

Distribution of Global Measures of Deviation Between the Empirical Distribution Function and Its Concave Majorant

Vladimir N. Kulikov · Hendrik P. Lopuhaä

Received: 17 October 2006 / Revised: 19 April 2007 /
Published online: 15 June 2007
© Springer Science+Business Media, LLC 2007

Abstract We investigate the distribution of some global measures of deviation between the empirical distribution function and its least concave majorant. In the case that the underlying distribution has a strictly decreasing density, we prove asymptotic normality for several L_k -type distances. In the case of a uniform distribution, we also establish their limit distribution together with that of the supremum distance. It turns out that in the uniform case, the measures of deviation are of greater order and their limit distributions are different.

Keywords Empirical process · Least concave majorant · Central limit theorem · Brownian motion with parabolic drift · L_k distance

1 Introduction

Let X_1, \dots, X_n be a sample from a non-increasing density f on $[0, \infty)$ with corresponding concave distribution function F . Denote by F_n the empirical distribution function constructed from the sample and let \hat{F}_n be the least concave majorant of F_n , by which we mean the smallest concave function on $[0, \infty)$ that lies above F_n . This paper deals with the distribution of global measures of deviation between \hat{F}_n and F_n .

The distance between \hat{F}_n and F_n has been studied at several places in the literature. Kiefer and Wolfowitz [8] proved that, under additional assumptions that require $f' < 0$, both F_n and \hat{F}_n are asymptotic minimax in the class of concave distribution functions, i.e., $\sup |\hat{F}_n - F_n| = o(n^{-2/3} \log n)$ with probability one; closer inspection of their argument shows that this distance is of the order $\mathcal{O}((n^{-1} \log n)^{2/3})$ almost

V.N. Kulikov
ING Financial Markets, Amsterdam, The Netherlands

H.P. Lopuhaä (✉)
Delft Institute of Applied Mathematics, Delft University of Technology, Delft, The Netherlands
e-mail: h.p.lopuhaa@ewi.tudelft.nl

surely. This means that \hat{F}_n is essentially no better than F_n , except that it is concave. Similar results were obtained by Wang [14, 15] who studied supremum distances between the empirical cumulative hazard C_n and its least concave majorant \hat{C}_n , and between the corresponding distribution functions, for distributions with an increasing failure rate. Carolan [1] proved that $\hat{F}_n - F_n = o_p(1/\sqrt{n})$ at the endpoints of the longest interval where F is linear.

The first distributional result is due to Wang [16], who obtained the limit distribution of $n^{2/3}(\hat{F}_n(t) - F_n(t))$, for $t > 0$ being a fixed point. This was extended to process convergence in [12]. In the regression setting, Durot and Tocquet [4] also obtained the pointwise limit distribution and established asymptotic normality of the L_k -distance between \hat{F}_n and F_n . Durot [3], motivated by designing a test for monotonicity of the regression curve, obtained the limit distribution of the supremum distance under uniformity.

In this paper, we study the L_k -distance $\|\hat{F}_n - F_n\|_k$ and the supremum distance $\|\hat{F}_n - F_n\|_\infty$, for distributions with a non-increasing density with compact support, say $[0, 1]$. Our main result is that for any continuous g ,

$$n^{1/6} \left(\int_0^1 n^{2k/3} (\hat{F}_n(t) - F_n(t))^k g(t) dt - \mu \right) \tag{1.1}$$

is asymptotically normal. This implies that $n^{1/6}(n^{2/3}\|\hat{F}_n - F_n\|_k - \mu^{1/k})$ is asymptotically normal. This result is similar to the one in [4] for the regression setup, which has been obtained independently of our efforts. One of the main differences between the regression setting and our setup is the embedding of the empirical process. In the regression setting the empirical process can be embedded directly into Brownian motion itself, whereas in our setup it can only be embedded in the process

$$s \mapsto W(n^{1/3}(F(t + n^{-1/3}s) - F(t))).$$

This introduces an additional difficulty of approximating the value of the concave majorant of this process at zero by the corresponding value of the process $s \mapsto W(f(t)s)$. Although the maximum difference between the two processes is too large, the key observation that makes things work is that the values of the concave majorants at zero are sufficiently close.

We prove asymptotic normality of $\|\hat{F}_n - F_n\|_k$ under the assumption that $f' < 0$. In Sect. 2 we state a central limit theorem for (1.1) and briefly sketch an outline of the proof. Next, asymptotic normality of $\|\hat{F}_n - F_n\|_k$ and of $\int (\hat{F}_n - F_n)^k dF_n$ can be deduced from this result. The assumption $f' < 0$ rules out the uniform distribution. This assumption is essential, because the limit behavior of $\|\hat{F}_n - F_n\|_k$ under uniformity differs from the case $f' < 0$. Groeneboom [6] also encountered this phenomenon when studying the L_1 -error for the derivative of \hat{F}_n . In Sect. 3 we investigate the behavior of several global functionals under uniformity. This includes $\|\hat{F}_n - F_n\|_k$ and $\|\hat{F}_n - F_n\|_\infty$, for which we prove that they converge in distribution to corresponding functionals of Brownian motion at rate \sqrt{n} . For $\int (\hat{F}_n - F_n)^k dF_n$ we establish a similar result, which may be of interest in view of designing a test for uniformity. Also, for this functional the uniform distribution is shown to be least favorable among all non-increasing densities on $[0, 1]$, a property that is shared with $\|\hat{F}_n - F_n\|_\infty$ (for the latter see also [3]). All proofs are postponed to Sect. 4.

2 Asymptotic Normality of L_k -Functionals

Let X_1, X_2, \dots, X_n be a sample from a non-increasing density f and denote F as the corresponding distribution function. Suppose that f has bounded support, which then without loss of generality may be taken to be the interval $[0, 1]$. Let \hat{F}_n be the least concave majorant of the empirical distribution function F_n on $[0, 1]$. Consider the process

$$A_n(t) = n^{2/3}(\hat{F}_n(t) - F_n(t)), \quad t \in [0, 1]. \tag{2.1}$$

The limiting distribution of A_n , can be described in terms of the mapping CM_I that maps a function $h : \mathbb{R} \rightarrow \mathbb{R}$ into the least concave majorant of h on the interval $I \subset \mathbb{R}$. If we define the process

$$Z(t) = W(t) - t^2, \tag{2.2}$$

where W denotes standard two-sided Brownian motion originating from zero, then it is shown in [16] that, for $t \in (0, 1)$ fixed, $A_n(t)$ converges in distribution to $c_1(t)\zeta(0)$, where $c_1(t)$ is defined in (2.5), and

$$\zeta(t) = [CM_{\mathbb{R}}Z](t) - Z(t). \tag{2.3}$$

This was extended to process convergence in [12], where it is proved that, for $t \in (0, 1)$ fixed and $t + c_2(t)sn^{-1/3} \in (0, 1)$, the process

$$\zeta_{nt}(s) = c_1(t)A_n(t + c_2(t)sn^{-1/3}) \tag{2.4}$$

converges in distribution to the process ζ in the space $D(\mathbb{R})$ of cadlag functions on \mathbb{R} , where

$$c_1(t) = \left(\frac{|f'(t)|}{2f^2(t)}\right)^{1/3} \quad \text{and} \quad c_2(t) = \left(\frac{4f(t)}{|f'(t)|^2}\right)^{1/3}. \tag{2.5}$$

Our main result is a central limit theorem for L_k -functionals $\int_0^1 A_n(t)^k g(t)dt$, where g is continuous.

Theorem 2.1 *Suppose that f satisfies conditions*

- (A1) f is a twice continuous differentiable decreasing density with support on $[0, 1]$;
- (A2) $0 < f(1) \leq f(t) \leq f(s) \leq f(0) < \infty$ for $0 \leq s \leq t \leq 1$;
- (A3) $\inf_{t \in [0,1]} |f'(t)| > 0$.

Let g be a continuous function on $[0, 1]$ and let A_n be defined by (2.1). Then for all $k \geq 1$, with

$$\mu = E[\zeta(0)^k] \int_0^1 \frac{2^{k/3} f(t)^{2k/3}}{|f'(t)|^{k/3}} g(t)dt,$$

$n^{1/6}(\int_0^1 A_n(t)^k g(t)dt - \mu)$ converges in distribution to a normal random variable with mean zero and variance

$$\sigma^2 = \int_0^1 \frac{2^{(2k+5)/3} f(t)^{(4k+1)/3}}{|f'(t)|^{(2k+2)/3}} g(t)^2 dt \int_0^\infty \text{cov}(\zeta(0)^k, \zeta(s)^k) ds,$$

where ζ is defined in (2.3).

Let us first briefly sketch the line of reasoning how we establish this result. First observe that, up to constants, A_n is the image of F_n under the mapping D_I , that maps a function $h : \mathbb{R} \rightarrow \mathbb{R}$ into the difference between the least concave majorant of h on the interval I and h itself:

$$D_I h = \text{CM}_I h - h.$$

We can therefore write $A_n = n^{2/3} [D_{[0,1]} F_n]$. We will approximate F_n by means of a Brownian motion version and use its image under $D_{[0,1]}$ to approximate A_n . To this end, let E_n denote the empirical process $\sqrt{n}(F_n - F)$ and let B_n be a Brownian bridge constructed on the same probability space as the uniform empirical process $E_n \circ F^{-1}$ via the Hungarian embedding [9]. Let ξ_n be a $N(0, 1)$ distributed random variable independent of B_n and define versions W_n of Brownian motion by

$$W_n(t) = B_n(t) + \xi_n t, \quad t \in [0, 1]. \tag{2.6}$$

Write $F_n^E = F_n$ and let F_n^W be its Brownian approximation defined by

$$F_n^W(t) = F(t) + n^{-1/2} W_n(F(t)), \quad t \in [0, 1], \tag{2.7}$$

where W_n is defined in (2.6), and let

$$A_n^W(t) = n^{2/3} [D_{[0,1]} F_n^W](t). \tag{2.8}$$

For $J = E, W$, we first approximate the process $s \mapsto F_n^J(t + n^{-1/3}s)$ by the process $Y_{nt}(s) + L_{nt}^J(s)$, where

$$Y_{nt}(s) = n^{1/6} (W_n(F(t + n^{-1/3}s)) - W_n(F(t))) + \frac{1}{2} f'(t) s^2, \tag{2.9}$$

and L_{nt}^J denotes a linear drift (see Lemma 4.1). Since the mapping D_I is invariant under addition of linear terms, the moments of $A_n^J(t)$ can be approximated by the moments of $[D_I Y_{nt}](0)$ (see Lemma 4.2). The process Y_{nt} has the same distribution as the process

$$W(n^{1/3}(F(t + n^{-1/3}s) - F(t))) + \frac{1}{2} f'(t) s^2, \tag{2.10}$$

and by uniform continuity of Brownian motion on compacta, this process is close to the process $W(f(t)s) + f'(t)s^2/2$ on $|s| \leq \log n$. In view of Theorem 2.1 the difference between the corresponding concave majorants at zero must be of smaller order than $n^{-1/6}$. Unfortunately, it does not suffice to simply bound this difference by the maximum distance of the concave majorants on $|s| \leq \log n$, since this will be of order $\mathcal{O}(n^{-1/6} \log n)$. However, the key observation is that at zero the two concave

majorants are sufficiently close (see Lemma 4.3). This leads to the following key result:

$$EA_n^J(t)^k = \left(\frac{2f(t)^2}{|f'(t)|} \right)^{k/3} E\zeta(0)^k + o(n^{-1/6}),$$

where ζ is defined in (2.3) (see Lemma 4.4). A direct consequence is that the difference between the processes $A_n(t)^k$ and $A_n^W(t)^k$ is of smaller order than $n^{-1/6}$ (see Lemma 4.5). This means that it suffices to prove asymptotic normality for L_k -functionals

$$n^{1/6} \int_0^1 (A_n^W(t)^k - EA_n^W(t)^k)g(t)dt.$$

The fact that Brownian motion has independent increments will ensure that the process A_n^W is mixing (see Lemma 4.6). This allows us to approximate the integral by a sum of independent integral terms, which then leads to Theorem 2.1.

From Theorem 2.1 and an application of the delta method, asymptotic normality of L_k -distances $\|\hat{F}_n - F_n\|_k = (\int_0^1 (\hat{F}_n(t) - F_n(t))^k dt)^{1/k}$ can be established immediately.

Corollary 2.1 *Suppose that conditions (A1)–(A3) of Theorem 2.1 hold. Then $n^{1/6}(n^{2/3}\|\hat{F}_n - F_n\|_k - \mu^{1/k})$ converges in distribution to a normal random variable with mean zero and variance $\sigma^2\mu^{(2/k)-2}/k^2$, where μ and σ^2 are defined in Theorem 2.1 with $g = 1$.*

Another corollary of Theorem 2.1 is that $\int_0^1 (\hat{F}_n(t) - F_n(t))^k dF_n(t)$ has similar limit behavior. In Sect. 3 we establish the limit distribution of this L_k -functional at the uniform distribution, at which it attains its maximum value among all non-increasing densities on $[0, 1]$.

Corollary 2.2 *Suppose that conditions (A1)–(A3) of Theorem 2.1 hold. Then*

$$n^{1/6} \left(\int_0^1 [n^{2/3}(\hat{F}_n(t) - F_n(t))]^k dF_n(t) - \mu \right)$$

converges in distribution to a normal random variable with mean zero and variance σ^2 given in Theorem 2.1 with $g = f$.

Remark 2.1 The condition (A1) in Theorem 2.1 can be relaxed somewhat. At the cost of additional technicalities, the theorem remains true if we require $|f'(x) - f'(y)| \leq C|x - y|^{1/2+\epsilon}$, for some $\epsilon > 0$ and $C > 0$ not depending on f .

Remark 2.2 Note that the asymptotic variance of $\|\hat{F}_n - F_n\|_k$ tends to zero at a faster rate (i.e., at rate $n^{-5/3}$) than the variance of $\hat{F}_n(t) - F_n(t)$ at a fixed point t . In the latter case the rate is $n^{-4/3}$ (see [16]).

3 Behavior Under Uniformity

Within the class of distributions with a non-increasing density on $[0, 1]$, the greatest difference between \hat{F}_n and F_n is attained at the uniform distribution. This can be seen as follows. With $U_i = F(X_i)$, it holds that $G_n(F(t)) = F_n(t) \leq \hat{F}_n(t)$, where G_n denotes the empirical distribution function of the U_i 's. Let \hat{G}_n be the least concave majorant of G_n on $[0, 1]$. Because F is concave, $\hat{G}_n(F(t))$ is also concave, and it lies above $G_n(F(t)) = F_n(t)$. Since $\hat{F}_n(t)$ is the least concave function that lies above F_n , it follows that $\hat{F}_n(t) \leq \hat{G}_n(F(t))$. We find that for all $t \in [0, 1]$,

$$\hat{F}_n(t) - F_n(t) \leq \hat{G}_n(F(t)) - G_n(F(t)). \tag{3.11}$$

This property also follows from Theorem 1 in [1], and extends to global measures of deviation such as

$$S_n = \|\hat{F}_n - F_n\|_\infty = \sup_{t \in [0,1]} (\hat{F}_n(t) - F_n(t)). \tag{3.12}$$

See [3], for a similar result in the regression context. This property is however not true for $\|\hat{F}_n - F_n\|_k$, but only for weighted integrals such as

$$R_n = \int_0^1 (\hat{F}_n(t) - F_n(t))^k dF(t),$$

$$T_n = \int_0^1 (\hat{F}_n(t) - F_n(t))^k dF_n(t). \tag{3.13}$$

Proposition 3.1 *Let f be non-increasing on $[0, 1]$ and let S_n, R_n and T_n be defined as in (3.12) and (3.13). For a sample X_1, X_2, \dots, X_n from f and random variables U_1, U_2, \dots, U_n defined by $U_i = F(X_i)$, it holds that $S_n(X_1, \dots, X_n) \leq S_n(U_1, \dots, U_n)$ and similarly for R_n and T_n .*

The uniform distribution does not satisfy the conditions of Theorem 2.1. In fact, the third assumption $f' < 0$ is essential in the sense that the limit behavior of $\|\hat{F}_n - F_n\|_k$ under uniformity differs from that in Corollary 2.1. When $f' < 0$, the limit distribution of R_n can be deduced from Theorem 2.1, because in that case R_n corresponds to the choice $g = f$, and the limit distribution of T_n is given in Corollary 2.2. The behavior of these L_k -functionals and S_n in the uniform case is given in the following theorem. Let \hat{W} denote the least concave majorant of W on $[0, 1]$.

Theorem 3.1 *If $f = 1$, then*

- (i) $\sqrt{n} \|\hat{F}_n - F_n\|_k \rightarrow \|\hat{W} - W\|_k$ in distribution.
- (ii) $\sqrt{n} \|\hat{F}_n - F_n\|_\infty \rightarrow \sup_{t \in [0,1]} (\hat{W}(t) - W(t))$ in distribution.
- (iii) $n^{k/2} \int_0^1 (\hat{F}_n(t) - F_n(t))^k dF_n(t) \rightarrow \int_0^1 (\hat{W}(t) - W(t))^k dt$ in distribution.

In contrast with $\|\hat{F}_n - F_n\|_k$, for which we have established the limit distribution both in the uniform case and the case $f' < 0$, the exact limit behavior of $\|\hat{F}_n - F_n\|_\infty$

in the case $f' < 0$ is still unknown. The main difference with the uniform case is that when f is strictly decreasing, one only has local convergence to a limiting process, i.e., the process

$$s \mapsto c_1(t)n^{2/3}\{\hat{F}_n(t + c_2(t)sn^{-1/3}) - F_n(t + c_2(t)sn^{-1/3})\},$$

converges in distribution to the process $\zeta(s)$ as defined in (2.3). It is however clear that if f is strictly decreasing, $\|\hat{F}_n - F_n\|_\infty$ is of smaller order than in the uniform case. This follows immediately from [8], who showed that, if f is twice continuously differentiable with $f' < 0$, then $(\log n)^{-1}n^{2/3}\|\hat{F}_n - F_n\|_\infty$ tends to zero with probability one (in fact, they prove $\|\hat{F}_n - F_n\|_\infty = \mathcal{O}((n^{-1} \log n)^{2/3})$).

4 Proofs

We first show that for $J = E, W$, a properly scaled version of F_n^J can be approximated by the process Y_{nt} as defined in (2.9) plus linear term.

Lemma 4.1 *Suppose that f satisfies conditions (A1)–(A3). Let $F_n^E = F_n$ and let F_n^W be defined as in (2.7). Then for $t \in (0, 1)$ fixed, $J = E, W$ and $s \in [-tn^{1/3}, (1 - t)n^{1/3}]$:*

$$n^{2/3} F_n^J(t + n^{-1/3}s) = Y_{nt}(s) + L_{nt}^J(s) + R_{nt}^J(s),$$

where Y_{nt} is defined in (2.9), $L_{nt}^J(s)$ is linear in s , and where for all $k \geq 1$,

$$E \sup_{|s| \leq \log n} |R_{nt}^J(s)|^k = \mathcal{O}(n^{-k/3}(\log n)^{3k}),$$

uniformly in $t \in (0, 1)$.

Proof Taylor expansion together with (2.7) yields that

$$n^{2/3} F_n^W(t + n^{-1/3}s) = Y_{nt}(s) + L_{nt}^W(s) + R_{nt}^W(s),$$

with Y_{nt} as defined in (2.9), $L_{nt}^W(s)$ is linear in s :

$$L_{nt}^W(s) = n^{2/3} F(t) + n^{1/6} W_n(F(t)) + n^{1/3} f(t)s,$$

and

$$\begin{aligned} R_{nt}^W(s) &= n^{2/3}[F(t + n^{-1/3}s) - F(t) - f(t)n^{-1/3}s - f'(t)s^2n^{-2/3}] \\ &= \frac{1}{6}n^{-1/3} f''(\theta_1)s^3, \end{aligned}$$

for some $|\theta_1 - t| \leq n^{-1/3}|s|$. Similarly, with (2.6)

$$\begin{aligned} n^{2/3} F_n^E(t + n^{-1/3}s) &= n^{2/3} F_n^W(t + n^{-1/3}s) \\ &\quad + n^{1/6}\{E_n(t + n^{-1/3}s) - B_n(F(t + n^{-1/3}s))\} \end{aligned}$$

$$\begin{aligned}
 & -n^{1/6}\xi_n\{F(t) + f(t)n^{-1/3}s\} \\
 & -n^{1/6}\xi_n\{F(t + n^{-1/3}s) - F(t) - f(t)n^{-1/3}s\} \\
 & = Y_{nt}(s) + L_{nt}^E(s) + R_{nt}^E(s),
 \end{aligned}$$

where $L_{nt}^E(s) = L_{nt}^W(s) - n^{1/6}\xi_n F(t) - n^{-1/6}\xi_n f(t)s$ is linear in s , and

$$\begin{aligned}
 R_{nt}^E(s) &= R_{nt}^W(s) + n^{1/6}\{E_n(t + n^{-1/3}s) - B_n(F(t + n^{-1/3}s))\} \\
 &\quad - \frac{1}{2}n^{-1/2}\xi_n f'(\theta_2)s^2,
 \end{aligned}$$

for some $|\theta_2 - t| \leq n^{-1/3}|s|$. It follows immediately from conditions (A1)–(A3) that:

$$\sup_{|s| \leq \log n} |R_{nt}^W(s)|^k = \mathcal{O}(n^{-k/3}(\log n)^{3k}). \tag{4.14}$$

Note that

$$\sup_{|s| \leq \log n} |R_{nt}^E(s)| \leq \sup_{|s| \leq \log n} |R_{nt}^W(s)| + n^{1/6}S_n + \frac{1}{2} \sup |f'|n^{-1/2}(\log n)^2|\xi_n|,$$

where $S_n = \sup_{s \in \mathbb{R}} |E_n(s) - B_n(F(s))|$. From [9] we have that

$$P\{S_n \geq n^{-1/2}(C \log n + x)\} \leq Ke^{-\lambda x},$$

for positive constants C, K , and λ . This implies that for all $k \geq 1$,

$$ES_n^k = \mathcal{O}(n^{-k/2}(\log n)^k). \tag{4.15}$$

Next use that for all $a, b > 0$ and $k \geq 1$

$$(a + b)^k \leq 2^k(a^k + b^k). \tag{4.16}$$

Then from conditions (A1)–(A3) together with (4.15) and (4.14) we find that

$$E \sup_{|s| \leq \log n} |R_{nt}^E(s)|^k = \mathcal{O}(n^{-k/3}(\log n)^{3k}).$$

This proves the lemma. □

The next step is to approximate the moments of $A_n^J(t)$ by corresponding moments of the concave majorant of the process Y_{nt} . For this we need to show that the concave majorants of F_n^J on $[0,1]$, and of Y_{nt} on a neighborhood of t , are equal at t .

Lemma 4.2 *Suppose that f satisfies conditions (A1)–(A3). For $t \in (0, 1)$ fixed, let Y_{nt} be defined as in (2.9). Let $A_n^E(t) = A_n(t)$ and $A_n^W(t)$ be defined in (2.1) and (2.8). Define $H_{nt} = [-n^{1/3}t, n^{1/3}(1 - t)] \cap [-\log n, \log n]$. Then for all $k \geq 1$, and for $J = E, W$*

$$EA_n^J(t)^k = E[D_{H_{nt}}Y_{nt}](0)^k + o(n^{-1/6})$$

uniformly for $t \in (0, 1)$.

Proof For $d > 0$, let $I_{nt}(d) = [0, 1] \cap [t - dn^{-1/3}, t + dn^{-1/3}]$ and $N_{nt}^J(d)$ be the event that $[CM_{[0,1]}F_n^J](s)$ and $[CM_{I_{nt}(d)}F_n^J](s)$ are equal for $s \in I_{nt}(d/2)$. Write $I_{nt} = I_{nt}(\log n)$ and $N_{nt}^J = N_{nt}^J(\log n)$. Then on the event N_{nt}^J , the concave majorants $CM_{[0,1]}F_n^J$ and $CM_{I_{nt}}F_n^J$ coincide on $[t - n^{-1/3} \log n/2, t + n^{-1/3} \log n/2]$. In particular, they coincide at t , so that

$$A_n^J(t)1_{N_{nt}^J} = n^{2/3}[D_{I_{nt}}F_n^J](t)1_{N_{nt}^J} \quad \text{for } J = E, W. \tag{4.17}$$

By definition $|A_n^E(t)| \leq 2n^{2/3}$ and

$$A_n^W \leq 2n^{2/3} \left(1 + n^{-1/2} \sup_{s \in [0,1]} |W_n(s)| \right),$$

so that

$$\begin{aligned} E|A_n^J(t)^k - n^{2k/3}[D_{I_{nt}}F_n^J](t)^k|1_{(N_{nt}^J)^c} \\ \leq 2^{k+1}n^{2k/3}E \left(1 + n^{-1/2} \sup_{s \in [0,1]} |W_n(s)| \right)^k 1_{(N_{nt}^J)^c} \\ \leq 2^{k+1}n^{2k/3} \left\{ E \left(1 + n^{-1/2} \sup_{s \in [0,1]} |W_n(s)| \right)^{2k} \right\}^{1/2} \{P((N_{nt}^J)^c)\}^{1/2}. \end{aligned}$$

Next, use (4.16) together with the fact that all moments of $\sup_{s \in [0,1]} |W_n(s)|$ are finite. Then it follows from

$$P\{N_{nt}^J(d)^c\} \leq 8 \exp(-Cd^3), \tag{4.18}$$

where $C > 0$ does not depend on d, t and n (see [12]), that

$$\begin{aligned} E A_n^J(t)^k &= n^{2k/3} E[D_{I_{nt}}F_n^J](t)^k + E(A_n^J(t)^k - n^{2k/3}[D_{I_{nt}}F_n^J](t)^k)1_{(N_{nt}^J)^c} \\ &= n^{2k/3} E[D_{I_{nt}}F_n^J](t)^k + n^{2k/3} \mathcal{O}(e^{-C(\log n)^3/2}), \end{aligned}$$

uniformly for $t \in (0, 1)$. From Lemma 4.1 we have for $s \in H_{nt} = n^{1/3}(I_{nt} - t)$:

$$n^{2/3}[D_{I_{nt}}F_n^J](t) = [D_{H_{nt}}(Y_{nt} + R_{nt}^J)](0) = [D_{H_{nt}}Y_{nt}](0) + \Delta_{nt},$$

where $\Delta_{nt} = [D_{H_{nt}}(Y_{nt} + R_{nt}^J)](0) - [D_{H_{nt}}Y_{nt}](0)$. We find that

$$E A_n^J(t)^k = E[D_{H_{nt}}Y_{nt}](0)^k + \epsilon_{nt} + n^{2k/3} \mathcal{O}(e^{-C(\log n)^3/2}), \tag{4.19}$$

where, by application of the mean value theorem,

$$|\epsilon_{nt}| \leq k E|\theta_{nt}|^{k-1} |\Delta_{nt}| \leq k \{E|\theta_{nt}|^{2k-2}\}^{1/2} \{E|\Delta_{nt}|^2\}^{1/2}, \tag{4.20}$$

with $|\theta_{nt} - [D_{H_{nt}}Y_{nt}](0)| \leq |\Delta_{nt}|$. Since $H_{nt} \subset [-\log n, \log n]$, by application of (4.16),

$$E|\theta_{nt}|^{2k-2} \leq 4^{2k-2} \left(E \sup_{|s| \leq \log n} |Y_{nt}(s)|^{2k-2} + E|\Delta_{nt}|^{2k-2} \right), \tag{4.21}$$

where with Lemma 4.1, for all $k \geq 1$

$$E|\Delta_{nt}|^k \leq 2^k E \sup_{|s| \leq \log n} |R_{nt}^J(s)|^k = \mathcal{O}(n^{-k/3}(\log n)^{3k}), \tag{4.22}$$

uniformly for $t \in (0, 1)$. Furthermore, for $|s| \leq \log n$, there exist constants $C_1, C_2 > 0$ that only depend on f , such that

$$\begin{aligned} \sup_{|s| \leq \log n} |Y_{nt}(s)| &\leq \sup_{|s| \leq C_1 \log n} |W_n(s)| + C_2(\log n)^2 \\ &\stackrel{d}{=} (C_1 \log n)^{1/2} \sup_{|s| \leq 1} |W(s)| + C_2(\log n)^2. \end{aligned}$$

Because all moments of $\sup_{|s| \leq 1} |W(s)|$ are finite, from (4.20), (4.21) and (4.22) we conclude that $\epsilon_{nt} = \mathcal{O}(n^{-1/3}(\log n)^{2k+1})$. Together with (4.19) this proves the lemma. □

The process $Y_{nt}(s)$ has the same distribution as the process (2.10) and by uniform continuity of Brownian motion on compacta, this process is close to the process $W(f(t)s) + f'(t)s^2/2$. Unfortunately, it does not suffice to simply bound the maximum distance of the concave majorants on growing intervals containing zero. However, the next lemma ensures that the two concave majorants at zero are sufficiently close. We only need this lemma for continuous g , but with a little more effort a similar result can be obtained for non-continuous g .

Lemma 4.3 *Let g be a continuous function on an interval $B \subset \mathbb{R}$. Let $0 \in B^\circ$ and let $\phi : \mathbb{R} \rightarrow \mathbb{R}$ be invertible with $\phi(0) = 0$. Let $\sup_B g < \infty$ and suppose there exists an $\alpha \in [0, 1/2]$ such that*

$$1 - \alpha \leq \frac{\phi(t)}{t} \leq 1 + \alpha, \tag{4.23}$$

for all $t \in B \setminus \{0\}$. Then

$$|[\text{CM}_{\phi^{-1}(B)}(g \circ \phi)](0) - [\text{CM}_B g](0)| \leq 4\alpha \left\{ \sup_B g - [\text{CM}_B g](0) \right\}.$$

Proof Consider the function $h(t) = g(t) - \sup_B g$. For $a < b$, let $[a, b] \subset B$ be an interval containing zero. With property (4.23), t and $\phi(t)$ have the same sign. Hence, $\phi^{-1}(a) < \phi^{-1}(b)$ and $0 \in [\phi^{-1}(a), \phi^{-1}(b)]$. This yields the following inequalities

$$\begin{aligned} \frac{1 + \alpha}{1 - \alpha} \cdot \frac{h(a)b - h(b)a}{b - a} &\leq \frac{h(a)\phi^{-1}(b) - h(b)\phi^{-1}(a)}{\phi^{-1}(b) - \phi^{-1}(a)} \\ &\leq \frac{1 - \alpha}{1 + \alpha} \cdot \frac{h(a)b - h(b)a}{b - a}. \end{aligned} \tag{4.24}$$

First assume that both $\text{CM}_B h$ and $\text{CM}_{\phi^{-1}(B)}(h \circ \phi)$ have non-empty segments containing zero. Let $[\tau_1, \tau_2] \subset B$, with $\tau_1 < \tau_2$, be the segment of $\text{CM}_B h$ that contains

zero. Similarly, let $[\xi_1, \xi_2] \subset \phi^{-1}(B)$ be the segment of $CM_{\phi^{-1}(B)}(h \circ \phi)$ that contains zero, with $\xi_1 < \xi_2$. Denote $t_i = \phi^{-1}(\tau_i)$ and $x_i = \phi(\xi_i)$, for $i = 1, 2$, so that $t_1 < t_2$ and $x_1 < x_2$. Consider the line between $(x_1, h(x_1))$ and $(x_2, h(x_2))$. Since $[x_1, x_2] \subset B$, the intercept at zero of this line must be below $[CM_B h](0)$:

$$\frac{h(x_1)x_2 - h(x_2)x_1}{x_2 - x_1} \leq [CM_B h](0) = \frac{h(\tau_1)\tau_2 - h(\tau_2)\tau_1}{\tau_2 - \tau_1}. \tag{4.25}$$

Similarly, consider the line between $(t_1, (h \circ \phi)(t_1))$ and $(t_2, (h \circ \phi)(t_2))$. Since $[t_1, t_2] \subset \phi^{-1}(B)$, the intercept at zero of this line must be below $[CM_{\phi^{-1}(B)}(h \circ \phi)](0)$:

$$\begin{aligned} \frac{(h \circ \phi)(t_1)t_2 - (h \circ \phi)(t_2)t_1}{t_2 - t_1} &\leq [CM_{\phi^{-1}(B)}(h \circ \phi)](0) \\ &= \frac{(h \circ \phi)(\xi_1)\xi_2 - (h \circ \phi)(\xi_2)\xi_1}{\xi_2 - \xi_1}, \end{aligned}$$

or equivalently,

$$\begin{aligned} \frac{h(\tau_1)\phi^{-1}(\tau_2) - h(\tau_2)\phi^{-1}(\tau_1)}{\phi^{-1}(\tau_2) - \phi^{-1}(\tau_1)} &\leq [CM_{\phi^{-1}(B)}(h \circ \phi)](0) \\ &= \frac{h(x_1)\phi^{-1}(x_2) - h(x_2)\phi^{-1}(x_1)}{\phi^{-1}(x_2) - \phi^{-1}(x_1)}. \end{aligned}$$

Together with (4.25) and (4.24), this implies that

$$\begin{aligned} \frac{1 + \alpha}{1 - \alpha} [CM_B h](0) &\leq \frac{h(\tau_1)\phi^{-1}(\tau_2) - h(\tau_2)\phi^{-1}(\tau_1)}{\phi^{-1}(\tau_2) - \phi^{-1}(\tau_1)} \leq [CM_{\phi^{-1}(B)}(h \circ \phi)](0) \\ &\leq \frac{1 - \alpha}{1 + \alpha} \cdot \frac{h(x_1)x_2 - h(x_2)x_1}{x_2 - x_1} \leq \frac{1 - \alpha}{1 + \alpha} [CM_B h](0). \end{aligned}$$

Now use that $1 - 4\alpha \leq (1 - \alpha)/(1 + \alpha) \leq (1 + \alpha)/(1 - \alpha) \leq 1 + 4\alpha$, for $\alpha \in [0, 1/2]$, and the fact that $h \circ \phi = g \circ \phi - \sup_B g$. This finishes the proof for the case that both $CM_B h$ and $CM_{\phi^{-1}(B)}(h \circ \phi)$ have non-empty segments containing zero.

If this is not the case, for $\epsilon > 0$ sufficiently small, such that $[-\epsilon, \epsilon] \subset B$, define

$$g_\epsilon(t) = \begin{cases} g(t) & \text{if } t \in B \setminus [-\epsilon, \epsilon], \\ g(0) + (g(0) - g(-\epsilon))t/\epsilon & \text{if } t \in [-\epsilon, 0], \\ g(0) + (g(\epsilon) - g(0))\phi^{-1}(t)/\phi^{-1}(\epsilon) & \text{if } t \in [0, \epsilon]. \end{cases}$$

Then g_ϵ is continuous and linear on $[-\epsilon, 0]$ and the function $g_\epsilon \circ \phi$ is linear on $[0, \phi^{-1}(\epsilon)]$. This implies that for the corresponding function $h_\epsilon = g_\epsilon - \sup_B g_\epsilon$, both $CM_B h$ and $CM_{\phi^{-1}(B)}(h \circ \phi)$ have non-empty segments containing zero. Next, let $\delta > 0$ arbitrary and choose $\epsilon > 0$ sufficiently small such that $\sup |g_\epsilon - g| \leq \delta$. Then it follows that $|[CM_B g](0) - [CM_B g_\epsilon](0)| \leq \sup_{t \in [-\epsilon, \epsilon]} |g(t) - g_\epsilon(t)| \leq \delta$, and sim-

ilarly

$$\begin{aligned} & |[CM_{\phi^{-1}(B)}(g \circ \phi)](0) - [CM_{\phi^{-1}(B)}(g_\epsilon \circ \phi)](0)| \\ & \leq \sup_{t \in [\phi^{-1}(-\epsilon), \phi^{-1}(\epsilon)]} |(g \circ \phi)(t) - (g_\epsilon \circ \phi)(t)| \\ & = \sup_{t \in [-\epsilon, \epsilon]} |g(t) - g_\epsilon(t)| \leq \delta, \end{aligned}$$

where $\delta > 0$ can be chosen arbitrarily small. □

Lemma 4.4 *Suppose that f satisfies conditions (A1)–(A3). Let $t \in (0, 1)$ and let ζ be defined as in (2.3). Let $A_n^E(t) = A_n(t)$ and $A_n^W(t)$ be defined in (2.1) and (2.8). Then for all $k \geq 1$, and for $J = E, W$,*

$$EA_n^J(t)^k = \left(\frac{2f(t)^2}{|f'(t)|}\right)^{k/3} E\zeta(0)^k + o(n^{-1/6}),$$

uniformly in $t \in (n^{-1/3} \log n, 1 - n^{-1/3} \log n)$, and

$$EA_n^J(t)^k \leq \left(\frac{2f(t)^2}{|f'(t)|}\right)^{k/3} E\zeta(0)^k + o(n^{-1/6}),$$

uniformly in $t \in (0, 1)$.

Proof For $t \in (0, 1)$ fixed let Y_{nt} be defined as in (2.9) and let

$$Z_{nt}(s) = W(f(t)s) + \frac{1}{2}f'(t)s^2. \tag{4.26}$$

Let $a_{nt} = \max(0, t - n^{-1/3} \log n)$ and $b_{nt} = \min(1, t + n^{-1/3} \log n)$. Define the interval $J_{nt} = [n^{1/3}(F(a_{nt}) - F(t))/f(t), n^{1/3}(F(b_{nt}) - F(t))/f(t)]$ and the mapping

$$\phi_{nt}(s) = \frac{n^{1/3}(F(t + n^{-1/3}s) - F(t))}{f(t)}.$$

Let H_{nt} be the interval defined in Lemma 4.2. Then $H_{nt} = \phi_{nt}^{-1}(J_{nt}) = [n^{1/3}(a_{nt} - t), n^{1/3}(b_{nt} - t)]$, and there exists a constant $C_1 > 0$ only depending on f , such that for all $s \in H_{nt}$, we have $1 - \alpha_n \leq \phi_{nt}(s)/s \leq 1 + \alpha_n$, where $\alpha_n = C_1 n^{-1/3} \log n$. By definition

$$(Z_{nt} \circ \phi_{nt})(s) = \tilde{Y}_{nt}(s) + \frac{1}{2}f'(t)s^2 \left(\frac{\phi_{nt}(s)^2}{s^2} - 1\right),$$

where \tilde{Y}_{nt} is the process in (2.10), which has the same distribution as Y_{nt} . Since $H_{nt} \subset [-\log n, \log n]$, there exists a constant $C_2 > 0$ only depending on f , such that

$$|[D_{H_{nt}} \tilde{Y}_{nt}](0) - [D_{H_{nt}}(Z_{nt} \circ \phi_{nt})](0)| \leq C_2 n^{-1/3} (\log n)^3. \tag{4.27}$$

Now apply Lemma 4.3 with $g = Z_{nt}$, $\phi = \phi_{nt}$, $\alpha = \alpha_n$ and $B = J_{nt}$. This yields that

$$|[D_{H_{nt}}(Z_{nt} \circ \phi_{nt})](0) - [D_{J_{nt}}Z_{nt}](0)| \leq 8\alpha_n \sup_{s \in \mathbb{R}} |Z_{nt}(s)|.$$

Together with (4.27) we conclude that there exists a constant $C > 0$ only depending on f , such that

$$|[D_{H_{nt}}\tilde{Y}_{nt}](0) - [D_{J_{nt}}Z_{nt}](0)| \leq \frac{C \log n}{n^{1/3}} \left((\log n)^2 + \sup_{s \in \mathbb{R}} |Z_{nt}(s)| \right). \tag{4.28}$$

Similar to the proof of Lemma 4.2, this implies that

$$E[D_{H_{nt}}\tilde{Y}_{nt}](0)^k = E[D_{J_{nt}}Z_{nt}](0)^k + \epsilon_{nt}, \tag{4.29}$$

where $|\epsilon_{nt}| \leq k\{E|\theta_{nt}|^{2k-2}\}^{1/2}\{E|\Delta_{nt}|^2\}^{1/2}$, and $\Delta_{nt} = [D_{H_{nt}}\tilde{Y}_{nt}](0) - [D_{J_{nt}}Z_{nt}](0)$ and $|\theta_{nt} - [D_{H_{nt}}\tilde{Y}_{nt}](0)| \leq |\Delta_{nt}|$. Note that with $c_1(t)$ and $c_2(t)$ as defined in (2.5), by Brownian scaling one has

$$c_1(t)Z_{nt}(c_2(t)s) \stackrel{d}{=} Z(s) \tag{4.30}$$

where Z is defined in (2.2). Since $P\{\sup(W(t) - t^2) > x\} \leq 4 \exp(-x^3/2)$ (see for instance [10]), it follows that for all $k \geq 1$,

$$E\left(\sup_{s \in \mathbb{R}} |Z_{nt}(s)|\right)^k \leq CE\left(\sup_{s \in \mathbb{R}} |Z(s)|\right)^k < \infty,$$

for a constant $C > 0$ only depending on f . From (4.28) we conclude that for all $k \geq 1$,

$$E|\Delta_{nt}|^k = \mathcal{O}(n^{-k/3}(\log n)^{3k}). \tag{4.31}$$

Similar to the proof of Lemma 4.2, using an inequality similar to (4.21), together with (4.31), we find that $\epsilon_{nt} = \mathcal{O}(n^{-1/3}(\log n)^{2k+1})$, so that from (4.29) we get

$$E[D_{H_{nt}}\tilde{Y}_{nt}](0)^k = E[D_{J_{nt}}Z_{nt}](0)^k + \mathcal{O}(n^{-1/3}(\log n)^{2k+1}).$$

Together with Lemma 4.2 and scaling property (4.30), we find that

$$EA_n^J(t)^k = c_1(t)^{-k} E[D_{\mathbb{R}}Z](0)^k + c_1(t)^{-k} E\left([D_{I_{nt}}Z](0)^k - [D_{\mathbb{R}}Z](0)^k\right) + o(n^{-1/6}), \tag{4.32}$$

where $I_{nt} = c_2(t)^{-1}J_{nt}$. First note that for any $t \in (0, 1)$, on the interval I_{nt} , the concave majorant $CM_{I_{nt}}Z$ always lies below $CM_{\mathbb{R}}Z$. Because I_{nt} contains 0, this implies that

$$EA_n^J(t)^k \leq c_1(t)^{-k} E[D_{\mathbb{R}}Z](0)^k + o(n^{-1/6}),$$

uniformly for $t \in (0, 1)$.

When $t \in (n^{-1/3} \log n, 1 - n^{-1/3} \log n)$, there exist an $M > 0$, only depending on f , such that $[-M \log n, M \log n] \subset I_{nt}$. Note that on the interval $[-M \log n, M \log n]$

we always have $CM_{[-M \log n, M \log n]}Z \leq CM_{I_{nt}}Z \leq CM_{\mathbb{R}}Z$. Write $N_{nM} = N(M \log n)$, where $N(d)$ is the event that $[CM_{\mathbb{R}}Z](s)$ and $[CM_{[-d, d]}Z](s)$ are equal for $s \in [-d/2, d/2]$. On the event N_{nM} , we have $[CM_{[-M \log n, M \log n]}Z](0) = [CM_{I_{nt}}Z](0) = [CM_{\mathbb{R}}Z](0)$. Hence

$$\begin{aligned} |E([D_{I_{nt}}Z](0)^k - [D_{\mathbb{R}}Z](0)^k)| &\leq E|[D_{I_{nt}}Z](0)^k - [D_{\mathbb{R}}Z](0)^k|1_{N_{nM}^c} \\ &\leq 2^{k+1}E\left(\sup_{s \in \mathbb{R}}|Z(s)|\right)^k 1_{N_{nM}^c} \\ &\leq 2^{k+1}\left\{E\left(\sup_{s \in \mathbb{R}}|Z(s)|\right)^{2k}\right\}^{1/2}\{P(N_{nM}^c)\}^{1/2}. \end{aligned}$$

Since $E(\sup|Z|)^{2k} < \infty$, together with (see [12])

$$P(N(d)^c) \leq \exp(-d^3/2^7) \tag{4.33}$$

it follows that $E([D_{I_{nt}}Z](0)^k - [D_{\mathbb{R}}Z](0)^k) = o(n^{-1/6})$. Together with (4.32) and the fact that $\zeta = D_{\mathbb{R}}Z$ this proves the lemma. \square

Lemma 4.5 *Suppose that f satisfies conditions (A1)–(A3). Let $A_n^E = A_n$ and A_n^W be defined by (2.1) and (2.8). Then for all $k \geq 1$, we have $E|A_n^E(t)^k - A_n^W(t)^k| = o(n^{-1/6})$ uniformly for $t \in (0, 1)$.*

Proof Let I_{nt} and N_{nt}^J be defined as in the proof of Lemma 4.2, and let $K_{nt} = N_{nt}^E \cap N_{nt}^W$. Then according to (4.17):

$$\begin{aligned} E|A_n^E(t)^k - A_n^W(t)^k| &= n^{2k/3}E|[D_{I_{nt}}F_n^E](t)^k - [D_{I_{nt}}F_n^W](t)^k|1_{K_{nt}} \\ &\quad + E|A_n^E(t)^k - A_n^W(t)^k|1_{K_{nt}^c}. \end{aligned} \tag{4.34}$$

We first bound $E|A_n^E(t)^k - A_n^W(t)^k|1_{K_{nt}^c}$:

$$\begin{aligned} E|A_n^E(t)^k - A_n^W(t)^k|1_{K_{nt}^c} &\leq EA_n^E(t)^k 1_{K_{nt}^c} + EA_n^W(t)^k 1_{K_{nt}^c} \\ &\leq \{EA_n^E(t)^{2k}\}^{1/2}\{P(K_{nt}^c)\}^{1/2} + \{EA_n^W(t)^{2k}\}^{1/2}\{P(K_{nt}^c)\}^{1/2}, \end{aligned}$$

where, according to (4.18), $P(K_{nt}^c) \leq 16e^{-C(\log n)^3}$ uniformly in $t \in (0, 1)$. Since from Lemma 4.4 we know that $EA_n^J(t)^{2k}$ are bounded uniformly in n and $t \in (0, 1)$, we conclude that

$$E|A_n^E(t)^k - A_n^W(t)^k|1_{K_{nt}^c} = \mathcal{O}(e^{-C(\log n)^3/2}), \tag{4.35}$$

uniformly in $t \in (0, 1)$.

To bound the first expectation in (4.34), apply the mean value theorem to write

$$\begin{aligned}
 & n^{2k/3} |[D_{I_{nt}} F_n^E](t)^k - [D_{I_{nt}} F_n^W](t)^k| 1_{K_{nt}} \\
 & \leq k |\theta_{nt}|^{k-1} n^{2/3} |[D_{I_{nt}} F_n^E](t) - [D_{I_{nt}} F_n^W](t)| 1_{K_{nt}} \\
 & \leq k (A_n^E(t)^{k-1} + A_n^W(t)^{k-1}) n^{2/3} |[D_{I_{nt}} F_n^E](t) - [D_{I_{nt}} F_n^W](t)|. \tag{4.36}
 \end{aligned}$$

By Lemma 4.1

$$n^{2/3} |[D_{I_{nt}} F_n^E](t) - [D_{I_{nt}} F_n^W](t)| \leq 2 \left(\sup_{|s| \leq \log n} |R_n^E(s)| + \sup_{|s| \leq \log n} |R_n^W(s)| \right).$$

Hence, together with (4.36), the first expectation in (4.34) can be bounded by

$$k \{E(A_n^E(t)^{k-1} + A_n^W(t)^{k-1})^2\}^{1/2} \left\{ E \left(\sup_{|s| \leq \log n} |R_n^E| + \sup_{|s| \leq \log n} |R_n^W| \right)^2 \right\}^{1/2}.$$

From Lemma 4.4 together with (4.16), it follows that the first expectation is bounded uniformly for $t \in (0, 1)$. According to Lemma 4.1, the second expectation is of the order $\mathcal{O}(n^{-1/3}(\log n)^3)$. Together with (4.35) this proves the lemma. \square

Lemma 4.6 *Suppose that f satisfies conditions (A1)–(A3). The process $\{A_n^W(t) : t \in [0, 1]\}$ is strong mixing. More specifically, for $d > 0$,*

$$\sup |P(A \cap B) - P(A)P(B)| \leq \alpha_n(d) = 48e^{-Cnd^3},$$

where $C > 0$ only depends on f and where the supremum is taken over all sets $A \in \sigma\{A_n^W(s) : 0 < s \leq t\}$ and $B \in \sigma\{A_n^W(u) : t + d \leq u < 1\}$.

Proof Let $t \in (0, 1)$ arbitrary and take $0 < s_1 \leq s_2 \leq \dots \leq s_k = t < t + d = u_1 \leq u_2 \leq \dots \leq u_l < 1$. Consider events

$$\begin{aligned}
 E_1 &= \{A_n^W(s_1) \in B_1, \dots, A_n^W(s_k) \in B_k\}, \\
 E_2 &= \{A_n^W(u_1) \in C_1, \dots, A_n^W(u_l) \in C_l\},
 \end{aligned}$$

for Borel sets B_1, \dots, B_k and C_1, \dots, C_l of \mathbb{R} . Note that cylinder sets of the form E_1 and E_2 generate the σ -algebras $\sigma\{A_n^W(s) : 0 < s \leq t\}$ and $\sigma\{A_n^W(u) : t + d \leq u < 1\}$, respectively. Define the event

$$\begin{aligned}
 S &= \{[CM_{[0,1]} F_n^W](u) = [CM_{[0,t+d/2]} F_n^W](u) \text{ for } u \in [0, t] \\
 &\text{and } [CM_{[0,1]} F_n^W](u) = [CM_{[t+d/2,1]} F_n^W](u) \text{ for } u \in [t + d, 1]\}.
 \end{aligned}$$

Let $E'_1 = E_1 \cap S$ and $E'_2 = E_2 \cap S$. Then E'_1 only depends on the process F_n^W before time $t + d/2$ and E'_2 only depends on the process F_n^W after time $t + d/2$. Hence, by independency of the increments of the process F_n^W the events E'_1 and E'_2 are independent. Therefore by means of (4.18),

$$|P(E_1 \cap E_2) - P(E_1)P(E_2)| \leq 3P(S^c) \leq 48e^{-Cd^3n}$$

for some constant $C > 0$ that only depends on f . This proves the lemma. \square

From Lemmas 4.4 and 4.5 it follows immediately that for proving asymptotic normality of $n^{1/6} \int_0^1 (A_n(t)^k - EA_n(t)^k)g(t)dt$, it suffices to prove that its Brownian version

$$T_n^W = n^{1/6} \int_0^1 (A_n^W(t)^k - EA_n^W(t)^k)g(t)dt, \tag{4.37}$$

is asymptotically normal. The proof runs along the lines of the proof of Theorem 4.1 in [7]. We first derive the asymptotic variance of T_n^W . To this end we introduce the Brownian version of the process ζ_{nt} defined in (2.4). For $t \in (0, 1)$ fixed and $t + c_2(t)sn^{-1/3} \in (0, 1)$,

$$\zeta_{nt}^W(s) = c_1(t)A_n^W(t + c_2(t)sn^{-1/3}), \tag{4.38}$$

where A_n^W is defined in (2.8) and $c_1(t)$ and $c_2(t)$ are defined in (2.5). From the fact that ζ_{nt} converges to ζ in distribution (see Theorem 1.1 in [12]) and Lemma 4.5, it follows immediately that the process

$$\{\zeta_{nt}^W(s) : s \in \mathbb{R}\} \rightarrow \{\zeta(s) : s \in \mathbb{R}\} \text{ in distribution.} \tag{4.39}$$

Furthermore, note that Lemma 4.4 implies that for every $m = 1, 2, \dots$ there exists a constant $M > 0$ such that $EA_n^W(t)^{km} < M$, uniformly in $n = 1, 2, \dots$ and $t \in (0, 1)$. Hence it follows from Markov’s inequality, that for all $m = 1, 2, \dots$ there exists a constant $M' > 0$

$$P\{|\zeta_{nt}^W(s)|^k > y\} \leq \frac{M'}{y^m},$$

uniformly in $n = 1, 2, \dots, t \in (0, 1)$ and $t + c_2(t)sn^{-1/3} \in (0, 1)$. This guarantees uniform integrability of the sequence $\zeta_{nt}^W(s)^k$ for s, t and k fixed, so that together with (4.39) it implies convergence of moments of $(\zeta_{nt}^W(0)^k, \zeta_{nt}^W(s)^k)$ to the corresponding moments of $(\zeta(0)^k, \zeta(s)^k)$. This leads to the following lemma.

Lemma 4.7 *Suppose that f satisfies conditions (A1)–(A3). Then for any function g that is continuous on $[0, 1]$, and any $k \geq 1$,*

$$\begin{aligned} & \text{var}\left(n^{1/6} \int_0^1 A_n^W(t)^k g(t)dt\right) \\ & \rightarrow \int_0^1 \frac{2^{(2k+5)/3} f(t)^{(4k+1)/3}}{|f'(t)|^{(2k+2)/3}} g(t)^2 dt \int_0^\infty \text{cov}(\zeta(0)^k, \zeta(s)^k) ds. \end{aligned}$$

Proof We have with ζ_{nt}^W as defined in (4.38),

$$\begin{aligned} & \text{var}\left(n^{1/6} \int_0^1 A_n^W(t)^k g(t) dt\right) \\ &= 2n^{1/3} \int_0^1 \int_u^1 \text{cov}(A_n^W(t)^k, A_n^W(u)^k) g(t) g(u) dt du \\ &= 2 \int_0^1 \frac{c_2(t)}{c_1(t)^{2k}} \int_0^{n^{1/3}(1-t)/c_2(t)} \text{cov}(\zeta_{nt}^W(0)^k, \zeta_{nt}^W(s)^k) \\ & \quad \times g(t) g(t + c_2(t)sn^{-1/3}) dt ds, \end{aligned}$$

by change of variables of integration $u = t + c_2(t)sn^{-1/3}$. As noted above for s and t fixed,

$$\text{cov}(\zeta_{nt}^W(0)^k, \zeta_{nt}^W(s)^k) \rightarrow \text{cov}(\zeta(0)^k, \zeta(s)^k).$$

Lemma 4.4 implies that $E|\zeta_{nt}^W(0)|^{3k} \leq C_1$ and $E|\zeta_{nt}^W(s)|^{3k} \leq C_2$, uniformly in $n = 1, 2, \dots, s$ and t . Hence (see for instance Lemma 3.2 in [7]),

$$\text{cov}(\zeta_{nt}^W(0)^k, \zeta_{nt}^W(s)^k) \leq C_3 \alpha_n (n^{-1/3} c_2(t)s)^{1/3} \leq D_1 \exp(-D_2|s|^3),$$

where $C_3 > 0$ only depends on C_1, C_2 and $D_1, D_2 > 0$ do not depend on n, s and t . Substituting $c_1(t), c_2(t)$ as defined in (2.5), and using that g is uniformly bounded on $[0, 1]$, it follows by dominated convergence that

$$\begin{aligned} & \text{var}\left(n^{1/6} \int_0^1 A_n^W(t)^k g(t) dx\right) \\ & \rightarrow \int_0^1 \frac{2^{(2k+5)/3} f(t)^{(4k+1)/3}}{|f'(t)|^{(2k+2)/3}} g(t)^2 dt \int_0^\infty \text{cov}(\zeta(0)^k, \zeta(s)^k) ds. \quad \square \end{aligned}$$

Proof of Theorem 2.1 It suffices to prove the statement for T_n^W as defined in (4.37). Define

$$X_n(t) = (A_n^W(t)^k - EA_n^W(t)^k)g(t).$$

Let

$$L_n = n^{-1/3}(\log n)^3, \quad M_n = n^{-1/3} \log n, \quad N_n = \left\lceil \frac{1}{L_n + M_n} \right\rceil,$$

where $[x]$ denotes the integer part of x . We divide $[0, 1]$ into blocks of alternating length

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

where $1 \leq j \leq N_n$. Now write $T_n^W = S'_n + S''_n + R_n$, where

$$S'_n = n^{1/6} \sum_{j=1}^{N_n} \int_{A_j} X_n(t) dt, \quad S''_n = n^{1/6} \sum_{j=1}^{N_n} \int_{B_j} X_n(t) dt,$$

$$R_n = n^{1/6} \int_{N_n(L_n+M_n)}^1 X_n(t) dt.$$

According to Lemma 4.4 and the Cauchy–Schwarz inequality, for all $s, t \in (0, 1)$,

$$E|X_n(s)X_n(t)| \leq C, \tag{4.40}$$

where C is uniform with respect to s, t and n . Together with the fact that the length of the interval of integration for R_n is $O(n^{-1/3}(\log n)^3)$ this shows $E|R_n| \rightarrow 0$ and hence $R_n = o_p(1)$.

Next we show that contribution of integrals over small blocks is negligible. To this end consider

$$E(S''_n)^2 = n^{1/3} \sum_{j=1}^{N_n} E \left(\int_{B_j} X_n(t) dt \right)^2 + n^{1/3} \sum_{i \neq j} \int_{B_i} \int_{B_j} E X_n(s) X_n(t) ds dt.$$

As in the proof of Lemma 4.7, we have that

$$|E X_n(s) X_n(t)| = |g(s)g(t)| |\text{cov}(A_n^W(s)^k, A_n^W(t)^k)| \leq D_1 e^{-D_2 n|s-t|^3},$$

where $D_1, D_2 > 0$ do not depend s, t and n , using the fact that g is uniformly bounded on $[0, 1]$. Moreover, for $s \in B_i$ and $t \in B_j$, we have $|s - t| \geq n^{-1/3}(\log n)^3$. Since $N_n = O(n^{1/3}/(\log n)^3)$ this implies that

$$\left| n^{1/3} \sum_{i \neq j} \int_{B_i} \int_{B_j} E X_n(s) X_n(t) ds dt \right| \leq n^{1/3} N_n^2 M_n^2 D_1 e^{-D_2 (\log n)^9} \rightarrow 0.$$

Hence, using (4.40) we obtain $E(S''_n)^2 = O(n^{1/3} N_n M_n^2) + o(1) \rightarrow 0$, so that the contribution of the small blocks is negligible. Define

$$Y_j = n^{1/6} \int_{A_j} X_n(t) dt \quad \text{and} \quad \sigma_n^2 = \text{var} \left(\sum_{j=1}^{N_n} Y_j \right),$$

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

$$\left| E \exp \left\{ \frac{i u}{\sigma_n} \sum_{j=1}^{N_n} Y_j \right\} - \prod_{j=1}^{N_n} E \exp \left\{ \frac{i u}{\sigma_n} Y_j \right\} \right|$$

$$\leq \sum_{k=2}^{N_n} \left| E \exp \left\{ \frac{i u}{\sigma_n} \sum_{j=1}^k Y_j \right\} - E \exp \left\{ \frac{i u}{\sigma_n} \sum_{j=1}^{k-1} Y_j \right\} E \exp \left\{ \frac{i u}{\sigma_n} Y_k \right\} \right|$$

$$\leq 4(N_n - 1) \alpha_n(M_n),$$

where α_n is defined in Lemma 4.6. For the last inequality, see for instance Lemma 3.1 in [7]. Observe that $(N_n - 1)\alpha_n(M_n) \rightarrow 0$, which means that we can apply the central limit theorem to independent copies of Y_j . Asymptotic normality of S'_n follows if we can show that the independent copies of the Y_j 's satisfy the Lindeberg condition. Note that $EY_j^2 1_{\{|Y_j| > \epsilon\sigma_n\}} \leq E|Y_j|^3 / (\epsilon\sigma_n)$. Again by Cauchy–Schwarz and uniform boundedness of the moments of $|X_n(t)|$ 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

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

Note that $\sigma_n^2 = \text{var}(S'_n) = \text{var}(T_n^W) + \text{var}(S''_n + R_n) - 2ET_n^W(S''_n + R_n)$. Using $E(S''_n)^2 = o(1)$ and $ER_n^2 = o(1)$, together with Cauchy–Schwarz, we conclude that $\text{var}(S''_n + R_n) = E(S''_n)^2 + ER_n^2 + 2E(S''_n R_n) \rightarrow 0$. According to the Lemma 4.7, $ET_n^W(S''_n + R_n) \leq (E(T_n^W)^2 \text{var}(S''_n + R_n))^{1/2} \rightarrow 0$. So we find that $\sigma_n^2 = \text{var}(S'_n) = \sigma^2 + o(1)$, which implies

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

Proof of Corollary 2.2 It suffices to prove $n^{1/6} \int_0^1 A_n^k d(F_n - F) = o_p(1)$. Define the p -variation of a function g on an interval $J = [a, b]$ as $v_p(g; J) = \sup\{\sum_{j=1}^m |g(x_j) - g(x_{j-1})|^p\}$, where the supremum is taken over all point partitions $a = x_0 < x_1 < \dots < x_{m-1} < x_m = b$ of J . Let $V_p(g; J) = v_p(g; J)^{1/p}$ and let $V_{p,\infty}(g; J) = V_p(g; J) + \sup\{|g(x)|; x \in J\}$ be the p -variation norm of g . Then, for $p, q > 0, 1/p + 1/q > 1$,

$$\left| \int_0^1 A_n(t)^k d(F_n - F)(t) \right| \leq C_{p,q} V_{p,\infty}(A_n^k; [0, 1]) V_q(F_n - F; [0, 1]), \quad (4.41)$$

where $C_{p,q}$ is a universal constant (see for instance Theorem II.3.27 in [2]). According to Theorems I.6.1 and I.6.2 in [2],

$$V_{q,\infty}(F_n - F; [0, 1]) = \begin{cases} O_p(n^{(1-q)/q} (LLn)^{1/2}) & \text{for } 1 \leq q \leq 2, \\ O_p(n^{-1/2}) & \text{for } q > 2, \end{cases} \quad (4.42)$$

where $LLn = \max(1, \log n)$. Since, $V_{p,\infty}(g^k; J) \leq k \sup_J |g|^{k-1} V_{p,\infty}(g; J)$, and because according to [8], $\sup |A_n| = o(\log n)$ with probability one, it remains to determine $V_p(A_n; [0, 1]) = n^{2/3} V_p(\hat{F}_n - F_n; [0, 1])$. Let $\tau_1, \dots, \tau_{k_n}$ denote the points where \hat{F}_n has a change of slope and define $\tau_0 = 0$ and $\tau_{k_n+1} = 1$. Then, since $\hat{F}_n - F_n$ is positive and zero at the τ_i 's,

$$v_p(\hat{F}_n - F_n; [0, 1]) = \sum_{i=0}^{k_n} v_p(\hat{F}_n - F_n; [\tau_i, \tau_{i+1}]).$$

Note that for $t \in [\tau_i, \tau_{i+1}]$, we can write $\hat{F}_n(t) - F_n(\tau_i) - f(\tau_i)(t - \tau_i) = (\hat{f}_n(\tau_{i+1}) - f(\tau_i))(t - \tau_i)$, so that by Jensen’s inequality

$$\begin{aligned} v_p(\hat{F}_n - F_n; [\tau_i, \tau_{i+1}]) &\leq 3^{p-1} |\hat{f}_n(\tau_{i+1}) - f(\tau_i)|^p (\tau_{i+1} - \tau_i)^p \\ &\quad + 3^{p-1} v_p(F_n - F; [\tau_i, \tau_{i+1}]) \\ &\quad + 3^{p-1} v_p(g; [\tau_i, \tau_{i+1}]), \end{aligned} \tag{4.43}$$

where $g(t) = F(t) - F(\tau_i) - f(\tau_i)(t - \tau_i)$. By means of Theorem 2.1 in [7] it can be shown that $\sup |\tau_{i+1} - \tau_i| = O_p(n^{-1/3} \log n)$. See for instance, the proof of Lemma 4.1 in [11]. Since $|g(t) - g(s)|^p \leq \sup |f'|^p |\tau_{i+1} - \tau_i|^{2p-1} |t - s|$, it follows that

$$\sum_{i=0}^{k_n} v_p(g; [\tau_i, \tau_{i+1}]) = O_p(n^{-(2p-1)/3} (\log n)^{2p-1}). \tag{4.44}$$

Furthermore,

$$\begin{aligned} \sum_{i=0}^{k_n} v_p(F_n - F; [\tau_i, \tau_{i+1}]) &\leq v_p(F_n - F; [0, 1]) \\ &\leq V_{p,\infty}(F_n - F; [0, 1])^p, \end{aligned} \tag{4.45}$$

where the order of the right hand side can be deduced from (4.42). Finally,

$$\sum_{i=0}^{k_n} |\hat{f}_n(\tau_{i+1}) - f(\tau_i)|^p (\tau_{i+1} - \tau_i)^p \leq \sup |\tau_{i+1} - \tau_i|^{p-1} \int_0^1 |\hat{f}_n(t) - \tilde{f}_n(t)|^p dt,$$

where $\tilde{f}_n(t) = f(\tau_i)$, for $t \in [\tau_i, \tau_{i+1}]$, $i = 0, 1, \dots, k_n$. Note that according to Theorem 1.1 in [11], $\int |\hat{f}_n - f|^p dt = O_p(n^{-p/3})$ for $1 \leq p < 2.5$, and that

$$\begin{aligned} \int_0^1 |f(t) - \tilde{f}_n(t)|^p dt &= \sum_{i=0}^{k_n} \int_{\tau_i}^{\tau_{i+1}} (f(t) - f(\tau_i))^p dt \\ &\leq \sup |f'|^p \sup |\tau_{i+1} - \tau_i|^p = O_p(n^{-p/3} (\log n)^p). \end{aligned}$$

This means that

$$\sum_{i=0}^{k_n} |\hat{f}_n(\tau_{i+1}) - f(\tau_i)|^p (\tau_{i+1} - \tau_i)^p = O_p(n^{-(2p-1)/3} (\log n)^{2p-1}). \tag{4.46}$$

Hence, according to decomposition (4.43), and bounds (4.44), (4.45), (4.46), and because $\sup A_n = o(\log n)$ with probability one, it follows that

$$V_{p,\infty}(A_n; [0, 1]) = \begin{cases} O_p(n^{-1/3+1/p} (LLn)^{1/2}) & \text{for } 1 \leq p < 2, \\ O_p(n^{1/6} (\log n)^{3/2}) & \text{for } p = 2, \\ O_p(n^{1/6}) & \text{for } 2 < p < 2.5. \end{cases}$$

Now choose $2 < p < 2.5$ and $5/3 < q < p/(p - 1) < 2$. Then,

$$\left| \int_0^1 A_n(t)^k d(F_n - F)(t) \right| = o_p(n^{-5/6+1/q} (\log n)^{k-1} (LLn)^{1/2}),$$

and since $q > 5/3$, this finishes the proof. □

Proof of Proposition 3.1 The inequality for S_n follows immediately from (3.11). Using $F_n(t) = G_n(F(t))$ in addition, we find

$$\begin{aligned} \int_0^1 (\hat{F}_n(t) - F_n(t))^k dF_n(t) &\leq \int_0^1 (\hat{G}_n(F(t)) - G_n(F(t)))^k dG_n(F(t)) \\ &= \int_0^1 (\hat{G}_n(s) - G_n(s))^k dG_n(s) \end{aligned}$$

which proves the inequality for T_n . The argument for R_n is similar. □

Proof of Theorem 3.1 When $f = 1$, then $F(t) = t$, so that

$$n^{1/2}(\hat{F}_n(t) - F_n(t)) = [D_{[0,1]}E_n](t), \tag{4.47}$$

where $E_n(t) = \sqrt{n}(F_n(t) - t)$ is the uniform empirical process. Since the mapping $h \mapsto \|[D_{[0,1]}h]\|_k$ is continuous, it follows that $\sqrt{n}\|\hat{F}_n - F_n\|_k$ converges in distribution to $\|[D_{[0,1]}B]\|_k$, where B denotes Brownian bridge. Because $B(t)$ has the same distribution as $W(t) - tW(1)$ and $D_{[0,1]}$ is invariant under addition of linear functions, this proves (i). The argument for (ii) is similar using (4.47) and the fact that the mapping $h \mapsto \sup_{t \in [0,1]} [D_{[0,1]}h](t)$ is continuous.

For (iii) it suffices to prove $\int D_n(t)^k d(F_n - F)(t) = o_p(1)$, where $D_n = \sqrt{n}(\hat{F}_n - F_n)$. Similar to (4.41) we have

$$\left| \int_0^1 D_n^k d(F_n - F) \right| \leq C_{p,q} k (\sup |D_n|)^{k-1} V_{p,\infty}(D_n; [0, 1]) V_q(F_n - F; [0, 1]),$$

where according to Marshall’s lemma (see for instance [13]), $\sup |D_n| \leq 2 \sup |\sqrt{n}(F_n - F)| = O_p(1)$. Furthermore, as in (4.43), we can write

$$\begin{aligned} \sum_{i=0}^{k_n} v_p(\hat{F}_n - F_n; [\tau_i, \tau_{i+1}]) &\leq 2^{p-1} \sum_{i=0}^{k_n} |\hat{f}_n(\tau_{i+1}) - 1|^p (\tau_{i+1} - \tau_i)^p \\ &\quad + 2^{p-1} v_p(F_n - F; [0, 1]). \end{aligned}$$

Take $p = 2$ and $1 < q < 2$. Then $v_p(F_n - F; [0, 1]) = O_p(n^{-1}LLn)$, according to (4.42), and

$$\sum_{i=0}^{k_n} |\hat{f}_n(\tau_{i+1}) - 1|^2 (\tau_{i+1} - \tau_i)^2 \leq \int_0^1 (\hat{f}_n(t) - 1)^2 dt = O_p(n^{-1}(\log n)),$$

according to Theorem 3.2 in [5]. Because $\sup |D_n| = O_p(1)$, it follows that

$$\left| \int_0^1 D_n^k d(F_n - F) \right| = O_p(n^{(1-q)/q} (\log n)^{1/2} (LLn)^{1/2}),$$

Since $q > 1$, this proves the theorem. \square

References

1. Carolan, C.A.: The least concave majorant of the empirical distribution function. *Can. J. Stat.* **30**, 317–328 (2002)
2. Dudley, R.M., Norvaiša, R.: Differentiability of Six Operators on Nonsmooth Functions and p -Variation. *Lecture Notes in Mathematics*, vol. 1703. Springer, Berlin (1999)
3. Durot, C.: A Kolmogorov-type test for monotonicity of regression. *Stat. Probab. Lett.* **63**, 425–433 (2003)
4. Durot, C., Tocquet, A.S.: On the distance between the empirical process and its concave majorant in a monotone regression framework. *Ann. Inst. H. Poincaré Probab. Stat.* **39**, 217–240 (2003)
5. Groeneboom, P.: The concave majorant of Brownian motion. *Ann. Probab.* **11**(4), 1016–1027 (1983)
6. Groeneboom, P.: Estimating a monotone density. In: Le, L.M. (ed.) *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer*, vol. II, pp. 539–555. Wadsworth, Belmont (1985)
7. Groeneboom, P., Hooghiemstra, G., Lopuhaä, H.P.: Asymptotic normality of the L_1 error of the Grenander estimator. *Ann. Stat.* **27**, 1316–1347 (1999)
8. Kiefer, J., Wolfowitz, J.: Asymptotically minimax estimation of concave and convex distribution functions. *Z. Wahrsch. Verw. Geb.* **34**, 73–85 (1976)
9. Kórnlos, J., Major, P., Tusnády, G.: An approximation of partial sums of independent RV's and the sample DF. *Z. Wahrsch. Verw. Geb.* **32**, 111–131 (1975)
10. Kulikov, V.N.: Direct and indirect use of maximum likelihood. PhD Thesis, TU Delft (2003)
11. Kulikov, V.N., Lopuhaä, H.P.: Asymptotic normality of the L_k -error of the Grenander estimator. *Ann. Stat.* **33**(5), 2228–2255 (2005)
12. Kulikov, V.N., Lopuhaä, H.P.: The limit process of the difference between the empirical distribution function and its concave majorant. *Stat. Probab. Lett.* **76**(16), 1781–1786 (2006)
13. Robertson, T., Wright, F.T., Dykstra, R.L.: *Order Restricted Inference*. Wiley, New York (1988)
14. Wang, J.L.: Asymptotically minimax estimators for distributions with increasing failure rate. *Ann. Stat.* **14**(3), 1113–1131 (1986)
15. Wang, J.L.: Estimators of a distribution function with increasing failure rate average. *J. Stat. Plann. Inference* **16**(3), 415–427 (1987)
16. Wang, Y.: The limit distribution of the concave majorant of an empirical distribution function. *Stat. Probab. Lett.* **20**, 81–84 (1994)