator is $\hat f_n^{\mathrm{S},2}(0)=\hat f_n(n^{-1/5})$. It is a consistent estimator of $f(0)$, converging at rate $n^{2/5}$, and has the following limit behavior:

$$A_{22}n^{2/5}\left\{\hat f_n^{\mathrm{S},2}(0)-f(0)\right\}\to D_R\left[W(t)-t^3\right](1/B_{22}),$$

where A_{22} and B_{22} are defined in Theorem 3.1(ii). Again, we also propose an adaptive estimator $\hat f_n^{\mathrm{A},2}(0)=\hat f_n(c_2^*\hat B_{22}n^{-1/5})$ for $f(0)$, where $\hat B_{22}$ is an estimate for the constant $B_{22}=36^{1/5}f(0)^{1/5}|f''(0)|^{-2/5}$ in Theorem 3.1(ii), and $c_2^*\approx 0.345$ is the value that minimizes $E(D_R[W(t)-t^3](c))^2$. For $\hat B_{22}$ we take $\hat B_{22}=36^{1/5}\hat f_n^{\mathrm{S},2}(0)^{1/5}|\tilde f_n''(0)|^{-2/5}$, where we estimate $f''(0)$ by $\tilde f_n''(0)=\min(2n^{1/4}(\hat f_n(n^{-1/8})-\hat f_n(n^{-1/5})),-n^{-1/5})$. As we have seen above, $\hat f_n^{\mathrm{S},2}(0)$ is consistent for $f(0)$, and according to Theorem 3.1, $\tilde f_n''(0)$ is consistent for $f''(0)$ with rate $n^{1/8}$ if f is three times continuously differentiable. Therefore $\hat B_{22}$ is a consistent estimator for B_{22} and $\hat f_n^{\mathrm{A},2}(0)$ is a consistent estimator of $f(0)$, converging with rate $n^{2/5}$:

$$A_{22}n^{2/5}\left\{\hat f_n^{\mathrm{A},2}(0)-f(0)\right\}\to D_R\left[W(t)-t^3\right](c_2^*),$$

where A_{22} is defined in Theorem 3.1(ii).

We simulated 10000 samples of sizes $n=50,100,200$, and 10000 from a half-normal distribution. For each sample, the values of $n^{2/5}\{\hat f_n^{\mathrm{S},2}(0)-f(0)\}$ and $n^{2/5}\{\hat f_n^{\mathrm{A},2}(0)-f(0)\}$ were computed. Sun and Woodroofe (1996) do not consider the possibility of constructing a special estimator for the case $k=2$, though we believe that this is also possible with a penalization technique. In Table 2 we listed simulated values for the mean, variance and

n	$n^{2/5}\{\hat{f}_n^{\text{S},2}(0)-f(0)\}$			$n^{2/5}\{\hat{f}_n^{\text{A},2}(0)-f(0)\}$		
	Mean	Variance	MSE	Mean	Variance	MSE
50	−0.429	0.371	0.555	−0.252	0.459	0.523
100	−0.437	0.402	0.592	−0.278	0.502	0.579
200	−0.440	0.440	0.634	−0.373	0.549	0.688
10000	−0.419	0.559	0.735	−0.326	0.747	0.853

Table 2: Simulated mean, variances and mean squared error for both estimators at the half-normal distribution.

mean squared error of both estimators. The simple estimator is more biased but its variance is smaller than variance of the adaptive one.

If it is not known in advance that $k=2$, then application of estimators $\hat{f}_n^{\text{S},2}(0)$ and $\hat{f}_n^{\text{A},2}(0)$ is undesirable. If in fact $k=1$, they are still consistent, but their convergence rate will be $n^{1/5}$. On the other hand, when $k=2$, then $\hat{f}_n^{\text{S}}(0)$, $\hat{f}_n^{\text{A}}(0)$, and $f_n^{\text{P}}(\hat{\alpha}_n, 0)$ are still applicable. In that case, according to Theorem 3.1(i), $\hat{f}_n^{\text{S}}(0)$ is a consistent estimator of $f(0)$ converging at rate $n^{1/3}$, such that

$$n^{1/3}\left\{\hat{f}_n^{\text{S}}(0)-f(0)\right\} \to \sqrt{f(0)}D_R\left[W(t)\right](1).$$

Also $\hat{f}_n^{\text{A}}(0)$ is still consistent for $f(0)$ in case $k=2$, but now at rate $n^{7/18}$. This can be seen as follows. Since $f'(0)=0$, it follows that

$$n^{1/6}\tilde{f}_n'(0) \to -\sqrt{f(0)}D_R\left[W(t)\right](1)+\frac{f''(0)}{2}.$$

As $\hat{f}_n^{\text{S}}(0)=f(0)+\mathcal{O}_p(n^{-1/3})$, this implies that $\hat{B}_{21}n^{-1/3}=\mathcal{O}_p(n^{-2/9})$. Application of Theorem 3.1(i) yields that $\hat{f}_n^{\text{A}}(0)=f(0)+\mathcal{O}_p(n^{-7/18})$. Sun and Woodroofe (1996) also propose to use $\hat{f}_n^{\text{P}}(\hat{\alpha}_n, 0)$ as an estimate of $f(0)$ in the case $k\geq 2$. They prove that in that case $n^{1/3}\{\hat{f}_n^{\text{P}}(\hat{\alpha}_n,0)-f(0)\}\to 0$.

We simulated 10000 samples of sizes $n=50, 100, 200$, and 10000 from a standard half-normal distribution. For each sample the values of $n^{1/3}\{\hat{f}_n^{\text{S}}(0)-f(0)\}$, $n^{1/3}\{\hat{f}_n^{\text{A}}(0)-f(0)\}$ and $n^{1/3}\{\hat{f}_n^{\text{P}}(\hat{\alpha}_n,0)-f(0)\}$ were computed. In Table 3 we listed simulated values for the mean, variance and mean squared error of the three estimators. The simple estimator has the smallest variance, but as the sample size increases it becomes more biased. Nevertheless,

n	$n^{1/3}\{\hat{f}_n^{\text{S}}(0)-f(0)\}$			$n^{1/3}(\hat{f}_n^{\text{A}}(0)-f(0)\}$			$n^{1/3}\{\hat{f}_n^{\text{P}}(\hat{\alpha}_n,0)-f(0)\}$		
	Mean	Variance	MSE	Mean	Variance	MSE	Mean	Variance	MSE
50	0.012	0.320	0.320	0.046	0.475	0.477	0.331	0.659	0.768
100	0.058	0.317	0.320	0.073	0.406	0.412	0.336	0.742	0.855
200	0.104	0.316	0.327	0.091	0.383	0.391	0.338	0.812	0.926
10000	0.269	0.296	0.368	0.204	0.319	0.361	0.279	0.714	0.792

Table 3: Simulated mean, variances and mean squared error for the three estimators at the half-normal distribution.

Estimator	Exponential			Half-normal		
	Mean	Variance	MSE	Mean	Variance	MSE
$n^{1/3}\{\hat{f}_n^{\mathrm{S}}(0) - f(0)\}$	−0.885	0.805	1.591	0.336	0.316	0.429
$n^{1/3}\{\hat{f}_n(c_1^* B_{21} n^{-1/3}) - f(0)\}$	−0.298	1.043	1.131	0	0	0
$n^{1/3}\{\hat{f}_n^{\mathrm{P}}(\hat{\alpha}_n, 0) - f(0)\}$	−0.349	1.096	1.218	0	0	0
$n^{2/5}\{\hat{f}_n^{\mathrm{S},2}(0) - f(0)\}$	$-\infty$	∞	∞	−0.415	0.670	0.842
$n^{2/5}\{\hat{f}_n(c_2^* B_{22} n^{-1/5}) - f(0)\}$	$-\infty$	∞	∞	−0.140	0.718	0.737

Table 4: Theoretical limiting mean, variances and mean squared error for the three estimators.

it is stable for already small sample sizes. The adaptive estimator becomes more biased with growing sample size, but with smaller MSE. The penalized MLE is most biased, also having a much larger variance than its simple and adaptive competitors.

Finally, in Table 4 we listed the true limiting values for the mean, variance, and MSE, for all estimators at the exponential and half-normal distribution. The finite sample behavior of the simple estimators $\hat{f}_n^{\mathrm{S}}(0)$ (see Tables 1 and 3) and $\hat{f}_n^{\mathrm{S},2}(0)$ (see Table 2) reasonably matches the theoretical behavior. The adaptive estimators exhibit larger deviations from their theoretical values. This is probably explained by the fact that even for larger sample sizes, the estimation of the derivatives of f in B_{2k} still has a large influence.

One might prefer a scale-equivariant version of the above estimators. One possibility is $\hat{f}_n(X_{m:n})$, where $X_{m:n}$ denotes the mth order statistic. The sequence $m = m(n)$ should be chosen in such a way that $m(n) \to \infty$ and $m(n)/n \to 0$, e.g., $m(n) = \lfloor an^{2/3} \rfloor$. In that case, one can show that $\hat{f}_n(X_{m:n})$ is asymptotically equivalent to $\hat{f}_n(af(0)^{-1}n^{-1/3})$. Its limiting distribution can be obtained from Theorem 3.1 and the AMSE optimal choice a^* will depend on $f(0)$ and $f'(0)$. For this choice, $\hat{f}_n(a^* f(0)^{-1} n^{-1/3})$ has the same behavior as $\hat{f}_n(c_1^* B_{21} n^{-1/3})$. Another possibility is to estimate $f(0)$ by means of a numerical derivative of F_n:

$$\hat{f}_n^{\mathrm{D}}(0) = \frac{F_n(X_{m:n})}{X_{m:n}} = \frac{m/n}{X_{m:n}},$$

where $m = m(n)$ as above. For this estimator it can be shown that $n^{1/3}\{\hat{f}_n^{\mathrm{D}}(0) - f(0)\}$ is asymptotically normal with mean $-|f'(0)|a/(2f(0))$ and variance $f(0)^2/a$. This implies that the minimal AMSE is a multiple of $(f(0)|f'(0)|)^{2/3}$, which also holds for $\hat{f}_n^{\mathrm{S}}(0)$ and $\hat{f}_n^{\mathrm{A}}(0)$ (see Theorem 3.1(ii) for the case $k = 1$). Computer simulations show that the AMSE of $\hat{f}_n^{\mathrm{A}}(0)$ is always the smallest of the three.

Acknowledgements: The authors wish to thank two unknown referees for their fruitful suggestions that has lead to a substantial improvement of the original manuscript. We thank the first referee for his/her careful reading of the manuscript and for providing a neat closed expression of the limit distribution in Theorem 3.1(ii) and suggesting to incorporate vanishing derivatives at zero. We thank the second referee for mentioning the scale-equivariant alternatives and suggesting to incorporate the constant c in Theorem 3.1, which inspired us to take a closer look at the problem of estimating $f(0)$.

6 Appendix

PROOF OF LEMMA 2.1:
Let $x_0 = \text{argmax}_{x \in K} f(x)$. If $x_0 = \infty$, there is nothing left to prove, therefore assume that $x_0 < \infty$.

(i) By definition of x_0 and the fact that g is non-increasing, for $x \geq x_0$, we must have $f(x) + g(x) \leq f(x_0) + g(x_0)$. Hence, we must have

$$\underset{x \in K}{\text{argmax}} \{f(x) + g(x)\} \leq x_0 = \underset{x \in K}{\text{argmax}}\, f(x).$$

This proves (i).

(ii) If $(C + x_0, \infty) \cap K = \emptyset$, the statement is trivially true, so only consider the case $(C + x_0, \infty) \cap K \neq \emptyset$. Then by definition $f(x) \leq f(x_0)$, for all $x \in (C + x_0, \infty) \cap K$, and by the property of g, we also have $g(x) \leq g(x_0)$, for $x \in (C + x_0, \infty) \cap K$. This implies $f(x) + g(x) \leq f(x_0) + g(x_0)$, for all $x \in (C + x_0, \infty) \cap K$. Hence, we must have

$$\underset{x \in K}{\text{argmax}} \{f(x) + g(x)\} \leq C + x_0 = C + \underset{x \in K}{\text{argmax}}\, f(x).$$

This proves the lemma. ■

PROOF OF LEMMA 3.1:
Decompose the process Z_{n1} as follows,

$$\begin{aligned} Z_{n1}(x, t) = n^{\alpha/2} W_n\left(F(tn^{-\alpha})\right) + n^{(1+\alpha)/2}\left\{F(tn^{-\alpha}) - f(0)tn^{-\alpha}\right\} - xt \\ - n^{\alpha/2} F(tn^{-\alpha}) W_n(1) + n^{\alpha/2} H_n(tn^{-\alpha}), \end{aligned} \tag{6.1}$$

where $H_n(t) = E_n(t) - B_n(F(t))$. By Brownian scaling, the process $n^{\alpha/2} W_n\left(F(tn^{-\alpha})\right)$ has the same distribution as the process $W\left(n^{\alpha} F(tn^{-\alpha})\right)$, and by uniform continuity of Brownian motion on compacta,

$$W\left(n^{\alpha} F(tn^{-\alpha})\right) - W\left(f(0)t\right) \to 0,$$

uniformly for t in compact sets. Since $\alpha > 1/(2k + 1)$ we have that

$$n^{(1+\alpha)/2}\left\{F(tn^{-\alpha}) - f(0)tn^{-\alpha}\right\} = n^{(1+\alpha)/2} \frac{f^{(k)}(\theta_t)}{(k+1)!}(tn^{-\alpha})^{k+1} \to 0,$$

uniformly for t in compact sets. Because $n^{\alpha/2} F(tn^{-\alpha}) W_n(1) = \mathcal{O}_p(n^{-\alpha/2})$, together with (2.2) this proves (i).

In case (ii), first note that if $f^{(k)}(0) \neq 0$, then we must have $f^{(k)}(0) < 0$. Since otherwise $f^{(k-1)}$ is increasing in a neighborhood of zero, which implies that $f^{(k-2)}$ is increasing in a neighborhood of zero, and so on, which eventually would imply that f is increasing in a neighborhood of zero. Since $\alpha = 1/(2k + 1)$, the only difference with the proof of (i) is the behavior of the deterministic term

$$n^{(k+1)/(2k+1)}\left\{F(tn^{-1/(2k+1)}) - f(0)tn^{-1/(2k+1)}\right\} \to -\frac{|f^{(k)}(0)|}{(k+1)!}t^{k+1},$$

uniformly for t in compact sets. Similar to the proof of (i), using Brownian scaling and uniform continuity of Brownian motion on compacta this proves (ii).

For case (iii) the process Z_{n2} can be written as

$$\begin{aligned}&n^{b/2}\Big\{W_n\big(F(cn^{-\alpha}+tn^{-b})\big)-W_n\big(F(cn^{-\alpha})\big)\Big\}\\&+n^{(b+1)/2}\Big\{F(cn^{-\alpha}+tn^{-b})-F(cn^{-\alpha})-f(cn^{-\alpha})tn^{-b}\Big\}-xt\\&-n^{b/2}\Big\{F(cn^{-\alpha}+tn^{-b})-F(cn^{-\alpha})\Big\}W_n(1)+n^{b/2}H_n(cn^{-\alpha}+tn^{-b})-n^{b/2}H_n(cn^{-\alpha}).\end{aligned}$$

The process $n^{b/2}\{W_n\big(F(cn^{-\alpha}+tn^{-b})\big)-W_n\big(cF(n^{-\alpha})\big)\}$ has the same distribution as the process $W\big(n^b(F(cn^{-\alpha}+tn^{-b})-F(cn^{-\alpha}))\big)$, and by uniform continuity of Brownian motion on compacta,

$$W\big(n^b(F(cn^{-\alpha}+tn^{-b})-F(cn^{-\alpha}))\big)-W\big(f(0)t\big)\to 0,$$

uniformly for t in compact sets. Finally, for some $\theta_1\in[cn^{-\alpha},cn^{-\alpha}+tn^{-b}]$ and for some $\theta_2\in[0,cn^{-\alpha}+tn^{-b}]$, it holds that

$$\begin{aligned}&n^{(b+1)/2}\Big\{F(cn^{-\alpha}+tn^{-b})-F(cn^{-\alpha})-f(cn^{-\alpha})tn^{-b}\Big\}\\&=n^{(1-3b)/2}\frac{f'(\theta_1)}{2}t^2=n^{(1-3b)/2}\frac{f^{(k)}(\theta_2)}{2(k-1)!}\theta_1^{k-1}t^2\to-\frac{|f^{(k)}(0)|}{2(k-1)!}c^{k-1}t^2,\end{aligned}$$

uniformly for t in compact sets. Since $n^{b/2}\Big\{F(n^{-\alpha}+tn^{-b})-F(n^{-\alpha})\Big\}W_n(1)=\mathcal{O}_p(n^{-b/2})$, together with (2.2) this proves (iii). ■

To verify condition (iii) of Theorem 2.1 we need that $F(c+t)-F(c)-f(c)t$ is suitably bounded. The next lemma guarantees that this is the case.

Lemma 6.1 *Suppose that f satisfies (C2). Then there exists a value $t_0>0$, such that $\inf|f^{(k)}|=\inf_{0\le s\le t_0}|f^{(k)}(s)|>0$. For any $0\le c\le t_0/2$ we can bound $F(c+t)-F(c)-f(c)t$ by*

(i) $-\dfrac{\inf|f^{(k)}|}{(k+1)!}t^{k+1}$, *for* $0\le t\le t_0/2$,

(ii) $-\dfrac{\inf|f^{(k)}|}{(k+1)!}(t_0/2)^k\,t$, *for* $t>t_0/2$,

(iii) $-\dfrac{\inf|f^{(k)}|}{2(k-1)!}(c/2)^{k-1}\,t^2$, *for* $-c/2<t<t_0/2$.

Furthermore, for small enough c and for $-c<t<-c/2$

(iv) $F(c+t)-F(c)-f(c)t\le -C_1c^{k+1}$, *where $C_1>0$ does not depend on c and t.*

Proof: The existence of $t_0>0$ follows directly from condition (C2). According to the proof of Lemma 3.1, under condition (C2) we must have $f^{(k)}(0)<0$. This in turn implies that $f^{(i)}(s)<0$ for $0\le s\le t_0$ and $i=1,2,\ldots,k$. Hence, for $0\le t\le t_0/2$, the inequality for $F(c+t)-F(c)-f(c)t$ is a direct consequence of a Taylor expansion, where all negative terms except for the last one are omitted.

For $t > t_0/2$, write

$$\begin{aligned} F(c+t) - F(c) - f(c)t &= F(c+t_0/2) - F(c) - f(c)t_0/2 \\ &\quad + (f(c+t_0/2) - f(c))(t - t_0/2) \\ &\quad + F(c+t) - F(c+t_0/2) - f(c+t_0/2)(t-t_0/2), \end{aligned}$$

where $F(c+t) - F(c+t_0/2) - f(c+t_0/2)(t-t_0/2) \le 0$, because f is non-increasing. By the same argument as above, $F(c+t_0/2) - F(c) - f(c)t_0/2 \le f^{(k)}(\theta_1)\,(t_0/2)^{k+1}/(k+1)!$ and $(f(c+t_0/2) - f(c)) \le f^{(k)}(\theta_2)\,(t_0/2)^k/k!$, for some $c < \theta_1, \theta_2 < c + t_0/2$. This implies that for $t > t_0/2$, we can bound $F(c+t) - F(c) - f(c)t$ from above by

$$\begin{aligned} & -\frac{(t_0/2)^{k+1}}{(k+1)!} \inf |f^{(k)}| - \frac{(t_0/2)^k}{k!} \inf |f^{(k)}|\,(t - t_0/2) \\ & \le -\frac{(t_0/2)^k}{(k+1)!} \inf |f^{(k)}|\,(t_0/2 + t - t_0/2) = -\frac{(t_0/2)^k}{(k+1)!} \inf |f^{(k)}| t. \end{aligned}$$

For $-c/2 < t < t_0/2$, first write $F(c+t) - F(c) - f(c)t = f'(\theta_4)t^2/2$, for $c/2 < \theta_4 < c + t_0/2$. By condition (C2), $f'(\theta_4) = f^{(k)}(\theta_5)\theta_4^{k-1}/(k-1)!$, for some $0 < \theta_5 < \theta_4$. This means that

$$F(c+t) - F(c) - f(c)t = \frac{\theta_4^{k-1}}{2(k-1)!} f^{(k)}(\theta_5)t^2 \le -\frac{(c/2)^{k-1}}{2(k-1)!} \inf |f^{(k)}| t^2.$$

Finally, for $-c < t < -c/2$, first note that $f(c+t) - f(c) \ge 0$, so that $F(c+t) - F(c) - f(c)t$ is non-decreasing in t. Write

$$F(c+t) - F(c) - f(c)t = \frac{f^{(k)}(\theta_6)}{(k+1)!}(c+t)^{k+1} - \frac{f^{(k)}(\theta_7)}{(k+1)!}c^{k+1} - \frac{f^{(k)}(\theta_8)}{k!}c^k t,$$

for $0 < \theta_6 < c + t$ and $0 < \theta_7, \theta_8 < c$. Because this expression is non-decreasing for $-c < t < -c/2$, and since $f^{(k)}(\theta_i) - f^{(k)}(0) = o(1)$, for $i = 6, 7, 8$, uniformly in $-c < t < -c/2$, we conclude that

$$F(c+t) - F(c) - f(c)t \le \frac{f^{(k)}(0)}{(k+1)!} c^{k+1} \left(\frac{1}{2^{k+1}} - 1 + \frac{k+1}{2} \right) (1 + o(1)),$$

as $c \downarrow 0$. Since $f^{(k)}(0) < 0$, this proves the lemma. ■

Proof of Lemma 3.2:
(i) Decompose Z_{n1} as in (6.1). Let $0 < \epsilon < x$ and define $X_{n1}(t) = n^{\alpha/2} H_n(tn^{-\alpha}) - \epsilon t/2$, where $H_n(t) = E_n(t) - B_n(F(t))$. Next, consider the event

$$A_{n1} = \{X_{n1}(s) \ge X_{n1}(t), \text{ for all } s, t \in [0, \infty), \text{ such that } t - s \ge \delta_n\}. \tag{6.2}$$

Then with $\delta_n = n^{-(1-\alpha)/2}(\log n)^2$, by using (2.2) we have that

$$P(A_{n1}) \ge P\left\{ \sup_{t \in [0,\infty)} |H_n(t)| \le \frac{\epsilon}{4} n^{-1/2} (\log n)^2 \right\} \to 1.$$

Also define the process $X_{n2}(t) = -n^{\alpha/2} F(tn^{-\alpha}) W_n(1) - \epsilon t/2$, and consider the event

$$A_{n2} = \{X_{n2}(s) \ge X_{n2}(t), \text{ for all } 0 \le s \le t < \infty\}. \tag{6.3}$$

Then, since every sample path of the process X_{n2} is differentiable, we have

$$P(A_{n2}) \geq P\left\{-f(tn^{-\alpha})W_n(1) - \frac{\epsilon}{2}n^{\alpha/2} \leq 0, \text{ for all } t \in [0,\infty)\right\} \to 1.$$

Hence, if $A_n = A_{n1} \cap A_{n2}$, then $P(A_n) \to 1$. Since for any $\eta > 0$,

$$P\left\{\operatorname*{argmax}_{t\in[0,\infty)} Z_{n1}(t)1_{A_n^c} > \eta\right\} \leq P(A_n^c) \to 0,$$

we conclude that $(\operatorname{argmax}_t Z_{n1}(t))\, 1_{A_n^c} = \mathcal{O}_p(1)$. This means that we only have to consider $(\operatorname{argmax}_t Z_{n1}(t))\, 1_{A_n}$. From Lemma 2.1, we have

$$\left(\operatorname*{argmax}_{t\in[0,\infty)} Z_{n1}(t)\right) 1_{A_n} \leq \operatorname*{argmax}_{t\in[0,\infty)} S_{n1}(t) + \delta_n, \tag{6.4}$$

where

$$S_{n1}(t) = n^{\alpha/2}W_n\left(F(tn^{-\alpha})\right) - (x-\epsilon)t + n^{(1+\alpha)/2}\left(F(tn^{-\alpha}) - f(0)tn^{-\alpha}\right).$$

Since $F(tn^{-\alpha}) - f(0)tn^{-\alpha}$ is non-increasing for $t \geq 0$, according to Lemma 2.1,

$$\begin{aligned}\operatorname*{argmax}_{t\in[0,\infty)} S_{n1}(t) &\leq \operatorname*{argmax}_{t\in[0,\infty)}\left\{n^{\alpha/2}W_n\big(F(tn^{-\alpha})\big) - (x-\epsilon)t\right\} \\ &\leq \sup\left\{t \geq 0 : n^{\alpha/2}W_n\big(F(tn^{-\alpha})\big) - (x-\epsilon)t \geq 0\right\}.\end{aligned} \tag{6.5}$$

By change of variables $u = G(t) = n^{\alpha}F(tn^{-\alpha})$, and using that for $u \in [0, n^{\alpha}]$,

$$\frac{u}{f(0)} \leq G^{-1}(u) \leq \frac{u}{f\big(F^{-1}(un^{-\alpha})\big)}, \tag{6.6}$$

we find that the right hand side of (6.5) is bounded by

$$G^{-1}\left(\sup\left\{u \geq 0 : n^{\alpha/2}W_n(un^{-\alpha}) - \frac{x-\epsilon}{f(0)}u \geq 0\right\}\right).$$

By Brownian scaling (2.3),

$$\sup\left\{u \geq 0 : n^{\alpha/2}W_n(un^{-\alpha}) - \frac{x-\epsilon}{f(0)}u \geq 0\right\} \stackrel{d}{=} \frac{f(0)^2}{(x-\epsilon)^2}\sup\left\{u \geq 0 : W(u) - u \geq 0\right\},$$

which is of order $\mathcal{O}_p(1)$. The latter can be seen for instance from (3.3). Because $\delta_n = n^{-(1-\alpha)/2}(\log n)^2 = o(1)$, together with (6.4), (6.5) and (6.6), it follows that

$$0 \leq \operatorname*{argmax}_{t\in[0,\infty)} Z_{n1}(t) \leq \left(\operatorname*{argmax}_{t\in[0,\infty)} Z_{n1}(t)\right) 1_{A_n} + \mathcal{O}_p(1) \leq \frac{\mathcal{O}_p(1)}{f\big(F^{-1}(\mathcal{O}_p(n^{-\alpha}))\big)} + \mathcal{O}_p(1),$$

which proves (i).

(ii) In this case $\alpha = 1/(2k+1)$, so that the argument up to (6.4) is the same. Let $\epsilon > 0$ and $A_n = A_{n1} \cap A_{n2}$, where A_{n1} as defined in (6.2) with $\delta_n = n^{-k/(k+1)}(\log n)^2$ and A_{n2} as defined in (6.3). We now find that

$$\left(\operatorname*{argmax}_{t\in[0,\infty)} Z_{n1}(t)\right) 1_{A_n} \leq \operatorname*{argmax}_{t\in[0,\infty)} S_{n1}(t) + \delta_n \leq \sup\{t \geq 0 : S_{n1}(t) \geq 0\} + \delta_n. \tag{6.7}$$

Let t_0 be the value from Lemma 6.1 and consider the event

$$D_{n1} = \left\{n^{-\alpha} \sup\{t \geq 0 : S_{n1}(t) \geq 0\} \leq t_0/2\right\}.$$

If $S_{n1}(t) \geq 0$, then according to Lemma 6.1(ii), for $tn^{-\alpha} > t_0/2$ and n sufficiently large, we find that

$$\begin{aligned}
0 &\leq n^{\alpha/2} W_n\big(F(tn^{-\alpha})\big) - (x-\epsilon)t + n^{(1+\alpha)/2}\left(F(tn^{-\alpha}) - f(0)tn^{-\alpha}\right) \\
&\leq n^{\alpha/2} \sup_{0\leq u\leq 1} |W_n(u)| - (x-\epsilon)t - n^{(1-\alpha)/2}\frac{(t_0/2)^k}{(k+1)!} \inf |f^{(k)}| t \\
&\leq n^{\alpha/2} \sup_{0\leq u\leq 1} |W_n(u)| - n^{(1-\alpha)/2} C_1 t \left(1 + \frac{x-\epsilon}{n^{(1-\alpha)/2} C_1}\right) \\
&\leq n^{\alpha/2} \left\{\sup_{0\leq u\leq 1} |W_n(u)| - C_1 n^{1/2} t_0/4\right\},
\end{aligned}$$

where $C_1 = \inf |f^{(k)}|(t_0/2)^k/(k+1)!$. Therefore

$$P\left(D_{n1}^c\right) \leq P\left(\sup_{0\leq u\leq 1} |W(u)| \geq C_1 n^{1/2} t_0/4\right) \to 0.$$

This means we can restrict ourselves to the event $A_n \cap D_{n1}$, so that by analogous reasoning as before, from (6.7) we get

$$\begin{aligned}
\left(\operatorname*{argmax}_{t\in[0,\infty)} Z_{n1}(t)\right) 1_{A_n\cap D_{n1}} &\leq \sup\Big\{t \geq 0 : S_{n1}(t) \geq 0\Big\} 1_{D_{n1}} + \delta_n \\
&\leq \sup\Big\{0 \leq t \leq n^{\alpha} t_0/2 : S_{n1}(t) \geq 0\Big\} + \delta_n.
\end{aligned}$$

According to Lemma 6.1(i), for $0 \leq tn^{-\alpha} \leq t_0/2$ and using that $\alpha = 1/(2k+1)$, we get

$$n^{(1+\alpha)/2}\left(F(tn^{-\alpha}) - f(0)tn^{-\alpha}\right) \leq -\frac{\inf|f^{(k)}|}{(k+1)!} t^{k+1},$$

so that

$$\begin{aligned}
0 &\leq \left(\operatorname*{argmax}_{t\in[0,\infty)} Z_{n1}(t)\right) 1_{A_n\cap D_{n1}} \\
&\leq \sup\left\{0 \leq t \leq n^{\alpha} t_0/2 : n^{\alpha/2} W_n\big(F(tn^{-\alpha})\big) - (x-\epsilon)t - \frac{\inf|f^{(k)}|}{(k+1)!} t^{k+1} \geq 0\right\} + \delta_n.
\end{aligned} \tag{6.8}$$

Next, distinguish between

(A) $-(x-\epsilon)t - \inf|f^{(k)}|t^{k+1}/(2(k+1)!) \geq 0$,

(B) $-(x-\epsilon)t - \inf|f^{(k)}|t^{k+1}/(2(k+1)!) < 0$.

Since $t \geq 0$, case (A) can only occur when $x - \epsilon < 0$, in which case we have $0 \leq t \leq \big(2(k+1)!(\epsilon - x)/\inf|f^{(k)}|\big)^{1/k}$, which is of order $\mathcal{O}(1)$. In case (B), it follows that

$$n^{\alpha/2}W_n\big(F(n^{-\alpha}t)\big) - \frac{\inf|f^{(k)}|}{2(k+1)!}t^{k+1} \geq 0.$$

We conclude from (6.8) that

$$\begin{aligned} 0 &\leq \left(\operatorname*{argmax}_{t\in[0,\infty)} Z_{n1}(t)\right) 1_{A_n\cap D_{n1}} \\ &\leq \sup\left\{0 \leq t \leq n^{\alpha}t_0/2 : n^{\alpha/2}W_n\big(F(tn^{-\alpha})\big) - \frac{\inf|f^{(k)}|}{2(k+1)!}t^{k+1} \geq 0\right\} + \mathcal{O}_p(1) + \delta_n \\ &\leq \sup\left\{t \in [0,\infty) : n^{\alpha/2}W_n\big(F(tn^{-\alpha})\big) - \frac{\inf|f^{(k)}|}{2(k+1)!}t^{k+1} \geq 0\right\} + \mathcal{O}_p(1). \end{aligned} \tag{6.9}$$

Similar to the proof of (i), by change of variables $u = G(t) = n^{\alpha}F(tn^{-\alpha})$ and using (6.6) with $\alpha = 1/(2k+1)$, we find that the argmax on the right hand side of (6.9) is bounded from above by

$$G^{-1}\left(\sup\left\{u \in [0,\infty) : n^{\alpha/2}W_n(un^{-\alpha}) - \frac{\inf|f^{(k)}|}{2(k+1)!f(0)^{k+1}}u^{k+1} \geq 0\right\}\right) + \mathcal{O}_p(1).$$

By Brownian scaling (2.3), we obtain that the supremum in the first term has the same distribution as

$$\left(\frac{2(k+1)!f(0)^{k+1}}{\inf|f^{(k)}|}\right)^{2/(2k+1)} \sup\left\{u \geq 0 : W(u) - u^{k+1} \geq 0\right\}.$$

Again by using (3.3), this is of order $\mathcal{O}_p(1)$. Similar to the proof of (i), from (6.6) and (6.9) we find that

$$0 \leq \operatorname*{argmax}_{t\in[0,\infty)} Z_{n1}(t) \leq \left(\operatorname*{argmax}_{t\in[0,\infty)} Z_{n1}(t)\right) 1_{A_n\cap D_{n1}} + \mathcal{O}_p(1) \leq \frac{\mathcal{O}_p(1)}{f\left(F^{-1}(\mathcal{O}_p(n^{-\alpha}))\right)} + \mathcal{O}_p(1),$$

which proves (ii).

(iii) Decompose Z_{n2} as in the proof of Lemma 3.1. Let $\epsilon > 0$ and $A_n = A_{n1} \cap A_{n2}$, with A_{n1} defined similar to (6.2) with $\delta_n = n^{-(1-b)/2}(\log n)^2$, where b is the same as in Lemma 3.1, and A_{n2} is defined similar to (6.3). By the same argument as in the proof of (i) and (ii), it suffices to consider $(\operatorname{argmax}_t Z_{n2}(t))\, 1_{A_n}$. We find

$$\left(\operatorname*{argmax}_{t\in[-cn^{b-\alpha},\infty)} Z_{n2}(t)\right) 1_{A_n} \leq \operatorname*{argmax}_{t\in[-cn^{b-\alpha},\infty)} M_{n2}(t) + \delta_n \leq \sup\{t \geq 0 : M_{n2}(t) \geq 0\} + \delta_n,$$

where $M_{n2}(t)$ has the same distribution as

$$\begin{aligned} S_{n2}(t) &= n^{b/2}W\Big(F(cn^{-\alpha}+tn^{-b})-F(cn^{-\alpha})\Big) \\ &\quad +n^{(b+1)/2}\Big(F(cn^{-\alpha}+tn^{-b})-F(cn^{-\alpha})-f(cn^{-\alpha})tn^{-b}\Big)-(x-\epsilon)t. \end{aligned}$$

As in the proof of (ii), consider $D_{n2}=\big\{n^{-b}\sup\{t\geq 0: S_{n2}(t)\geq 0\}\leq t_0/2\big\}$, where t_0 is the value from Lemma 6.1. By the same reasoning as used in the proof of (ii), it again follows from Lemma 6.1(ii) that $P\left(D_{n2}^c\right)\to 0$, so we only have to consider $\sup\{t\geq 0: S_{n2}(t)\geq 0\}\,1_{D_{n2}}$. Hence, similar to the proof of (ii) we get

$$\sup\{t\geq 0: S_{n2}(t)\geq 0\}\,1_{D_{n2}}\leq \sup\Big\{0\leq t\leq n^bt_0/2: S_{n2}(t)\geq 0\Big\}.$$

Since, $b>1/(2k+1)$, for $k\geq 2$, we cannot proceed as in the proof of (ii) by using Lemma 6.1(i) to bound the drift term. However, according to Lemma 6.1(iii), for $0\leq t\leq n^bt_0/2$,

$$n^{(b+1)/2}\Big(F(cn^{-\alpha}+tn^{-b})-F(cn^{-\alpha})-f(cn^{-\alpha})tn^{-b}\Big)\leq -\frac{\inf|f^{(k)}|}{2^k(k-1)!}t^2,$$

so that $\sup\{0\leq t\leq n^bt_0/2: S_{n2}(t)\geq 0\}$ is bounded from above by

$$\sup\Big\{0\leq t\leq n^bt_0/2: n^{b/2}W\big(F(cn^{-\alpha}+tn^{-b})-F(cn^{-\alpha})\big)-(x-\epsilon)t-\frac{\inf|f^{(k)}|}{2^k(k-1)!}t^2\geq 0\Big\}.$$

Similar to (6.9), we conclude that $\sup\{t\geq 0: S_{n2}(t)\geq 0\}\,1_{D_{n2}}$ is bounded from above by

$$\sup\Big\{t\geq 0: n^{b/2}W_n\Big(F(cn^{-\alpha}+tn^{-b})-F(cn^{-\alpha})\Big)-\frac{\inf|f^{(k)}|}{2^{k+1}(k-1)!}t^2\geq 0\Big\}+\mathcal{O}_p(1). \qquad (6.10)$$

Next, change variables $u=G(t)=n^b\big(F(cn^{-\alpha}+tn^{-b})-F(cn^{-\alpha})\big)$. Then for any $u\in[0,n^b(1-F(cn^{-\alpha}))]$, it follows that

$$\frac{u}{f(0)}\leq G^{-1}(u)\leq \frac{u}{f\left(F^{-1}(un^{-b}+F(cn^{-\alpha}))\right)}, \qquad (6.11)$$

so that (6.10) is bounded from above by

$$G^{-1}\left(\sup\Big\{u\geq 0: n^{b/2}W(un^{-b})-\frac{\inf|f^{(k)}|}{2^{k+1}(k-1)!f(0)^2}u^2\geq 0\Big\}\right)+\mathcal{O}_p(1).$$

As in the proof of (ii), by Brownian scaling (2.3) together with (6.11), we find that

$$\begin{aligned} \mathop{\mathrm{argmax}}_{t\in[-cn^{b-\alpha},\infty)} Z_{n2}(t) &\leq \left(\mathop{\mathrm{argmax}}_{t\in[-cn^{b-\alpha},\infty)} Z_{n2}(t)\right)1_{A_n\cap D_{n2}}+\mathcal{O}_p(1) \\ &\leq \frac{\mathcal{O}_p(1)}{f\left(F^{-1}(\mathcal{O}_p(n^{-b})+F(cn^{-\alpha}))\right)}+\mathcal{O}_p(1)=\mathcal{O}_p(1). \end{aligned} \qquad (6.12)$$

To obtain a lower bound for the left hand side of (6.12), first note that

$$\mathop{\mathrm{argmax}}_{t\in[-cn^{b-\alpha},\infty)} Z_{n2}(t)\geq \mathop{\mathrm{argmax}}_{t\in[-cn^{b-\alpha},0]} Z_{n2}(t)=-\mathop{\mathrm{argmax}}_{t\in[0,cn^{b-\alpha}]} Z_{n2}(-t). \qquad (6.13)$$

¿From here, the argument runs along the same lines as for the upper bound. Let $\epsilon > 0$ and, similar to (6.2) and (6.3), define the events A_{n1} and A_{n2} with

$$\begin{aligned} X_{n1}(t) &= n^{b/2} H_n(cn^{-\alpha} - tn^{-b}) - \epsilon t/2, \\ X_{n2}(t) &= -n^{b/2} F(cn^{-\alpha} - tn^{-b}) - \epsilon t/2. \end{aligned}$$

With $A_n = A_{n1} \cap A_{n2}$, as before we get $(\text{argmax}_t\, Z_{n2}(-t))\, 1_{A_n}^c = \mathcal{O}_p(1)$ and

$$\left(\underset{t}{\text{argmax}}\, Z_{n2}(-t)\right) 1_{A_n} \leq \underset{t\in[0,cn^{b-\alpha})}{\text{argmax}}\, M_{n3}(t) + \delta_n,$$

where $M_{n3}(t)$ has the same distribution as

$$\begin{aligned} S_{n3}(t) &= n^{b/2} W\left(F(cn^{-\alpha} - tn^{-b}) - F(cn^{-\alpha})\right) \\ &\quad + n^{(b+1)/2}\left(F(cn^{-\alpha} - tn^{-b}) - F(cn^{-\alpha}) + f(cn^{-\alpha})tn^{-b}\right) + (x+\epsilon)t \\ &\leq n^{b/2} \sup\{|W(u)| : 0 \leq u \leq f(0)cn^{-\alpha}\} \\ &\quad + n^{(b+1)/2}\left(F(cn^{-\alpha} - tn^{-b}) - F(cn^{-\alpha}) + f(cn^{-\alpha})tn^{-b}\right) + (x+\epsilon)t. \end{aligned}$$

Consider $D_{n3} = \{n^{-b} \sup\{0 \leq t \leq cn^{b-\alpha} : \; S_{n3}(t) \geq 0\} \leq cn^{-\alpha}/2\}$, and note that by Brownian scaling $\sup\{|W(u)| : 0 \leq u \leq f(0)cn^{-\alpha}\}$ has the same distribution as $n^{-\alpha/2} \sup\{|W(u)| : 0 \leq u \leq cf(0)\}$. Reasoning as in the proof of (ii), using Lemma 6.1(iv), we obtain that for $cn^{-\alpha}/2 \leq n^{-b}t \leq cn^{-\alpha}$ and n sufficiently large

$$\begin{aligned} 0 &\leq n^{(b-\alpha)/2} \sup_{0\leq u\leq cf(0)} |W(u)| + n^{(b+1)/2}\left(F(cn^{-\alpha} - tn^{-b}) - F(cn^{-\alpha}) + f(cn^{-\alpha})tn^{-b}\right) \\ &\qquad + (x+\epsilon)t \\ &\leq n^{(b-\alpha)/2}\left(\sup_{0\leq u\leq cf(0)} |W(u)| - C_1 n^{(1-(2k+1)\alpha)/2}\left(1 + \frac{x+\epsilon}{C_1 n^{(b+1)/2-(k+1)\alpha}}\right)\right) \\ &\leq n^{(b-\alpha)/2}\left(\sup_{0\leq u\leq cf(0)} |W(u)| - C_1 n^{(1-(2k+1)\alpha)/2}/2\right). \end{aligned}$$

Therefore, $P(D_{n3}^c) \to 0$, so we only have to consider $(\text{argmax}_t\, S_{n3}(t))1_{D_{n3}}$. Hence, similar to the proof of (ii), we get

$$\underset{t\in[0,cn^{b-\alpha})}{\text{argmax}}\, S_{n3}(t)1_{D_{n3}} + \delta_n \leq \sup\left\{0 \leq t \leq cn^{b-\alpha}/2 : \; S_{n3}(t) \geq 0\right\} + \delta_n.$$

According to Lemma 6.1(iii), for $0 \leq tn^{-b} \leq cn^{-\alpha}/2$ we have

$$n^{(b+1)/2}\left(F(cn^{-\alpha} - tn^{-b}) - F(cn^{-\alpha}) + f(cn^{-\alpha})tn^{-b}\right) \leq -\frac{\inf|f^{(k)}|}{2^k(k-1)!}t^2. \tag{6.14}$$

Similar to (ii), separate cases and obtain that $\text{argmax}_{t\in[0,cn^{b-\alpha})}\, S_{n3}(t)1_{D_{n3}} + \delta_n$ is bounded from above by

$$\sup\left\{0 \leq t \leq cn^{b-\alpha}/2 : n^{b/2} W\left(F(cn^{-\alpha} - tn^{-b}) - F(cn^{-\alpha})\right) - \frac{\inf|f^{(k)}|}{2^{k+1}(k-1)!}t^2 \geq 0\right\} + \mathcal{O}_p(1).$$

After change of variables $u = G(t) = n^b\big(F(cn^{-\alpha} - tn^{-b}) - F(cn^{-\alpha})\big)$, and using that for $u \in [-n^b F(cn^{-\alpha}), 0]$, one has

$$-\frac{u}{f(0)} \leq G^{-1}(u) \leq -\frac{u}{f(cn^{-\alpha})},$$

we now find that

$$\underset{t\in[0,cn^{b-\alpha})}{\text{argmax}}\; S_{n3}(t) + \delta_n \leq \frac{1}{f(cn^{-\alpha})} \sup\left\{u \leq 0 : W_n(u) - \frac{\inf|f^{(k)}|}{2^{k+1}(k-1)!f(0)^2}u^2 \geq 0\right\} + \mathcal{O}_p(1).$$

As above, by Brownian scaling (2.3) together with (6.13), it follows that

$$\underset{t\in[-cn^{b-\alpha},\infty)}{\text{argmax}}\; Z_{n2}(t) \geq \frac{\mathcal{O}_p(1)}{f(cn^{-\alpha})} + \mathcal{O}_p(1) = \mathcal{O}_p(1).$$

Together with (6.12) this proves the lemma. ■

Proof of Theorem 4.1: To prove case (i), similar to the proof of Theorem 3.1(i), it suffices to consider $n^{(1-\alpha)/2}(f(1) - \hat{f}_n(1 - n^{-\alpha}))$. For $x > 0$, according to (2.1),

$$P\left\{n^{(1-\alpha)/2}\left(f(1) - \hat{f}_n(1 - cn^{-\alpha})\right) \leq x\right\} = P\left\{n^{\alpha}(1 - U_n(f(1) - xn^{-(1-\alpha)/2})) \leq c\right\},$$

where according to Lemma 4.2(i),

$$n^{\alpha}\left(1 - U_n(f(1) - xn^{-(1-\alpha)/2})\right) = \underset{t\in[0,n^{\alpha}]}{\text{argmax}}\, Y_{n1}(x,t) = \mathcal{O}_p(1).$$

From here on, the proof proceeds in completely the same manner as that of Theorem 3.1(i). We conclude that for $x > 0$,

$$\begin{aligned} P\left\{n^{(1-\alpha)/2}\left(f(1) - \hat{f}_n(1 - cn^{-\alpha})\right) \leq x\right\} &= P\left\{\underset{0\leq t\leq n^{\alpha}}{\text{argmax}}\, Y_{n1}(t) \leq c\right\} \\ &\to P\left\{\underset{t\geq 0}{\text{argmax}}\, \{W(f(1)t) - xt\} \leq c\right\} \\ &= P\left\{\underset{t\geq 0}{\text{argmax}}\left\{W(t) - \frac{xc^{1/2}t}{f(1)^{1/2}}\right\} \leq 1\right\}. \end{aligned}$$

Similar to switching relation (2.1), the right hand side equals $P\{(f(1)/c)^{1/2} D_R[W(t)](1) \leq x\}$. Furthermore, similar to the proof of Theorem 3.1(i) it follows that

$$P\left\{n^{(1-\alpha)/2}\left(f(1) - \hat{f}_n(1 - n^{-\alpha})\right) \leq 0\right\} \to 0.$$

This proves (i).

For (ii), first note that if $f^{(k)}(1) \neq 0$, then we must have $(-1)^{k+1}f^{(k)}(1) < 0$. Since otherwise $(-1)^k f^{(k-1)}(1-t)$ is increasing in a neighborhood of zero, which implies that $(-1)^{k-1}f^{(k-2)}(1-t)$ is increasing in a neighborhood of zero, and so on, which eventually would imply that f is increasing in a neighborhood of one. Hence, from (C4) it follows that

$$\begin{aligned} &n^{k/(2k+1)}\left(f(1 - c\tilde{B}_{2k}n^{-1/(2k+1)}) - \hat{f}_n(1 - c\tilde{B}_{2k}n^{-1/(2k+1)})\right) - |f^{(k)}(1)|(c\tilde{B}_{2k})^k/k! \\ &= n^{k/(2k+1)}\left(f(1) - \hat{f}_n(1 - c\tilde{B}_{2k}n^{-1/(2k+1)})\right) + o(1), \end{aligned}$$

and that according to (2.1), $P\left\{n^{k/(2k+1)}\left(f(1)-\hat{f}_n(1-c\tilde{B}_{2k}n^{-1/(2k+1)})\right)\le x\right\}$ is equal to

$$P\left\{\tilde{B}_{2k}^{-1}n^{1/(2k+1)}\left(1-U_n\left(f(1)-xn^{-k/(2k+1)}\right)\right)\le c\right\}.$$

With Y_{n1} being the process defined in Lemma 4.1 with $\alpha=1/(2k+1)$, we get

$$\tilde{B}_{2k}^{-1}n^{1/(2k+1)}\left(1-U_n\left(f(1)-xn^{-k/(2k+1)}\right)\right)=\underset{t\in[0,n^\alpha)}{\operatorname{argmax}}\left\{Y_{n1}(x,\tilde{B}_{2k}t)\right\}=\mathcal{O}_p(1).$$

The rest of the proof is completely similar to that of Theorem 3.1(ii). For (iii), note that according to (2.1), we have

$$\begin{aligned}&P\left\{n^{(1-b)/2}\left(f(1-cn^{-\alpha})-\hat{f}_n(1-cn^{-\alpha})\right)\le x\right\}\\&\quad=P\left\{n^b\left(1-U_n(f(1-cn^{-\alpha})-xn^{-(1-b)/2})-cn^{-\alpha}\right)\le 0\right\},\end{aligned}\tag{6.15}$$

and with Y_{n2} as defined in Lemma 4.2(iii), we get

$$n^b\left(U_n(f(1-cn^{-\alpha})-xn^{-(1-b)/2})-cn^{-\alpha}\right)=\underset{t\in[-cn^{b-\alpha},n^b-cn^{b-\alpha})}{\operatorname{argmax}}Y_{n2}(x,t)=\mathcal{O}_p(1).$$

From here the proof is completely similar to that of Theorem 3.1(iii). ■

Proof of Lemma 4.1:
Similar to the proof of Lemma 3.1, the process $Y_{n1}(x,t)$ can be written as

$$\begin{aligned}&n^{\alpha/2}\left\{W_n(F(1-tn^{-\alpha}))-W_n(1)\right\}\\&\quad+n^{(1+\alpha)/2}\left\{F(1-tn^{-\alpha})-1+f(1)tn^{-\alpha}\right\}-xt\\&\quad\quad-n^{\alpha/2}\left\{F(1-tn^{-\alpha})-1\right\}W_n(1)+n^{\alpha/2}H_n(1-tn^{-\alpha}).\end{aligned}$$

First note that the process $n^{\alpha/2}\left\{W_n(F(1-tn^{-\alpha}))-W_n(1)\right\}$ has the same distribution as the process $W(n^\alpha(1-F(1-tn^{-\alpha})))$, which can be approximated by the process $W(f(1)t)$, using uniform continuity of Brownian motion on compacta. Since $\alpha>1/(2k+1)$ we have that

$$n^{(1+\alpha)/2}\left\{F(1-tn^{-\alpha})-F(1)+f(1)tn^{-\alpha}\right\}=n^{(1+\alpha)/2}(-1)^{k+1}\frac{f^{(k)}(\theta_t)}{(k+1)!}(tn^{-\alpha})^{k+1}\to 0,$$

uniformly for t in compact sets. Because $n^{\alpha/2}(F(1-tn^{-\alpha})-F(1))W_n(1)=\mathcal{O}_p(n^{-\alpha/2})$, together with (2.2) this proves (i). In case (ii), where $\alpha=1/(2k+1)$, the only difference is the behavior of the deterministic term

$$n^{(k+1)/(2k+1)}\left\{F(1-tn^{-1/(2k+1)})-F(1)+f(1)tn^{-1/(2k+1)}\right\}\to-\frac{|f^{(k)}(1)|}{(k+1)!}t^{k+1},$$

uniformly for t in compact sets, where we use that $(-1)^{k+1}f^{(k)}(1)<0$ (see the proof of Theorem 4.1). Similar to the proof of (i), using Brownian scaling and uniform continuity of

Brownian motion on compacta this proves (ii). For case (iii) the process Y_{n2} can be written as

$$
\begin{aligned}
& n^{b/2}\left\{W_n\big(F(1-cn^{-\alpha}-tn^{-b})\big)-W_n\big(F(1-cn^{-\alpha})\big)\right\} \\
& +n^{(b+1)/2}\left\{F(1-cn^{-\alpha}-tn^{-b})-F(1-cn^{-\alpha})+f(1-cn^{-\alpha})tn^{-b}\right\}-xt \\
& -n^{b/2}\left\{F(1-cn^{-\alpha}-tn^{-b})-F(1-cn^{-\alpha})\right\}W_n(1) \\
& +n^{b/2}H_n(1-cn^{-\alpha}-tn^{-b})-n^{b/2}H_n(1-cn^{-\alpha}).
\end{aligned}
$$

The process $n^{b/2}\left\{W_n\big(F(1-cn^{-\alpha}-tn^{-b})\big)-W_n\big(1-cF(n^{-\alpha})\big)\right\}$ has the same distribution as the process $W\big(n^b(F(1-cn^{-\alpha})-F(1-cn^{-\alpha}-tn^{-b}))\big)$, which can be approximated by the process $W(f(1)t)$, again by using uniform continuity of Brownian motion on compacta. Finally, for some $\theta_1 \in [1-cn^{-\alpha}-tn^{-b}, 1-cn^{-\alpha}]$ and for some $\theta_2 \in [1-cn^{-\alpha}-tn^{-b}, 1]$ it holds that

$$
\begin{aligned}
& n^{(b+1)/2}\left\{F(1-cn^{-\alpha}-tn^{-b})-F(1-cn^{-\alpha})+f(1-cn^{-\alpha})tn^{-b}\right\} \\
& = n^{(1-3b)/2}\frac{f'(\theta_1)}{2}t^2 = n^{(1-3b)/2}\frac{f^{(k)}(\theta_2)}{2(k-1)!}(\theta_1-1)^{k-1}t^2 \to -\frac{|f^{(k)}(1)|}{2(k-1)!}c^{k-1}t^2,
\end{aligned}
$$

uniformly for t in compact sets. Since $n^{b/2}\left\{F(1-n^{-\alpha}-tn^{-b})-F(1-n^{-\alpha})\right\}W_n(1) = \mathcal{O}_p(n^{-b/2})$, together with (2.2) this proves (iii). ∎

To verify condition (iii) of Theorem 2.1 we need that $F(1-c-t)-F(1-c)+f(1-c)t$ is suitably bounded. The next lemma guarantees that this is the case.

Lemma 6.2 *Suppose that f satisfies (C4). Then there exists a value $t_1 > 0$, such that $\inf|f^{(k)}| = \inf_{1-t_1\le s\le 1}|f^{(k)}(s)| > 0$. For any $0 \le c \le t_1/2$ we can bound $F(1-c-t)-F(1-c)+f(1-c)t$ by*

(i) $-\dfrac{\inf|f^{(k)}|}{(k+1)!}t^{k+1}$, *for* $0 \le t \le t_1/2$

(ii) $-\dfrac{\inf|f^{(k)}|}{(k+1)!}(t_1/2)^k\, t$, *for* $t_1/2 < t < 1-c$

(iii) $-\dfrac{\inf|f^{(k)}|}{2(k-1)!}(c/2)^{k-1}\, t^2$, *for* $-c/2 < t < t_1/2$.

Furthermore, for small enough c and for $-c < t < -c/2$

(iv) $F(1-c-t)-F(1-c)-f(1-c)t \le -C_1c^{k+1}$, *where $C_1 > 0$ does not depend on c and t.*

Proof: The existence of $t_1 > 0$ follows directly from condition (C4). According to the proof of Theorem 4.1, under condition (C4) we must have $(-1)^{k+1}f^{(k)}(1) < 0$, which in turn implies that $(-1)^{i+1}f^{(i)}(1-s) < 0$ for $0 \le s \le t_1$ and $i = 1, 2, \ldots, k$. Hence, for $0 \le t \le t_1/2$, the inequality for $F(1-c-t)-F(1-c)+f(1-c)t$ is a direct consequence of a Taylor expansion, where all negative terms except for the last one are omitted.

For $t > t_1/2$, write

$$\begin{aligned}F(1-c-t) - F(1-c) + f(1-c)t = {} & F(1-c-t_1/2) - F(1-c) + f(1-c)t_1/2 \\ & + \left(f(1-c) - f(1-c-t_1/2)\right)(t - t_1/2) \\ & + F(1-c-t) - F(1-c-t_1/2) + f(1-c-t_1/2)(t-t_1/2),\end{aligned}$$

where $F(1-c-t)-F(1-c-t_1/2)+f(1-c-t_1/2)(t-t_1/2) \leq 0$, because f is non-increasing. By the same argument as above,

$$F(1-c-t_1/2) - F(1-c) + f(1-c)t_1/2 \leq (-1)^{k+1} f^{(k)}(\theta_1) \left(t_1/2\right)^{k+1} / (k+1)!$$

and

$$\left(f(1-c) - f(1-c-t_1/2)\right) \leq (-1)^{k+1} f^{(k)}(\theta_2) \left(t_1/2\right)^{k} / k!,$$

for some $c < \theta_1, \theta_2 < c + t_1/2$. This implies that for $t > t_1/2$, we can bound

$$\begin{aligned}& F(1-c-t) - F(1-c) + f(1-c)t \\ & \quad \leq -\frac{(t_1/2)^{k+1}}{(k+1)!} \inf|f^{(k)}| - \frac{(t_1/2)^{k}}{k!} \inf|f^{(k)}|\,(t - t_1/2) \\ & \quad \leq -\frac{(t_1/2)^{k}}{(k+1)!} \inf|f^{(k)}|\,(t_1/2 + t - t_1/2) = -\frac{(t_1/2)^{k}}{(k+1)!} \inf|f^{(k)}|t.\end{aligned}$$

For $-c/2 < t < t_1/2$, first write $F(1-c-t) - F(1-c) + f(1-c)t = f'(\theta_4)t^2/2$, for $1 - c/2 - t_1/2 < \theta_4 < 1 - c$. By condition (C4), $f'(\theta_4) = (-1)^{k-1} f^{(k)}(\theta_5)(1-\theta_4)^{k-1}/(k-1)!$, for some $\theta_4 < \theta_5 < 1$. This means that

$$F(1-c-t) - F(1-c) + f(1-c)t = (-1)^{k-1} \frac{(1-\theta_4)^{k-1}}{2(k-1)!} f^{(k)}(\theta_5) t^2 \leq -\frac{(c/2)^{k-1}}{2(k-1)!} \inf|f^{(k)}|t^2.$$

Finally, for $-c < t < -c/2$, first note that $f(1-c) - f(1-c-t) \geq 0$, so that $F(1-c-t) - F(1-c) + f(1-c)t$ is non-decreasing in t. Write

$$\begin{aligned}& F(1-c-t) - F(1-c) + f(1-c)t \\ & \quad = (-1)^{k+1} \frac{f^{(k)}(\theta_6)}{(k+1)!}(c+t)^{k+1} - (-1)^{k+1} \frac{f^{(k)}(\theta_7)}{(k+1)!} c^{k+1} - (-1)^{k} \frac{f^{(k)}(\theta_8)}{k!} c^{k} t,\end{aligned}$$

for $1 - c - t < \theta_6 < 1$ and $1 - c < \theta_7, \theta_8 < 1$. Because this expression is non-decreasing for $-c < t < -c/2$, and since $f^{(k)}(\theta_i) - f^{(k)}(1) = o(1)$, for $i = 6, 7, 8$, uniformly in $-c < t < -c/2$, we conclude that

$$F(1-c-t) - F(1-c) + f(1-c)t \leq (-1)^{k+1} \frac{f^{(k)}(1)}{(k+1)!} c^{k+1} \left(\frac{1}{2^{k+1}} - 1 + \frac{k+1}{2} \right) (1 + o(1))$$

as $c \downarrow 0$. Since $(-1)^{k+1} f^{(k)}(1) < 0$, this proves the lemma. ∎

Proof of Lemma 4.2: The proof mimics the proof of Lemma 3.2. Let $0 < \epsilon < x$. Define processes

$$\begin{aligned}X_{n1}(t) &= n^{\alpha/2} H_n(1 - tn^{-\alpha}) - \frac{1}{2}\epsilon t, \\ X_{n2}(t) &= -n^{\alpha/2} \left\{ F(1 - tn^{-\alpha}) - 1 \right\} W_n(1) - \frac{1}{2}\epsilon t,\end{aligned}$$

and define the event A_n as in the proof of Lemma 3.2(i). It follows that $(\text{argmax}_t\, Y_{n1}(t))1_{A_n^c} = O_p(1)$, so that we only have to deal with $(\text{argmax}_t\, Y_{n1}(t))1_{A_n}$. From Lemma 2.1, we have

$$\left(\underset{t\in[0,n^\alpha)}{\text{argmax}}\, Y_{n1}(t)\right) 1_{A_n} \leq \underset{t\in[0,n^\alpha)}{\text{argmax}}\, M_{n1}(t) + \delta_n, \tag{6.16}$$

where $M_{n1}(t)$ is the process

$$n^{\alpha/2}\left\{W_n(F(1-tn^{-\alpha})) - W_n(1)\right\} - (x-\epsilon)t + n^{(1+\alpha)/2}\left(F(1-tn^{-\alpha}) - F(1) + f(1)tn^{-\alpha}\right).$$

By Brownian scaling M_{n1} has the same distribution as

$$S_{n1}(t) = n^{\alpha/2}W\left(F(1-tn^{-\alpha}) - 1\right) - (x-\epsilon)t + n^{(1+\alpha)/2}\left(F(1-tn^{-\alpha}) - 1 + f(1)tn^{-\alpha}\right).$$

Proceeding as in the proof of Lemma 3.2(i), using that the function $F(1-tn^{-\alpha})-1+f(1)tn^{-\alpha}$ is non-increasing, we find that

$$0 \leq \left(\underset{t\in[0,n^\alpha]}{\text{argmax}}\, S_{n1}(t)\right) 1_{A_n} \quad\leq\quad \underset{t\in[0,n^\alpha]}{\text{argmax}}\left\{W\left(n^\alpha(F(1-tn^{-\alpha})-1)\right) - (x-\epsilon)t\right\} + \delta_n,$$

where $\delta_n = n^{-(1-\alpha)/2}(\log n)^2$. Finally, by change of variables $u = H(t) = n^\alpha(1-F(1-tn^{-\alpha}))$, and the fact that for any $u \in [0, n^\alpha]$,

$$\frac{u}{f(0)} \leq H^{-1}(u) \leq \frac{u}{f(1)}, \tag{6.17}$$

we find that

$$\begin{aligned}
&\underset{t\in[0,n^\alpha]}{\text{argmax}}\left\{W\left(n^\alpha(F(1-tn^{-\alpha})-1)\right) - (x-\epsilon)t\right\} \\
&\quad\leq \sup\left\{t\in[0,n^\alpha] : W\left(n^\alpha(F(1-tn^{-\alpha})-1)\right) - (x-\epsilon)t \geq 0\right\} \\
&\quad\leq \frac{1}{f(1)}\sup\left\{u\in[0,\infty) : W(-u) - \frac{x-\epsilon}{f(0)}u \geq 0\right\} = \mathcal{O}_p(1),
\end{aligned}$$

which proves (i).

(ii) As in the proof of Lemma 3.2(ii), similar to (6.7) we obtain

$$\left(\underset{t\in[0,n^\alpha)}{\text{argmax}}\, Y_{n1}(t)\right) 1_{A_n} \leq \underset{t\in[0,n^\alpha)}{\text{argmax}}\, M_{n1}(t) + \delta_n \leq \sup\left\{0\leq t\leq n^\alpha : M_{n1}(t)\geq 0\right\} + \delta_n,$$

where M_{n1} has the same distribution as S_{n1} with $\alpha = 1/(2k+1)$. As in the proof of Lemma 3.2(ii), restrict to $D_{n1} = \{n^{-\alpha}\sup\{t\geq 0 : S_{n1}(t)\geq 0\} \leq t_1/2\}$, with t_1 being the value from Lemma 6.2. Then $P(D_{n1}^c) \to 0$. Then by application of Lemma 6.2(ii), similar to (6.8) we find that $(\text{argmax}_t\, S_{n1}(t))\, 1_{A_n\cap D_{n1}}$ is bounded from above by

$$\sup\left\{0\leq t\leq n^\alpha t_1/2 : n^{\alpha/2}W\big(F(1-tn^{-\alpha})-1\big) - (x-\epsilon)t - \frac{\inf|f^{(k)}|}{(k+1)!}t^{k+1} \geq 0\right\} + \delta_n.$$

Proceeding as in the proof of Lemma 3.2(ii), similar to (6.9) this supremum is bounded by

$$\sup\left\{t\in[0,n^\alpha): n^{\alpha/2}W\big(F(1-tn^{-\alpha})-1\big)-\frac{\inf|f^{(k)}|}{2(k+1)!}t^{k+1}\geq 0\right\}+\mathcal{O}_p(1).$$

By change of variables $u=H(t)=n^\alpha(1-F(1-tn^{-\alpha}))$, and using (6.17), we find that this argmax is bounded by

$$\frac{1}{f(1)}\sup\left\{u\in[0,\infty): n^{\alpha/2}W(-un^{-\alpha})-\frac{\inf|f^{(k)}|}{2(k+1)!f(0)^{k+1}}u^{k+1}\geq 0\right\}+\mathcal{O}_p(1).$$

By Brownian scaling (2.3), we obtain that the supremum in the first term has the same distribution as

$$\left(\frac{2(k+1)!f(0)^{k+1}}{\inf|f^{(k)}|}\right)^{2/(2k+1)}\sup\left\{u\geq 0: W(u)-u^{k+1}\geq 0\right\}.$$

Again by using (3.3), this is of order $\mathcal{O}_p(1)$, which proves (ii).

For case (iii), decompose Y_{n2} as in the proof of Lemma 4.1. Let $\epsilon>0$ and let A_n be same event as in the proof of (i) and (ii) with $\delta_n=n^{-1/3}(\log n)^2$. Write $I_n=[-n^{1/3-\alpha},n^{1/3}(1-n^{-\alpha})]$, then by the same argument as in the proof of (i) and (ii), we find that

$$\left(\operatorname*{argmax}_{t\in I_n}Y_{n2}(t)\right)1_{A_n}\leq\operatorname*{argmax}_{t\in I_n}M_{n2}(t)+\delta_n\leq\sup\{0\leq t\leq n^{-b}-cn^{b-\alpha}: M_{n2}(t)\geq 0\}+\delta_n,$$

where $M_{n2}(t)$ has the same distribution as

$$\begin{aligned}S_{n2}(t) &= n^{b/2}W\left(F(1-cn^{-\alpha}-tn^{-b})-F(1-cn^{-\alpha})\right)\\ &\quad+n^{(b+1)/2}\left(F(1-cn^{-\alpha}-tn^{-b})-F(1-cn^{-\alpha})+f(1-cn^{-\alpha})tn^{-b}\right)-(x-\epsilon)t.\end{aligned}$$

According to Lemma 6.2,

$$n^{(b+1)/2}\left(F(1-cn^{-\alpha}-tn^{-b})-F(1-cn^{-\alpha})+f(1-cn^{-\alpha})tn^{-b}\right)\leq-\frac{\inf|f^{(k)}|}{2^k(k-1)!}t^2,$$

for $0\leq t\leq n^bt_1/2$, so that $\sup\{0\leq t\leq n^bt_1/2: S_{n2}(t)\geq 0\}$ is bounded from above by

$$\sup\left\{0\leq t\leq n^bt_1/2: n^{b/2}W\big(F(1-cn^{-\alpha})-F(1-cn^{-\alpha}-tn^{-b})\big)-(x-\epsilon)t-\frac{\inf|f^{(k)}|}{2^k(k-1)!}t^2\geq 0\right\}.$$

Then by change of variables $u=H(t)=n^b\big(F(1-cn^{-\alpha})-F(1-cn^{-\alpha}-tn^{-b})\big)$, and using that for any $u\in[0,n^bF(1-cn^{-\alpha})]$,

$$\frac{u}{f(0)}\leq H^{-1}(u)\leq\frac{u}{f(1)},$$

it follows that this argmax is bounded by

$$\frac{1}{f(1)}\sup\left\{u\geq 0: n^{b/2}W(un^{-b})-\frac{\inf|f^{(k)}|}{2^{k+1}(k-1)!f(0)^2}u^2\geq 0\right\}+\mathcal{O}_p(1).$$

The lower bound for $\operatorname{argmax}_t Y_{n2}(t)$ is obtained by the same type of argument as for the lower bound in the proof of Lemma 3.2(iii). This proves the lemma. ■

References

COX, D.R. (1969). Some sampling problems in technology, In *New development in survey sampling* (Edited by Johnson, N.L. and Smith, H.), p.506–527, Wiley, New York.

FELLER, W. (1971). *An Introduction to Probability Theory and its Applications*, 2nd edition, John Wiley.

GRENANDER, U. (1956). On the theory of mortality measurements, Part II, *Skand. Akt. Tid.* **39**, p.125–153.

GROENEBOOM, P. (1985). Estimating a monotone density, In *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer* (Edited by L.LeCam and R.Olshen) **2**, p.539–555.

GROENEBOOM, P., HOOGHIEMSTRA, G. AND LOPUHAÄ, H.P. (1999). Asymptotic normality of the L_1-error of the Grenander estimator, *Ann. Statist.* **27**, No.4, p.1316–1347.

HAMPEL, F. (1987). Design, modelling and analysis of some biological data sets, In *Design, data and analysis, by some friends of Cuthbert Daniel* (C.L.Mallows, editor) p.111–115, Wiley, New York.

KEIDING, N., KAJSA, K., HELLE, H. AND MADS, T. (2002). Estimating time to pregnancy from current duration in a cross-sectional sample, *Biostatistics* **3**, No.4, p.565–578.

KIM, J. AND POLLARD, D. (1990). Cube root asymptotics, *Ann. Statist.* **18**, No.1, p.191–219.

KOMLOS, J., MAJOR, P. AND TUSNADY, G. (1975). An approximation of partial sums of independent rv'-s, and the sample DF. I, *Z. Wahrsch. Verw. Gebiete* **32**, p.111–131.

KULIKOV, V.N. AND LOPUHAÄ, H.P. (2004A). Asymptotic normality of the L_k-error of the Grender estimator. To appear in *Annals of Statistics*.

LYNDEN-BELL, D. (1992). Eddington Malmquist bias, streaming motions and the distribution of galaxies, In *Statistical Challenges in Modern Astronomy* (edited by Eric D. Feigelson, Gutti J. Babu), p.201–221, Springer-Verlag, New York.

POLLARD, D. (1984). *Convergence of Stochastic Processes*, Springer-Verlag.

PRAKASA RAO, B.L.S. (1969). Estimation of a unimodal density, *Sankhya Ser. A* **31**, p.23–36.

SUN, J. AND WOODROOFE, M. (1996). Adaptive smoothing for a penalized NPMLE of a non-increasing density, *J. Statist. Planning and Inference* **52**, p.143–159.

VARDI, Y. (1989). Multiple censoring, renewal process, deconvolution and decreasing density: Nonparametric estimation, *Biometrika* **76**, No.4, p.751–761.

WATSON, G.S. (1971). Estimating functionals of particle size distribution, *Biometrika* **58**, No.3, p.483–490.

WOODROOFE, M. AND SUN, J. (1993). A penalized maximum likelihood estimate of $f(0+)$ when f is non-increasing, *Statist. Sinica* **3**, p.501–515.