Asymptotic normality of the L_1 error of the Grenander estimator

Piet Groeneboom, Gerard Hooghiemstra and Hendrik P. Lopuhaä[1]

Faculty ITS
Department of Mathematics
Delft University of Technology
Mekelweg 4
2628 CD Delft

July 4, 2000

[1]Corresponding author

Abstract

In Groeneboom (1985, 1989) a jump process was introduced that can be used (among other things) to study the asymptotic properties of the Grenander estimator of a monotone density. In this paper we derive the asymptotic normality of a suitably rescaled version of the L_1 error of the Grenander estimator, using properties of this jump process.

AMS 1991 *Subject classifications.* Primary 62E20, 62G05; Secondary 60J65, 60J75

Key words and phrases. Brownian motion with quadratic drift, central limit theorem, concave majorant, isotonic estimation, jump process, L_1-norm, monotone density.

1 Introduction

Let f be a decreasing density with support $[0,1]$. Denote by F_n the empirical distribution function of a sample $X_1,\ldots,X_n$ from f. Let $\hat{F}_n$ be the *concave majorant* of F_n on [0,1], by which we mean the smallest concave function such that

$$\hat{F}_n(t) \geq F_n(t),\ t \in [0,1],\ \text{and}\ \hat{F}_n(0)=0,\ \hat{F}_n(1)=1.$$

The Grenander estimator $\hat{f}_n$ is defined as the left derivative of $\hat{F}_n$.

In Groeneboom (1985) the asymptotic behavior of $\hat{f}_n$ was investigated. Instead of studying the process $\{\hat{f}_n(t), t \in (0,1)\}$ itself, the better tractable *(inverse) process* $\{U_n(a) : a \in [f(1), f(0)]\}$ was studied, where $U_n(a)$ is defined as the last time that the process $F_n(t) - at$ attains its maximum:

$$U_n(a) = \sup\{t \in [0,1] : F_n(t) - at \text{ is maximal}\}. \tag{1.1}$$

A new proof, based on the inverse process U_n, was given of a result in Prakasa Rao (1969) on pointwise weak convergence of $\hat{f}_n$. In Groeneboom (1985) also analytical properties of the weak limit of the locally rescaled process $U_n(a)$ were discussed and it was indicated how the *process* U_n together with a Hungarian embedding technique could be used to prove asymptotic normality of the L_1 error

$$\|\hat{f}_n - f\|_1 = \int_0^1 |\hat{f}_n(t) - f(t)|\, dt. \tag{1.2}$$

The analytical properties of the limit process $a \mapsto V(a)$ were made rigorous in Groeneboom (1989) and at the same time it was mentioned that a rigorous treatment of the asymptotic normality of the L_1 error would appear elsewhere. This paper fulfills that promise.

We feel that this result is important, since the problem of estimating a monotone density is closely related to several other (inverse) problems, e.g., estimation of the distribution function of interval censored data (see, e.g. Groeneboom and Wellner (1992)), and estimation of a monotone hazard function, and since the result was referred to by several authors, see, for instance, Devroye and Györfi (1985), pp. 213 and 214, Devroye (1987), p. 145, Csörgö and Horvath (1988), Birgé (1989), and Wang (1992). Recently, the result has been taken up again in the context of nonparametric regression, see Durot (1996). In fact, the methods used by Durot (1996), whose work was done independently, are closer in spirit to the methods, suggested in Groeneboom (1985), than our present paper, which relies on ideas, developed in Groeneboom (1989). In both settings, the proof relies heavily on the fact that Brownian motion has independent increments. One of the main differences between the model, considered in Durot (1996), and the present paper is that in the regression setting one can make a direct embedding into Brownian motion, whereas in our case we can only make such a embedding into the Brownian bridge and we need rather delicate arguments to make the transition to Brownian motion (Corollary 3.3 in the present paper).

The main result can be stated as follows. Define

$$V(c) = \sup\{t : W(t) - (t-c)^2 \text{ is maximal}\}, \tag{1.3}$$

where $\{W(t) : -\infty < t < \infty\}$ denotes standard two-sided Brownian motion on $I\!R$ originating from zero (i.e. $W(0) = 0$).

Theorem 1.1 *(*MAIN THEOREM*) Let f be a twice differentiable decreasing density on [0,1], satisfying*

(A1) $0 < f(1) \le f(t) \le f(s) \le f(0) < \infty$, *for* $0 \le s \le t \le 1$.

(A2) $0 < \inf_{t\in(0,1)} |f'(t)| \le \sup_{t\in(0,1)} |f'(t)| < \infty$.

(A3) $\sup_{t\in(0,1)} |f''(t)| < \infty$.

Then with $\mu = 2E|V(0)| \int_0^1 |\frac{1}{2}f'(t)f(t)|^{1/3}\,dt$,

$$n^{1/6}\left\{n^{1/3}\int_0^1 |\hat f_n(t) - f(t)|\,dt - \mu\right\}$$

converges in distribution to a normal random variable with mean zero and variance $\sigma^2 = 8\int_0^\infty cov(|V(0)|, |V(c) - c|)\,dc$.

Actually, this is precisely the theorem, as stated in GROENEBOOM (1985) (with the same conditions). In that paper, however, a sketch of proof of two pages was given, whereas, unfortunately, we need a lot more pages to write down all the details (an experience shared with Cécile Durot in her work on the regression problem). The difficulty in proving a result of this type stems from the fact that the Grenander estimator is a non-linear functional of the empirical distribution function. For this reason methods of proof are needed that are very different from those used in, e.g., CSÖRGÖ AND HORVATH (1988), where the linearity of the kernel estimators is used in an essential way.

In Section 2 we show

$$\|\hat f_n - f\|_1 = \int_{f(1)}^{f(0)} |U_n(a) - g(a)|\,da + o_p(n^{-1/2}), \tag{1.4}$$

where g denotes the inverse of f (see Corollary 2.1). In this section we also obtain an exponential upper bound for the tail probabilities of $V_n^E(a) = n^{1/3}(U_n(a) - g(a))$.

In Section 3 the process $a \mapsto V_n^E(a)$ is approximated (using Hungarian embedding) by a process $a \mapsto V_n^B(a)$, defined for the Brownian bridge. The process V_n^B is in turn approximated by a similar process $a \mapsto V_n^W(a)$, defined for Brownian motion. A key tool for the results in this section is Lemma 3.4, showing that the probability of a jump of V_n^B and V_n^W in an interval of length $hn^{-1/3}$ is of order h, if h is not too small. We suspect that the restriction "not too small" is actually not needed, but this restriction arises naturally in the present approach. The methods in this section are motivated by results that hold in the "canonical setting" of the process V, studied in GROENEBOOM (1989).

Another key observation that makes things work in Section 3 is that, although we cannot construct a Brownian motion and a Brownian bridge which are close in the supremum distance on $[0,1]$, we have that, if

$$W(F(t)) = B(F(t)) + \xi F(t),$$

where B is the Brownian bridge on $[0,1]$, and ξ is a standard normal random variable, independent of B, the associated processes of locations of maxima V_n^B and V_n^W, defined for $B \circ F$ and $W \circ F$, respectively, are very close indeed.

The results in Section 3 imply that it is sufficient to prove that

$$n^{1/6} \int_{f(1)}^{f(0)} \left(|V_n^W(a)| - E|V_n^W(a)| \right) da$$

tends in distribution to a normal distribution with expectation 0 and variance σ^2, where σ^2 is given in Theorem 1.1. In Section 3 the process V_n^W is also shown to be strongly mixing. This leads to a central limit theorem which is proved in Section 4 by using Bernstein's method of big blocks and small blocks. Throughout, it will be assumed that conditions (A1) to (A3) hold.

2 Localization.

In this section we show that the distribution of the random variables

$$V_n^E(a) = n^{1/3}(U_n(a) - g(a)) \tag{2.1}$$

have exponentially fast decreasing tails. This will enable us to compare the process U_n locally with a similar process, defined for the Brownian bridge. For $s \le t$, we use the following abbreviations:

$$\begin{aligned} F_n(s,t) &= F_n(t) - F_n(s), \\ F(s,t) &= F(t) - F(s). \end{aligned}$$

Lemma 2.1 *Let $a \in [f(1), f(0)]$ and let $t_0 = g(a)$. Then*

$$P\{V_n^E(a) > x\} \le P\left\{ \sup_{t \in [t_0 + xn^{-1/3}, 1]} \frac{F_n(t_0,t)}{F(t_0,t)} \ge \frac{f(t_0)xn^{-1/3}}{F(t_0, t_0 + xn^{-1/3})} \right\},$$

for each x such that $t_0 < t_0 + xn^{-1/3} \le 1$, and

$$P\{V_n^E(a) < -x\} \le P\left\{ \inf_{t \in [0, t_0 - xn^{-1/3}]} \frac{F_n(t,t_0)}{F(t,t_0)} \le \frac{f(t_0)xn^{-1/3}}{F(t_0 - xn^{-1/3}, t_0)} \right\},$$

for each x such that $0 \le t_0 - xn^{-1/3} < t_0$.

Proof: For each x, such that $t_0 < t_0 + xn^{-1/3} \le 1$, we have

$$P\{V_n^E(a) > x\} \le P\{F_n(t_0,t) - a(t-t_0) \ge 0, \text{ for some } t \in (t_0 + xn^{-1/3}, 1]\}, \tag{2.2}$$

and for each x such that $0 \le t_0 - xn^{-1/3} < t_0$:

$$P\{V_n^E(a) < -x\} \le P\{F_n(t,t_0) - a(t_0-t) \le 0, \text{ for some } t \in [0, t_0 - xn^{-1/3})\}. \tag{2.3}$$

The probability on the right-hand side of (2.2) can be written as

$$P\left\{\frac{F_n(t_0,t)}{F(t_0,t)} \ge \frac{f(t_0)(t-t_0)}{F(t_0,t)}, \text{ for some } t \in (t_0 + xn^{-1/3}, 1]\right\}. \tag{2.4}$$

Since the function

$$\gamma(t) = \frac{f(t_0)(t-t_0)}{F(t_0,t)},$$

is increasing for $t \in (t_0, 1)$ (using the monotonicity of f), it follows that (2.4) is bounded above by

$$P\left\{\sup_{t \in (t_0 + xn^{-1/3}, 1]} \frac{F_n(t_0,t)}{F(t_0,t)} \ge \frac{f(t_0)xn^{-1/3}}{F(t_0, t_0 + xn^{-1/3})}\right\}.$$

Similarly, the probability on the right-hand side of (2.3) can be bounded from above by

$$P\left\{\inf_{t \in [0, t_0 - xn^{-1/3})} \frac{F_n(t,t_0)}{F(t,t_0)} \le \frac{f(t_0)xn^{-1/3}}{F(t_0 - xn^{-1/3}, t_0)}\right\}.$$

□

To bound the probabilities given in Lemma 2.1 we will apply Doob's inequality to suitably chosen martingales. These martingales are given in the next lemma.

Lemma 2.2 *Let $0 \le t_0 \le 1$. Consider, for n fixed, the processes*

$$t \mapsto M_{1n}(t) = \frac{F_n(t_0,t)}{F(t_0,t)}, \quad t \in (t_0, 1]$$

and

$$t \mapsto M_{2n}(t) = \frac{F_n(t,t_0)}{F(t,t_0)}, \quad t \in [0, t_0).$$

Let $\mathcal{F}_s = \sigma\{F_n(t) : t \in [s,1]\}$ and $\mathcal{G}_s = \sigma\{F_n(t) : t \in [0,s]\}$. Then, conditionally on $F_n(t_0)$, the process M_{1n} is a reverse time martingale with respect to the filtration $\{\mathcal{F}_s : s \in (t_0, 1]\}$ and M_{2n} is a forward time martingale with respect to the filtration $\{\mathcal{G}_s : s \in [0, t_0)\}$.

Proof: Note that conditionally on $F_n(t_0)$ and $F_n(t_0, s)$, for $t_0 < t < s < 1$, the random variable $nF_n(t_0,t)$ has a binomial distribution with parameter $nF_n(t_0,s)$ and probability of success $p = F(t_0,t)/F(t_0,s)$. This shows that for $t < s$:

$$E_0\left[F_n(t_0,t) \mid \mathcal{F}_s\right] = F_n(t_0,s)\frac{F(t_0,t)}{F(t_0,s)},$$

where $E_0(\cdot) = E[\cdot \mid F_n(t_0)]$. This implies that for $t_0 < t < s < 1$, we have that

$$E_0\left[M_{1n}(t) \mid \mathcal{F}_s\right] = M_{1n}(s).$$

Similarly, conditionally on $F_n(t_0)$ and $F_n(s,t_0)$, for $0 < s < t < t_0$, the random variable $nF_n(t,t_0)$ has a binomial distribution with parameters $nF_n(s,t_0)$ and $p = F(t,t_0)/F(s,t_0)$. This leads to

$$E_0\left[M_{2n}(t) \mid \mathcal{G}_s\right] = M_{2n}(s).$$

□

We have the following bounds for the martingales in Lemma 2.2.

Lemma 2.3 *Let $h(y) = 1 - y + y\log y$, $y > 0$. Then, for $t_0 \in [0,1)$, $y \geq 1$ and $\delta > 0$ such that $t_0 + \delta < 1$:*

$$P\left\{\sup_{t\in[t_0+\delta,1]} M_{1n}(t) \geq y\right\} \leq \exp\left\{-nF(t_0,t_0+\delta)h(y)\right\}$$

and for $t_0 \in (0,1]$, $0 < y \leq 1$ and $\delta > 0$ such that $t_0 - \delta > 0$:

$$P\left\{\inf_{t\in[0,t_0-\delta]} M_{2n}(t)) \leq y\right\} \leq \exp\left\{-nF(t_0-\delta,t_0)h(y)\right\}.$$

Proof: We start with the proof of the first inequality. According to Lemma 2.2 we have that for each $r > 0$, conditionally on $F_n(t_0)$, the process $\exp\{rM_{1n}(t)\}$ is a reverse time submartingale. Hence, by Doob's inequality,

$$\begin{aligned}
P\left\{\sup_{t\in[t_0+\delta,1]} M_{1n}(t) \geq y\right\} &= E\left[P\left\{\sup_{t\in[t_0+\delta,1]} M_{1n}(t) \geq y \middle| F_n(t_0)\right\}\right] \\
&= E\left[P\left\{\sup_{t\in[t_0+\delta,1]} e^{rM_{1n}(t)} \geq e^{ry} \middle| F_n(t_0)\right\}\right] \\
&\leq E\left[e^{-ry}E\left(e^{rM_{1n}(t_0+\delta)} \middle| F_n(t_0)\right)\right] \\
&= e^{-ry}Ee^{rM_{1n}(t_0+\delta)}.
\end{aligned}$$

Using that $nF_n(t_0,t_0+\delta)$ has a binomial distribution with parameters n and $p = F(t_0,t_0+\delta)$, we see that the last expression is equal to:

$$e^{-ry}\left(1 + p(e^{r/np} - 1)\right)^n \leq e^{-ry}\exp\left(np(e^{r/np} - 1)\right) = e^{-nph(y)},$$

by putting $r = np\log y$ in the last equality. This proves the first exponential bound.

For the proof of the second inequality we note that, for $y \in (0,1]$:

$$\begin{aligned} P\left\{\inf_{t\in[0,t_0-\delta]} M_{2n}(t) \le y\right\} &= E\left[P\left\{\sup_{t\in[0,t_0-\delta]} -M_{2n}(t) \ge -y \middle| F_n(t_0)\right\}\right] \\ &\le E\left[e^{ry} E\left(e^{-rM_{2n}(t_0-\delta)}\middle| F_n(t_0)\right)\right] \\ &= e^{ry} E e^{-rM_{2n}(t_0-\delta)}, \end{aligned}$$

where again Doob's inequality is used. Taking $p = F(t_0-\delta, t_0)$ and $r = -np\log y$, we get

$$e^{ry} E e^{-rM_{2n}(t_0-\delta)} \le e^{-nph(y)}.$$

$\square$

Remark. The function $y \mapsto h(y)$, used in Lemma 2.3, but also in the sequel, is a well-known function in large deviation theory. It is non-negative and convex on $(0,\infty)$. Its minimum 0 is attained at $y=1$. Actually $h(y) = \int_1^y \log u\, du$, $y > 0$.

We are now ready to prove the following theorem.

Theorem 2.1 *Let $V_n^E(a)$ be defined by (2.1). Then there exists a constant $C > 0$, only depending on f, such that for all $n \ge 1$, $a \in [f(1), f(0)]$ and $x > 0$,*

$$P\{|V_n^E(a)| > x\} \le 2e^{-Cx^3}.$$

Proof: We will write $\delta_n = xn^{-1/3}$. First consider the probability

$$P\{V_n^E(a) > x\}. \tag{2.5}$$

If $g(a) + \delta_n \ge 1$, this probability is zero, in which case there is nothing to prove, so we can restrict ourselves to values of $x > 0$, such that $g(a) + \delta_n < 1$. Let

$$y_n = \frac{f(t_0)\delta_n}{F(t_0, t_0+\delta_n)},$$

where $t_0 = g(a)$. Note that $y_n > 1$, since f is strictly decreasing. We also have, using assumption (A1),

$$y_n = \frac{f(t_0)\delta_n}{F(t_0, t_0+\delta_n)} \le \frac{f(t_0)}{f(t_0+\delta_n)} \le \frac{f(0)}{f(1)} < \infty.$$

Hence $1 < y_n < c_1$, for a constant $c_1 > 0$, independent of x such that $t_0 + \delta_n < 1$. By Lemma 2.1, the probability in (2.5) is bounded above by

$$P\left\{\sup_{t\in[t_0+\delta_n,1]} M_{1n}(t) \ge y_n\right\}.$$

According to Lemma 2.3 this probability is bounded by

$$\exp\left\{-nF(t_0, t_0+\delta_n)h(y_n)\right\}. \tag{2.6}$$

Using a Taylor expansion with a Lagrangian remainder term of the convex function $u \mapsto h(u)$ at $u = 1$, we get

$$h(y_n) = \tfrac{1}{2}h''(\xi_n)(y_n - 1)^2 \geq \tfrac{1}{2}c_1^{-1}(y_n - 1)^2, \tag{2.7}$$

where $1 \leq \xi_n \leq c_1$. But

$$|y_n - 1| \geq \frac{\delta_n \inf_{u\in(0,1)} |f'(u)|}{2f(0)},$$

and hence, by (2.7),

$$h(y_n) \geq c_2\delta_n^2,$$

for a constant $c_2 > 0$, independent of x such that $t_0+\delta_n < 1$. Since $F(t_0, t_0+\delta_n) \geq f(1)\delta_n$, it now follows that (2.6) is bounded above by $\exp(-Cx^3)$.

Now consider the probability

$$P\{V_n^E(a) < -x\}. \tag{2.8}$$

If $g(a) - xn^{-1/3} \leq 0$, this probability is zero, so we can restrict ourselves to consider an $x > 0$ such that $g(a) - xn^{-1/3} > 0$. Define

$$y_n = \frac{f(t_0)\delta_n}{F(t_0 - \delta_n, t_0)}.$$

The fact that f is strictly decreasing this time implies that $y_n < 1$. Using Lemma 2.1 it is seen that (2.8) is bounded above by

$$P\left\{\inf_{t\in[0,t_0-\delta_n]} M_{2n}(t) \leq y_n\right\},$$

which, by Lemma 2.3, leads to the upper bound

$$\exp\left\{-nf(1)\delta_n h(y_n)\right\}.$$

We have, using $h''(x) \geq 1$, $x \in (0, 1]$:

$$h(y_n) = \tfrac{1}{2}h''(\xi_n)(y_n - 1)^2 \geq \tfrac{1}{2}(y_n - 1)^2,$$

where in this case $0 < \xi_n \leq 1$. Following the same line of argument as above, we get the upper bound $\exp\{-Cx^3\}$. □

Lemma 2.3 also enables us to show that the difference between the L_1 risk in (1.2) and the integral

$$\int_{f(1)}^{f(0)} |U_n(a) - g(a)|\, da,$$

defined in terms of the inverse process, is of order $o_p(n^{-1/2})$.

Corollary 2.1 *Let $\hat{f}_n$ be the Grenander estimator and let U_n be defined in (1.1). Then*

$$\int_0^1 |\hat{f}_n(t) - f(t)|\,dt - \int_{f(1)}^{f(0)} |U_n(a) - g(a)|\,da = \mathcal{O}_p(n^{-2/3}). \tag{2.9}$$

Proof: The difference on the left-hand side of (2.9) can be written as

$$\int_0^1 [\hat{f}_n(t) - f(0)]^+\,dt + \int_0^1 [f(1) - \hat{f}_n(t)]^+\,dt,$$

where $x^+ = \max(0, x)$, $x \in I\!R$. We will show that the first term is $\mathcal{O}_p(n^{-2/3})$. The second term can be treated similarly.

We have that

$$\begin{aligned}\int_0^1 \left[\hat{f}_n(t) - f(0)\right]^+ dt &= \int_0^{U_n(f(0))} \left(\hat{f}_n(t) - f(0)\right) dt = F_n(U_n(f(0))) - f(0)U_n(f(0)) \\ &= F_n(U_n(f(0))) - F(U_n(f(0))) + F(U_n(f(0))) - f(0)U_n(f(0)).\end{aligned}$$

According to Theorem 2.1, for the second difference on the right-hand side we have

$$|F(U_n(f(0))) - f(0)U_n(f(0))| \leq \frac{1}{2} \sup |f'|\, U_n(f(0))^2 = \mathcal{O}_p(n^{-2/3}). \tag{2.10}$$

Let $Z_n = F_n(U_n(f(0))) - F(U_n(f(0)))$ and $\delta_n = n^{-1/3}\log n$. Then write

$$Z_n = Z_n 1_{\{U_n(f(0)) > \delta_n\}} + Z_n 1_{\{U_n(f(0)) \leq \delta_n\}}.$$

Then according to Theorem 2.1

$$E|Z_n| 1_{\{U_n(f(0)) > \delta_n\}} \leq 2P\{U_n(f(0)) > \delta_n\} \leq 4e^{-C(\log n)^3}.$$

Hence by the Markov inequality we can conclude that

$$Z_n 1_{\{U_n(f(0)) > \delta_n\}} = o_p(n^{-2/3}). \tag{2.11}$$

Let (B_n) be a sequence of Brownian bridges given by the Hungarian embedding approximating $n^{1/2}(F_n - F)$, cf. Komlos, Major and Tusnády (1975). Then

$$|Z_n| 1_{\{U_n(f(0)) \leq \delta_n\}} \leq n^{-1/2} \sup_{t \in [0, F(\delta_n)]} |B_n(t)| + \mathcal{O}_p(n^{-1}\log n).$$

Since $B_n(t) \stackrel{d}{=} W(t) + tW(1)$, where W denotes Brownian motion, the right hand side can be bounded by a random variable that has the same distribution as

$$n^{-1/2} \sup_{t \in [0, F(\delta_n)]} |W(t)| + n^{-1/2} F(\delta_n) |W(1)| + \mathcal{O}_p(n^{-1}\log n).$$

Note that $F(\delta_n)|W(1)| = \mathcal{O}_p(\delta_n)$. Furthermore, since for any $\epsilon > 0$,

$$P\left\{\sup_{t \in [0, F(\delta_n)]} |W(t)| > \epsilon\right\} \leq 4P\left\{W(1) \geq \frac{\epsilon}{F(\delta_n)^{1/2}}\right\},$$

we have that

$$n^{-1/2} \sup_{t\in[0,F(\delta_n)]} |W(t)| = o_p(n^{-2/3}),$$

which implies that $Z_n 1_{\{U_n(f(0))\le\delta_n\}} = o_p(n^{-2/3})$. Together with (2.10) and (2.11) this proves that

$$\int_0^1 \left[\hat{f}_n(t) - f(0)\right]^+ dt = \mathcal{O}_p(n^{-2/3}).$$

□

3 Brownian motion approximation

In this section we show that it is sufficient to prove Theorem 1.1 for a similar process, with Brownian motion replacing the empirical process. Let E_n denote the empirical process $\sqrt{n}(F_n - F)$ and let $V_n^E(a)$ be defined as in (2.1). Then we have, for fixed $a \in (f(1), f(0))$,

$$V_n^E(a) = \operatorname*{argmax}_t \left\{D_n^E(a,t) - n^{1/3}at\right\}, \tag{3.1}$$

where $t \mapsto D_n^E(a,t)$ is the drifting empirical process

$$\begin{aligned} D_n^E(a,t) &= n^{1/6}\left\{E_n(g(a)+n^{-1/3}t) - E_n(g(a))\right\} \\ &\quad + n^{2/3}\left\{F(g(a)+n^{-1/3}t) - F(g(a))\right\}, \end{aligned}$$

and where the argmax is taken over all values of t such that $g(a) + n^{-1/3}t \in [0,1]$. Here the argmax function is the supremum of the times at which the maximum is attained (in order to have a well-defined functional also on sets of probability zero).

Let Brownian bridge B_n and the uniform empirical process $E_n \circ F^{-1}$ be constructed on the same probability space via the Hungarian embedding of Komlos, Major and Tusnády (1975). Let

$$V_n^B(a) = \operatorname*{argmax}_t \left\{D_n^B(a,t) - n^{1/3}at\right\}, \tag{3.2}$$

where

$$\begin{aligned} D_n^B(a,t) &= n^{1/6}\left\{B_n(F(g(a)+n^{-1/3}t)) - B_n(F(g(a)))\right\} \\ &\quad + n^{2/3}\left\{F(g(a)+n^{-1/3}t) - F(g(a))\right\}. \end{aligned} \tag{3.3}$$

Then (3.1) suggests that $V_n^E(a)$ is close to $V_n^B(a)$. We will show that this is indeed the case. We define versions W_n of Brownian motion by

$$W_n(t) = B_n(t) + \xi_n t, \quad t \in [0,1], \tag{3.4}$$

where ξ_n is a standard normal random variable, independent of B_n. Moreover, let

$$V_n^W(a) = \operatorname*{argmax}_t \left\{D_n^W(a,t) - n^{1/3}at\right\}, \tag{3.5}$$

where

$$D_n^W(a,t) = n^{1/6}\left\{W_n(F(g(a)+n^{-1/3}t)) - W_n(F(g(a)))\right\} + n^{2/3}\left\{F(g(a)+n^{-1/3}t) - F(g(a))\right\}. \tag{3.6}$$

Note that $V_n^B(a)$ and $V_n^W(a)$ are defined in the same way as $V_n^E(a)$, but with E_n replaced by $B_n \circ F$ and $W_n \circ F$, respectively. For $J = E, B, W$, the argmax $V_n^J(a)$ can be seen as the t-coordinate of the point that is touched first when dropping a line with slope $n^{1/3}a$ on the process $t \mapsto D_n^J(a,t)$. Furthermore, note that for every fixed $a, b \in (f(1), f(0))$, we have the following property

$$V_n^J(b) + n^{1/3}(g(b) - g(a)) = \operatorname*{argmax}_t \left\{D_n^J(a,t) - n^{1/3}bt\right\}, \tag{3.7}$$

where as before the argmax is taken over values of t such that $g(a)+n^{-1/3}t \in [0,1]$. Hence (3.7) is the t-coordinate of the point that is touched first when dropping a line with slope $n^{1/3}b$ on the process $t \mapsto D_n^J(a,t)$. Moreover, note that

$$c \mapsto V_n^J(c) + n^{1/3}(g(c) - g(a)) \text{ jumps at } b \text{ if and only if } c \mapsto V_n^J(c) \text{ jumps at } b. \tag{3.8}$$

We have the following results for $V_n^B(a)$ and $V_n^W(a)$, analogous to Theorem 2.1.

Theorem 3.1 *Let $V_n^B(a)$ and $V_n^W(a)$ be defined by (3.2) and (3.5), respectively. Then there exist a constant $C > 0$, only depending on f, such that for all $n \geq 1$, $a \in (f(1), f(0))$ and $x > 0$,*

$$P\{|V_n^W(a)| > x\} \leq 2e^{-Cx^3} \quad \text{and} \quad P\{|V_n^B(a)| > x\} \leq 4e^{-Cx^3}.$$

Proof: Let $a \in (f(1), f(0))$ and let $t_0 = g(a)$. We first consider $P\{V_n^W(a) > x\}$. If $t_0 + xn^{-1/3} \geq 1$, this probability is zero, so we may assume $t_0 + xn^{-1/3} < 1$. Let the process $t \mapsto X_n^W(a,t)$ be defined by

$$X_n^W(a,t) = n^{1/6}\{W_n(F(g(a)+n^{-1/3}t)) - W_n(F(g(a)))\}, \quad t \in [0, n^{1/3}(1-g(a))], \tag{3.9}$$

and let, for $r \in \mathbb{R}$, the process Y_n be defined by

$$Y_n(t) = \frac{e^{rX_n^W(a,t)}}{Ee^{rX_n^W(a,t)}}, \quad t \in [0, n^{1/3}(1-t_0)]. \tag{3.10}$$

Then Y_n is a martingale with respect to the filtration induced by $t \mapsto X_n^W(a,t)$, and

$$Ee^{rX_n^W(a,t)} = \exp\left\{\tfrac{1}{2}r^2 n^{1/3} F(t_0, t_0 + n^{-1/3}t)\right\}.$$

We now define the stopping time τ_n by

$$\tau_n = \inf\{t \in [x, n^{1/3}(1-t_0)] : Z_n^W(a,t) \geq 0\},$$

where $Z_n^W(a,t) = D_n^W(a,t) - n^{1/3}at$, with D_n^W defined in (3.6). If $Z_n^W(a,t) < 0$ for all $t \in [x, n^{1/3}(1-t_0)]$, we define $\tau_n = \infty$. By the optional stopping theorem (cf. Rogers and Williams (1997), p.189) we have

$$EY_n(\tau_n \wedge n^{1/3}(1-t_0)) = EY_n(0) = 1.$$

On the other hand,

$$\begin{aligned}
&EY_n(\tau_n \wedge n^{1/3}(1-t_0)) \geq EY_n(\tau_n)1_{\{\tau_n<\infty\}} \\
&\geq E\exp\Big\{-n^{2/3}rF(t_0, t_0+n^{-1/3}\tau_n) + n^{1/3}ra\tau_n - \tfrac{1}{2}r^2n^{1/3}F(t_0, t_0+n^{-1/3}\tau_n)\Big\}1_{\{\tau_n<\infty\}} \\
&\geq E\exp\Big\{c_1r\tau_n^2 - c_2r^2\tau_n\Big\}1_{\{\tau_n<\infty\}},
\end{aligned}$$

where $c_1 = \frac{1}{2}\inf_{t\in(0,1)}|f'(t)|$ and $c_2 = \frac{1}{2}f(0)$. If we take $r = c_1x/(2c_2)$ and $C = c_1^2/(4c_2)$, we conclude that

$$1 = EY_n(\tau_n \wedge n^{1/3}(1-t_0)) \geq E\exp\{Cx\tau_n(2\tau_n - x)\}1_{\{\tau_n<\infty\}} \geq \exp\{Cx^3\}P\{\tau_n < \infty\}.$$

Hence we find

$$P\{V_n^W(a) > x\} \leq P\left\{\sup_{t\in[x,n^{1/3}(1-t_0)]} Z_n^W(a,t) \geq 0\right\} = P\{\tau_n < \infty\} \leq \exp\left\{-Cx^3\right\}.$$

For the opposite inequality we note that

$$P\{V_n^W(a) < -x\} \leq P\left\{\sup_{t\in[x,n^{1/3}t_0]} Z_n^W(a,-t) \geq 0\right\}.$$

This can be bounded in the same way as before, by introducing the stopping time

$$\tilde{\tau}_n = \inf\{t \in [x, n^{1/3}t_0] : Z_n^W(a,-t) \geq 0\},$$

and applying the optional stopping argument to the backward time martingale

$$\tilde{Y}_n(t) = \frac{e^{rX_n^W(a,-t)}}{Ee^{rX_n^W(a,-t)}}, \quad t \in [0, n^{1/3}t_0].$$

For the argmax associated with the Brownian bridge we have with (3.4),

$$V_n^B(a) = \underset{t}{\operatorname{argmax}}\left\{Z_n^W(a,t) - n^{1/6}F(t_0, t_0+n^{-1/3}t)\xi_n\right\}.$$

Now choose $\delta > 0$ in such a way that $\delta f(0) < \frac{1}{4}\inf_{t\in(0,1)}|f'(t)|$, and note that for $x < n^{1/3}$,

$$P\{|\xi_n| > \delta n^{1/6}x\} \leq \exp\{-\tfrac{1}{2}\delta^2n^{1/3}x^2\} \leq \exp\{-\tfrac{1}{2}\delta^2x^3\}.$$

Hence

$$
\begin{aligned}
&P\{V_n^B(a) > x\} \\
&\le P\left\{\sup_{t\in[x,n^{1/3}(1-t_0)]} \left(Z_n^W(a,t) + \delta x n^{1/3} F(t_0, t_0 + n^{-1/3}t)\right) \ge 0\right\} + e^{-\frac{1}{2}\delta^2 x^3} \\
&\le P\left\{\sup_{t\in[x,n^{1/3}(1-t_0)]} \left(X_n^W(a,t) - c_1' t^2\right) \ge 0\right\} + e^{-\frac{1}{2}\delta^2 x^3},
\end{aligned}
$$

with $c_1' = \frac{1}{4}\inf_{t\in(0,1)} |f'(t)|$. Repeating the above optional stopping argument with τ_n replaced by the stopping time

$$
\tau_n' = \inf\left\{t \in [x, n^{1/3}(1-t_0)] : X_n^W(a,t) - c_1' t^2 \ge 0\right\}, \tag{3.11}
$$

the first probability in the last expression is bounded from above by $e^{-C'x^3}$, where $C' = (c_1')^2/(4c_2)$, with c_2 as before. It follows that

$$
P\{V_n^B(a) > x\} \le 2e^{-Cx^3},
$$

for all $x > 0$ and some $C > 0$, only depending on f. Similarly,

$$
P\{V_n^B(a) < -x\} \le P\left\{\sup_{t\in[x,n^{1/3}t_0]} \left(X_n^W(a,-t) - c_1' t^2\right) \ge 0\right\} + e^{-\frac{1}{2}\delta^2 x^3}.
$$

The bound on $P\{V_n^B(a) < -x\}$ is obtained by using the stopping time

$$
\tilde{\tau}_n' = \inf\left\{t \in [x, n^{1/3}t_0] : X_n^W(a,-t) - c_1' t^2 \ge 0\right\},
$$

and applying the optional stopping argument to the backward time martingale $\tilde{Y}_n(t)$. $\Box$

Remark 3.1 Theorem 3.1 for V_n^W holds more general. Let $L_n(a)$ be the location of the maximum of the process $t \mapsto X_n^W(a,t) - \Delta_n(a,t)$, where X_n^W is defined in (3.9) and $\Delta_n(a,t) \ge c_1 t^2$, uniformly for $t \in [0, n^{1/3}(t_0 \vee (1-t_0))]$. By the same argument as in the proof of Theorem 3.1, it follows that $P\{|L_n(a)| > x\} \le 2e^{-Cx^3}$, where C only depends on c_1.

The following theorem shows that properly normalized versions of $V_n^J(a)$ converge in distribution to a centered version of (1.3). For $a \in (f(1), f(0))$, let

$$
J_n(a) = \left\{c : a - \phi_2(a)cn^{-1/3} \in (f(1), f(0))\right\},
$$

and for $J = E, B, W$ and $c \in J_n(a)$, we define,

$$
V_{n,a}^J(c) = \phi_1(a) V_n^J(a - \phi_2(a)cn^{-1/3}), \tag{3.12}
$$

where

$$\begin{aligned}
\phi_1(a) &= \frac{|f'(g(a))|^{2/3}}{(4a)^{1/3}} > 0, \\
\phi_2(a) &= (4a)^{1/3}|f'(g(a))|^{1/3} > 0.
\end{aligned}$$

For $c \in I\!R$, let

$$\xi(c) = V(c) - c, \tag{3.13}$$

with $V(c)$ defined in (1.3).

Theorem 3.2 *For $J = E, B, W$, $d \geq 1$, $a \in (f(1), f(0))$ and $c \in J_n(a)^d$,we have joint distributional convergence of $(V^J_{n,a}(c_1), \ldots, V^J_{n,a}(c_d))$ to the random vector $(\xi(c_1), \ldots, \xi(c_d))$.*

Proof: First consider $V^W_{n,a}(c)$ in the case $d = 1$. Using (3.7) with $b = a - \phi_2(a)cn^{-1/3}$, we have that

$$\tilde{V}^W_{n,a}(c) = \phi_1(a)V^W_n(a - \phi_2(a)cn^{-1/3}) + \phi_1(a)n^{1/3}\left\{g(a - \phi_2(a)cn^{-1/3}) - g(a)\right\},$$

is the argmax of the process $t \mapsto Z^W_{n,a}(c,t)$, where

$$\begin{aligned}
Z^W_{n,a}(c,t) &= \frac{\phi_1(a)^{1/2}}{a^{1/2}}n^{1/6}\left\{W_n(F(g(a) + n^{-1/3}\phi_1(a)^{-1}t)) - W_n(F(g(a)))\right\} \\
&\quad + \frac{\phi_1(a)^{1/2}}{a^{1/2}}n^{2/3}\left\{F(g(a) + n^{-1/3}\phi_1(a)^{-1}t) - F(g(a)) - n^{-1/3}a\phi_1(a)^{-1}t\right\} \\
&\quad + 2ct.
\end{aligned}$$

Note that $\phi_1(a)n^{1/3}(g(a - \phi_2(a)cn^{-1/3}) - g(a))$ converges to c, as $n \to \infty$. By using Brownian scaling, a simple Taylor expansion and the uniform continuity of Brownian motion on compacta, for each $k = 1, 2, \ldots$ and each $c \in J_n(a)$ we have

$$\sup_{|t|\leq k} |Z^W_{n,a}(c,t) - Z(c,t)| \xrightarrow{P} 0, \quad \text{as } n \to \infty,$$

where

$$Z(c,t) = \left(\frac{\phi_1(a)}{a}\right)^{1/2} W\left(\frac{at}{\phi_1(a)}\right) - (t^2 - 2ct) \stackrel{d}{=} W(t) - t^2 + 2ct.$$

Now let $d \geq 1$ and note that for $t = (t_1, \ldots, t_d)$,

$$\begin{aligned}
(\tilde{V}^W_{n,a}(c_1), \ldots, \tilde{V}^W_{n,a}(c_d)) &= \operatorname*{argmax}_t \sum_{i=1}^d Z^W_{n,a}(c_i, t_i), \\
(V(c_1), \ldots, V(c_d)) &= \operatorname*{argmax}_t \sum_{i=1}^d Z(c_i, t_i).
\end{aligned}$$

Finally, because

$$\sup_{\|t\|\le k}\left|\sum_{i=1}^d Z_{n,a}^W(c_i,t_i)-\sum_{i=1}^d Z(c_i,t_i)\right|\le\sum_{i=1}^d\sup_{|t_i|\le k}|Z_{n,a}^W(c_i,t_i)-Z(c_i,t_i)|,$$

we conclude that the process $t\mapsto\sum_{i=1}^d Z_{n,a}^W(c_i,t_i)$ converges in the uniform topology on compacta to the process $t\mapsto\sum_{i=1}^d Z(c_i,t_i)$. The result for V_n^W follows from Theorem 2.7 in Kim and Pollard (1990).

Using (3.4) we can prove the same result for V_n^B by repeating the above steps, since $n^{-1/6}\xi_n t\to 0$ in probability, uniformly in t on compacta of $\mathbb{R}$. Finally, by using $\sup_{t\in\mathbb{R}}|D_n^E(a,t)-D_n^B(a,t)|=\mathcal{O}_p(n^{-1/2}\log n)$, the same result follows for V_n^E. □

We will need some independence structure for the process $\{U_n^W(a),a\in(f(1),f(0))\}$, where

$$U_n^W(a)=\underset{t\in[0,1]}{\operatorname{argmax}}\{W_n(F(t))+\sqrt{n}(F(t)-at)\}.$$

The mixing property of the process U_n^W can be argued intuitively in the following way. Observe that the event $\{U_n^W(a)=x\}$ is equivalent to

$$\begin{aligned}W_n(F(x))-W_n(F(t))&\ge\sqrt{n}(F(t)-F(x))+a\sqrt{n}(x-t),\quad t<x,\\W_n(F(x))-W_n(F(t))&>\sqrt{n}(F(t)-F(x))+a\sqrt{n}(x-t),\quad t>x.\end{aligned}$$

These are conditions on increments of $W_n\circ F$. Since for large M, the event $|U_n^W(a)-g(a)|<n^{-1/3}M$ has a probability close to 1, we can restrict t and x to $n^{-1/3}M$-neighborhoods of $g(a)$. The mixing property then follows from the fact that Brownian motion has independent increments.

Theorem 3.3 *The process $\{U_n^W(a)) : a\in(f(1),f(0))\}$ is strong mixing with mixing function:*

$$\alpha_n(d)=12e^{-C_1nd^3},\tag{3.14}$$

where the constant $C_1>0$ only depends on f. More specifically, for arbitrary $a\in(f(1),f(0))$ and $a+d\in(f(1),f(0))$:

$$\sup|P(A\cap B)-P(A)P(B)|\le\alpha_n(d),$$

where the supremum is taken over all sets $A\in\sigma\{U_n^W(c):f(1)<c\le a\}$ and $B\in\sigma\{U_n^W(c):a+d\le c<f(0)\}$.

Proof: Let $a\in(f(1),f(0))$ be arbitrary and take $f(1)<a_1\le a_2\le\dots\le a_k=a<a+d=c_1\le c_2\le\dots\le c_l<f(0)$ and consider the events

$$\begin{aligned}E_1&=\{U_n^W(a_1)\in A_1,\dots,U_n^W(a_k)\in A_k\},\\E_2&=\{U_n^W(c_1)\in B_1,\dots,U_n^W(c_l)\in B_l\},\end{aligned}$$

for Borel sets $A_1, \ldots, A_k$ and $B_1, \ldots, B_l$ of $\mathbb{R}$. Note that cylinder sets of the form E_1 and E_2 generate the σ-algebras $\sigma\{U_n^W(c) : f(1) < c \le a\}$ and $\sigma\{U_n^W(c) : a + d \le c < f(0)\}$, respectively. Now take $M_n = \frac{1}{4} dn^{1/3} \inf_{u \in (0,1)} |g'(u)|$ and consider the events

$$\begin{aligned} E_1' &= E_1 \cap \{U_{n,M_n}^W(a) = U_n^W(a)\}, \\ E_2' &= E_2 \cap \{U_{n,M_n}^W(a+d) = U_n^W(a+d)\}, \end{aligned}$$

where

$$U_{n,M_n}^W(c) = \operatorname{argmax}\{n^{1/3}|t - g(c)| \le M_n : W_n(F(t)) + \sqrt{n}(F(t) - ct)\}.$$

By monotonicity of U_n^W it follows that the event E_1' depends only on the increments of Brownian motion beyond time $F(g(a) - n^{-1/3}M_n)$ (note that g is decreasing) and that the event E_2' is only depending on the increments of Brownian motion before time $F(g(a+d) + n^{-1/3}M_n)$. By definition of M_n, it follows that E_1' and E_2' are independent. Since for all $a \in (f(1), f(0))$ we have that $V_n^W(a) = n^{1/3}(U_n^W(a) - g(a))$, according to Theorem 3.1,

$$\begin{aligned} &|P(E_1 \cap E_2) - P(E_1)P(E_2)| \\ &\quad \le 3P\{U_{n,M_n}^W(a) \ne U_n^W(a)\} + 3P\{U_{n,M_n}^W(a+d) \ne U_n^W(a+d)\} \\ &\quad = 3P\{n^{1/3}|U_n^W(a) - g(a)| > M_n\} + 3P\{n^{1/3}|U_n^W(a+d) - g(a+d)| > M_n\} \\ &\quad \le 12e^{-CM_n^3}, \end{aligned}$$

which proves the theorem. □

Apart from this exponential bound on the mixing function we will need the following two lemmas. The lemmas are analogous to Theorems 17.2.1 and 17.2.2 in Ibragimov and Linnik (1971) and can be proven similarly, since in the quoted Theorems 17.2.1 and 17.2.2 the stationarity is not essential.

Lemma 3.1 *If X is measurable with respect to $\sigma\{U_n^W(c) : f(1) < c \le a\}$ and Y is measurable with respect to $\sigma\{U_n^W(c) : a + d \le c < f(0)\}$ ($d > 0$), and if $|X| \le C_2$, $|Y| \le C_3$ a.s., then*

$$|E(XY) - E(X)E(Y)| \le 4C_2C_3\alpha_n(d).$$

Lemma 3.2 *If X is measurable with respect to $\sigma\{U_n^W(c) : f(1) < c \le a\}$ and Y is measurable with respect to $\sigma\{U_n^W(c) : a + d \le c < f(0)\}$ ($d > 0$), and suppose that for some $\delta > 0$,*

$$E|X|^{2+\delta} \le C_4, \quad E|Y|^{2+\delta} \le C_5,$$

then

$$|E(XY) - E(X)E(Y)| \le C_6(\alpha_n(d))^{\delta/(2+\delta)},$$

where $C_6 > 0$ only depends on C_4 and C_5.

In the following, we shall need some properties of the process V, which are contained in GROENEBOOM (1989) and HOOGHIEMSTRA AND LOPUHAÄ (1998). They are stated in the following lemma.

Lemma 3.3 *Let $V(0)$ be defined in (1.3) and for $b, c \in \mathbb{R}$, let $V_b(c)$ be defined by*

$$V_b(c) = \operatorname*{argmax}_t \{W(t) - b(t-c)^2\}. \tag{3.15}$$

Then,

(i) $V(0)$ has a bounded symmetric density.

(ii) for $x \to \infty$, $P\{|V(0)| > x\} \sim \lambda x^{-1} e^{-\frac{2}{3}x^3 - \kappa x}$, where $\lambda, \kappa > 0$.

(iii) for $h \downarrow 0$, $P\{V_b \text{ jumps in } (a-h, a+h)\} \leq \beta_1 h(1+o(1))$, where the constant $\beta_1 > 0$ is independent of a.

Proof: ad(i)-(ii). The first statement follows immediately from the representation for the density of $V(0)$ given in GROENEBOOM (1989). The second statement is Lemma 2.1 in HOOGHIEMSTRA AND LOPUHAÄ (1998).

ad(iii). Let $A_h = \{V \text{ jumps in } [0, h)\}$. Since the process $c \mapsto \xi(c)$ is stationary and has jumps at the same points as the process $c \mapsto V(c)$, we have that

$$\begin{aligned} P\{V \text{ jumps in } (a-h, a+h)\} &= P\{V \text{ jumps in } (-h, h)\} \\ &\leq 2\int_{-\infty}^{\infty} P\{A_h \mid V(0) = x\}\, f_{V(0)}(x)\, dx, \end{aligned}$$

where we also use the fact that $-V(-c) \stackrel{d}{=} V(c)$. In the proof of Theorem 3.1 in HOOGHIEMSTRA AND LOPUHAÄ (1998) it is derived, that

$$\lim_{h\downarrow 0} \frac{P\{A_h \mid V(0) = x\}}{h} = 2\int_0^{\infty} \frac{g_1(u+x)}{g_1(x)} up(u)\, du$$

(see GROENEBOOM (1989) or HOOGHIEMSTRA AND LOPUHAÄ (1998) for the exact definitions of the functions g_1 and p) and moreover that the right hand side is bounded uniformly in x. This implies that

$$P\{V \text{ jumps in } (a-h, a+h)\} \leq \beta_1' h + o(h), \quad h \downarrow 0,$$

where the constant β_1' is independent of a. By Brownian scaling we have that

$$V_b(c) \stackrel{d}{=} b^{-2/3} V(cb^{2/3}), \tag{3.16}$$

so that

$$P\{V_b \text{ jumps in } (a-h, a+h)\} \leq b^{2/3}\beta_1' h + o(h), \quad h \downarrow 0,$$

which proves (iii). □

Leaving the setting of the process V, it seems intuitively clear that the processes V_n^B and V_n^W have the same qualitative behavior, and will in particular satisfy a property analogous to Lemma 3.3(iii). This will be proved in the following lemma.

Lemma 3.4 *Let the interval J_n be defined by*

$$J_n = [f(1) + n^{-1/3}(\log n)^2, f(0) - n^{-1/3}(\log n)^2].$$

Then there exists a constant $\beta_2 > 0$, independent of $a \in J_n$, such that for $J = B, W$ and for all $h \in (0,1)$,

$$P\left\{V_n^J \text{ jumps in } (a - hn^{-1/3}, a + hn^{-1/3})\right\} \leq \beta_2 \delta_{n,h} + o(\delta_{n,h})$$

as $\delta_{n,h} \downarrow 0$, where $\delta_{n,h} = h \vee (n^{-1/3}(\log n)^2)$.

Proof: We first show the statement for V_n^W. Let $t_0 = g(a)$. For notational convenience define for $|c| \leq 1$,

$$V_n^W(a,c) = V_n^W(a + n^{-1/3}c) + n^{1/3}\{g(a + n^{-1/3}c) - g(a)\}.$$

Define the event $A_n = \{|V_n^W(a,c)| \leq \log n, \text{ for all } |c| \leq 1\}$. From (3.7) it follows that the process $c \mapsto V_n^W(a,c)$ is nonincreasing. Therefore,

$$P\{A_n^c\} \leq P\{V_n^W(a,-1) > \log n\} + P\{V_n^W(a,1) < -\log n\}.$$

Since $n^{1/3}|g(a \pm n^{-1/3}) - g(a)| \leq \sup_{u \in (0,1)} |g'(u)|$, it follows from conditions (A1)-(A3) and Theorem 3.1 that $P\{A_n^c\} = \mathcal{O}(e^{-C(\log n)^3})$. Hence we can restrict ourselves to A_n.

In order to transform $t \mapsto W_n(F(t_0 + n^{-1/3}t))$ into a process $y \to W_n(F(t_0) + n^{-1/3}y)$, define H_n by

$$H_n(y) = n^{1/3}\{H(F(t_0) + n^{-1/3}y) - t_0\}, \quad y \in [-n^{1/3}F(t_0), n^{1/3}(1 - F(t_0))], \tag{3.17}$$

where H is the inverse of F. Consider the process V_n^W as defined in (3.5), with t replaced by $H_n(y)$. Then by property (3.7) it follows that

$$V_n^W(a,c) = \sup\left\{H_n(y) \in [-n^{1/3}t_0, n^{1/3}(1 - t_0)] : \tilde{W}_n(a,y) - p_n(c,y) \text{ is maximal}\right\},$$

where

$$\tilde{W}_n(a,y) = n^{1/6}\{W_n(F(g(a)) + n^{-1/3}y) - W_n(F(g(a)))\}, \tag{3.18}$$

and

$$p_n(c,y) = -n^{1/3}y + n^{1/3}(a + n^{-1/3}c)H_n(y). \tag{3.19}$$

Conditions (A1)-(A3) imply that there exists a constant $K_1 > 0$, only depending on f, such that on A_n we have

$$\left|H_n^{-1}\left(V_n^W(a,c)\right)\right| \leq K_1 \log n.$$

Suppose that the process $c \mapsto V_n^W(c)$ jumps in the interval $(a - hn^{-1/3}, a + hn^{-1/3})$. Then from (3.8) if follows that the process $c \mapsto V_n^W(a,c)$ has a jump at some $c_0 \in (-h,h)$. This means that if we drop the function $y \mapsto p_n(c_0,y) + \beta$, for varying $\beta \in I\!R$, on the process $y \mapsto \tilde{W}_n(a,y)$, it first touches $\tilde{W}_n(a,y)$ simultaneously in two points (y_1,w_1) and (y_2,w_2). Note that on the event A_n, we have $|y_1 - y_2| \leq 2K_1 \log n$. We first show that for each y_i,

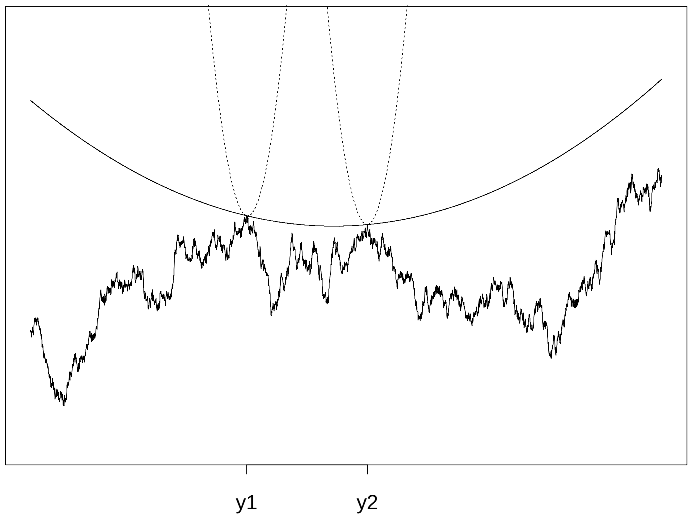

Figure 1: *The function* $p_n(c_0, y)$ *(straight line) and parabolas* $\pi_n(b_1, y)$ *and* $\pi_n(b_2, y)$ *(dotted) touching the process* $y \mapsto \tilde{W}_n(a, y)$ *at* y_1 *and* y_2.

$i = 1, 2$, we can construct a parabola that lies above $p_n(c_0, y)$ for all $|y| \le K_1 \log n$, and that touches $p_n(c_0, y)$ at (y_i, w_i).

To this end consider the second derivative of $p_n(c, y)$. Conditions (A1)-(A3) imply that for $|c| < 1$, there exists a constant $K_2 > 0$, only depending on f, such that

$$p_n''(c, y) = \frac{d^2 p_n(c, y)}{dy^2} \le aH''(F(t_0)) \left\{1 + K_2 n^{-1/3} |1 + y|\right\}.$$

Choose $M > K_2$ and define the parabola

$$\pi_n(c, y) = ca^{-1} y + \alpha_n y^2, \tag{3.20}$$

where $\alpha_n = \frac{1}{2} aH''(F(t_0)) \left\{1 + Mn^{-1/3}(1 + K_1 \log n)\right\}$. Then it follows immediately that for all $|y| \le K_1 \log n$, $|c| < 1$ and $b \in I\!R$:

$$\pi_n''(b, y) > p_n''(c, y).$$

If we choose b_1 such that $b_1 a^{-1} + 2\alpha_n y_1 = p_n'(c_0, y_1)$, then $\pi_n(b_1, y)$ and $p_n(c_0, y)$ have the same tangent at y_1. If we also take $\beta_1 = p_n(c_0, y_1) - \pi_n(b_1, y_1)$, then it follows that the parabola $\pi_n(b_1, y) + \beta_1$ lies above $p_n(c_0, y)$ and touches $p_n(c_0, y)$ at y_1. This implies that if we drop $\pi_n(b_1, y) + \beta$, for varying $\beta \in I\!R$, on the process $y \mapsto \tilde{W}_n(a, y)$ it first touches $\tilde{W}_n(a, y)$ at y_1. A similar construction holds at y_2 with a suitable choice for b_2 (see figure 1). Hence if we define

$$V_n^\pi(c) = \sup\left\{y \in [-n^{1/3} F(t_0), n^{1/3}(1 - F(t_0))] : \tilde{W}_n(a, y) - \pi_n(c, y) \text{ is maximal}\right\},$$

then from the above construction, it follows that the process $c \mapsto V_n^\pi(c)$ has a jump in the interval $[b_1, b_2]$ of maximal size $|y_1 - y_2| \le 2K_1 \log n$. Since $p_n'(c_0, y_i) = \pi_n'(b_i, y_i)$, for

$i = 1, 2$, it follows from conditions (A1)-(A3) that there exists a constant $K_3 > 0$, only depending on f, such that

$$|b_i - c_0| \leq K_3 y_i n^{-1/3} \log n, \quad i = 1, 2.$$

Because $c_0 \in (-h, h)$, this means that the interval $[b_1, b_2]$ is contained in

$$I_n = (-K_4(h \vee n^{-1/3}(\log n)^2), K_4(h \vee n^{-1/3}(\log n)^2)).$$

for some $K_4 > (1 \vee K_1 K_3)$. We conclude that, on the event A_n, we have that if $c \mapsto V_n^W(c)$ jumps in the interval $(a - hn^{-1/3}, a + hn^{-1/3})$, then the process $c \mapsto V_n^\pi(c)$ jumps in the interval I_n. However, the process $y \mapsto \tilde{W}_n(a, y)$ is distributed like Brownian motion W, so $V_n^\pi(c)$ is distributed as

$$\sup\Big\{y \in [-n^{1/3}F(t_0), n^{1/3}(1 - F(t_0))] : W(y) - ca^{-1}y - \alpha_n y^2 \text{ is maximal}\Big\}.$$

On the event A_n, this random variable is only different from

$$V_n(c) = \underset{y \in \mathbb{R}}{\operatorname{argmax}} \left\{ W(y) - \alpha_n \left(y + \frac{c}{2a\alpha_n} \right)^2 \right\},$$

if $V_n(c)$ is outside $[-K_1 \log n, K_1 \log n]$. Hence

$$P\{V_n^\pi \text{ jumps in } I_n, A_n\} \leq P\{V_n \text{ jumps in } I_n, A_n\} + P\left\{ \sup_{c \in I_n} |V_n(c)| > K_1 \log n, A_n \right\}.$$

According to Lemma 3.3, the first probability is of the order $h \vee (n^{-1/3}(\log n)^2)$. From the monotonicity of the process $c \mapsto V_n(c)$, property (3.16), the stationarity of the process $c \mapsto \xi(c)$ and Lemma 3.3, it follows that the second probability is of smaller order. This proves the result for V_n^W.

Turning to the Brownian bridge and the process $c \mapsto V_n^B(c)$, for $|c| \leq 1$ let

$$V_n^B(a, c) = V_n^B(a + n^{-1/3}c) + n^{1/3}\{g(a + n^{-1/3}c) - g(a)\}$$

and

$$\tilde{B}_n(a, y) = n^{1/6}\{B_n(F(g(a)) + n^{-1/3}y) - B_n(F(g(a)))\}.$$

Then

$$V_n^B(a, c) = \sup\Big\{H_n(y) \in [-n^{1/3}t_0, n^{1/3}(1 - t_0)] : \tilde{B}_n(a, y) - p_n(c, y) \text{ is maximal}\Big\}$$

where $p_n(c, y)$ is defined in (3.19). Now define $\psi_n(c)$ by

$$\psi_n(c) = \sup\Big\{y \in [-n^{1/3}F(t_0), n^{1/3}(1-F(t_0))] : \tilde{B}_n(a, y) - p_n(c - n^{-1/6}a\xi_n, y) \text{ is maximal}\Big\}.$$

Then $V_n^B(a, c) = H_n(\psi_n(c + n^{-1/6}a\xi_n))$. Using (3.4), we have

$$\psi_n(c) = \sup\Big\{y \in [-n^{1/3}F(t_0), n^{1/3}(1 - F(t_0))] : \tilde{W}_n(a, y) - q_n(c, y) \text{ is maximal}\Big\},$$

where $\tilde{W}_n$ is defined in (3.18) and

$$q_n(c,y) = n^{-1/6}\xi_n y - n^{1/3}y + n^{1/3}(a + n^{-1/3}c - n^{-1/2}a\xi_n)H_n(y).$$

Consider the event $A_n' \cap A_n''$ where

$$A_n' = \Big\{|V_n^B(a,c)| \le \log n, \text{ for all } c \in (-h,h)\Big\} \quad \text{and} \quad A_n'' = \{|\xi_n| \le n^{1/6}/\log n\}.$$

Similar to the event A_n, we have that $P\{(A_n')^c\}$ is of the order $e^{-C(\log n)^3}$. Furthermore, $P\{(A_n'')^c\} = 2(1 - \Phi(n^{1/6}/\log n))$, which is of smaller order than $n^{-1/3}(\log n)^2$. Hence we can restrict ourselves to the event $A_n' \cap A_n''$. Now suppose that $c \mapsto V_n^B(c)$ jumps in the interval $(a - hn^{-1/3}, a + hn^{-1/3})$. This means that the process $c \mapsto \psi_n(c)$ jumps in the interval $(-h + n^{-1/6}a\xi_n, h + n^{-1/6}a\xi_n)$. In that case a completely similar argument as before, involving a comparison of the derivatives of $q_n(c,y)$ and the parabola $\pi_n(c,y)$ defined in (3.20), yields that there exists a constant $K_5 > 0$, only depending on f, such that the process $c \mapsto V_n^\pi(c)$ jumps in the interval

$$I_n' = [-K_5(h \vee n^{-1/3}(\log n)^2), K_5(h \vee n^{-1/3}(\log n)^2)].$$

Hence on the event $A_n' \cap A_n''$, it follows that the probability that the process $c \mapsto V_n^\pi(c)$ has a jump in the interval I_n', is bounded by a probability of the order $h \vee (n^{-1/3}(\log n)^2)$. The result for V_n^B now follows. □

Corollary 3.1 *Let V_n^E be defined as in (3.1) and let V_n^B be defined as in (3.2). Then*

$$\int_{f(1)}^{f(0)} |V_n^E(a) - V_n^B(a)|\,da = \mathcal{O}_p(n^{-1/3}(\log n)^3).$$

Proof: Let the empirical process E_n and the Brownian bridge B_n be constructed on the same probability space. Then by the Hungarian embedding, we may assume

$$\sup_{t\in[0,1]} |E_n(t) - B_n(F(t))| = \mathcal{O}_p(n^{-1/2}\log n).$$

If K_n denotes the event $\{\sup_{t\in[0,1]} |E_n(t) - B_n(F(t))| \le n^{-1/2}(\log n)^2\}$, then $P\{K_n\} \to 1$, as $n \to \infty$. Also let

$$A_n = \Big\{|V_n^E(a)| \le \log n,\, |V_n^B(a)| \le \log n\Big\}$$

and write $A_n' = K_n \cap A_n$. Then by Theorem 2.1 and 3.1, we have $P\{K_n \cap A_n^c\} \le 6e^{-C(\log n)^3}$. Hence, since $|V_n^E(a) - V_n^B(a)| \le 2n^{1/3}$, we have for $a \in (f(1), f(0))$,

$$E|V_n^E(a) - V_n^B(a)|1_{K_n} \le E|V_n^E(a) - V_n^B(a)|1_{A_n'} + 12n^{1/3}e^{-C(\log n)^3}.$$

Now define $\epsilon_n = n^{-1/3}(\log n)^3$ and note that

$$\begin{aligned} E|V_n^E(a) - V_n^B(a)|1_{A_n'} &\le \int_0^{\epsilon_n} P\{|V_n^E(a) - V_n^B(a)| > x,\, A_n'\}\, dx \\ &\quad + \int_{\epsilon_n}^{2\log n} P\{|V_n^E(a) - V_n^B(a)| > x,\, A_n'\}\, dx. \end{aligned}$$

The first term on the right hand side is bounded by ϵ_n. To bound the second probability, consider the process $t \mapsto Z_n^B(a,t)$ be defined by

$$Z_n^B(a,t) = D_n^B(a,t) - n^{1/3}at, \quad t \in [-n^{1/3}g(a), n^{1/3}(1-g(a))],$$

where D_n^B is defined in (3.3), and let $\delta_n = n^{-1/3}(\log n)^2$. Since $n^{1/6}|E_n(t) - B_n(F(t))| \le \delta_n$ on the event A_n', we can only have $|V_n^E(a) - V_n^B(a)| > x$, if

$$|Z_n^B(a, V_n^B(a)) - Z_n^B(a,t)| \le 2\delta_n \tag{3.21}$$

for some $t \in [-n^{1/3}g(a), n^{1/3}(1-g(a))]$, such that $|t - V_n^B(a)| > x$.

Consider the line through the points $(V_n^B(a), D_n^B(a, V_n^B(a)))$ and $(t, D_n^B(a,t))$. This line has slope

$$n^{1/3}b = \frac{D_n^B(a,t) - D_n^B(a, V_n^B(a))}{t - V_n^B(a)} = \frac{Z_n^B(a,t) - Z_n^B(a, V_n^B(a))}{t - V_n^B(a)} + n^{1/3}a.$$

Hence it follows that

$$|b - a| \le 2n^{-1/3}\frac{\delta_n}{x}.$$

This means that if we drop a line with slope $n^{1/3}b$, it either first touches the process $s \mapsto D_n^B(a,s)$ simultaneously in the two (different) points t and $V_n^B(a)$, or in a third point different from both t and $V_n^B(a)$. Property (3.7) implies that the process

$$c \mapsto V_n^B(c) + n^{1/3}(g(c) - g(a))$$

must have a jump in the interval $I_n(x) = [a - 2n^{-1/3}\delta_n/x, a + 2n^{-1/3}\delta_n/x]$, and according to property (3.8) this means that the process $c \mapsto V_n^B(c)$ jumps in the interval $I_n(x)$. Hence, we get from Lemma 3.4,

$$\begin{aligned} E|V_n^E(a) - V_n^B(a)|1_{A_n'} &\le \epsilon_n + \int_{\epsilon_n}^{2\log n} P\{V_n^B \text{ jumps in } I_n(x),\, A_n'\}\, dx \\ &\le \epsilon_n + \beta_2\delta_n \int_{\epsilon_n}^{2\log n} \left(\frac{2}{x} \vee 1\right) dx = \mathcal{O}(n^{-1/3}(\log n)^3), \end{aligned}$$

where the term $\mathcal{O}(n^{-1/3}(\log n)^3)$ is uniform in $a \in (f(1), f(0))$. The result now follows from the Markov inequality. □

The following corollary will enable us to replace $E\int |V_n^W(a)|\, da$ by the asymptotic expectation μ, given in Theorem 1.1.

Corollary 3.2 *Let V_n^W be defined by (3.5), and let μ be defined as in Theorem 1.1. Moreover, let $V(0)$ be defined by (1.3). Then,*

(i) for all a such that

$$n^{1/3}\{F(g(a)) \wedge (1 - F(g(a)))\} \geq \log n, \tag{3.22}$$

we have

$$E|V_n^W(a)| = E|V(0)| \frac{(4a)^{1/3}}{|f'(g(a))|^{2/3}} + \mathcal{O}(n^{-1/3}(\log n)^4),$$

where the term $\mathcal{O}(n^{-1/3}(\log n)^4)$ is uniform in all a, satisfying (3.22).

(ii)

$$\lim_{n\to\infty} n^{1/6}\left\{\int_{f(1)}^{f(0)} E|V_n^W(a)|\,da - \mu\right\} = 0.$$

Proof: ad (i). Write $t_0 = g(a)$, so that

$$V_n^W(a) = \sup\left\{t \in [-n^{1/3}t_0, n^{1/3}(1-t_0)] : Z_n^W(a,t) \text{ is maximal}\right\},$$

where $Z_n^W(a,t) = D_n^W(a,t) - n^{1/3}at$, with D_n^W as defined in (3.6). Let $\tilde{V}_n^\pi(a)$ be the argmax defined by

$$\tilde{V}_n^\pi(a) = \sup\left\{t \in [-n^{1/3}t_0, n^{1/3}(1-t_0)] : Z_n^\pi(a,t) \text{ is maximal}\right\}$$

where

$$Z_n^\pi(a,t) = X_n^W(a,t) - n^{2/3}\frac{|f'(g(a))|}{2a^2}\Big(F(g(a) + n^{-1/3}t) - F(g(a))\Big)^2,$$

with X_n^W as defined in (3.9). It follows immediately that

$$\sup_{|t|\leq \log n} |Z_n^W(a,t) - Z_n^\pi(a,t)| \leq \delta_n, \tag{3.23}$$

where $\delta_n = K_1 n^{-1/3}(\log n)^3$, with $K_1 > 0$ only depending on f. Let A_n be the event $A_n = \{|\tilde{V}_n^\pi(a)| \leq \log n,\, |V_n^W(a)| \leq \log n\}$. Since $P\{|\tilde{V}_n^\pi(a)| > x\} \leq 2e^{-Cx^3}$, which can be seen by using the exponential martingale Y_n from (3.10) and a stopping time similar to (3.11), it follows, also using Theorem 3.1, that $P(A_n^c) = \mathcal{O}(e^{-C(\log n)^3})$. Hence

$$E|V_n^W(a) - \tilde{V}_n^\pi(a)| \leq E|V_n^W(a) - \tilde{V}_n^\pi(a)|1_{A_n} + \mathcal{O}(n^{1/3}e^{-C(\log n)^3}),$$

where the term $\mathcal{O}(n^{1/3}e^{-C(\log n)^3})$ is uniform in $a \in (f(1), f(0))$. Write $\epsilon_n = n^{-1/3}(\log n)^4$, and

$$\begin{aligned} E|V_n^W(a) - \tilde{V}_n^\pi(a)|1_{A_n} &= \int_0^{\epsilon_n} P\{|V_n^W(a) - \tilde{V}_n^\pi(a)| > x, A_n\}\,dx \\ &\quad + \int_{\epsilon_n}^{2\log n} P\{|V_n^W(a) - \tilde{V}_n^\pi(a)| > x, A_n\}\,dx. \end{aligned}$$

The first term on the right hand side is bounded by ϵ_n. Because (3.23) applies on A_n, we obtain, using the same argument as used in the proof of Corollary 3.1, that

$$\sup_{a\in(f(1),f(0))} E|V_n^W(a)-\tilde V_n^\pi(a)| = \mathcal{O}(n^{-1/3}(\log n)^4). \tag{3.24}$$

By change of variables $t = H_n(y)$, with H_n defined in (3.17), we have that

$$\tilde V_n^\pi(a) = \sup\left\{H_n(y)\in[-n^{1/3}t_0, n^{1/3}(1-t_0)] : X_n^W(a,H_n(y)) - \frac{|f'(t_0)|}{2f(t_0)^2}y^2 \text{ is maximal}\right\}.$$

Since $y\mapsto X_n^W(a,H_n(y))$ is distributed like Brownian motion W, we find that $\tilde V_n^\pi(a)$ is distributed as $H_n(V_{n,b})$, where with $b=\frac{1}{2}|f'(t_0)|/f(t_0)^2$,

$$V_{n,b} = \sup\Big\{y\in[-n^{1/3}F(t_0), n^{1/3}(1-F(t_0))] : W(y)-by^2 \text{ is maximal}\Big\}.$$

Now consider $V_b(0)$ as defined in (3.15), and write

$$E|H_n(V_{n,b})| = E|H_n(V_b(0))| + E(|H_n(V_{n,b})|-|H_n(V_b(0))|). \tag{3.25}$$

From conditions (A1)-(A3) and relation (3.16) we find that

$$E|H_n(V_b(0))| = a^{-1}E|V_b(0)| + \mathcal{O}(n^{-1/3}) = E|V(0)|\,\frac{(4a)^{1/3}}{|f'(g(a))|^{2/3}} + \mathcal{O}(n^{-1/3}).$$

Since $n^{1/3}\{F(t_0)\wedge(1-F(t_0))\}\ge\log n$, the location $V_{n,b}$ can only be different from $V_b(0)$ if $|V_b(0)|>\log n$. By using (3.16) and Lemma 3.3 we find that $P\{|V_b(0)|>\log n\}\le Ke^{-\frac{2}{3}(\log n)^3}$, where $K>0$ only depends on f. Hence from (3.25) we conclude that

$$E|\tilde V_n^\pi(a)| = E|H_n(V_{n,b})| = E|V(0)|\,\frac{(4a)^{1/3}}{|f'(g(a))|^{2/3}} + \mathcal{O}(n^{-1/3}).$$

Together with (3.24) this proves (i).

ad (ii). This follows immediately from (i), since the values of a for which (3.22) does not hold only give a contribution of order $n^{-1/3}\log n$ to the integral

$$\int_{f(1)}^{f(0)} E|V_n^W(a)|\,da,$$

and since

$$\mu = E|V(0)|\int_{f(1)}^{f(0)} \frac{(4a)^{1/3}}{|f'(g(a))|^{2/3}}\,da. \qquad \square$$

The following result shows that we only have to prove the asymptotic normality result for the process V_n^W.

Corollary 3.3 *Let V_n^B and V_n^W be defined as by (3.2) and (3.5), respectively. Then*

$$n^{1/6}\int_{f(1)}^{f(0)}\Big(|V_n^B(a)|-|V_n^W(a)|\Big)\,da=o_p(1).$$

Proof: Let, as before, W_n and B_n be linked by (3.4). Consider D_n^B and D_n^W as defined in (3.3) and (3.6). Let A_n be the event

$$A_n=\{|\xi_n|\le\log n,|V_n^W(a)|\le\log n,|V_n^B(a)|\le\log n\}.$$

Then on the event A_n, for all $|t|\le\log n$, we have $|D_n^W(a,t)-D_n^B(a,t)|\le K_1n^{-1/6}(\log n)^2$, for some constant $K_1>0$ only depending on f. By a similar argument as in the proof of Corollaries 3.1 and 3.2, we get

$$\sup_{a\in(f(1),f(0))}E|V_n^B(a)-V_n^W(a)|=\mathcal{O}(n^{-1/6}(\log n)^3). \tag{3.26}$$

This shows that in this way we cannot find a sufficiently small bound for the integral $n^{1/6}\int\{|V_n^B(a)|-|V_n^W(a)|\}\,da$.

Therefore, for a belonging to the set

$$J_n=\{a:\text{both }a\text{ and }a(1-\xi_nn^{-1/2})\in(f(1),f(0))\},$$

we introduce

$$V_n^B(a,\xi_n)=V_n^B(a-an^{-1/2}\xi_n)+n^{1/3}\left\{g(a-an^{-1/2}\xi_n)-g(a)\right\}.$$

By property (3.7) we have that

$$V_n^B(a,\xi_n)=\sup\left\{t\in[-n^{1/3}t_0,n^{1/3}(1-t_0)]:Z_n^\xi(a,t)\text{ is maximal}\right\},$$

where

$$Z_n^\xi(a,t)=Z_n^W(a,t)-n^{1/6}\xi_n\left\{F(t_0+n^{-1/3}t)-F(t_0)\right\}+n^{-1/6}\xi_nf(t_0)t.$$

Let the event A_n' be defined by

$$A_n'=\{|\xi_n|\le n^{1/6},\ |V_n^W(a)|\le\log n,\ |V_n^B(a,\xi_n)|\le\log n\}.$$

Then on A_n', for all $|t|\le\log n$ we have that $|Z_n^W(a,t)-Z_n^\xi(a,t)|\le K_2n^{-1/3}(\log n)^2$, for some constant $K_2>0$ not depending on a. Again by a similar argument as in the proof of Corollaries 3.1 and 3.2, we get

$$\sup_{a\in J_n}E|V_n^B(a,\xi_n)-V_n^W(a)|=\mathcal{O}(n^{-1/3}(\log n)^3),$$

With A_n' as defined in the manuscript, we have that

$$P\{(A_n')^c\} \leq P\{|\xi_n| > n^{1/6}\} + P\{|V_n^W(a)| > \log n\} + P\{|V_n^B(a,\xi_n)| > \log n\}.$$

Since $n^{1/3}|g(a - an^{-1/2}\xi_n) - g(a)| \leq \sup|g'|an^{-1/6}\xi_n$, for n sufficiently large

$$\begin{aligned} P\{|V_n^B(a,\xi_n)| > \log n\} &\leq P\{|V_n^B(a - an^{-1/2}\xi_n)| > \tfrac{1}{2}\log n\} \\ &= \int P\{|V_n^B(a - an^{-1/2}y)| > \tfrac{1}{2}\log n\}\phi(y)\,dy \\ &\leq 4e^{-C(\frac{1}{2}\log n)^3}. \end{aligned}$$

Hence

$$E|V_n^B(a,\xi_n) - V_n^W(a)| \leq E|V_n^B(a,\xi_n) - V_n^W(a)|1_{A_n'} + \mathcal{O}(n^{1/3}e^{-C(\log n)^3}),$$

where the term $\mathcal{O}(n^{1/3}e^{-C(\log n)^3})$ is uniform in $a \in (f(1), f(0))$. Write $\epsilon_n = n^{1/3}(\log n)^3$, and

$$\begin{aligned} E|V_n^B(a,\xi_n) - V_n^W(a)|1_{A_n'} &= \int_0^{\epsilon_n} P\{|V_n^B(a,\xi_n) - V_n^W(a)| > x, A_n'\}\,dx \\ &\quad + \int_{\epsilon_n}^{2\log n} P\{|V_n^B(a,\xi_n) - V_n^W(a)| > x, A_n'\}\,dx \end{aligned}$$

The first term on the right hand side is bounded by ϵ_n. Note that on A_n',

$$\sup_{|t|\leq \log n} |Z_n^W(a,t) - Z_n^\xi(a,t)| \leq \delta_n$$

where $\delta_n = K_2 n^{-1/3}(\log n)^2$, with K_2 not depending on a.

If $|V_n^B(a,\xi_n) - V_n^W(a)| > x$, then for some $t \in [-n^{1/3}g(a), n^{1/3}(1-g(a))]$ we must have $|t - V_n^W(a)| > x$. Similar as in the proof of Corollary 3.1, if follows that for such a t,

$$|Z_n^W(a, V_n^W(a)) - Z_n^W(a,t)| \leq 2\delta_n.$$

Consider the line through the points $(V_n^W(a), D_n^W(a, V_n^W(a)))$ and $(t, D_n^W(a,t))$. This line has slope

$$n^{1/3}b = \frac{D_n^W(a,t) - D_n^W(a, V_n^W(a))}{t - V_n^W(a)} = \frac{Z_n^W(a,t) - Z_n^W(a, V_n^W(a))}{t - V_n^W(a)} + n^{1/3}a$$

Hence it follows that

$$|b - a| \leq 2n^{-1/3}\frac{\delta_n}{x}.$$

This means that if we slide down a line with slope $n^{1/3}b$, it either first touches the process $s \mapsto D_n^W(a,s)$ simultaneously in two different points t and $V_n^W(a)$, or in a third point different from t and $V_n^W(a)$. According to property (3.7), this implies that the process

$$c \mapsto V_n^W(c) + n^{1/3}(g(c) - g(a))$$

must have a jump in the interval $I_n(x) = [a - 2n^{-1/3}\delta_n/x, a + 2n^{-1/3}\delta_n/x]$, and from property (3.8) this means that the process $c \mapsto V_n^W(c)$ has a jump in the interval $I_n(x)$. Hence, we get from Lemma 3.4, with $h = 2\delta_n/x = 2K_2 n^{-1/3}(\log n)^2/x$:

$$\begin{aligned} \int_{\epsilon_n}^{2\log n} P\{|V_n^W(a) - \tilde{V}_n^\pi(a)| > x, A_n\}\,dx &= \int_{\epsilon_n}^{2\log n} P\{V_n^W \text{ has a jump in } I_n(x),\, A_n\}\,dx \\ &\leq \beta_3 n^{-1/3}(\log n)^2 \int_{\epsilon_n}^{2\log n} \left(2K_2 x^{-1} \vee 1\right)\,dx \\ &= \mathcal{O}(n^{-1/3}(\log n)^3), \end{aligned}$$

where the term $\mathcal{O}(n^{-1/3}(\log n)^3)$ is uniform in $a \in (f(1), f(0))$.

and hence

$$n^{1/6}\int_{a\in J_n} E|V_n^B(a,\xi_n)-V_n^W(a)|\,da=o(1).$$

From Theorem 3.1 it follows that $E|V_n^B(a)|=\mathcal{O}(1)$ and $E|V_n^W(a)|=\mathcal{O}(1)$ uniformly in $a\in(f(1),f(0))$. Hence the contribution of the integrals over $[f(1),f(0)]\setminus J_n$ is negligible, and it remains to show that

$$n^{1/6}\int_{a\in J_n}\{|V_n^B(a,\xi_n)|-|V_n^B(a)|\}\,da=o_p(1). \tag{3.27}$$

Note that for the same reason

$$n^{1/6}\int_{a\in J_n}|V_n^W(a)|\,da=n^{1/6}\int_{f(1)}^{f(0}|V_n^W(a)|\,da+\mathcal{O}_p(n^{-1/3}),$$

and that by change of variables we get

$$n^{1/6}\int_{a\in J_n}|V_n^B(a,\xi_n)|\,da=n^{1/6}\int_{f(1)}^{f(0)}|V_n^B(a)-ag'(a)\xi_n n^{-1/6}|\,da+\mathcal{O}_p(n^{-1/3}).$$

Therefore

$$\begin{aligned}
&n^{1/6}\int_{a\in J_n}\{|V_n^B(a,\xi_n)|-|V_n^B(a)|\}\,da\\
&\quad=n^{1/6}\int_{f(1)}^{f(0)}\{|V_n^B(a)-ag'(a)\xi_n n^{-1/6}|-|V_n^B(a)|\}\,da+\mathcal{O}_p(n^{-1/3}).
\end{aligned}$$

Let $\epsilon>0$. Then

$$\begin{aligned}
&n^{1/6}\int_{f(1)}^{f(0)}\{|V_n^B(a)-ag'(a)n^{-1/6}\xi_n|-|V_n^B(a)|\}\,da \qquad (3.28)\\
&=n^{1/6}\int_{f(1)}^{f(0)}\{|V_n^B(a)-ag'(a)n^{-1/6}\xi_n|-|V_n^B(a)|\}1_{[0,\epsilon]}(|V_n^B(a)|)\,da\\
&\quad+n^{1/6}\int_{f(1)}^{f(0)}\{|V_n^B(a)-ag'(a)n^{-1/6}\xi_n|-|V_n^B(a)|\}1_{(\epsilon,\infty)}(|V_n^B(a)|)\,da.
\end{aligned}$$

We clearly have, using the independence of ξ_n and V_n^B, that the expectation of first term on the right hand side of (3.28) is bounded above by

$$E|\xi_n|\int_{f(1)}^{f(0)}|ag'(a)|P\{|V_n^B(a)|\le\epsilon\}\,da.$$

According to Theorem 3.2 it follows that $P\{|V_n^B(a)|\le\epsilon\}\to P\{|\xi(0)|\le\phi_1(a)\epsilon\}$. Hence from Lemma 3.3 and conditions (A1)-(A3) it follows that there exists a $K_3>0$, such that for any $\epsilon>0$,

$$\limsup_{n\to\infty} En^{1/6}\int_{f(1)}^{f(0)}\Big||V_n^B(a)-ag'(a)n^{-1/6}\xi_n|-|V_n^B(a)|\Big|1_{[0,\epsilon]}(|V_n^B(a)|)\,da\le K_3\epsilon. \tag{3.29}$$

For the second term on the right hand side of (3.28) we have

$$
\begin{aligned}
n^{1/6} & \int_{f(1)}^{f(0)} \{|V_n^B(a) - ag'(a)n^{-1/6}\xi_n| - |V_n^B(a)|\} 1_{(\epsilon,\infty)}(|V_n^B(a)|)\, da \\
&= \int_{f(1)}^{f(0)} 1_{(\epsilon,\infty)}(|V_n^B(a)|)\, \frac{-2\xi_n ag'(a)V_n^B(a) + n^{-1/6}\xi_n^2(ag'(a))^2}{|V_n^B(a) - ag'(a)n^{-1/6}\xi_n| + |V_n^B(a)|}\, da \\
&= \int_{f(1)}^{f(0)} 1_{(\epsilon,\infty)}(|V_n^B(a)|)\, \frac{-2\xi_n ag'(a)V_n^B(a)}{|V_n^B(a) - ag'(a)n^{-1/6}\xi_n| + |V_n^B(a)|}\, da + \mathcal{O}_p(n^{-1/6}) \\
&= -\xi_n \int_{f(1)}^{f(0)} ag'(a)\mathrm{sign}(V_n^B(a)) 1_{(\epsilon,\infty)}(|V_n^B(a)|)\, da + \mathcal{O}_p(n^{-1/6}), \qquad (3.30)
\end{aligned}
$$

using that for $|V_n^B(a)| > \epsilon$,

$$
\left| \frac{2V_n^B(a)}{|V_n^B(a) - ag'(a)n^{-1/6}\xi_n| + |V_n^B(a)|} - \frac{V_n^B(a)}{|V_n^B(a)|} \right| \leq \frac{|ag'(a)n^{-1/6}\xi_n|}{\epsilon} = \mathcal{O}_p(n^{-1/6}).
$$

For $a \in (f(1), f(0))$, let $S_n^B(a) = \mathrm{sign}(V_n^B(a)) 1_{(\epsilon,\infty)}(|V_n^B(a)|)$ and similarly, let $S_n^W(a) = \mathrm{sign}(V_n^W(a)) 1_{(\epsilon,\infty)}(|V_n^W(a)|)$. Then

$$
E\left\{ \xi_n \int_{f(1)}^{f(0)} ag'(a) S_n^B(a)\, da \right\}^2 = 2 \int\int_{f(1)<a<b<f(0)} abg'(a)g'(b) E\, S_n^B(a) S_n^B(b)\, da\, db.
$$

Furthermore

$$
|E\, S_n^B(a)S_n^B(b) - E\, S_n^W(a)S_n^W(b)| \leq E|S_n^B(a) - S_n^W(a)| + E|S_n^B(b) - S_n^W(b)|.
$$

Note that for every $a \in (f(1), f(0))$,

$$
E|S_n^B(a) - S_n^W(a)| \leq 2P\{|V_n^B(a) - V_n^W(a)| > 2\epsilon\} + P\{|V_n^B(a)| \leq \epsilon\} + P\{|V_n^W(a)| \leq \epsilon\}.
$$

By using the Markov inequality together with (3.26), the first probability on the right hand side tends to zero, uniformly in $a \in (f(1), f(0))$. According to Theorem 3.2 both $P\{|V_n^B(a)| \leq \epsilon\}$ and $P\{|V_n^W(a)| \leq \epsilon\}$ tend to $P\{|\xi(0)| \leq \phi_1(a)\epsilon\}$, which is $\mathcal{O}(\epsilon)$ according to Lemma 3.3 and conditions (A1)-(A3). Hence, there exists a $K_4 > 0$ such that for any $\epsilon > 0$,

$$
\begin{aligned}
\limsup_{n\to\infty} & E\left\{ \xi_n \int_{f(1)}^{f(0)} ag'(a) S_n^B(a)\, da \right\}^2 \\
& \leq \limsup_{n\to\infty} 2 \int\int_{f(1)<a<b<f(0)} abg'(a)g'(b) E\, S_n^W(a) S_n^W(b)\, da\, db + K_4\epsilon.
\end{aligned}
$$

Finally, write

$$
E\, S_n^W(a)S_n^W(b) = \mathrm{cov}(S_n^W(a), S_n^W(b)) + E\, S_n^W(a)\, E\, S_n^W(b).
$$

According to Lemma 3.1 and Theorem 3.3, for every $f(1) < a < b < f(0)$ we get that

$$|\mathrm{cov}(S_n^W(a), S_n^W(b))| \le 48 e^{-C_1 n(b-a)} \to 0.$$

Also for every $a \in (f(1), f(0))$, according to Theorem 3.2,

$$\begin{aligned} E\, S_n^W(a) &= P\{V_n^W(a) > \epsilon\} - P\{V_n^W(a) < -\epsilon\} \\ &\to P\{\xi(0) > \phi_1(a)\epsilon\} - P\{\xi(0) < -\phi_1(a)\epsilon\} = 0, \end{aligned}$$

because the distribution of $\xi(0)$ is symmetric (Lemma 3.3). It follows that there exists a $K_4 > 0$ such that for any $\epsilon > 0$,

$$\limsup_{n\to\infty} E\left\{\xi_n \int_{f(1)}^{f(0)} a g'(a) S_n^B(a)\, da\right\}^2 \le K_4 \epsilon.$$

Together with (3.29) and (3.28), this proves (3.27) and the result follows. □

4 Asymptotic normality

From Section 3 it follows that for proving Theorem 1.1, it suffices to prove that

$$T_n^W = n^{1/6} \int_{f(1)}^{f(0)} \left(|V_n^W(a)| - E|V_n^W(a)|\right) da \tag{4.1}$$

is asymptotically normal. We first derive the asymptotic variance of T_n^W. Theorem 3.2 together with Theorem 3.1, which guarantees the uniform integrability of the sequence $V_{n,a}^W(c)$ for a and c fixed, imply convergence of moments of $(V_{n,a}^W(0), V_{n,a}^W(c))$ to the corresponding moments of $(\xi(0), \xi(c))$. This leads to the following lemma.

Lemma 4.1 *For $n \to \infty$,*

$$\mathrm{var}\left(n^{1/6} \int_{f(1)}^{f(0)} |V_n^W(a)|\, da\right) \to 8 \int_0^\infty \mathit{cov}(|\xi(0)|, |\xi(c)|)\, dc.$$

Proof: We have that

$$\begin{aligned} &\mathrm{var}\left(n^{1/6} \int_{f(1)}^{f(0)} |V_n^W(a)|\, da\right) \\ &= 2n^{1/3} \int_{f(1)}^{f(0)} \int_a^{f(0)} \mathrm{cov}(|V_n^W(a)|, |V_n^W(b)|)\, db\, da \\ &= 8 \int_{f(1)}^{f(0)} a g'(a) \int_0^{n^{1/3}\phi_2(a)^{-1}(a-f(0))} \mathrm{cov}(|V_{n,a}^W(0)|, |V_{n,a}^W(c)|)\, dc\, da, \end{aligned}$$

by change of integration variables $b = a - \phi_2(a) c n^{-1/3}$. As noted above we have for a and c fixed

$$\mathrm{cov}(|V_{n,a}^W(0)|, |V_{n,a}^W(c)|) \to \mathrm{cov}(|\xi(0)|, |\xi(c)|).$$

Theorem 3.1 and the assumptions (A1)-(A2) also imply that, uniformly in n, a and c,

$$E|V_{n,a}^W(0)|^3 \leq C_4 \quad \text{and} \quad E|V_{n,a}^W(c)|^3 \leq C_5.$$

Hence by Lemma 3.2 together with (3.14), it follows that

$$|\text{cov}(|V_{n,a}^W(0)|, |V_{n,a}^W(c)|)| \leq C_6(\alpha_n(|n^{-1/3}\phi_2(a)c|))^{1/3} \leq D_1 e^{-D_2|c|^3},$$

where $D_1, D_2 > 0$ do not depend on n, a and c. It follows by dominated convergence that

$$\text{var}\left(n^{1/6}\int_{f(1)}^{f(0)} |V_n^W(a)|\, da\right) \to 8\int_{f(1)}^{f(0)} ag'(a)\int_0^{-\infty} \text{cov}(|\xi(0)|, |\xi(c)|)\, dc\, da.$$

Since the process $c \mapsto \xi(c)$ is stationary,

$$\int_0^{-\infty} \text{cov}(|\xi(0)|, |\xi(c)|)\, dc = -\int_0^{\infty} \text{cov}(|\xi(0)|, |\xi(c)|)\, dc.$$

Furthermore

$$-\int_{f(1)}^{f(0)} ag'(a)\, da = -\int_1^0 f(x)\, dx = 1.$$

This proves the lemma. □

Theorem 4.1 *Let T_n^W be defined by (4.1). Then*

$$T_n^W \xrightarrow{d} \mathcal{N}(0, \sigma^2),$$

where

$$\sigma^2 = 8\int_0^{\infty} \mathit{cov}(|\xi(0)|, |\xi(c)|)\, dc.$$

Proof: Define

$$W_n'(a) = |V_n^W(a)| - E|V_n^W(a)|.$$

Let

$$\begin{aligned}
L_n &= (f(0) - f(1))n^{-1/3}(\log n)^3, \\
M_n &= (f(0) - f(1))n^{-1/3}\log n, \\
N_n &= [(f(0) - f(1))/(L_n + M_n)] = \left[\frac{n^{1/3}}{\log n + (\log n)^3}\right],
\end{aligned}$$

where $[x]$ denotes the integer part of x. We divide the interval $(f(1), f(0))$ into blocks of alternating lengths

$$\begin{aligned}
A_j &= (f(1) + (j-1)(L_n + M_n), f(1) + (j-1)(L_n + M_n) + L_n), \\
B_j &= (f(1) + (j-1)(L_n + M_n) + L_n, f(1) + j(L_n + M_n)),
\end{aligned}$$

where $1 \leq j \leq N_n$. Now write

$$T'_n = S'_n + S''_n + R_n,$$

where

$$\begin{aligned} S'_n &= n^{1/6} \sum_{j=1}^{N_n} \int_{A_j} W'_n(a)\, da, \\ S''_n &= n^{1/6} \sum_{j=1}^{N_n} \int_{B_j} W'_n(a)\, da, \\ R_n &= n^{1/6} \int_{f(1)+N_n(L_n+M_n)}^{f(0)} W'_n(a)\, da. \end{aligned}$$

According to Theorem 3.1, and the Cauchy-Schwarz inequality, for all $a, b \in (f(1), f(0))$

$$E|W'_n(a)W'_n(b)| < K \tag{4.2}$$

where the constant $K > 0$ is uniformly in n, a and b. Together with the fact that

$$f(0) - f(1) - N_n(L_n + M_n) \leq L_n + M_n = \mathcal{O}(n^{-1/3}(\log n)^3),$$

this shows that $ER_n^2 \to 0$, and hence $R_n = o_p(1)$.

Next we show that the contribution of the small blocks (of length M_n) is negligible. To this end consider

$$E(S''_n)^2 = n^{1/3} \sum_{j=1}^{N_n} E\left(\int_{B_j} W'_n(a)\, da\right)^2 + n^{1/3} \sum_{i \neq j} \int_{B_i} \int_{B_j} EW'_n(a)W'_n(b)\, da\, db.$$

We have

$$|EW'_n(a)W'_n(b)| = |\mathrm{cov}(|V_n^W(a)|, |V_n^W(b)|)| \leq D_3 e^{-D_4 n|b-a|^3}$$

where $D_3, D_4 > 0$ only depend on f, by using Lemma 3.2 and (3.14). For $a \in B_i$ and $b \in B_j$, $i \neq j$, we have that $|b - a| \geq n^{-1/3}(\log n)^3$. Since $N_n \sim n^{1/3}/(\log n)^3$, this implies that

$$\left| n^{1/3} \sum_{i \neq j} \int_{B_i} \int_{B_j} EW'_n(a)W'_n(b)\, da\, db \right| \leq n^{1/3} N_n^2 M_n^2 D_3 e^{-D_4(\log n)^9} \to 0.$$

Hence

$$E(S''_n)^2 = n^{1/3} \sum_{j=1}^{N_n} E\left(\int_{B_j} W'_n(a)\, da\right)^2 + o(1).$$

Using (4.2) we obtain

$$E(S''_n)^2 = \mathcal{O}(n^{1/3} N_n M_n^2) \to 0,$$

and hence that the contribution of the small blocks is negligible.

Put

$$Y_j = n^{1/6} \int_{A_j} W'_n(a) da \quad \text{and} \quad \sigma_n^2 = \mathrm{var}\left(\sum_{j=1}^{N_n} Y_j\right),$$

so that $S'_n = \sum_{j=1}^{N_n} Y_j$ and $\sigma_n^2 = \mathrm{var}(S'_n)$. We have

$$\begin{aligned}&\left| E\exp\left\{\frac{iu}{\sigma_n}\sum_{j=1}^{N_n} Y_j\right\} - \prod_{j=1}^{N_n} E\exp\left\{\frac{iu}{\sigma_n}Y_j\right\}\right| \\ &\quad\leq \sum_{k=2}^{N_n}\left| E\exp\left\{\frac{iu}{\sigma_n}\sum_{j=1}^{k} Y_j\right\} - E\exp\left\{\frac{iu}{\sigma_n}\sum_{j=1}^{k-1} Y_j\right\} E\exp\left\{\frac{iu}{\sigma_n}Y_k\right\}\right| \\ &\quad\leq 4(N_n-1)\alpha_n(M_n),\end{aligned}$$

where the last inequality follows from Lemma 3.1. Observe that $N_n\alpha_n(M_n)\to 0$, which means that we can in fact apply the central limit theorem to independent copies of Y_j. The asymptotic normality of S'_n, hence follows if we show that the contributions of the large blocks, Y_j, satisfy the Lindeberg condition e.g., for each $\varepsilon > 0$,

$$\frac{1}{\sigma_n^2}\sum_{j=1}^{N_n} EY_j^2 1_{\{|Y_j|>\varepsilon\sigma_n\}} \to 0, \quad n\to\infty.$$

Note that by the Markov inequality

$$EY_j^2 1_{\{|Y_j|>\varepsilon\sigma_n\}} \leq \frac{1}{\varepsilon\sigma_n}E(|Y_j|^3).$$

Again using Cauchy-Schwarz and the uniform boundedness of the moments of $|W'_n(a)|$ we obtain

$$\sup_{1\leq j\leq N_n} E(|Y_j|^3) = n^{1/2}\mathcal{O}(|A_j|^3) = \mathcal{O}(n^{-1/2}(\log n)^9).$$

Hence

$$\frac{1}{\sigma_n^2}\sum_{j=1}^{N_n} EY_j^2 1_{\{|Y_j|>\varepsilon\sigma_n\}} \leq \frac{1}{\varepsilon\sigma_n^3}N_n \sup_{1\leq j\leq N_n} E(|Y_j|^3) = \mathcal{O}(\sigma_n^{-3}n^{-1/6}(\log n)^6).$$

Note that

$$\sigma_n^2 = \mathrm{var}(S'_n) = \mathrm{var}(T'_n) + \mathrm{var}(S''_n + R_n) - 2ET'_n(S''_n+R_n).$$

Using the obtained limit results for $E(S''_n)^2$ and ER_n^2 and the inequality of Cauchy-Schwarz we conclude that

$$\mathrm{var}(S''_n+R_n) = E(S''_n)^2 + ER_n^2 + 2E(S''_nR_n) \to 0,$$

and that according to Lemma 4.1

$$ET'_n(S''_n+R_n) \leq \sqrt{E(T'_n)^2\mathrm{var}(S''_n+R_n)} \to 0.$$

So we find that

$$\sigma_n^2 = \mathrm{var}(S'_n) = \sigma^2 + o(1),$$

which implies that

$$\frac{1}{\sigma_n^2}\sum_{j=1}^{N_n} EY_j^2 1_{\{|Y_j|>\varepsilon\sigma_n\}} = \mathcal{O}(n^{-1/6}(\log n)^6) \to 0.$$

This proves the theorem. □

Acknowledgement. We want to thank Jon Wellner, who suggested to us the martingales of Lemma 2.2, leading to the exponential bounds in Section 2. We also like to thank Vladimir Koulikov and an anonymous referee for their detailed comments.

References

Birgé, L. (1989). The Grenander estimator: a nonasymptotic approach. *Ann. of Statist.*, **17**, 1532–1549.

Csörgö, M. and Horvath, L. (1988). Central limit theorems for L_p-norms of density estimators. *Probab. Th. Rel. Fields*, **80**, 269-291.

Devroye, L. and Györfi, L. (1985). *Nonparametric density estimation: the L_1 view.* John Wiley, New York.

Devroye, L. (1987). *A course in density estimation.* Birkhäuser, Boston.

Durot, C. (1996). *Sharp asymptotics for isotonic regression.* Part of forthcoming Ph.D. thesis, Université Paris Sud, Orsay.

Groeneboom, P. (1985). Estimating a monotone density. In *Proceedings of the Berkeley Conference in Honor of Jerzey Neyman and Jack Kiefer*, L.M. Le Cam and R.A. Olshen, eds., **2**, 539–555.

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

Groeneboom, P. and Wellner, J.A. (1992). *Information bounds and nonparametric maximum likelihood estimation*, Birkhäuser, Boston.

Hooghiemstra, G. and Lopuhaä, H.P. (1998) An extremal limit theorem for the argmax process of Brownian motion minus a parabolic drift. *Extremes*, **1**, 215-240.

Ibragimov, I.A. and Linnik, Y.V. (1971). *Independent and stationary sequences of random variables.* Wolters-Noordhoff, Groningen.

Kim, J. and Pollard, D. (1990). Cube root asymptotics. *Ann. of Statist.*, **18**, 191-219.

Komlos, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent RV's, and the sample DF, part I. *Z. Wahrscheinlichkeitstheor. Verw. Geb.*, **32**, 111–131.

Pollard, D. (1984). *Convergence of stochastic processes.* Springer, New York.

Prakasa Rao, B.L.S. (1969). Estimation of a unimodal density. *Sankhȳa Ser. A*, **31**, 23-36.

Rogers, L.C.G. and Williams, D. (1997) *Diffusions, Markov processes, and Martingales, Volume 1.* Wiley, New York.

Wang, Y. (1992). *The L_1 theory of estimation of monotone and unimodal densities.* Part of unpublished Ph.D. thesis, University of California at Berkeley.