The Annals of Statistics
2012, Vol. 40, No. 3, 1578–1608
DOI: 10.1214/12-AOS1015

THE LIMIT DISTRIBUTION OF THE L_∞-ERROR OF GRENANDER-TYPE ESTIMATORS

BY CÉCILE DUROT, VLADIMIR N. KULIKOV AND HENDRIK P. LOPUHAÄ

University of Nanterre, ASR Nederland and Delft University of Technology

Let f be a nonincreasing function defined on [0, 1]. Under standard regularity conditions, we derive the asymptotic distribution of the supremum norm of the difference between f and its Grenander-type estimator on sub-intervals of [0, 1]. The rate of convergence is found to be of order $(n/\log n)^{-1/3}$ and the limiting distribution to be Gumbel.

1. Introduction. After the derivation of the nonparametric maximum likelihood estimator (NPMLE) of a monotone density and a monotone failure rate by Grenander [10], and the least squares estimator of a monotone regression function by Brunk [4], it has taken some time before the distribution theory for such estimators entered the literature. The limiting distribution of the NPMLE of a decreasing density on $[0, \infty)$ at a fixed point in the interior of the support, has been established by Prakasa Rao [27]. Similar results were obtained for the NPMLE of a monotone failure rate in [28] and for an estimator of a monotone regression function in [5]. Woodroofe and Sun [32] showed that the NPMLE of a decreasing density is inconsistent at zero. The behavior at the boundary has been further investigated in [2, 23]. Smooth estimation has been studied in [25], for monotone regression curves, and in [31] for monotone densities; see also [9] and [1]. The limit distribution of the NPMLE of a decreasing function in the Gaussian white noise model was obtained in [33]. Related likelihood ratio based techniques have been investigated in [3, 26].

Groeneboom [11] reproved the result in [27] by introducing a new approach based on inverses. This approach has become a cornerstone in deriving pointwise asymptotics of several shape constrained nonparametric estimators, for example, for the distribution function of interval censored observations (see [15]) or for estimators of a monotone density and a monotone hazard under random censoring (see [17]); see also [18] for the limiting distribution of the NPMLE of a monotone density under random censoring and [24] for similar results on isotonic estimators for a monotone baseline hazard in Cox proportional hazards model. The limit distribution of these estimators involves an argmax process $\{\zeta(c) : c \in \mathbb{R}\}$ connected with two-sided Brownian motion with a parabolic drift. This process has been

Received November 2011; revised May 2012.
MSC2010 subject classifications. Primary 62E20, 62G20; secondary 62G05, 62G07.
Key words and phrases. Supremum distance, extremal limit theorem, least concave majorant, monotone density, monotone regression, monotone failure rate.

studied extensively in [12], where it is also claimed that the approach based on inverses should be sufficiently general to deal with global measures of deviation, such as the L_1-distance or the supremum distance between the estimator and the monotone function of interest. Indeed, the limiting distribution of the L_1-distance between a decreasing density and its NPMLE was obtained in [14], and a similar result can be found in [6] in the monotone regression setup. These results were extended to general L_k-distances in [22] and [7]. In [7], the limiting distribution of L_k-distances is obtained in a very general framework that includes, among others, the monotone density case, monotone regression and monotone failure rate.

Little to nothing is known about the behavior of the supremum distance. In [19], the rate of the supremum distance is established in a semi-parametric model for censored observations, and it is suggested that the same rate should hold in the monotone density case. In [16] an extremal limit theorem has been obtained for suprema of the process $\zeta(c)$ over increasing intervals. However, a long-standing open problem remains, although this problem has important statistical applications: what is the limiting distribution of the supremum distance between a monotone function and its isotonic estimator? Indeed, while pointwise confidence intervals for a decreasing density, a monotone regression function or a monotone hazard are available using the limiting distribution of the isotonic estimator at the fixed point, nonparametric confidence bands have remained a formidable challenge; they could be built if the limiting distribution of the supremum distance between a monotone function and its isotonic estimator were known. It is the purpose of this paper to settle this question in the same general framework as considered in [7]. The precise construction of a nonparametric confidence band requires additional technicalities that are beyond the scope of the present paper. It is only briefly discussed here, and details are deferred to a separate paper.

We consider Grenander type estimators $\widehat{f}_n$ for decreasing functions f with compact support, say [0, 1]. These are estimators that are defined as the left-hand slope of the least concave majorant of an estimator for the primitive of f. This setup includes Grenander's [10] estimator of a monotone density, Brunk's [4] estimator for a monotone regression function, as well as the estimator for a monotone failure rate under random censoring, considered in [17]. We obtain the rate of convergence for the supremum of $|\widehat{f}_n - f|$ over subintervals of [0, 1]. The rate is shown to be of the order $(n/\log n)^{-1/3}$, even on subintervals that grow toward [0, 1], as long as one stays away sufficiently far from the boundaries, so that the inconsistency at the boundaries (see, e.g., [32]) is not going to dominate the supremum. The rate that we obtain coincides with the one suggested in [19] for Grenander's [10] estimator for a decreasing density, but it is now proven rigorously in a more general setting under optimal conditions on the boundaries of the intervals over which $\sup|\widehat{f}_n - f|$ is taken. Moreover, we show that the rate $(n/\log n)^{-1/3}$ is sharp. Our main result is Theorem 2.2, in which we show that a suitably standardized supremum of $|\widehat{f}_n - f|$ converges in distribution to a standard Gumbel random variable.

Our results are obtained following the same sort of approach as that used in [6, 7, 11, 14, 17], among others. We first establish corresponding results for the supremum of the inverses of $\widehat{f}_n$ and f, and then transfer them to the supremum of $\widehat{f}_n$ and f themselves. A major difference with deriving asymptotics of L_k-distances is, that in these cases one can benefit from the linearity of the integral and handle several approximations pointwise with Markov's inequality. This is no longer possible with suprema. With suprema, to transfer results for inverses to results for $\widehat{f}_n$, a key ingredient is a precise uniform bound on the spacings between consecutive jump points of $\widehat{f}_n$.

The paper is organized as follows. In Section 2, we list the assumptions under which our results can be obtained and state our main results concerning the rate of convergence and the limiting distribution of $\sup|\widehat{f}_n - f|$. We also briefly discuss the construction of confidence bands. We formulate corresponding results for the supremum distance between the inverses of $\widehat{f}_n$ and f in Section 3. This is the heart of the proof, which is carried out in Section 4. Finally, in Section 5, we provide a uniform bound on the spacings between consecutive jump points of $\widehat{f}_n$ and then transfer the results obtained in Section 3 for the inverses of $\widehat{f}_n$ and f to the supremum distance between the functions themselves.

To limit the length of the paper, the rigorous proofs of several preliminary results needed for the proofs in Sections 4 and 5 have been put in a supplement [8].

2. Assumptions and main results. Based on $n \geq 2$ independent observations, we aim at estimating a function $f : [0, 1] \to \mathbb{R}$ subject to the constraint that it is nonincreasing. Assume we have at hand a cadlag (right continuous with finite left-hand limits at every point) stepwise estimator F_n of

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

with finitely many jump points. In the case of i.i.d. observations with a common density function f, a typical example is the empirical distribution function with n discontinuity points located at the observations. In the following, we shall consider the monotone estimator $\widehat{f}_n$ of f as defined in [7], that is, the estimator $\widehat{f}_n$ is the left-hand slope of the least concave majorant of F_n with

$$\widehat{f}_n(0) = \lim_{t \downarrow 0} \widehat{f}_n(t).$$

As detailed in Section 2.1 below, this definition generalizes well-known monotone estimators, such as the Grenander estimator of a nonincreasing density, or the least-squares estimator of a monotone regression function. It should be noted that $\widehat{f}_n$ is nonincreasing, left-continuous and piecewise constant. We are interested in the limiting behavior of the supremum distance between the monotone estimator and the function f.

2.1. *Uniform rate of convergence.* We first show that the rate of convergence of $\widehat{f}_n$ to f in terms of the supremum distance is of order $(\log n/n)^{1/3}$. To this end, we make the following assumptions. Unless stated otherwise, for a function h defined on [0, 1], we write $\|h\|_\infty = \sup_{t\in[0,1]} |h(t)|$.

(A1) The function f is decreasing and differentiable on [0, 1] with

$$\inf_{t\in[0,1]} |f'(t)| > 0 \quad \text{and} \quad \sup_{t\in[0,1]} |f'(t)| < \infty.$$

(A2) Let B_n be either a Brownian bridge or a Brownian motion. There exist $q \geq 4$, $C_q > 0$, $L : [0, 1] \to \mathbb{R}$ and versions of F_n and B_n such that

$$\mathbb{P}\big(n^{1-1/q} \|F_n - F - n^{-1/2} B_n \circ L\|_\infty > x\big) \leq C_q x^{-q}$$

for all $x \in (0, n]$. Moreover, L is increasing and differentiable on [0, 1] with $\inf_{t\in[0,1]} L'(t) > 0$ and $\sup_{t\in[0,1]} L'(t) < \infty$.

(A3) There exists $C_0 > 0$ such that for all $x > 0$ and $t = 0, 1$,

$$\mathbb{E}\Big[\sup_{u\in[0,1], x/2\leq|t-u|\leq x} \big(F_n(u) - F(u) - F_n(t) + F(t)\big)^2\Big] \leq \frac{C_0 x}{n}.$$

These conditions are similar to the ones used in [7]. Assumption (A1) is completely the same as the one in [7]. Assumption (A2) is similar to (A4) in [7], but now we only require $q \geq 4$ and bounds on the first derivative of L. Here we can relax the condition on q, because in the current situation the error terms have to be of smaller order than $(n/\log n)^{1/3}$ instead of $n^{1/2}$ in [7]. The existence of L'', as imposed in (A4) in [7], is not needed to establish Theorem 2.1. Finally, assumption (A3) is equal to (A2′) in [7]. Assumption (A2) in [7] is no longer needed, since we are able to obtain sufficient bounds on particular tail probabilities with our current assumptions (A1)–(A2). See Lemma 6.4 and also the proof of Lemma 6.10 in [8].

A typical example that falls into the above framework is the problem of estimating a nonincreasing density f on [0, 1]. Assume we observe i.i.d. random variables $X_1, X_2, \ldots, X_n$ with common nonincreasing density function $f : [0, 1] \to \mathbb{R}$, and let F_n be the corresponding empirical distribution function. In this case, the monotone estimator $\widehat{f}_n$ of f coincides with the Grenander estimator. Assumption (A1) is equal to the ones in [7, 14, 22], and is standard when studying L_k-distances between $\widehat{f}_n$ and f. The existence of a second derivative of f is not needed to obtain Theorem 2.1. In the monotone density model, assumption (A2) is satisfied for all $q > 0$, with $L = F$ being the distribution function corresponding to f and B_n a Brownian bridge, due to the Hungarian embedding of [20]. From Theorem 6 in [7] it follows that assumption (A3) holds in the monotone density model. Another example that falls into the above framework is the problem of estimating a monotone regression function. Assume for instance that we observe $y_i = f(i/n) + \varepsilon_i$, $i = 1, 2, \ldots, n$, where the ε_i's are i.i.d. centered random variables with a finite variance σ^2, and $f : [0,1] \to \mathbb{R}$ is nonincreasing. Let F_n be the partial sum process given by

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} y_i \mathbb{1}_{i \le nt}.$$

In this case, the monotone estimator $\widehat{f}_n$ of f coincides with the Brunk estimator. Assumption (A1) is equal to the ones in [6, 7] and is standard when studying L_k-distances in this model. Assumption (A2) is satisfied for all $q \ge 2$ such that $\mathbb{E}|\varepsilon_i|^q < \infty$ with $L(t) = \sigma^2 t$ and B_n a Brownian motion, due to embedding of [29]. Thus, (A2) is satisfied in the above regression model provided $\mathbb{E}|\varepsilon_i|^4 < \infty$. From Theorem 5 in [7] it follows that assumption (A3) holds in the above regression model. Other examples of statistical models that fall in the above framework, with corresponding q and L, are discussed in [7].

The uniform rate of convergence of $\widehat{f}_n$ to f for general Grenander-type estimators is given in the following theorem.

THEOREM 2.1. *Assume* (A1), (A2) *and* (A3). *Let* $(\alpha_n)_n$ *and* $(\beta_n)_n$ *be sequences of positive numbers such that*

$$\alpha_n \ge K_1 n^{-1/3}(\log n)^{-2/3} \quad \textit{and} \quad \beta_n \ge K_2 n^{-1/3}(\log n)^{-2/3} \tag{1}$$

for some $K_1, K_2 > 0$ *that do not depend on* n. *Then*,

$$\sup_{t\in(\alpha_n, 1-\beta_n]} \left|\widehat{f}_n(t) - f(t)\right| = O_p\left(\frac{\log n}{n}\right)^{1/3}.$$

The rate in Theorem 2.1 coincides with the one found for the maximum likelihood estimator in a semi-parametric model for censored data by Jonker and van der Vaart [19], who suggest that this rate should also hold for Grenander's [10] estimator for a decreasing density. They consider $\alpha_n \gg n^{-1/3}(\log n)^{1/3}$ and β_n constant, which is a slightly stronger assumption than the one in Theorem 2.1. Note that condition (1) in Theorem 2.1 is sharp. If $\alpha_n = n^{-\gamma}$, for some $1/3 < \gamma < 1$, then $n^{(1-\gamma)/2}(\widehat{f}_n(\alpha_n) - f(\alpha_n))$ converges in distribution, according to Theorem 3.1(i) in [23], so that

$$(n/\log n)^{1/3}\left|\widehat{f}_n(\alpha_n) - f(\alpha_n)\right| \to \infty.$$

In fact, for sequences $(\alpha_n)_n$ such that $n^{1/3}(\log n)^{2/3}\alpha_n \to 0$, it can be shown similarly that $(n\alpha_n)^{1/2}\{\widehat{f}_n(\alpha_n) - f(\alpha_n)\}$ converges in distribution, which would yield $(n/\log n)^{1/3}|\widehat{f}_n(\alpha_n) - f(\alpha_n)| \to \infty$.

2.2. *Limiting distribution.* Whereas the previous theorem only provides a bound on the rate of convergence, it is nevertheless crucial for deriving the actual asymptotics of the supremum norm of $\widehat{f}_n - f$ on suitable intervals. For this purpose, we need an additional Hölder assumption on f' and L''.

(A4) The function L in (A2) is twice differentiable and there exist $C_0 > 0$ and $\sigma \in (0, 1]$ such that for all $t, u \in [0, 1]$,

$$|f'(u) - f'(t)| \le C_0|u - t|^\sigma \quad \text{and} \quad |L''(u) - L''(t)| \le C_0|u - t|^\sigma. \tag{2}$$

The condition on L'' in assumption (A4) is a bit stronger than the one in [7]. This is needed to guarantee that the difference between the values of L'' at t and its nearest point of jump of $\widehat{f}_n$ is negligible. The condition on f' in assumption (A4) is the same as (4) in [7], who already observed that the existence of f'', as assumed in [14, 22], is no longer needed. Note that in the monotone density model $L'' = f'$, in which case (A4) reduces to a Hölder condition on f' only. In the monotone regression model, L is linear so that (A4) again reduces to a Hölder condition on f' only.

In order to formulate the limit distribution, we need the following definition:

$$\zeta(c) = \operatorname*{argmax}_{t\in\mathbb{R}}\{W(t + c) - t^2\} \qquad \text{for all } c \in \mathbb{R}, \tag{3}$$

where W is a standard two-sided Brownian motion on $\mathbb{R}$ originating from zero, and argmax denotes the greatest location of the maximum. For fixed $t \in (0, 1)$, properly scaled versions of $n^{1/3}(\widehat{f}_n(t) - f(t))$ converge in distribution to the random variable $\zeta(0)$ (see, e.g., [27] or [11]). Moreover, ζ serves as the limit process for properly scaled versions of $n^{1/3}(\widehat{U}_n - g)$ (see, e.g., Theorem 3.2 in [14]), where $\widehat{U}_n$ and g are the inverse functions of $\widehat{f}_n$ and f respectively, as defined in Section 3 below. Properties of the process $\{\zeta(c), c \in \mathbb{R}\}$ can be found in [12]; for example, the process $\{\zeta(c), c \in \mathbb{R}\}$ is a stationary process. According to Corollary 3.4 in [12], the tails of the density μ of $\zeta(0)$ satisfy the following expansion:

$$\mu(t) \sim 2\lambda|t| \exp(-2|t|^3/3 - \kappa|t|) \tag{4}$$

as $|t| \to \infty$, where κ and λ are positive constants.

We now present the main result of this paper. It states that the limit distribution of the supremum distance between $\widehat{f}_n$ and f, if properly normalized, is Gumbel. By $x_n \gg y_n$ we mean $x_n/y_n \to \infty$, as $n \to \infty$.

THEOREM 2.2. *Assume that* (A1), (A2), (A3) *and* (A4) *hold. Consider* $0 \le u < v \le 1$ *fixed. Then, for any sequence of real numbers* $(\alpha_n)_n$ *and* $(\beta_n)_n$ *both satisfying*

$$\alpha_n \to 0, \qquad \beta_n \to 0 \quad \textit{and} \quad 1 - v + \beta_n, u + \alpha_n \gg n^{-1/3}(\log n)^{-2/3}, \tag{5}$$

we have that for any $x \in \mathbb{R}$,

$$\mathbb{P}\left(\log n\left\{\left(\frac{n}{\log n}\right)^{1/3} \sup_{t\in(u+\alpha_n,v-\beta_n]} \frac{|\widehat{f}_n(t) - f(t)|}{|2f'(t)L'(t)|^{1/3}} - \mu_n\right\} \le x\right) \to \exp\{-\mathrm{e}^{-x}\}$$

as $n \to \infty$, *where*

$$\mu_n = 1 - \frac{\kappa}{2^{1/3}(\log n)^{2/3}} + \frac{1}{\log n}\left[\frac{1}{3}\log\log n + \log(\lambda C_{f,L})\right], \tag{6}$$

with

$$C_{f,L} = 2\int_u^v \left(\frac{|f'(t)|^2}{L'(t)}\right)^{1/3} \mathrm{d}t,$$

and λ *and* κ *taken from* (4).

Note that from Theorem 2.2, with $u = 0$ and $v = 1$, it follows that for convenient α_n and β_n,

$$\left(\frac{n}{\log n}\right)^{1/3} \sup_{t\in(\alpha_n,1-\beta_n]} \frac{|\widehat{f}_n(t) - f(t)|}{|2f'(t)L'(t)|^{1/3}} = 1 + o_p(1).$$

Since both f' and L' are bounded from above and bounded away from zero, this proves that there are positive numbers C_1, C_2 that depend only on f' and L' such that

$$C_1 + o_p(1) \le \left(\frac{n}{\log n}\right)^{1/3} \sup_{t\in(\alpha_n,1-\beta_n]} \left|\widehat{f}_n(t) - f(t)\right| \le C_2 + o_p(1).$$

This means that the rate in Theorem 2.1 is sharp.

2.3. *Confidence bands*. Our main motivation for proving Theorem 2.2 is to build confidence bands for a monotone function f. Indeed, this theorem ensures that for any $x \in \mathbb{R}$, with probability tending to $\exp(-\mathrm{e}^{-x})$, we have

$$\left|\widehat{f}_n(t) - f(t)\right| \le \left(\frac{\log n}{n}\right)^{1/3} \left|2f'(t)L'(t)\right|^{1/3}\left\{\mu_n + \frac{x}{\log n}\right\},$$

simultaneously for all $t \in (u + \alpha_n, v - \beta_n]$. Combining this with either plug-in estimators of f' and L' or bootstrap methods would provide a confidence band for f, at the price of additional technicalities. Indeed, the use of plug-in estimators for the derivatives f' and L' may lead to inaccurate intervals for small sample sizes n, so that bootstrap methods should be preferable. But it is known that the standard bootstrap typically does not work for Grenander-type estimators; see [21, 30]. Thus, we shall use a smoothed bootstrap, which will raise the question of the choice of the smoothing parameter. In view of all this, we believe that the precise construction of a confidence band is beyond the scope of the present paper and is deferred to a separate paper.

Note that the conditions of Theorem 2.2 do not cover the supremum distance over the whole interval [0, 1]. However, this is to be expected. For instance, consider the monotone density model. This model is one of the examples that is covered by our general setup (see Section 2.1) and it is well known that the Grenander estimator $\widehat{f}_n$ in this model is inconsistent at 0 and 1 (e.g., see [32]). Therefore, a distributional result can only be expected if the supremum is taken over subintervals of [0, 1] that do not include 0 and 1. Let us notice, however, that we can obtain a confidence band for f on any sub-interval $(u, v]$ with fixed $u, v \in (0, 1)$ (by considering $\alpha_n = \beta_n = 0$), and that the largest interval on which our result allows to build a confidence band is $(\alpha_n, 1 - \beta_n]$, where $\alpha_n \gg n^{-1/3}(\log n)^{-2/3}$ and similarly, $\beta_n \gg n^{-1/3}(\log n)^{-2/3}$. In order to obtain a confidence band on the whole interval [0, 1], we would have to slightly modify the Grenander-type estimator $\widehat{f}_n$ in order to make it consistent near the boundaries. For instance, we conjecture that, if we consider either the modified estimator in [23] or the penalized estimator in [32] instead of $\widehat{f}_n$, then the limit distribution of the supremum distance between this modified estimator and f over the whole interval [0, 1] is the same as the limit distribution of the supremum distance between $\widehat{f}_n$ and f over the largest interval allowed in Theorem 2.2. Thus, such modified estimators would provide a confidence band for f over the whole interval [0, 1]. As mentioned above, the precise construction of confidence bands is deferred to a separate paper, and we will do similarly with the precise study of modified estimators at the boundaries.

3. The inverse process. To establish Theorems 2.1 and 2.2, we use the same approach as in [6, 7, 11, 14]. We first obtain analogous results (i.e., rate of convergence and limit distribution) for the supremum between the inverses of $\widehat{f}_n$ and f, and then transfer them to the supremum between the functions $\widehat{f}_n$ and f themselves. Let F_n^+ be the upper version of F_n defined as follows: $F_n^+(0) = F_n(0)$ and for every $t \in (0, 1]$,

$$F_n^+(t) = \max\Big\{F_n(t), \lim_{u\uparrow t} F_n(u)\Big\}.$$

Let $\widehat{U}_n$ denote the (generalized) inverse of $\widehat{f}_n$, defined for $a \in \mathbb{R}$ by $\widehat{U}_n(a) = \sup\{t \in [0, 1] : \widehat{f}_n(t) \geq a\}$, with the convention that the supremum of an empty set is zero. This is illustrated in Figure 1 below. From Figure 1, it can be seen that the value $t = U_n(a)$ maximizes $F_n^+(t) - at$, so that

$$\widehat{U}_n(a) = \underset{t\in[0,1]}{\operatorname{argmax}}\{F_n^+(t) - at\}. \tag{7}$$

The advantage of characterizing the inverse process $\widehat{U}_n$ by (7), is that in this way, it is more tractable than the estimator $\widehat{f}_n$ itself, as being the argmax of a relatively

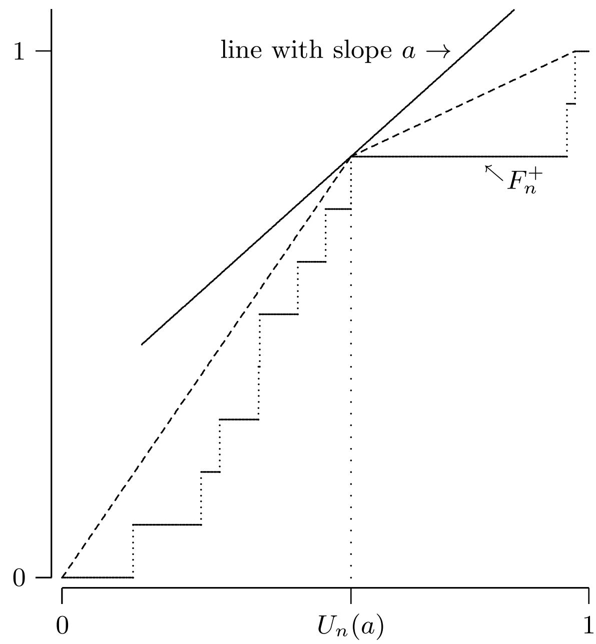

FIG. 1. *The function F_n^+, its concave majorant* (*dashed*) *and a line with slope a* (*solid*).

simple process. It is the purpose of this section to establish results analogous to Theorems 2.1 and 2.2 for the inverse process.

Let g denote the (generalized) inverse function of f. In Theorems 3.1 and 3.2, we give an upper bound for the rate of convergence of $\widehat{U}_n$ to g, and an extremal limit result for the supremum distance between $\widehat{U}_n$ and g. We derive the limit distribution of the supremum distance between $\widehat{U}_n$ and g in Corollary 3.1.

THEOREM 3.1. *Assume that* (A1) *and* (A2) *hold. Then*

$$\sup_{a\in\mathbb{R}}\left|\widehat{U}_n(a) - g(a)\right| = O_p\left(\frac{\log n}{n}\right)^{1/3}.$$

THEOREM 3.2. *Assume that* (A1), (A2) *and* (A4) *hold, and define for $a \in \mathbb{R}$ the normalizing function*

$$A(a) = \frac{|f'(g(a))|^{2/3}}{(4L'(g(a)))^{1/3}}. \tag{8}$$

Let $0 \le u < v \le 1$ fixed, and let $(\alpha_n)_n$ and $(\beta_n)_n$ be sequences such that $\alpha_n \to 0$, $\beta_n \to 0$ and $0 \le u + \alpha_n < v - \beta_n \le 1$ for n sufficiently large. Define

$$S_n = n^{1/3} \sup_{a\in[f(v-\beta_n), f(u+\alpha_n)]} A(a)\left|\widehat{U}_n(a) - g(a)\right|. \tag{9}$$

Then

$$\mathbb{P}(S_n \le u_n) \to \exp\left\{-2\tau \int_u^v \frac{|f'(t)|^{2/3}}{(4L'(t))^{1/3}}\,\mathrm{d}t\right\} \tag{10}$$

for any sequence $(u_n)_n$ *such that* $u_n \to \infty$ *in such a way that* $n^{1/3}\mu(u_n) \to \tau > 0$, *where* μ *denotes the density of* $\zeta(0)$, *as defined in* (3).

The expansion in (4) allows us to provide a precise expansion of u_n [see (34)] and to derive the following corollary from Theorem 3.2. According to this corollary, the limit distribution of S_n is Gumbel.

COROLLARY 3.1. *Assume that* (A1), (A2) *and* (A4) *hold. Let* S_n *be defined by* (9), *with* $0 \le u < v \le 1$, *and* α_n, β_n *satisfying the conditions of Theorem* 3.2. *Then, for all* $x \in \mathbb{R}$,

$$\mathbb{P}\left\{\log n\left\{\left(\frac{2}{\log n}\right)^{1/3} S_n - \mu_n\right\} \le x\right\} \to \exp\{-\mathrm{e}^{-x}\},$$

where μ_n *is defined by* (6).

In order to transfer the results for $\widehat{U}_n$ to $\widehat{f}_n$, we establish Lemma 5.2. This lemma does require conditions on sequences $s_n = u + \alpha_n$ and $t_n = 1 - v + \beta_n$ that are stronger than the ones in Theorem 2.2. However, once we have established the limit distribution for such sequences, we will show that Theorem 2.2 can be extended to more general sequences satisfying (5).

4. Proofs of Theorems 3.1 and 3.2 and Corollary 3.1. We suppose in the sequel that assumptions (A1) and (A2) are fulfilled, and we denote by C, C_1, $C_2, \ldots$ positive real numbers that depend only on q, C_q, f, L [and possibly on σ under the additional assumption (A4)]. These real numbers may change from one line to the other. We write $x \vee y = \max(x, y)$ and $x \wedge y = \min(x, y)$, for any real numbers x and y.

In order to deal simultaneously with the cases where B_n is a Brownian bridge or a Brownian motion [see assumption (A2)], we shall make use of the representation

$$B_n(t) = W_n(t) - \xi_n t, \qquad t \in [0, 1], \tag{11}$$

where W_n is a standard Brownian motion, $\xi_n \equiv 0$ if B_n is a Brownian motion and $\xi_n \equiv W_n(1)$, a standard Gaussian variable that is independent of B_n, in case B_n is a Brownian bridge. To prove Theorem 3.1, we need some preliminary results on the tail probabilities of $\widehat{U}_n - g$ and its supremum. These results can be found in Supplement B in [8]. A first result, which is similar to Lemmas 2, 3 and 4 in [7], is that there exist $C_1 > 0$ and $C_2 > 0$ such that for all $a \in \mathbb{R}$ and $x > 0$,

$$\mathbb{P}\left(n^{1/3}\left|\widehat{U}_n(a) - g(a)\right| > x\right) \le \frac{C_1 n^{1-q/3}}{x^{2q}} + 2\exp(-C_2 x^3). \tag{12}$$

In particular, for all $a \in \mathbb{R}$, this implies that $\widehat{U}_n(a) - g(a) = O_p(n^{-1/3})$. See Lemma 6.4 in [8]. This is not sufficient to obtain Theorem 3.1, but it will be used for its proof.

PROOF OF THEOREM 3.1. Recall that $g(a) = 1$ for all $a \le f(1)$, $g(a) = 0$ for $a \ge f(0)$ and $\widehat{U}_n$ is nonincreasing and takes values in [0, 1]. Hence, we can write

$$\sup_{a \le f(1)} \left|\widehat{U}_n(a) - g(a)\right| = \left|\widehat{U}_n\big(f(1)\big) - g\big(f(1)\big)\right| \tag{13}$$

and

$$\sup_{a \ge f(0)} \left|\widehat{U}_n(a) - g(a)\right| = \left|\widehat{U}_n\big(f(0)\big) - g\big(f(0)\big)\right|. \tag{14}$$

This means that

$$\sup_{a \in \mathbb{R}} \left|\widehat{U}_n(a) - g(a)\right| = \sup_{a \in [f(1), f(0)]} \left|\widehat{U}_n(a) - g(a)\right|.$$

Therefore, to prove Theorem 3.1 it suffices to show that

$$\sup_{a \in [f(1), f(0)]} \left|\widehat{U}_n(a) - g(a)\right| = O_p\left(\frac{\log n}{n}\right)^{1/3}.$$

According to Lemma 6.5 in [8], the bound in (12) can be extended such that for any $x > 0$,

$$\mathbb{P}\left(\sup_{a \in [f(1), f(0)]} \left|\widehat{U}_n(a) - g(a)\right| > x\left(\frac{\log n}{n}\right)^{1/3}\right)$$
$$\le \widetilde{C}_3 n^{1/3}\left(\frac{C_1 n^{1-q/3}}{x^{2q}(\log n)^{2q/3}} + 2n^{-C_2 x^3}\right),$$

where $\widetilde{C}_3 = C_3\{f(1) - f(0)\}$. The latter upper bound tends to zero as $n \to \infty$ for all $x > (3C_2)^{-1/3}$ since $q \ge 4$ by assumption. This completes the proof of Theorem 3.1. □

We suppose in the sequel that in addition to (A1) and (A2), assumption (A4) is fulfilled. The first step in proving Theorem 3.2 is to approximate an adequately normalized version of $\widehat{U}_n(a)$ by the location of the maximum of a Brownian motion with parabolic drift. To this end define

$$V_n(a) = n^{1/3}\big(L\big(\widehat{U}_n(a^{\xi})\big) - L\big(g(a)\big)\big), \tag{15}$$

where

$$a^{\xi} = a - n^{-1/2}\xi_n L'\big(g(a)\big) \qquad \text{for all } a \in \mathbb{R}, \tag{16}$$

with ξ_n taken from representation (11). Then for $0 \le u < v \le 1$ and α_n, β_n satisfying the conditions of Theorem 3.2, we obtain

$$S_n \vee O_p(1) = \sup_{a \in [f(v-\beta_n), f(u+\alpha_n)]} \frac{A(a)}{L'(g(a))} \bigl|V_n(a)\bigr| \vee O_p(1) \\ + O_p\bigl(n^{-\sigma/2}(\log n)^{2/3}\bigr) + O_p\bigl(n^{-1/6}\bigr),$$

where S_n is defined by (9), and $\sigma \in (0, 1]$ is taken from (A4). See Lemma 6.6 in [8].

Next, we proceed with localization. The purpose of this is that localized versions of $V_n(a)$ and $V_n(b)$, can be approximated by independent random variables, if a and b are in disjoint intervals that are suitably separated. First note that the location of the maximum of a process is invariant under addition of constants or multiplication by $n^{2/3}$. Therefore, from (7) it follows that for all $a \in \mathbb{R}$ we have

$$V_n(a) = \operatorname*{argmax}_{t \in I_n(a)} \bigl\{W_{g(a)}(t) + D_n(a,t) + R_n(a,t)\bigr\}, \tag{17}$$

where

$$I_n(a) = \bigl[n^{1/3}\bigl(L(0) - L\bigl(g(a)\bigr)\bigr), n^{1/3}\bigl(L(1) - L\bigl(g(a)\bigr)\bigr)\bigr] \tag{18}$$

for every $s \in [0, 1]$ fixed, W_s is the standard Brownian motion defined by

$$W_s(t) = n^{1/6}\bigl\{W_n\bigl(L(s) + n^{-1/3}t\bigr) - W_n\bigl(L(s)\bigr)\bigr\} \qquad \text{for } t \in \mathbb{R}, \tag{19}$$

with W_n defined by (11), and

$$\begin{aligned} D_n(a,t) &= n^{2/3}\bigl(F \circ L^{-1} - aL^{-1}\bigr)\bigl(L\bigl(g(a)\bigr) + n^{-1/3}t\bigr) \\ &\quad - n^{2/3}\bigl(F\bigl(g(a)\bigr) - ag(a)\bigr), \\ R_n(a,t) &= n^{2/3}\bigl(a - a^{\xi}\bigr)\bigl(L^{-1}\bigl(L\bigl(g(a)\bigr) + n^{-1/3}t\bigr) - g(a)\bigr) \\ &\quad - n^{-1/6}\xi_n t + \tilde{R}_n(a,t), \end{aligned} \tag{20}$$

where ξ_n is taken from representation (11), and for all a and t,

$$\bigl|\tilde{R}_n(a,t)\bigr| \le n^{2/3}\bigl\|F_n - F - n^{-1/2}B_n \circ L\bigr\|_\infty. \tag{21}$$

For all $a \in \mathbb{R}$, we define the localized version of $V_n(a)$ by

$$\tilde{V}_n(a) = \operatorname*{argmax}_{t \in I_n(a):|t| \le \log n} \bigl\{W_{g(a)}(t) + D_n(a,t) + R_n(a,t)\bigr\}. \tag{22}$$

We find that

$$\sup_{a \in [f(v-\beta_n), f(u+\alpha_n)]} \frac{A(a)}{L'(g(a))}\bigl|V_n(a)\bigr| \\ = \sup_{a \in [f(v-\beta_n), f(u+\alpha_n)]} \frac{A(b(a))}{L'(g(b(a)))}\bigl|\tilde{V}_n(a)\bigr| + o_p(\log n)^{-2/3}$$

for any $b(a) \in \mathbb{R}$ that satisfies $|a - b(a)| \le n^{-1/3}(\log n)^2$. See Lemma 6.7 in [8].

Finally, using the fact that, roughly speaking,

$$D_n(a,t) \approx -\frac{|f'(g(a))|}{2(L'(g(a)))^2}t^2 \approx -\frac{|f'(g(b))|}{2(L'(g(b)))^2}t^2$$

for all b close enough to a, we bound $|\tilde{V}_n(a)|$ from above and below by the absolute value of the following quantities:

$$\tilde{V}_n^+(a,b) = \operatorname*{argmax}_{t \in I_n(a):|t| \le \log n} \left\{ W_{g(a)}(t) - \left(\frac{|f'(g(b))|}{2(L'(g(b)))^2} - 2\epsilon_n \right) t^2 \right\} \tag{23}$$

and

$$\tilde{V}_n^-(a,b) = \operatorname*{argmax}_{t \in I_n(a):|t| \le \log n} \left\{ W_{g(a)}(t) - \left(\frac{|f'(g(b))|}{2(L'(g(b)))^2} + 2\epsilon_n \right) t^2 \right\}, \tag{24}$$

where $I_n(a)$ and $W_{g(a)}$ are defined in (18) and (19), b is chosen sufficiently close to a, and where $(\epsilon_n)_n$ is a sequence of positive numbers that converges to zero as $n \to \infty$, which is to be chosen suitably. The purpose of this is that when we will vary a over a small interval and fix b to be the midpoint of this interval, we will obtain variables $\tilde{V}_n^+(a,b)$ that are defined with the same drift,

$$-\left(\frac{|f'(g(b))|}{2(L'(g(b)))^2} - 2\epsilon_n \right) t^2,$$

and the Browian motion $W_{g(a)}$ only depending on a. The case of $\tilde{V}_n^-(a,b)$ is similar.

For $0 \le u < v \le 1$, and α_n, β_n satisfying the conditions of Theorem 3.2, we obtain

$$S_n \le \sup_{a \in [f(v-\beta_n), f(u+\alpha_n)]} \frac{A(b(a))}{L'(g(b(a)))} \left|\tilde{V}_n^+(a, b(a))\right| \vee O_p(1) + o_p(\log n)^{-2/3},$$

and

$$S_n \vee O_p(1) \ge \sup_{a \in [f(v-\beta_n), f(u+\alpha_n)]} \frac{A(b(a))}{L'(g(b(a)))} \left|\tilde{V}_n^-(a, b(a))\right| + o_p(\log n)^{-2/3}$$

for any $b(a) \in \mathbb{R}$ that satisfies $|a - b(a)| \le n^{-1/3}(\log n)^2$, where S_n is defined by (9) and $\epsilon_n = 1/\log n$ in (24) and (23). See Lemma 6.8 in [8].

Note that in order to obtain the above approximations, we use the following lemma, which is a variation on Lemma 2.1 in [23]. Although very simple, it turns out to be a very useful tool to compare locations of maxima.

LEMMA 4.1. *Let $I \subset \mathbb{R}$ be an interval. Let g and Z be real valued functions defined on I such that there exists $\gamma > 0$ with*

$$g(u) < g(v) \qquad \text{for all } u, v \text{ such that } |u| > |v| + \gamma.$$

Assume that both $\sup_{u\in I} Z(u)$ *and* $\sup_{u\in I}\{Z(u)+g(u)\}$ *are achieved. Denoting by* argmax *an arbitrary point where the maximum is achieved*, *we have*

$$\Big|\operatorname*{argmax}_{u\in I}\{Z(u)+g(u)\}\Big| \le \Big|\operatorname*{argmax}_{u\in I}\{Z(u)\}\Big| + \gamma.$$

PROOF. Suppose the maximum of Z is achieved at $v\in I$, so that $Z(u)\le Z(v)$ for all $u\in I$. It is assumed that for all $u\in I$ such that $|u|>|v|+\gamma$, we have $g(u)<g(v)$. Therefore,

$$Z(u)+g(u) < Z(v)+g(v)$$

for all $u\in I$ such that $|u|>|v|+\gamma$. It follows that the maximum of $Z+g$ cannot be achieved at such a point u, which means that

$$\Big|\operatorname*{argmax}_{u\in I}\{Z(u)+g(u)\}\Big| \le |v|+\gamma.$$

This completes the proof by definition of v. $\square$

To relate the suprema of $\tilde V_n^+$ and $\tilde V_n^-$ with maxima of independent random variables, we will partition the interval $[f(v-\beta_n), f(u+\alpha_n)]$ into a union of disjoint intervals A_i and B_i of alternating length, and a remainder interval R_n, in such a way that the length of the small blocks A_i is

$$l_n = \frac{2\|f'\|_\infty}{\inf_{t\in[0,1]} L'(t)} n^{-1/3}\log n, \tag{25}$$

and the length of the big blocks B_i is $L_n = 2n^{-1/3}(\log n)^2$. More precisely, for $i=1,2,\dots,K_n$, where

$$K_n = \left[\frac{f(u+\alpha_n)-f(v-\beta_n)}{l_n+L_n}\right]-1, \tag{26}$$

let

$$\begin{aligned} A_i &= \big[f(v-\beta_n)+(i-1)(l_n+L_n), f(v-\beta_n)+il_n+(i-1)L_n\big],\\ B_i &= \big[f(v-\beta_n)+il_n+(i-1)L_n, f(v-\beta_n)+i(l_n+L_n)\big], \end{aligned} \tag{27}$$

and let $R_n = [f(v-\beta_n)+K_n(l_n+L_n), f(u+\alpha_n)]$, so that $l_n+L_n\le |R_n| < 2(l_n+L_n)$ and

$$\big[f(v-\beta_n), f(u+\alpha_n)\big] = \left(\bigcup_{i=1}^{K_n} A_i\right)\cup\left(\bigcup_{i=1}^{K_n} B_i\right)\cup R_n. \tag{28}$$

Now, suppose that $0\le u<v\le 1$, and α_n,β_n satisfy the conditions of Theorem 3.2 and let $(\zeta_i)_{i\in\mathbb{N}}$ be a sequence of independent processes, all distributed like ζ given

in (3). Then, using scaling properties of the Brownian motion, we can build (possibly dependent) copies $(\zeta_j^{(1)})$, $(\zeta_j^{(2)})$ of $(\zeta_i)_{i\in\mathbb{N}}$, such that

$$S_B \le \frac{S_n \vee O_p(1)}{1+O(1/\log n)} \le S_B^{(1)} \vee S_A^{(2)} \vee o_p(\log n)^{1/3} + o_p(\log n)^{-2/3}, \tag{29}$$

where

$$S_B \stackrel{d}{=} \max_{1\le i\le K_n} \sup_{c\in[0,\Delta_{in}]} |\zeta_i(c)| \quad \text{and} \quad S_B^{(1)} \stackrel{d}{=} \max_{1\le i\le K_n} \sup_{c\in[0,\Delta_{in}]} |\zeta_i^{(1)}(c)|,$$

$$S_A^{(2)} \stackrel{d}{=} \max_{2\le i\le K_n} \sup_{c\in[0,\delta_{in}]} |\zeta_i^{(2)}(c)|,$$

with K_n defined in (26) and where uniformly in i,

$$\Delta_{in} = (1+o(1))(\log n)^2 \left|\frac{L'(g(b_i))f'(g(b_i))}{2}\right|^{-1/3},$$

and $0 \le \delta_{in} \le C\log n$, for some $C > 0$, where b_i denotes the midpoint of the interval B_i defined in (27). See Lemma 6.9 in [8] for a rigorous proof of (29). The fact that $i \ge 2$ in the definition of $S_A^{(2)}$ is due to the fact that the first small block A_1 has to be treated separately.

At this stage, we need a precise control of the tail probabilities of the supremum of the limiting process ζ over increasing intervals. Specifically, in Supplement A of [8], we obtain the following slight variation on Theorem 1.1 in [16]. Suppose $\delta_n \to \infty$, $\tau_n \to 0$ and $u_n \to \infty$, in such way that $u_n/\delta_n \to 0$, $\delta_n\mu(u_n)/\tau_n \to 1$, and $\log(\tau_n)/\delta_n^3 \to 0$. Then

$$\left|\frac{\log \mathbb{P}(\sup_{c\in[0,\delta_n]} |\zeta(c)| \le u_n)}{-2\tau_n} - 1\right| \to 0. \tag{30}$$

See Lemma 6.3 in [8] for a rigorous proof.

We are then in the position to establish Theorem 3.2 and Corollary 3.1.

PROOF OF THEOREM 3.2. Let $(u_n)_n$ be a sequence such that $u_n \to \infty$ in such a way that

$$n^{1/3}\mu(u_n) \to \tau > 0, \tag{31}$$

where μ is the density of $\zeta(0)$. We will bound $\mathbb{P}(S_n \le u_n)$, where S_n is defined by (9), from above and below by means of (29). Write

$$S_1 = \max_{1\le i\le K_n} \sup_{c\in[0,\Delta_{in}]} |\zeta_i^{(1)}(c)|,$$

$$S_2 = \max_{2\le i\le K_n} \sup_{c\in[0,\delta_{in}]} |\zeta_i^{(2)}(c)|.$$

Then, according to (29)

$$\mathbb{P}(S_n \le u_n) \ge \mathbb{P}\big(\big(1 + O(1/\log n)\big)\{S_1 \vee S_2 \vee Q_n\} + R_n \le u_n\big),$$

where $Q_n = o_p(\log n)^{1/3}$ and $R_n = o_p(\log n)^{-2/3}$. Define the event $E_n = \{(\log n)^{2/3}|R_n| \le 1\}$, then $\mathbb{P}(E_n^c) \to 0$, so that

$$\begin{aligned}\mathbb{P}(S_n \le u_n) &\ge \mathbb{P}(S_1 \vee S_2 \vee Q_n \le v_n) + o(1)\\ &= \mathbb{P}(S_1 \le v_n, S_2 \le v_n, Q_n \le v_n) + o(1),\end{aligned}$$

where

$$v_n = \frac{u_n - (\log n)^{-2/3}}{1 + O((\log n)^{-1})} \sim u_n - (\log n)^{-2/3} \qquad \text{as } n \to \infty.$$

From (4) and (31), it is easily verified that u_n is of order $(\log n)^{1/3}$ [see also the expansion (34) below] and that

$$n^{1/3}\mu(v_n) \to \tau. \tag{32}$$

Therefore, since $\mathbb{P}(Q_n \le v_n) \to 1$, we have

$$\mathbb{P}(S_n \le u_n) \ge \mathbb{P}(S_1 \le v_n, S_2 \le v_n) + o(1).$$

We will investigate $\mathbb{P}(S_1 \le v_n)$ and $\mathbb{P}(S_2 \le v_n)$ separately.

Since the processes $\zeta_i^{(1)}$ are independent copies of ζ,

$$\mathbb{P}(S_1 \le v_n) = \prod_{i=1}^{K_n} \mathbb{P}\Big(\sup_{c\in[0,\Delta_{in}]} |\zeta(c)| \le v_n\Big).$$

For each $i = 1, 2, \dots, K_n$ fixed, we apply (30), with

$$\Delta_{in} = \big(1 + o(1)\big)(\log n)^2 \left|\frac{L'(g(b_i)) f'(g(b_i))}{2}\right|^{-1/3},$$

which is of the order $(\log n)^2$ uniformly in i, and $\tau_{in} = \tau\Delta_{in} n^{-1/3}$, where the b_i are the midpoints of the K_n big blocks B_i. The b_i are equidistant at distance $l_n + L_n = 2n^{-1/3}(\log n)^2(1 + O(\log n)^{-1})$. Since $\tau_{in} \to 0$ uniformly in i and v_n is of order $(\log n)^{1/3}$, we conclude that

$$\begin{aligned}&\prod_{i=1}^{K_n} \mathbb{P}\Big(\sup_{c\in[0,\Delta_{in}]} |\zeta(c)| \le v_n\Big)\\ &\qquad = \prod_{i=1}^{K_n} \exp\big(-2\tau_{in}\big(1 + o(1)\big)\big),\end{aligned}$$

where the small o-term is uniform in i. Therefore,

$$\begin{aligned}
&\prod_{i=1}^{K_n} \mathbb{P}\Big(\sup_{c\in[0,\Delta_{in}]} \big|\zeta(c)\big| \le v_n\Big) \\
&\qquad = \exp\left\{-2\big(1+o(1)\big)\tau \sum_{i=1}^{K_n} \frac{2n^{-1/3}(\log n)^2}{|4L'(g(b_i))f'(g(b_i))|^{1/3}}\right\} \\
&\qquad = \exp\left\{-2\tau \int_{f(v)}^{f(u)} \frac{1}{|4L'(g(b))f'(g(b))|^{1/3}}\,\mathrm{d}b\right\} + o(1) \\
&\qquad = \exp\left\{-2\tau \int_u^v \frac{|f'(t)|^{2/3}}{(4L'(t))^{1/3}}\,\mathrm{d}t\right\} + o(1).
\end{aligned}$$

It follows that

$$\mathbb{P}(S_1 \le v_n) \to \exp\left\{-2\tau \int_u^v \frac{|f'(t)|^{2/3}}{(4L'(t))^{1/3}}\,\mathrm{d}t\right\}.$$

The probability $\mathbb{P}(S_2 \le v_n)$ can be treated in the same way:

$$\begin{aligned}
\mathbb{P}(S_2 \le v_n) &= \prod_{i=1}^{K_n} \mathbb{P}\Big(\sup_{c\in[0,\delta_{in}]} \big|\zeta(c)\big| \le v_n\Big) \\
&= \exp\left\{-2\big(1+o(1)\big)\tau \sum_{i=1}^{K_n} \delta_{in} n^{-1/3}\right\} \to 1,
\end{aligned}$$

since, according to (29) and (26),

$$\sum_{i=1}^{K_n} \delta_{in} n^{1/3} \le C n^{-1/3} K_n \log n = O(\log n)^{-1}.$$

This yields that

$$\liminf_{n\to\infty} \mathbb{P}(S_n \le u_n) \ge \exp\left\{-2\tau \int_u^v \frac{|f'(t)|^{2/3}}{(4L'(t))^{1/3}}\,\mathrm{d}t\right\}.$$

Similarly, with (29),

$$\mathbb{P}(S_n \le u_n) \le \mathbb{P}\Big(\max_{1\le i\le K_n} \sup_{c\in[0,\Delta_{in}]} \big|\zeta_i(c)\big| \le v_n\Big) + o(1),$$

where v_n satisfies (32). This probability can be treated completely similar to $\mathbb{P}(S_1 \le v_n)$, so that

$$\limsup_{n\to\infty} \mathbb{P}(S_n \le u_n) \le \exp\left\{-2\tau \int_u^v \frac{|f'(t)|^{2/3}}{(4L'(t))^{1/3}}\,\mathrm{d}t\right\}.$$

This proves the theorem. $\square$

PROOF OF COROLLARY 3.1. Let $(u_n)_n$ be a sequence such that $u_n \to \infty$ in such a way that $n^{1/3}\mu(u_n) \to \tau > 0$, as $n \to \infty$. Taking logarithms in (4), we conclude that $(u_n)_n$ should satisfy

$$\frac{1}{3}\log n + \log u_n - \frac{2}{3}u_n^3 - \kappa u_n = \log\frac{\tau}{2\lambda} + o(1) \qquad \text{as } n \to \infty. \tag{33}$$

This means that $-2u_n^3/3$ is the dominating term, which should compensate $(\log n)/3$. Therefore, if we write $u_n = 2^{-1/3}(\log n)^{1/3} + \delta_n$, where $\delta_n = o(\log n)^{1/3}$, and insert this in (33), we obtain

$$\begin{aligned}
&\frac{1}{3}\log n + \log\left\{\left(\frac{\log n}{2}\right)^{1/3} + \delta_n\right\} \\
&\quad - \frac{2}{3}\left\{\frac{\log n}{2} + 3\left(\frac{\log n}{2}\right)^{2/3}\delta_n + 3\left(\frac{\log n}{2}\right)^{1/3}\delta_n^2 + \delta_n^3\right\} - \kappa\left(\frac{\log n}{2}\right)^{1/3} - \kappa\delta_n \\
&\quad = \log\frac{\tau}{2\lambda} + o(1).
\end{aligned}$$

Tedious, but straightforward computations first yield that $\delta_n \to 0$ and then that

$$\begin{aligned}
\delta_n = &-\frac{\kappa}{4^{1/3}}(\log n)^{-1/3} + \frac{4^{1/3}}{6}(\log n)^{-2/3}\log\log n \\
&- (\log n)^{-2/3}\left[\frac{\log\tau}{2^{1/3}} - \frac{\log(2\lambda)}{2^{1/3}} + \frac{4^{1/3}}{6}\log 2\right] + o(\log n)^{-2/3}.
\end{aligned}$$

If we put $\tau 4^{-1/3}C_{f,L} = \mathrm{e}^{-x}$, or $-\log\tau = x + \log C_{f,L} - (2\log 2)/3$, this implies that

$$\begin{aligned}
u_n = &\frac{1}{2^{1/3}}(\log n)^{1/3} - \frac{\kappa}{4^{1/3}}(\log n)^{-1/3} + \frac{4^{1/3}}{6}(\log n)^{-2/3}\log\log n \\
&+ (\log n)^{-2/3}\left(\frac{x + \log C_{f,L}}{2^{1/3}} + \frac{\log\lambda}{2^{1/3}}\right) + o(\log n)^{-2/3}.
\end{aligned} \tag{34}$$

If we also write $u_n = x/a_n + b_n + o(\log n)^{-2/3}$, with

$$\begin{aligned}
a_n &= 2^{1/3}(\log n)^{2/3}, \\
b_n &= \frac{(\log n)^{1/3}}{2^{1/3}} - \frac{\kappa}{(4\log n)^{1/3}} + \frac{4^{1/3}\log\log n}{6(\log n)^{2/3}} + \frac{\log(\lambda C_{f,L})}{2^{1/3}(\log n)^{2/3}},
\end{aligned}$$

then (10) is equivalent to $\mathbb{P}\{a_n(S_n - b_n) + o(1) \le x\} \to \exp\{-\mathrm{e}^{-x}\}$. Finally, it is easy to see that

$$\begin{aligned}
a_n(S_n - b_n) &= \log n\left\{\left(\frac{2}{\log n}\right)^{1/3} S_n - \left(\frac{2}{\log n}\right)^{1/3} b_n\right\} \\
&= \log n\left\{\left(\frac{2}{\log n}\right)^{1/3} S_n - \mu_n\right\}.
\end{aligned}$$

This proves Corollary 3.1. $\square$

5. Proof of Theorems 2.1 and 2.2. We suppose in the sequel that assumptions (A1), (A2) and (A3) are fulfilled. As before, C, C_1, $C_2, \dots$ denote positive real numbers that depend only on q, C_q, f, L, C_0, and possibly also on σ under the additional assumption (A4). It follows from the definition of $\widehat{f}_n$ that it can be discontinuous only at the jump points of F_n. In particular, the number of jump points of $\widehat{f}_n$ is finite. In the sequel, we will denote this number by $N_n - 1$ (note that $N_n \geq 1$). Moreover, we set $\tau_0 = 0$, $\tau_{N_n} = 1$, and in the case where $\widehat{f}_n$ has at least one jump point, that is, $N_n \geq 2$, we denote by $\tau_1 < \cdots < \tau_{N_n - 1}$ the ordered jump points of $\widehat{f}_n$.

To prove Theorems 2.1 and 2.2, we need a precise uniform bound on the spacings between consecutive jump points of $\widehat{f}_n$. This is given by the following lemma.

LEMMA 5.1. *Assume* (A1) *and* (A2). *Then*

$$\max_{1 \leq i \leq N_n} |\tau_i - \tau_{i-1}| = O_p\left(\frac{\log n}{n}\right)^{1/3}. \tag{35}$$

PROOF. It follows from the definition of $\widehat{f}_n$ and $\widehat{U}_n$ that these functions are nonincreasing left-continuous step functions with finitely many jump points, and that the maximal length of the flat parts of $\widehat{f}_n$ is precisely the maximal height of the jumps of $\widehat{U}_n$. Therefore,

$$\max_{1 \leq i \leq N_n} |\tau_i - \tau_{i-1}| = \sup_{a \in \mathbb{R}} \Big| \lim_{b \downarrow a} \widehat{U}_n(b) - \widehat{U}_n(a) \Big|.$$

Using the triangle inequality, it follows that

$$\max_{1 \leq i \leq N_n} |\tau_i - \tau_{i-1}| \leq \sup_{a \in \mathbb{R}} \Big\{ \Big| \lim_{b \downarrow a} \widehat{U}_n(b) - g(a) \Big| + \big| \widehat{U}_n(a) - g(a) \big| \Big\}.$$

But g is continuous on $\mathbb{R}$, so that Theorem 3.1 implies that

$$\max_{1 \leq i \leq N_n} |\tau_i - \tau_{i-1}| \leq 2 \sup_{a \in \mathbb{R}} \big| \widehat{U}_n(a) - g(a) \big| = O_p\left(\frac{\log n}{n}\right)^{1/3},$$

which completes the proof. $\square$

REMARK 5.1. Lemma 5.1 together with the identity $1 = \sum_{i=1}^{N_n} (\tau_i - \tau_{i-1})$, implies that $1/N_n = O_p(n^{-1/3} (\log n)^{1/3})$. This gives some idea about the order of magnitude of the number of jumps of $\widehat{f}_n$. Further investigation is needed to obtain a sharp upper bound, and we conjecture that it is of order $n^{1/3}$. This rate is also claimed in Theorem 3.1 in [13].

We will also need a bound on the mean absolute error between $\widehat{f}_n$ and f. In Supplement C in [8], we reprove Theorem 1 in [7] under slightly weaker assumptions; that is, there exists $C > 0$ such that

$$\mathbb{E} \big| \widehat{f}_n(t) - f(t) \big| \leq C n^{-1/3} \tag{36}$$

for all $t \in [n^{-1/3}, 1-n^{-1/3}]$ and

$$\mathbb{E}\big|\widehat{f}_n(t)-f(t)\big| \le C\big[n\big(t\wedge(1-t)\big)\big]^{-1/2} \tag{37}$$

for all $t \in (0, n^{-1/3}] \cup [1-n^{-1/3}, 1)$. See Lemma 6.10 in [8].

Note that the number of jump points of $\widehat{U}_n$ is precisely the number of flat parts of $\widehat{f}_n$, that is N_n, and denoting by $\gamma_1 > \cdots > \gamma_{N_n}$ the jump points of $\widehat{U}_n$, we have

$$\gamma_i = \widehat{f}_n(\tau_i) \quad \text{and} \quad \tau_i = \widehat{U}_n(\gamma_i) \qquad \text{for } i = 1, 2, \dots, N_n. \tag{38}$$

We will show that in order to study the supremum of $|\widehat{f}_n - f|$ over an interval, we can restrict ourselves to the situation where the boundaries of the interval are jump points of $\widehat{f}_n$ and where the values of $\widehat{f}_n$ stay in $(f(1), f(0))$. Indeed, in order to relate the supremum of $|\widehat{f}_n - f|$ to the supremum of $|U_n - g|$, we need to employ the identity $\gamma_i = f(g(\gamma_i))$, for $\gamma_i = \widehat{f}_n(\tau_i)$, so we need to make sure that $\widehat{f}_n(\tau_i) \in (f(1), f(0))$. To this end, define for any $t \in (0,1)$

$$i_1(t) = \min\big\{i \in \{1, 2, \dots, N_n\} \text{ such that } \tau_i \ge t\big\} \tag{39}$$

and

$$i_2(t) = \max\big\{i \in \{0, 1, \dots, N_n - 1\} \text{ such that } \tau_i < 1-t\big\}. \tag{40}$$

For any t such that $n^{1/3}t \to \infty$ and $n^{1/3}(1-t) \to \infty$, we establish the order of the difference with neighboring points of jump of $\widehat{f}_n$, that is,

$$\tau_i = t + O_p\big(n^{-1/3}\big) \tag{41}$$

for $i = i_1(t)-1, i_1(t), i_1(t)+1$, and similarly for $1-t$,

$$\tau_i = 1-t + O_p\big(n^{-1/3}\big) \tag{42}$$

for $i = i_2(t)-1, i_2(t), i_2(t)+1$. See Lemma 6.11 in [8]. Note that if there are no jumps on the interval $[s, 1-t)$, then $\tau_{i_1(s)} > \tau_{i_2(t)}$. This may happen if the length $1-t-s$ of the interval tends to zero too fast. However, if

$$n^{1/3}s \to \infty, \qquad n^{1/3}t \to \infty \tag{43}$$

and

$$n^{1/3}(1-t-s) \to \infty, \tag{44}$$

then

$$\mathbb{P}(s \le \tau_{i_1(s)} \le \tau_{i_2(t)} < 1-t) \to 1. \tag{45}$$

See Lemma 6.12 in [8]. According to Lemma 6.13 in [8],

$$\begin{aligned} &\mathbb{P}\big(\gamma_i < f(0) \text{ for all } i \ge i_1(s)\big) \to 1,\\ &\mathbb{P}\big(\gamma_i > f(1) \text{ for all } i \le i_2(t)\big) \to 1, \end{aligned} \tag{46}$$

whenever (44) holds, which ensures that $\widehat{f}_n(\tau_i) \in (f(1), f(0))$ simultaneously for various i's, with probability tending to one.

We are then in the position to prove Theorem 2.1.

PROOF OF THEOREM 2.1. First, we establish the result for sequences $\alpha_n = s_n$ and $\beta_n = t_n$ that satisfy (43) and (44). For the sake of brevity, write $i_1 = i_1(s_n)$ and $i_2 = i_2(t_n)$. Define the event

$$\begin{aligned} E_n &= \{s_n \le \tau_{i_1} \le \tau_{i_2} < 1 - t_n\} \\ &\quad \cap \{\gamma_i \in (f(1), f(0)) \text{ for all } i = i_1, \dots, i_2\}. \end{aligned} \tag{47}$$

Then according to (45) and (46), we have $\mathbb{P}(E_n) \to 1$, so we can restrict ourselves to the event E_n. We have

$$\begin{aligned} &\sup_{u \in (s_n, 1-t_n]} |\widehat{f}_n(u) - f(u)| \\ &\qquad \le \max_{i=i_1,\dots,i_2} \sup_{u \in (\tau_{i-1}, \tau_i]} |\widehat{f}_n(u) - f(u)| + \sup_{u \in (\tau_{i_2}, 1-t_n]} |\widehat{f}_n(u) - f(u)|. \end{aligned}$$

Recall that $\widehat{f}_n$ is constant on every interval $(\tau_{i-1}, \tau_i]$, for $i = 1, 2, \dots, N_n - 1$. Moreover, f' is bounded. Using the triangle inequality, it follows that

$$\begin{aligned} \sup_{u \in (\tau_{i-1}, \tau_i]} |\widehat{f}_n(u) - f(u)| &= \sup_{u \in (\tau_{i-1}, \tau_i]} |\widehat{f}_n(\tau_i) - f(u)| \\ &\le |\widehat{f}_n(\tau_i) - f(\tau_i)| + \|f'\|_\infty |\tau_{i-1} - \tau_i| \end{aligned}$$

for all $i = 1, 2, \dots, N_n - 1$ and

$$\sup_{u \in (\tau_{i_2}, 1-t_n]} |\widehat{f}_n(u) - f(u)| \le |\widehat{f}_n(1 - t_n) - f(1 - t_n)| + \|f'\|_\infty |\tau_{i_2} - \tau_{i_2+1}|.$$

From (43) and (44), we have $1 - t_n \in [n^{-1/3}, 1 - n^{-1/3}]$, for large enough n, so (36) ensures that $\widehat{f}_n(1 - t_n) - f(1 - t_n) = O_p(n^{-1/3})$. Using (35) and (38), it follows that

$$\begin{aligned} \sup_{u \in (s_n, 1-t_n]} |\widehat{f}_n(u) - f(u)| &\le \max_{i=i_1,\dots,i_2} |\widehat{f}_n(\tau_i) - f(\tau_i)| + O_p\left(\frac{\log n}{n}\right)^{1/3} \\ &= \max_{i=i_1,\dots,i_2} |\gamma_i - f \circ \widehat{U}_n(\gamma_i)| + O_p\left(\frac{\log n}{n}\right)^{1/3}. \end{aligned}$$

On the event E_n, we have $\gamma_i = f \circ g(\gamma_i)$, for all $i = i_1, \dots, i_2$, and therefore

$$\begin{aligned} \sup_{u \in (s_n, 1-t_n]} |\widehat{f}_n(u) - f(u)| &\le \|f'\|_\infty \max_{i=i_1,\dots,i_2} |g(\gamma_i) - \widehat{U}_n(\gamma_i)| + O_p\left(\frac{\log n}{n}\right)^{1/3} \\ &\le \|f'\|_\infty \sup_{a \in \mathbb{R}} |\widehat{U}_n(a) - g(a)| + O_p\left(\frac{\log n}{n}\right)^{1/3}. \end{aligned}$$

Theorem 2.1, with $\alpha_n = s_n$ and $\beta_n = t_n$ satisfying (43) and (44) now follows from Theorem 3.1.

It remains to extend the result to more general sequences α_n and β_n. For this purpose, define $s_n = n^{-1/3}(\log n)^{1/6}$. In view of the foregoing results, we know that

$$\sup_{t\in(s_n,1-s_n]} \left|\widehat{f}_n(t) - f(t)\right| = O_p\left(\frac{\log n}{n}\right)^{1/3}. \tag{48}$$

Suppose α_n and β_n satisfy (1). Let us notice that $\sup_{t\in(\alpha_n,1-\beta_n]}|\widehat{f}_n(t) - f(t)|$ decreases when either α_n or β_n increases, so that we can restrict our attention to small values of α_n and β_n. Without loss of generality we may assume that $\alpha_n \le n^{-1/3} \le s_n$ and $\beta_n \le n^{-1/3}$.

We then use the following property of nonincreasing functions h_1 and h_2 on an interval $[a,b]$:

$$\begin{aligned} &\sup_{t\in[a,b]} \left|h_1(t) - h_2(t)\right| \\ &\qquad \le \left|h_1(a) - h_2(a)\right| \vee \left|h_1(b) - h_2(b)\right| + \left|h_2(a) - h_2(b)\right|. \end{aligned} \tag{49}$$

See Lemma 6.1 in [8]. Since $\widehat{f}_n$ and f are both nonincreasing, according to (49), we have

$$\begin{aligned} &\sup_{t\in(\alpha_n,s_n]} \left|\widehat{f}_n(t) - f(t)\right| \\ &\qquad \le \left|\widehat{f}_n(\alpha_n) - f(\alpha_n)\right| \vee \left|f(s_n) - \widehat{f}_n(s_n)\right| + \|f'\|_\infty(s_n - \alpha_n). \end{aligned}$$

Because $s_n \in [n^{-1/3}, 1 - n^{-1/3}]$, it follows from (36) and (37) that $f(s_n) - \widehat{f}_n(s_n) = O_p(n^{-1/3})$ and $\widehat{f}_n(\alpha_n) - f(\alpha_n) = O_p((n\alpha_n)^{-1/2})$, which is of the order $O_p(n^{-1/3}(\log n)^{1/3})$, as we have assumed that $\alpha_n \ge K_1 n^{-1/3}(\log n)^{-2/3}$. We conclude

$$\sup_{t\in(\alpha_n,s_n]} \left|\widehat{f}_n(t) - f(t)\right| = O_p\left(\frac{\log n}{n}\right)^{1/3}.$$

Similarly, we obtain

$$\sup_{t\in(1-s_n,1-\beta_n]} \left|\widehat{f}_n(t) - f(t)\right| = O_p\left(\frac{\log n}{n}\right)^{1/3}$$

and therefore,

$$\sup_{t\in(\alpha_n,1-\beta_n]} \left|\widehat{f}_n(t) - f(t)\right| = \sup_{t\in(s_n,1-s_n]} \left|\widehat{f}_n(t) - f(t)\right| \vee O_p\left(\frac{\log n}{n}\right)^{1/3}.$$

Theorem 2.1 now follows from (48). $\square$

To prove Theorem 2.2, similarly to the proof of Theorem 2.1, we first establish the result for sequences $s_n = u + \alpha_n$ and $t_n = v - \beta_n$ satisfying (43) and (44), and then extend the result to more general sequences. The first step is to prove that the behavior of supremum over the interval $(s_n, 1 - t_n]$ is dominated by that of the largest interval between two jump points of $\widehat{f}_n$ contained in $(s_n, 1 - t_n]$. For this task, we make use of the notation τ_i, γ_i, i_1 and i_2 as introduced in (38), (39) and (40), and for $t \in [0, 1]$, we define the normalizing function

$$B(t) = \left(4\left|f'(t)\right|L'(t)\right)^{-1/3}. \tag{50}$$

It is easy to see that under assumptions (A1), (A2) and (A4), there exists $C_0 > 0$ and $\sigma \in (0, 1]$ such that

$$\left|A(u) - A(v)\right| \le C_0|u - v|^\sigma \quad \text{and} \quad \left|B(u) - B(v)\right| \le C_0|u - v|^\sigma \tag{51}$$

for all $u, v \in [0, 1]$, where A is given by (8). Recall that by convention, the supremum of an empty set is equal to zero.

For s, t that satisfy conditions (44) and (43), we first obtain

$$\begin{aligned} &\sup_{u\in(s,1-t]} B(u)\left|\widehat{f}_n(u) - f(u)\right| \\ &\qquad = \sup_{u\in(\tau_{i_1(s)},\tau_{i_2(t)}]} B(u)\left|\widehat{f}_n(u) - f(u)\right| \vee O_p\left(n^{-1/3}\right). \end{aligned} \tag{52}$$

See Lemma 6.14 in Supplement C in [8]. We are then able to make the connection between $\widehat{U}_n$ and $\widehat{f}_n$.

LEMMA 5.2. *Assume* (A1), (A2), (A3) *and* (A4). *Let* $0 < s < 1 - t < 1$, *possibly depending on* n, *such that* s, t *satisfy conditions* (43) *and* (44). *Then*

$$\begin{aligned} &\sup_{u\in(s,1-t]} B(u)\left|\widehat{f}_n(u) - f(u)\right| \\ &\qquad = \sup_{a\in[f(1-t),f(s)]} A(a)\left|\widehat{U}_n(a) - g(a)\right| + O_p\left(\frac{\log n}{n}\right)^{(\sigma+1)/3} \end{aligned}$$

for some $\sigma \in (0, 1]$.

PROOF. Again write $i_1 = i_1(s)$ and $i_2 = i_2(t)$. We first decompose the supremum into maxima of suprema taken over intervals between succeeding jump points of $\widehat{f}_n$:

$$\sup_{u\in(\tau_{i_1},\tau_{i_2}]} B(u)\left|\widehat{f}_n(u) - f(u)\right| = \max_{i_1+1\le i\le i_2} \sup_{u\in(\tau_{i-1},\tau_i]} B(u)\left|\widehat{f}_n(u) - f(u)\right|.$$

Then, by Theorem 2.1 and (35), we have that

$$\sup_{u\in(\tau_{i_1},\tau_{i_2}]} \left|\widehat{f}_n(u) - f(u)\right| \le \sup_{u\in(s,1-t]} \left|\widehat{f}_n(u) - f(u)\right| = O_p\left(\frac{\log n}{n}\right)^{1/3}.$$

Thus, we obtain by means of (51) and the triangle inequality that

$$\begin{aligned}&\sup_{u\in(\tau_{i_1},\tau_{i_2}]} B(u)\big|\widehat{f}_n(u)-f(u)\big|\\&\qquad=\max_{i_1+1\le i\le i_2} B(\tau_i)\sup_{u\in(\tau_{i-1},\tau_i]}\big|\widehat{f}_n(u)-f(u)\big|+O_p\left(\frac{\log n}{n}\right)^{(\sigma+1)/3}.\end{aligned}$$

By monotonicity of f, we have for all $i_1+1\le i\le i_2$,

$$\sup_{u\in(\tau_{i-1},\tau_i]}\big|\widehat{f}_n(\tau_i)-f(u)\big|=\big|\widehat{f}_n(\tau_i)-f(\tau_i)\big|\vee\big|\widehat{f}_n(\tau_i)-f(\tau_{i-1})\big|.$$

Hence, with (38) we arrive at

$$\begin{aligned}&\sup_{u\in(\tau_{i_1},\tau_{i_2}]} B(u)\big|\widehat{f}_n(u)-f(u)\big|\\&\qquad=\max_{i_1+1\le i\le i_2} B(\tau_i)\big\{\big|\gamma_i-f(\tau_i)\big|\vee\big|\gamma_i-f(\tau_{i-1})\big|\big\}+O_p\left(\frac{\log n}{n}\right)^{(\sigma+1)/3}.\end{aligned}$$

On the event E_n of (47), we can write $\gamma_i=f(g(\gamma_i))$ for all $i=i_1+1,\dots,i_2$, which, in view of (38), implies that

$$\begin{aligned}\big|\gamma_i-f(\tau_i)\big|&=\big|g(\gamma_i)-\widehat{U}_n(\gamma_i)\big|\cdot\big|f'(\theta_{i1})\big|,\\\big|\gamma_i-f(\tau_{i-1})\big|&=\big|g(\gamma_i)-\widehat{U}_n(\gamma_{i-1})\big|\cdot\big|f'(\theta_{i2})\big|\end{aligned}$$

for some θ_{i1} between $g(\gamma_i)$ and $\widehat{U}_n(\gamma_i)$, and θ_{i2} between $g(\gamma_i)$ and $\widehat{U}_n(\gamma_{i-1})$. By (46), Theorem 3.1 and (2), it follows that

$$\big|\gamma_i-f(\tau_i)\big|=\big|g(\gamma_i)-\widehat{U}_n(\gamma_i)\big|\cdot\big|f'\big(g(\gamma_i)\big)\big|+O_p\left(\frac{\log n}{n}\right)^{(\sigma+1)/3}.\tag{53}$$

By (38), (35) and Theorem 3.1, we have that

$$\begin{aligned}\max_{i_1+1\le i\le i_2}\big|g(\gamma_i)-\widehat{U}_n(\gamma_{i-1})\big|&=\max_{i_1+1\le i\le i_2}\big|g(\gamma_i)-\widehat{U}_n(\gamma_i)+\tau_i-\tau_{i-1}\big|\\&\le\sup_{a\in\mathbb{R}}\big|g(a)-\widehat{U}_n(a)\big|+O_p\left(\frac{\log n}{n}\right)^{1/3}\\&=O_p\left(\frac{\log n}{n}\right)^{1/3},\end{aligned}\tag{54}$$

so that similarly as above,

$$\big|\gamma_i-f(\tau_{i-1})\big|=\big|g(\gamma_i)-\widehat{U}_n(\gamma_{i-1})\big|\cdot\big|f'\big(g(\gamma_i)\big)\big|+O_p\left(\frac{\log n}{n}\right)^{(\sigma+1)/3}.$$

It follows that

$$\sup_{u\in(\tau_{i_1},\tau_{i_2}]} B(u)\big|\widehat{f}_n(u)-f(u)\big|$$
$$= \max_{i_1+1\le i\le i_2} B(\tau_i)\big|f'\big(g(\gamma_i)\big)\big|\big\{\big|g(\gamma_i)-\widehat{U}_n(\gamma_i)\big|\vee\big|g(\gamma_i)-\widehat{U}_n(\gamma_{i-1})\big|\big\}$$
$$+ O_p\left(\frac{\log n}{n}\right)^{(\sigma+1)/3}.$$

In order to replace $B(\tau_i)$ by $B(g(\gamma_i))$, we first note that (51), (53) and Theorem 3.1 imply that uniformly in i,

$$\big|B(\tau_i)-B\big(g(\gamma_i)\big)\big|\le C_0\big|\tau_i-g(\gamma_i)\big|^{\sigma}$$
$$\le C_0\|g'\|_\infty^{\sigma}\big|f(\tau_i)-\gamma_i\big|^{\sigma}=O_p\left(\frac{\log n}{n}\right)^{\sigma/3}.$$

By definition of A and B, we have $A(a)=B(g(a))|f'(g(a))|$, for all $a\in\mathbb{R}$, so from Theorem 2.1 and (54), we conclude that

$$\sup_{u\in(\tau_{i_1},\tau_{i_2}]} B(u)\big|\widehat{f}_n(u)-f(u)\big|$$
$$= \max_{i_1+1\le i\le i_2} A(\gamma_i)\big|g(\gamma_i)-\widehat{U}_n(\gamma_i)\big|\vee \max_{i_1\le i\le i_2-1} A(\gamma_{i+1})\big|g(\gamma_{i+1})-\widehat{U}_n(\gamma_i)\big|$$
$$+ O_p\left(\frac{\log n}{n}\right)^{(\sigma+1)/3}.$$

By the triangle inequality, on the event E_n of (47) we can write

$$|\gamma_{i+1}-\gamma_i|\le\|f'\|_\infty\big\{\big|g(\gamma_{i+1})-\widehat{U}_n(\gamma_i)\big|+\big|g(\gamma_i)-\widehat{U}_n(\gamma_i)\big|\big\}$$

for all $i_1\le i\le i_2-1$, so that Theorem 3.1 together with (54) implies that

$$\max_{i_1\le i\le i_2-1}|\gamma_{i+1}-\gamma_i|=O_p\left(\frac{\log n}{n}\right)^{1/3}. \tag{55}$$

Together with (51) and (54), this allows us to replace $A(\gamma_{i+1})$ by $A(\gamma_i)$, so that

$$\sup_{u\in(\tau_{i_1},\tau_{i_2}]} B(u)\big|\widehat{f}_n(u)-f(u)\big|$$
$$= \max_{i_1+1\le i\le i_2} A(\gamma_i)\big|g(\gamma_i)-\widehat{U}_n(\gamma_i)\big|\vee \max_{i_1\le i\le i_2-1} A(\gamma_i)\big|g(\gamma_{i+1})-\widehat{U}_n(\gamma_i)\big|$$
$$+ O_p\left(\frac{\log n}{n}\right)^{(\sigma+1)/3}.$$

Now, recall that $\widehat{U}_n$ is constant on intervals $(\gamma_{i+1},\gamma_i]$, and g is monotone. This implies that

$$\sup_{a\in(\gamma_{i+1},\gamma_i]}\big|\widehat{U}_n(a)-g(a)\big|=\big|\widehat{U}_n(\gamma_i)-g(\gamma_i)\big|\vee\big|\widehat{U}_n(\gamma_i)-g(\gamma_{i+1})\big|.$$

Therefore, taken into account joint indices, we find that

$$\begin{aligned}
&\sup_{u\in(\tau_{i_1},\tau_{i_2}]} B(u)\big|\widehat{f}_n(u)-f(u)\big| \\
&\quad= \max_{i_1+1\le i\le i_2-1} A(\gamma_i)\sup_{a\in(\gamma_{i+1},\gamma_i]}\big|\widehat{U}_n(a)-g(a)\big| \\
&\qquad \vee A(\gamma_{i_2})\big|g(\gamma_{i_2})-\widehat{U}_n(\gamma_{i_2})\big| \vee A(\gamma_{i_1})\big|g(\gamma_{i_1+1})-\widehat{U}_n(\gamma_{i_1})\big| \\
&\qquad + O_p\left(\frac{\log n}{n}\right)^{(\sigma+1)/3}.
\end{aligned}$$

Next, consider the term $A(\gamma_{i_1})|g(\gamma_{i_1+1})-\widehat{U}_n(\gamma_{i_1})|$, and let $\epsilon>0$. According to (43) and (41), there exists $C>0$ such that $\mathbb{P}(I_n)>1-\epsilon$, for n sufficiently large, where $I_n=\{\tau_{i_1}-s\le Cn^{-1/3}\}$. By monotonicity, we have on this event that $\gamma_{i_1}=\widehat{f}_n(\tau_{i_1})$ is between $\widehat{f}_n(s+Cn^{-1/3})$ and $\widehat{f}_n(s)$, which are both equal to $f(s)+O_p(n^{-1/3})$ by (36). A similar argument holds for γ_{i_1+1}, so that

$$\begin{aligned}
\gamma_{i_1}&=f(s)+O_p\big(n^{-1/3}\big)\quad\text{and}\\
\gamma_{i_1+1}&=f(s)+O_p\big(n^{-1/3}\big).
\end{aligned}\tag{56}$$

Together with (38) and (46), this implies

$$\begin{aligned}
\big|g(\gamma_{i_1+1})-\widehat{U}_n(\gamma_{i_1})\big| &= \big|g(\gamma_{i_1+1})-g\big(f(\tau_{i_1})\big)\big| \\
&\le \|g'\|_\infty\big|\gamma_{i_1+1}-f(\tau_{i_1})\big| \\
&= \|g'\|_\infty\big|f(s)-f(\tau_{i_1})\big|+O_p\big(n^{-1/3}\big) \\
&\le \|g'\|_\infty\|f'\|_\infty|s-\tau_{i_1}|+O_p\big(n^{-1/3}\big)=O_p\big(n^{-1/3}\big).
\end{aligned}$$

Similarly, it follows that

$$\big|g(\gamma_{i_2})-\widehat{U}_n(\gamma_{i_2})\big|=O_p\big(n^{-1/3}\big),\tag{57}$$

since by the same arguments as above, $\gamma_{i_2}=\widehat{f}_n(\tau_{i_2})$ is between $\widehat{f}_n(1-t)$ and $\widehat{f}_n(1-t-Cn^{-1/3})$ with probability greater than $1-\epsilon$, and both terms are equal to $f(1-t)+O_p(n^{-1/3})$. Since A is bounded, we conclude that

$$\begin{aligned}
&\sup_{u\in(\tau_{i_1},\tau_{i_2}]} B(u)\big|\widehat{f}_n(u)-f(u)\big| \\
&\quad= \max_{i_1+1\le i\le i_2-1} A(\gamma_i)\sup_{a\in(\gamma_{i+1},\gamma_i]}\big|\widehat{U}_n(a)-g(a)\big| \vee O_p\big(n^{-1/3}\big) \\
&\qquad + O_p\left(\frac{\log n}{n}\right)^{(\sigma+1)/3}.
\end{aligned}$$

To replace $A(\gamma_i)$ by $A(a)$ for $a \in (\gamma_{i+1}, \gamma_i]$, we use (55), (51) and Theorem 3.1. Together with (52), we conclude that

$$\begin{aligned}
&\sup_{u\in(s,1-t]} B(u)\big|\widehat{f}_n(u) - f(u)\big| \\
&\qquad = \sup_{a\in(\gamma_{i_2},\gamma_{i_1+1}]} A(a)\big|\widehat{U}_n(a) - g(a)\big| \vee O_p\big(n^{-1/3}\big) + O_p\left(\frac{\log n}{n}\right)^{(\sigma+1)/3}.
\end{aligned}\tag{58}$$

It remains to extend the latter supremum to the interval $[f(1-t), f(s)]$. We have

$$\sup_{a\in[f(1-t),\gamma_{i_2}]} A(a)\big|\widehat{U}_n(a) - g(a)\big| \le \|A\|_\infty \sup_{a\in[f(1-t),\gamma_{i_2}]} \big|\widehat{U}_n(a) - g(a)\big|.$$

According to (49),

$$\begin{aligned}
\sup_{a\in[f(1-t),\gamma_{i_2}]} \big|\widehat{U}_n(a) - g(a)\big| &\le \big|\widehat{U}_n\big(f(1-t)\big) - g\big(f(1-t)\big)\big| \\
&\qquad \vee \big|\widehat{U}_n(\gamma_{i_2}) - g(\gamma_{i_2})\big| + \|g'\|_\infty\big|\gamma_{i_2} - f(1-t)\big|.
\end{aligned}$$

Similarly to (56), we can write $\gamma_{i_2} = f(1-t) + O_p(n^{-1/3})$. Together with (12) and (57) we obtain

$$\sup_{a\in[f(1-t),\gamma_{i_2}]} \big|\widehat{U}_n(a) - g(a)\big| = O_p\big(n^{-1/3}\big)$$

and likewise,

$$\sup_{a\in[\gamma_{i_1+1},f(s)]} \big|\widehat{U}_n(a) - g(a)\big| = O_p\big(n^{-1/3}\big).$$

From (58), we conclude that

$$\begin{aligned}
&\sup_{u\in(s,1-t]} B(u)\big|\widehat{f}_n(u) - f(u)\big| \\
&\qquad = \sup_{a\in[f(t),f(s)]} A(a)\big|\widehat{U}_n(a) - g(a)\big| \vee R_n + O_p\left(\frac{\log n}{n}\right)^{(\sigma+1)/3},
\end{aligned}\tag{59}$$

where $R_n = O_p(n^{-1/3})$. We have

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{a\in[f(t),f(s)]} A(a)\big|\widehat{U}_n(a) - g(a)\big| \vee R_n \ne \sup_{a\in[f(t),f(s)]} A(a)\big|\widehat{U}_n(a) - g(a)\big|\Big) \\
&\qquad \le \mathbb{P}\Big(R_n \ge \sup_{a\in[f(t),f(s)]} A(a)\big|\widehat{U}_n(a) - g(a)\big|\Big).
\end{aligned}$$

But it follows from Corollary 3.1 that

$$\left(\frac{\log n}{n}\right)^{-1/3} \sup_{a\in[f(t),f(s)]} A(a)\big|\widehat{U}_n(a) - g(a)\big| = 2^{-1/3} + o_p(1).\tag{60}$$

Since $R_n = o_p((n/\log n)^{-1/3})$, it follows that the latter probability tends to zero as $n \to \infty$. The lemma now follows from (59). $\square$

PROOF OF THEOREM 2.2. Let S_n be defined by (9), with $0 \le u < v \le 1$ fixed and α_n and β_n satisfying (5). Let

$$s_n = u + \alpha_n \quad \text{and} \quad t_n = 1 - v + \beta_n. \tag{61}$$

Then automatically s_n and t_n will always satisfy condition (44). If, in addition, s_n and t_n satisfy condition (43), then according to Lemma 5.2 together with (60),

$$\sup_{t\in(u+\alpha_n, v-\beta_n]} B(t)\big|\widehat{f}_n(t) - f(t)\big|$$

has the same limit distribution as

$$\sup_{a\in[f(v-\beta_n), f(u+\alpha_n)]} A(a)\big|\widehat{U}_n(a) - g(a)\big|,$$

so that Theorem 2.2 follows from Corollary 3.1. When $0 < u < v < 1$, then s_n and t_n automatically satisfy (43), so we only have to consider the cases where either $u = 0$ or $v = 1$. If $u = 0$ and $n^{1/3}\alpha_n \to \infty$, or if $v = 1$ and $n^{1/3}\beta_n \to \infty$, then s_n and t_n, as defined in (61), also satisfy condition (43). Therefore, we can restrict ourselves to the case $\alpha_n = O(n^{-1/3})$ and $\beta_n = O(n^{-1/3})$.

Define $a_n = n^{-1/3}(\log n)^{1/6}$, so that $u + \alpha_n < u + a_n < v - a_n < v - \beta_n$. By means of (49), we find

$$\begin{aligned}
&\sup_{t\in(u+\alpha_n, u+a_n]} \big|\widehat{f}_n(t) - f(t)\big| \\
&\quad \le \big|\widehat{f}_n(u+\alpha_n) - f(u+\alpha_n)\big| \vee \big|\widehat{f}_n(u+a_n) - f(u+a_n)\big| \\
&\qquad + \big|f(u+\alpha_n) - f(u+a_n)\big|.
\end{aligned}$$

By definition, $|f(u+\alpha_n) - f(u+a_n)| \le \|f'\|_\infty|\alpha_n - a_n| = O(n^{-1/3}(\log n)^{1/6})$, and according to (36) and (37), together with (5),

$$\begin{aligned}
\widehat{f}_n(u+\alpha_n) - f(u+\alpha_n) &= O_p\big((n\alpha_n)^{-1/2}\big) = o_p\big(n^{-1/3}(\log n)^{1/3}\big), \\
\widehat{f}_n(u+a_n) - f(u+a_n) &= O_p\big(n^{-1/3}\big).
\end{aligned}$$

Because $B(t)$ is uniformly bounded, it follows that

$$\sup_{t\in(u+\alpha_n, u+a_n]} B(t)\big|\widehat{f}_n(t) - f(t)\big| = o_p\left(\frac{\log n}{n}\right)^{1/3},$$

and likewise

$$\sup_{t\in(v-a_n, v-\beta_n]} B(t)\big|\widehat{f}_n(t) - f(t)\big| = o_p\left(\frac{\log n}{n}\right)^{1/3}.$$

This means that

$$\sup_{t\in(u+\alpha_n,v-\beta_n]} B(t)\big|\widehat{f}_n(t)-f(t)\big| = \sup_{t\in(u+a_n,v-a_n]} B(t)\big|\widehat{f}_n(t)-f(t)\big| \vee R_n,$$

where $R_n = o_p((n/\log n)^{-1/3})$. Because $u+a_n$ and $1-v+a_n$ satisfy the conditions of Lemma 5.2, together with (60), it follows that

$$\sup_{t\in(u+\alpha_n,v-\beta_n]} B(t)\big|\widehat{f}_n(t)-f(t)\big|$$

has the same limit distribution as

$$\sup_{c\in[f(v-a_n),f(u+a_n)]} A(c)\big|\widehat{U}_n(c)-g(c)\big|,$$

so that Theorem 2.2 follows from Corollary 3.1. □

Acknowledgments. The authors would like to thank Fadoua Balabdaoui for co-organizing a research visit in Paris concerning this topic and for helpful comments and suggestions in many stimulating discussions.

SUPPLEMENTARY MATERIAL

Supplement to "The limit distribution of the L_∞-error of Grenander-type estimators" (DOI: 10.1214/12-AOS1015SUPP; .pdf).

- Supplement A: The supremum of the limiting process.
- Supplement B: Preliminary results for the inverse process.
- Supplement C: Points of jump.

REFERENCES

[1] ANEVSKI, D. and HÖSSJER, O. (2006). A general asymptotic scheme for inference under order restrictions. *Ann. Statist.* **34** 1874–1930. MR2283721

[2] BALABDAOUI, F., JANKOWSKI, H., PAVLIDES, M., SEREGIN, A. and WELLNER, J. A. (2011). On the Grenander estimator at zero. *Statist. Sinica* **21** 873–899. MR2829859

[3] BANERJEE, M. and WELLNER, J. A. (2001). Likelihood ratio tests for monotone functions. *Ann. Statist.* **29** 1699–1731. MR1891743

[4] BRUNK, H. D. (1958). On the estimation of parameters restricted by inequalities. *Ann. Math. Statist.* **29** 437–454. MR0132632

[5] BRUNK, H. D. (1970). Estimation of isotonic regression. In *Nonparametric Techniques in Statistical Inference* (*Proc. Sympos., Indiana Univ., Bloomington, Ind.*, 1969) 177–197. Cambridge Univ. Press, London. MR0277070

[6] DUROT, C. (2002). Sharp asymptotics for isotonic regression. *Probab. Theory Related Fields* **122** 222–240. MR1894068

[7] DUROT, C. (2007). On the $\mathbb{L}_p$-error of monotonicity constrained estimators. *Ann. Statist.* **35** 1080–1104. MR2341699

[8] DUROT, C., KULIKOV, V. N. and LOPUHAÄ, H. P. (2012). Supplement to "The limit distribution of the L_∞-error of Grenander-type estimators." DOI:10.1214/12-AOS1015SUPP.

[9] EGGERMONT, P. P. B. and LARICCIA, V. N. (2000). Maximum likelihood estimation of smooth monotone and unimodal densities. *Ann. Statist.* **28** 922–947. MR1792794

[10] GRENANDER, U. (1956). On the theory of mortality measurement. II. *Skand. Aktuarietidskr.* **39** 125–153. MR0093415

[11] GROENEBOOM, P. (1985). Estimating a monotone density. In *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer, Vol. II* (*Berkeley, Calif.*, 1983) 539–555. Wadsworth, Belmont, CA. MR0822052

[12] GROENEBOOM, P. (1989). Brownian motion with a parabolic drift and Airy functions. *Probab. Theory Related Fields* **81** 79–109. MR0981568

[13] GROENEBOOM, P. (2011). Vertices of the least concave majorant of Brownian motion with parabolic drift. *Electron. J. Probab.* **16** 2234–2258. MR2861676

[14] GROENEBOOM, P., HOOGHIEMSTRA, G. and LOPUHAÄ, H. P. (1999). Asymptotic normality of the L_1 error of the Grenander estimator. *Ann. Statist.* **27** 1316–1347. MR1740109

[15] GROENEBOOM, P. and WELLNER, J. A. (1992). *Information Bounds and Nonparametric Maximum Likelihood Estimation*. *DMV Seminar* **19**. Birkhäuser, Basel. MR1180321

[16] 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. MR1814624

[17] HUANG, J. and WELLNER, J. A. (1995). Estimation of a monotone density or monotone hazard under random censoring. *Scand. J. Stat.* **22** 3–33. MR1334065

[18] HUANG, Y. and ZHANG, C.-H. (1994). Estimating a monotone density from censored observations. *Ann. Statist.* **22** 1256–1274. MR1311975

[19] JONKER, M. A. and VAN DER VAART, A. W. (2001). A semi-parametric model for censored and passively registered data. *Bernoulli* **7** 1–31. MR1811742

[20] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975). An approximation of partial sums of independent RV's and the sample DF. I. *Z. Wahrsch. Verw. Gebiete* **32** 111–131. MR0375412

[21] KOSOROK, M. R. (2008). Bootstrapping the Grenander estimator. In *Beyond Parametrics in Interdisciplinary Research*: *Festschrift in Honor of Professor Pranab K. Sen* (N. Balakrishnan, E. Pena and M. Silvapulle, eds.). *Inst. Math. Stat. Collect.* **1** 282–292. IMS, Beachwood, OH. MR2462212

[22] KULIKOV, V. N. and LOPUHAÄ, H. P. (2005). Asymptotic normality of the L_k-error of the Grenander estimator. *Ann. Statist.* **33** 2228–2255. MR2211085

[23] KULIKOV, V. N. and LOPUHAÄ, H. P. (2006). The behavior of the NPMLE of a decreasing density near the boundaries of the support. *Ann. Statist.* **34** 742–768. MR2283391

[24] LOPUHAÄ, H. P. and NANE, G. F. (2011). Shape constrained nonparametric estimators of the baseline distribution in Cox proportional hazards model. Unpublished manuscript.

[25] MAMMEN, E. (1991). Estimating a smooth monotone regression function. *Ann. Statist.* **19** 724–740. MR1105841

[26] PAL, J. K. (2009). End-point estimation for decreasing densities: Asymptotic behaviour of the penalized likelihood ratio. *Scand. J. Stat.* **36** 764–781. MR2573307

[27] PRAKASA RAO, B. L. S. (1969). Estimation of a unimodal density. *Sankhyā Ser. A* **31** 23–36. MR0267677

[28] PRAKASA RAO, B. L. S. (1970). Estimation for distributions with monotone failure rate. *Ann. Math. Statist.* **41** 507–519. MR0260133

[29] SAKHANENKO, A. I. (1985). Estimates in an invariance principle. In *Limit Theorems of Probability Theory*. *Trudy Inst. Mat.* **5** 27–44, 175. "Nauka" Sibirsk. Otdel., Novosibirsk. MR0821751

[30] SEN, B., BANERJEE, M. and WOODROOFE, M. (2010). Inconsistency of bootstrap: The Grenander estimator. *Ann. Statist.* **38** 1953–1977. MR2676880

[31] VAN DER VAART, A. W. and VAN DER LAAN, M. J. (2003). Smooth estimation of a monotone density. *Statistics* **37** 189–203. MR1986176

[32] WOODROOFE, M. and SUN, J. (1993). A penalized maximum likelihood estimate of $f(0+)$ when f is nonincreasing. *Statist. Sinica* **3** 501–515. MR1243398

[33] ZIL'BERBURG, K. L. (2007). Estimation of a monotone function observed in white noise. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.* (*POMI*) **341** 124–133, 231. MR2363588

C. DUROT
UNIVERSITÉ PARIS OUEST NANTERRE LA DÉFENSE
200, AVENUE DE LA RÉPUBLIQUE
92001 NANTERRE CEDEX
FRANCE
E-MAIL: cecile.durot@gmail.com

V. N. KULIKOV
ASR NEDERLAND
ARCHIMEDESLAAN 10
3584 BA UTRECHT
THE NETHERLANDS
E-MAIL: vladimir.kulikov@asr.nl

H. P. LOPUHAÄ
DIAM
DELFT UNIVERSITY OF TECHNOLOGY
MEKELWEG 4
2628 CD DELFT
THE NETHERLANDS
E-MAIL: h.p.lopuhaa@tudelft.nl