Central limit theorems for the Lp-error of smooth isotonic estimators

Lopuhaä, Hendrik P.; Musta, Eni

DOI
10.1214/19-EJS1550

Publication date
2019

Document Version
Final published version

Published in
Electronic Journal of Statistics

Electronic Journal of Statistics
Vol. 13 (2019) 1031–1098
ISSN: 1935-7524
https://doi.org/10.1214/19-EJS1550

Central limit theorems for the L_p-error of smooth isotonic estimators

Hendrik P. Lopuhaä and Eni Musta

DIAM, Delft University of Technology,
van Mourik Broekmanweg 6, 2628 XE, Delft, Netherlands
e-mail: h.p.lopuhaa@tudelft.nl*;* e.musta@tudelft.nl

Abstract: We investigate the asymptotic behavior of the L_p-distance between a monotone function on a compact interval and a smooth estimator of this function. Our main result is a central limit theorem for the L_p-error of smooth isotonic estimators obtained by smoothing a Grenander-type estimator or isotonizing the ordinary kernel estimator. As a preliminary result we establish a similar result for ordinary kernel estimators. Our results are obtained in a general setting, which includes estimation of a monotone density, regression function and hazard rate. We also perform a simulation study for testing monotonicity on the basis of the L_2-distance between the kernel estimator and the smoothed Grenander-type estimator.

MSC 2010 subject classifications: Primary 62G20; secondary 62G10.
Keywords and phrases: Kernel estimator, L_p loss, central limit theorem, smoothed Grenander-type estimator, isotonized kernel estimator, boundary corrections, Hellinger loss, testing monotonicity.

Received July 2018.

1. Introduction

The property of monotonicity plays an important role when dealing with survival data or regression relationships. For example, it is often natural to assume that increasing a factor X has a positive (negative) effect on a response Y or that the risk for an event to happen is increasing (decreasing) over time. In situations like these, incorporating monotonicity constraints in the estimation procedure leads to more accurate results. The first non-parametric monotone estimators were introduced in [20], [6], and [41], concerning the estimation of a monotone probability density, regression function, and failure rate. These estimators are all piecewise constant functions that exhibit a non-normal limit distribution at rate $n^{1/3}$.

On the other hand, under some more regularity assumptions on the function of interest, smooth non-parametric estimators can be used to achieve a faster rate of convergence to a Gaussian distributional law. Typically, these estimators are constructed by combining an isotonization step with a smoothing step. Estimators constructed by smoothing followed by an isotonization step have been considered in [7], [47], [18], and [44], for the regression setting, in [46] for estimating a monotone density, and in [17], who consider maximum smoothed likelihood estimators for monotone densities. Methods that interchange the smoothing step and the isotonization step, can be found in [42], [14], and [36], who study kernel smoothed isotonic estimators. Comparisons between isotonized smooth estimators and smoothed isotonic estimators are made in [40], [26] and [25].

A lot of attention has been given in the literature to the pointwise asymptotic behavior of smooth estimators and monotone estimators, separately. However, for example for goodness of fit tests, global errors of estimates are needed instead of pointwise results. For the Grenander estimator of a monotone density, a central limit theorem for the L_1-error was formulated in [21] and proven rigorously in [22]. A similar result was established in [12] for the regression context. Extensions to general L_p-errors can be found in [31] and in [13], where the latter provides a unified approach that applies to a variety of statistical models. On the other hand, central limit theorems for regular kernel density estimators have been obtained in [10] and [9].

In this paper we investigate the L_p-error of smooth isotonic estimators obtained by kernel smoothing the Grenander-type estimator or by isotonizing the ordinary kernel estimator. We consider the same general setup as in [13], which includes estimation of a probability density, a regression function, or a failure rate under monotonicity constraints (see Section 3 in [13] for more details on these models). An essential assumption in this setup is that the observed process of interest can be approximated by a Brownian motion or a Brownian bridge. Our main results are central limit theorems for the L_p-error of smooth isotonic estimators for a monotone function on a compact interval. However, since the behavior of these estimators is closely related to the behavior of ordinary kernel estimators, we first establish a central limit theorem for the L_p-error of ordinary kernel estimators for a monotone function on a compact interval. This extends the work by [10] on the L_p-error of densities that are smooth on the whole real line, but is also of interest by itself. The fact that we no longer have a smooth function on the whole real line, leads to boundary effects. Unexpectedly, different from [10], we find that the limit variance of the L_p-error changes, depending on whether the approximating process is a Brownian motion or a Brownian bridge. Such a phenomenon has also not been observed in other isotonic problems, where a similar embedding assumption was made. Usually, both approximations lead to the same asymptotic results (e.g., see [13] and [31]).

After establishing a central limit theorem for the L_p-error of ordinary kernel estimators, we transfer this result to the smoothed Grenander estimator (SG). The key ingredient here is the behavior of the process obtained as the difference between a naive estimator and its least concave majorant. For this we use results from [38]. As an intermediate result, we show that the L_p-distance between the smoothed Grenander-type estimator and the ordinary kernel estimator converges at rate $n^{2/3}$ to some functional of two-sided Brownian motion minus a parabolic drift.

The situation for the isotonized kernel estimator (GS) is much easier, because it can be shown that this estimator coincides with the ordinary kernel estimator on large intervals in the interior of the support, with probability tending to one. However, since the isotonization step is performed last, the estimator is inconsistent at the boundaries. For this reason, we can only obtain a central limit theorem for the L_p-error on a sub-interval that approaches the whole support, as n diverges to infinity. Finally, the results on the L_p-error can be applied immediately to obtain a central limit theorem for the Hellinger loss.

The paper is organized as follows. In Section 2 we describe the model, the assumptions and fix some notation that will be used throughout the paper. A central limit theorem for the L_p-error of the kernel estimator is obtained in Section 3. This result is used in Section 4 and 5 to obtain the limit distribution of the L_p-error of the SG and GS estimators. Section 6 is dedicated to corresponding asymptotics for the Hellinger distance. In Section 7 we provide a possible application of our results by considering a test for monotonicity. Details of some of the proofs are delayed to Section 8 and to additional technicalities have been put in the Appendix.

2. Assumptions and notations

Consider estimating a function $\lambda : [0,1] \to \mathbb{R}$ subject to the constraint that it is non-increasing. Suppose that on the basis of n observations we have at hand a cadlag step estimator Λ_n of

$$\Lambda(t) = \int_0^t \lambda(u)\,\mathrm{d}u, \quad t \in [0,1].$$

A typical example is the estimation of a monotone density λ on a compact interval. In this case, Λ_n is the empirical distribution function. Hereafter M_n denotes the process $M_n = \Lambda_n - \Lambda$, μ is a measure on the Borel sets of $\mathbb{R}$, and

k is a twice differentiable symmetric probability density with support $[-1,1]$. (1)

The rescaled kernel is defined as $k_b(u) = b^{-1}k\,(u/b)$, where the bandwidth $b = b_n \to 0$, as $n \to \infty$. In the sequel we will make use of the following assumptions.

(A1) λ is decreasing and twice continuously differentiable on $[0,1]$ and such that $\inf_t |\lambda'(t)| > 0$.

(A2) Let B_n be either a Brownian motion or a Brownian bridge. There exists $q > 5/2$, $C_q > 0$, $L : [0,1] \to \mathbb{R}$ and versions of M_n and B_n such that

$$\mathbb{P}\left(n^{1-1/q} \sup_{t\in[0,1]} \left| M_n(t) - n^{-1/2} B_n \circ L(t) \right| > x \right) \leq C_q x^{-q}$$

for all $x \in (0,n]$. Moreover, L is increasing and twice differentiable on $[0,1]$ with $\sup_t |L''(t)| < \infty$ and $\inf_t |L'(t)| > 0$.

(A3) $\mathrm{d}\mu(t) = w(t)\,\mathrm{d}t$, where $w(t) \geq 0$ is continuous on $[0,1]$.

In particular, the approximation of the process M_n by a Gaussian process, as in assumption (A2), is required also in [13]. It corresponds to a general setting which includes estimation of a probability density, regression function or a failure rate under monotonicity constraints (see Section 3 in [13] for more details on these models).

First we introduce some notation. We partly adopt the one used in [10] and briefly explain their appearance. Let $\tilde{\lambda}_n^s$ be the standard kernel estimator of λ, i.e.

$$\tilde{\lambda}_n^s(t) = \int_{t-b}^{t+b} k_b(t-u)\,\mathrm{d}\Lambda_n(u), \quad \text{for } t \in [b, 1-b]. \tag{2}$$

As usual we decompose into a random term and a bias term:

$$(nb)^{1/2}\left(\tilde{\lambda}_n^s(t) - \lambda(t)\right) = (nb)^{1/2}\int k_b(t-u)\,\mathrm{d}(\Lambda_n - \Lambda)(u) + g_{(n)}(t) \tag{3}$$

where

$$g_{(n)}(t) = (nb)^{1/2}\left(\lambda_{(n)}(t) - \lambda(t)\right), \qquad \lambda_{(n)}(t) = \int k_b(t-u)\lambda(u)\,\mathrm{d}u. \tag{4}$$

When $nb^5 \to C_0 > 0$, then $g_{(n)}(t)$ converges to

$$g(t) = \frac{1}{2}C_0\lambda''(t)\int k(y)y^2\,\mathrm{d}y. \tag{5}$$

After separating the bias term, the first term on the right hand side of (3) involves an integral of $k_b(t-u)$ with respect to the process M_n. Due to (A2), this integral will be approximated by an integral with respect to a Gaussian process. For this reason, the limiting moments of the L_p-error involve integrals with respect to Gaussian densities, such as

$$\begin{aligned}\phi(x) &= (2\pi)^{-1/2}\exp(-x^2/2),\\ \psi(u,x,y) &= \frac{1}{2\pi\sqrt{1-u^2}}\exp\left(-\frac{x^2-2uxy+y^2}{2(1-u^2)}\right) = \frac{1}{\sqrt{1-u^2}}\phi\left(\frac{x-uy}{\sqrt{1-u^2}}\right)\phi(y),\end{aligned} \tag{6}$$

and a Taylor expansion of $k_b(t-u)$ yields the following constants involving the kernel function:

$$D^2 = \int k(y)^2\,\mathrm{d}y, \qquad r(s) = \frac{\int k(z)k(s+z)\,\mathrm{d}z}{\int k^2(z)\,\mathrm{d}z}. \tag{7}$$

For example, the limiting means of the L_p-error and a truncated version are given by:

$$\begin{aligned}m_n(p) &= \int_{\mathbb{R}}\int_0^1 \left|\sqrt{L'(t)}Dx + g_{(n)}(t)\right|^p w(t)\phi(x)\,\mathrm{d}t\,\mathrm{d}x,\\ m_n^c(p) &= \int_{\mathbb{R}}\int_b^{1-b} \left|\sqrt{L'(t)}Dx + g_{(n)}(t)\right|^p w(t)\phi(x)\,\mathrm{d}t\,\mathrm{d}x,\end{aligned} \tag{8}$$

where D and $g_{(n)}$ are defined in (7) and (4). Depending on the rate at which $b \to 0$, the limiting variance of the L_p-error has a different form. When $nb^5 \to 0$, the limiting variance turns out to be

$$\sigma^2(p) = \sigma_1 D^{2p} \int_0^1 |L'(u)|^p \, w(u)^2 \, \mathrm{d}u, \tag{9}$$

where

$$\sigma_1 = \int_{\mathbb{R}} \left\{ \int_{\mathbb{R}} \int_{\mathbb{R}} |xy|^p \psi(r(s), x, y) \, \mathrm{d}x \, \mathrm{d}y - \int_{\mathbb{R}} \int_{\mathbb{R}} |xy|^p \phi(x)\phi(y) \, \mathrm{d}x \, \mathrm{d}y \right\} \mathrm{d}s, \tag{10}$$

with σ_1 representing p-th moments of bivariate Gaussian vectors, where D, ψ, and ϕ are defined in (7) and (6). When $nb^5 \to C_0 > 0$ and B_n in (A2) is a Brownian motion, the limiting variance of the L_p-error is

$$\begin{aligned}\theta^2(p) = \int_0^1 \int_{\mathbb{R}^3} \left| g(u)^2 + g(u)(x+y)\sqrt{L'(u)} D + D^2 L'(u) xy \right|^p \\ w^2(u) \Big(\psi(r(s), x, y) - \phi(x)\phi(y) \Big) \, \mathrm{d}s \, \mathrm{d}y \, \mathrm{d}x \, \mathrm{d}u,\end{aligned} \tag{11}$$

where g, D, ψ, and ϕ are defined in (5), (7) and (6), whereas, if B_n in (A2) is a Brownian bridge, the limiting variance is slightly different,

$$\tilde{\theta}^2(p) = \theta^2(p) - \frac{\theta_1^2(p)}{D^2 L(1)}, \tag{12}$$

with

$$\theta_1(p) = \int_0^1 \int_{\mathbb{R}} \left| \sqrt{L'(t)} D x + g(t) \right|^p x\phi(x) \, \mathrm{d}x \, \sqrt{L'(t)} w(t) \mathrm{d}t. \tag{13}$$

Finally, the following inequality will be used throughout this paper:

$$\begin{aligned}\int_A^B ||q(t)|^p - |h(t)|^p| \, \mathrm{d}\mu(t) \le p2^{p-1} \int_A^B |q(t) - h(t)|^p \, \mathrm{d}\mu(t) \\ + p2^{p-1} \left(\int_A^B |h(t)|^p \, \mathrm{d}\mu(t) \right)^{1-1/p} \left(\int_A^B |q(t) - h(t)|^p \, \mathrm{d}\mu(t) \right)^{1/p},\end{aligned} \tag{14}$$

where $p \in [1, \infty)$, $-\infty \le A < B \le \infty$ and $q, h \in L_p(A, B)$.

3. Kernel estimator of a decreasing function

We extend the results of [10] and [9] to the case of a kernel estimator of a decreasing function with compact support. Note that, since the function of interest cannot be twice differentiable on $\mathbb{R}$ (not even continuous), the kernel estimator is inconsistent at zero and one. Moreover we show that the contribution of the boundaries to the L_p-error is not negligible, so in order to avoid the L_p-distance to explode we have to restrict ourselves to the interval $[b, 1-b]$ or apply some boundary correction.

3.1. A modified L_p-distance of the standard kernel estimator

Let $\tilde{\lambda}_n^s$ be the standard kernel estimator of λ defined in (2). In order to avoid boundary problems, we start by finding the asymptotic distribution of a modification of the L_p-distance

$$J_n^c(p) = \int_b^{1-b} \left|\tilde{\lambda}_n^s(t) - \lambda(t)\right|^p \,\mathrm{d}\mu(t), \tag{15}$$

instead of

$$J_n(p) = \int_0^1 \left|\tilde{\lambda}_n^s(t) - \lambda(t)\right|^p \,\mathrm{d}\mu(t). \tag{16}$$

Theorem 3.1. *Assume that (A1)-(A3) hold. Let k satisfy (1) and let J_n^c be defined in (15). Suppose $p \geq 1$ and $nb \to \infty$.*

i) If $nb^5 \to 0$, then

$$(b\sigma^2(p))^{-1/2}\left\{(nb)^{p/2}J_n^c(p) - m_n^c(p)\right\} \xrightarrow{d} N(0,1);$$

ii) If $nb^5 \to C_0^2 > 0$, and B_n in Assumption (A2) is a Brownian motion, then

$$(b\theta^2(p))^{-1/2}\left\{(nb)^{p/2}J_n^c(p) - m_n^c(p)\right\} \xrightarrow{d} N(0,1);$$

iii) If $nb^5 \to C_0^2 > 0$, and B_n in Assumption (A2) is a Brownian bridge, then

$$(b\tilde{\theta}^2(p))^{-1/2}\left\{(nb)^{p/2}J_n^c(p) - m_n^c(p)\right\} \xrightarrow{d} N(0,1),$$

where $m_n^c(p)$, $\sigma^2(p)$, $\theta^2(p)$, $\tilde{\theta}^2(p)$ are defined in (8), (9), (11), and (12), respectively.

The proof goes along the same lines as in the one for the case of the L_p-norms for kernel density estimators on the whole real line (see [10] and [9]). The main idea is that by means of assumption (A2), it is sufficient to prove the central limit theorem for the approximating process. When B_n in (A2) is a Brownian motion, the latter one can be obtained by a big-blocks-small-blocks procedure using the independence of the increments of the Brownian motion. When B_n in (A2) is a Brownian bridge, we can still obtain a central limit theorem, but the limiting variance turns out to be different. The latter result differs from what is stated in [10]. In [10], the complete proof for both Brownian motion and Brownian bridge, is only given for the case $nb^5 \to 0$, and it is shown that the random variables obtained by using the Brownian motion and the Brownian bridge as approximating processes are asymptotically equivalent (see their Lemma 6). In fact, when dealing with a Brownian bridge, the rescaled L_p-error is asymptotically equivalent to the L_p-error that corresponds to the Brownian motion process plus an additional term which is equal to $CW(L(1))$, for a constant C proportional on $\theta_1(p)$ defined in (13). When the bandwidth is small, i.e., $nb^5 \to 0$, the bias term $g(t)$ in the definition of $\theta_1(p)$ disappears. Hence, by the symmetry property of the standard normal density, $\theta_1(p) = 0$ and as a consequence $C = 0$. This means that the additional term resulting from the fact that we are dealing with a Brownian bridge converges to zero. For details, see the proof of Lemma 8.1. When $nb^5 \to C_0^2 > 0$, only a sketch of the proof is given in [10] for B_n being a Brownian motion and it is claimed that again the limit distribution would be the same for B_n being a Brownian bridge. However, in our setting we find that the limit variances are different.

Various settings in which Brownian motion or Brownian bridge approximations arise are described in Section 3 of [13]. In particular, for the density model, which is also considered in [10], the approximating process is a Brownian bridge. Hence, the difference in the limiting variances is an important issue. In other models, such as random censorship, Poisson process model, or regression model with fixed design points, the approximating process is a Brownian motion.

Proof of Theorem 3.1. From the definition of $J_n^c(p)$ we have

$$(nb)^{p/2} J_n^c(p) = \int_b^{1-b} \left|(nb)^{1/2} \int k_b(t-u)\,\mathrm{d}(\Lambda_n - \Lambda)(u) + g_{(n)}(t)\right|^p \,\mathrm{d}\mu(t).$$

Let $\{W(t) : t \in \mathbb{R}\}$ be a Wiener process and define

$$\Gamma_n^{(1)}(t) = \int k\left(\frac{t-u}{b}\right) \mathrm{d}W(L(u)), \tag{17}$$

Hence, if B_n in assumption (A2) is a Brownian motion, then according to (14),

$$\begin{aligned}
&\left|(nb)^{p/2} J_n^c(p) - \int_b^{1-b} \left|b^{-1/2}\Gamma_n^{(1)}(t) + g_{(n)}(t)\right|^p \,\mathrm{d}\mu(t)\right| \\
&\quad\leq p2^{p-1} b^{-p/2} \int_b^{1-b} \left|\int k\left(\frac{t-u}{b}\right) \mathrm{d}(B_n \circ L(u) - n^{1/2} M_n(u))\right|^p \,\mathrm{d}\mu(t) \\
&\qquad + p2^{p-1} \left(b^{-p/2} \int_b^{1-b} \left|\int k\left(\frac{t-u}{b}\right) \mathrm{d}(B_n \circ L - n^{1/2} M_n)(u)\right|^p \,\mathrm{d}\mu(t)\right)^{1/p} \cdot \\
&\qquad\qquad \cdot \left(\int_b^{1-b} \left|b^{-1/2}\Gamma_n^{(1)}(t) + g_{(n)}(t)\right|^p \,\mathrm{d}\mu(t)\right)^{1-1/p}
\end{aligned}$$

We can write

$$\begin{aligned}
\left|\int k\left(\frac{t-u}{b}\right) \mathrm{d}(B_n \circ L - n^{1/2} M_n)(u)\right| &= \left|\int_{-1}^1 k(y)\,\mathrm{d}(B_n \circ L - n^{1/2} M_n)(t-by)\right| \\
&= \left|\int_{-1}^1 (B_n \circ L - n^{1/2} M_n)(t-by)\,\mathrm{d}k(y)\right| \\
&\leq C \sup_{t\in[0,1]} \left|B_n \circ L(t) - n^{1/2} M_n(t)\right|.
\end{aligned} \tag{18}$$

According to assumption (A2), the right hand side of (18) is of the order $O_P(n^{-1/2+1/q})$, and because

$$b^{-1/2}O_P(n^{-1/2+1/q}) = (nb^5)^{3/10}O_P(n^{-2/5+1/q}) = o_P(1)$$

we derive that

$$\left|(nb)^{p/2}J_n^c(p) - \int_b^{1-b}\left|b^{-1/2}\Gamma_n^{(1)}(t) + g_{(n)}(t)\right|^p \,\mathrm{d}\mu(t)\right| = o_P(1).$$

As a result, the statement follows from the fact that

$$(b\sigma^2(p))^{-1/2}\left\{\int_b^{1-b}\left|b^{-1/2}\Gamma_n^{(1)}(t) + g_{(n)}(t)\right|^p \,\mathrm{d}\mu(t) - m_n^c(p)\right\} \xrightarrow{d} N(0,1),$$

where $g_{(n)}$ and $m_n^c(p)$ are defined in (4) and (8), respectively. This result is a generalization of Lemmas 1-5 in [10] and the proof goes in the same way. However, for completeness we give all the details in the Appendix. See Lemma A.1.

Finally, if B_n is a Brownian bridge on $[0, L(1)]$, we use the representation $B_n(t) = W(t) - tW(L(1))/L(1)$. By replacing $\Gamma_n^{(1)}$ with

$$\Gamma_n^{(2)}(t) = \int k\left(\frac{t-u}{b}\right)\mathrm{d}\left(W(L(u)) - \frac{L(u)}{L(1)}W(L(1))\right) \tag{19}$$

in the previous reasoning, the statement follows from Lemma 8.1. □

When $nb^4 \to 0$, the centering constant $m_n(p)$ can be replaced by a quantity that does not depend on n.

Theorem 3.2. *Assume that (A1)-(A3) hold. Let k satisfy (1) and let J_n^c be defined in (15). Suppose $p \geq 1$ and $nb \to \infty$, such that $nb^4 \to 0$. Then*

$$(b\sigma^2(p))^{-1/2}\left\{(nb)^{p/2}J_n^c(p) - m(p)\right\} \xrightarrow{d} N(0,1),$$

where $\sigma^2(p)$ is defined in (9) and

$$m(p) = \int_{\mathbb{R}} |x|^p\phi(x)\,\mathrm{d}x\left(\int k^2(t)\,\mathrm{d}t\right)^{p/2}\int_0^1 |L'(t)|^{p/2}\,\mathrm{d}\mu(t).$$

Proof. The statement follows from Theorem 3.1, if $|m_n^c(p) - m(p)| = o(b^{1/2})$. First we note that $\int_0^b |L'(t)|^{p/2}\,\mathrm{d}\mu(t) = o(b^{1/2})$ and $\int_{1-b}^1 |L'(t)|^{p/2}\,\mathrm{d}\mu(t) = o(b^{1/2})$. Moreover, according to (14), for each $x \in \mathbb{R}$, we have

$$\begin{aligned}
&\int_b^{1-b}\left|\left|\sqrt{L'(t)}Dx + g_{(n)}(t)\right|^p - \left|\sqrt{L'(t)}Dx\right|^p\right|\mathrm{d}\mu(t) \\
&\quad\leq p2^{p-1}\int_b^{1-b}\left|g_{(n)}(t)\right|^p\,\mathrm{d}\mu(t) \\
&\qquad + p2^{p-1}\left(\int_b^{1-b}\left|\sqrt{L'(t)}Dx\right|^p\,\mathrm{d}\mu(t)\right)^{1-1/p}\left(\int_b^{1-b}\left|g_{(n)}(t)\right|^p\,\mathrm{d}\mu(t)\right)^{1/p},
\end{aligned}$$

where $g_{(n)}(t)$ is defined in (4). Hence, it suffices to prove

$$b^{-p/2}\int_b^{1-b}\left|g_{(n)}(t)\right|^p\,\mathrm{d}\mu(t)=o(1).$$

This follows, since $\sup_{t\in[0,1]}\left|g_{(n)}(t)\right| = O((nb)^{1/2}b^2)$ and $b^{-p/2}(nb)^{p/2}b^{2p} = (nb^4)^{p/2}\to 0$. □

3.2. Boundary problems of the standard kernel estimator

We show that, actually, we cannot extend the results of Theorem 3.1 to the whole interval $[0,1]$, because then the inconsistency at the boundaries dominates the L_p-error. A similar phenomenon was also observed in the case of the Grenander-type estimator (see [13] and [31]), but only for $p\geq 2.5$. In our case the contribution of the boundaries to the L_p-error is not negligible for all $p\geq 1$. This mainly has to do with the fact that the functions $g_{(n)}$, defined in (4), diverge to infinity. As a result, all the previous theory, which relies on the fact that $g_{(n)}=O(1)$ does not hold. For example, for $t\in[0,b)$, we have

$$\begin{aligned} g_{(n)}(t) &= (nb)^{1/2}\int_0^{t+b}k_b(t-u)\,\mathrm{d}\Lambda(u)-\lambda(t)\\ &= (nb)^{1/2}\int_{-1}^{t/b}k(y)[\lambda(t-by)-\lambda(t)]\,\mathrm{d}y-(nb)^{1/2}\lambda(t)\int_{t/b}^1 k(y)\,\mathrm{d}y \qquad (20)\\ &= (nb)^{1/2}\left\{\int_{-1}^{t/b}k(y)[\lambda(t-by)-\lambda(t)]\,\mathrm{d}y-\lambda(t)\int_{t/b}^1 k(y)\,\mathrm{d}y\right\}. \end{aligned}$$

For the first term within the brackets, we have

$$\left|\int_{-1}^{t/b}k(y)[\lambda(t-by)-\lambda(t)]\,\mathrm{d}y\right|\leq b\sup_{t\in[0,1]}|\lambda'(t)|\left|\int_{-1}^{t/b}k(y)y\,\mathrm{d}y\right|=O(b), \qquad (21)$$

whereas for any $0<c<1$ and $t\in[0,cb]$,

$$0<\inf_{t\in[0,1]}\lambda(t)\int_c^1 k(y)\,\mathrm{d}y\leq\lambda(t)\int_{t/b}^1 k(y)\,\mathrm{d}y\leq\lambda(0). \qquad (22)$$

Because $nb\to\infty$, this would mean that

$$\sup_{t\in[0,cb]}g_{(n)}(t)\to-\infty. \qquad (23)$$

What would solve the problem is to assume that λ is twice differentiable as a function defined on $\mathbb{R}$ (see [10] and [9]). This is not the case, because here we are considering a function which is positive and decreasing on $[0,1]$ and usually is zero outside this interval. This means that as a function on $\mathbb{R}$, λ is not monotone anymore and has at least one discontinuity point.

The following results indicate that inconsistency at the boundaries dominates the L_p-error, i.e., the expectation and the variance of the integral close to the end points of the support diverge to infinity. We cannot even approach the boundaries at a rate faster than b (as in the case of the Grenander-type estimator), because the kernel estimator is inconsistent on the whole interval $[0, b)$ (and $(1 - b, 1]$).

Proposition 3.3. *Assume that (A1)-(A3) hold and let $\tilde{\lambda}_n^s$ be defined in (2). Let k satisfy (1). Suppose that $p \geq 1$ and $nb \to \infty$.*

i) When $nb^3 \to \infty$, then for each $p \geq 1$,

$$(nb)^{p/2}\mathbb{E}\left[\int_0^b \left|\tilde{\lambda}_n^s(t) - \lambda(t)\right|^p \,\mathrm{d}\mu(t)\right] \to \infty;$$

ii) If $bn^{1-1/p} \to 0$, then

$$b^{-1/2}\left\{\int_0^b (nb)^{p/2}\left|\tilde{\lambda}_n^s(t) - \lambda(t)\right|^p \,\mathrm{d}\mu(t) - \int_0^b \left|g_{(n)}(t)\right|^p \,\mathrm{d}\mu(t)\right\} \to 0,$$

where $g_{(n)}$ is defined in (4);

iii) Let

$$Y_n(t) = b^{1/2}\int_0^{t+b} k_b(t-u)\,\mathrm{d}B_n(L(u)), \qquad t \in [0, b]. \tag{24}$$

If $b^{-1}n^{-1+1/q} = O(1)$ and $b^{p-1}n^{p-2+2/q} \to 0$, then

$$b^{-1/2}\left|\int_0^b (nb)^{p/2}\left|\tilde{\lambda}_n^s(t) - \lambda(t)\right|^p \,\mathrm{d}\mu(t) - \int_0^b \left|Y_n(t) + g_{(n)}(t)\right|^p \,\mathrm{d}\mu(t)\right| \to 0, \tag{25}$$

in probability and when $bn^{1-1/p} \to \infty$, then for all $0 < c < 1$,

$$b^{-1}\mathrm{Var}\left(\int_0^{cb} \left|Y_n(t) + g_{(n)}(t)\right|^p \,\mathrm{d}\mu(t)\right) \to \infty,$$

where $g_{(n)}$ is defined in (4).

The previous results also hold if we consider the integral on $(1 - b, 1]$ instead of $[0, b)$.

The proof can be found in Appendix A.

Remark 3.4. Note that, if $b \sim n^{-\alpha}$, for some $0 < \alpha < 1$, then for $\alpha < 1/3$, Proposition 3.3(i) shows that for all $p \geq 1$, the expectation of the boundary regions in the L_p-error tends to infinity. This holds in particular for the optimal choice $\alpha = 1/5$. For $p < 1/(1-\alpha)$, Proposition 3.3(ii) allows us to include the boundary regions in the central limit theorem for the L_p-error of the kernel estimator,

$$(b\sigma^2(p))^{-1/2}\left\{(nb)^{p/2}J_n(p) - \bar{m}_n(p)\right\} \xrightarrow{d} N(0,1),$$

with $J_n(p)$ defined in (16) and $\bar{m}_n(p) = \int_0^1 \left|g_{(n)}(t)\right|^p \,\mathrm{d}\mu(t)$. However, the bias term $\bar{m}_n(p)$ is not bounded anymore. On the other hand, if $p > 1/(1-\alpha)$, Proposition 3.3(iii) shows that the boundary regions in the L_p-error behave asymptotically as random variables whose variance tends to infinity.

Remark 3.5. The choice of the measure μ instead of the Lebesgue measure, in [10] and [9], is motivated by the fact that, for a particular $\mu(t) = w(t)\mathrm{d}t$, the normalizing constants $m(p)$ and $\sigma(p)$ in the CLT will not depend on the unknown function. In our case, a proper choice for μ can also be used to get rid of the boundary problems. This happens when μ puts less mass on the boundary regions in order to compensate the inconsistency of the kernel estimator. For example, if $\mu(t) = t^{2p}(1-t)^{2p}\mathrm{d}t$, then

$$\int_0^b |g_{(n)}(t)|^p \,\mathrm{d}\mu(t) + \int_{1-b}^1 |g_{(n)}(t)|^p \,\mathrm{d}\mu(t) \to 0$$

and, as a result, Theorem 3.1 also holds if we replace $J_n^c(p)$ with $J_n(p)$, defined in (16).

3.3. Kernel estimator with boundary correction

One way to overcome the inconsistency problems of the standard kernel estimator is to apply some boundary correction. Let now $\hat{\lambda}_n^s$ be the 'corrected' kernel estimator of λ, i.e.

$$\hat{\lambda}_n^s(x) = \int_{x-b}^{x+b} k_b^{(x)}(x-u)\,\mathrm{d}\Lambda_n(u), \quad \text{for } x \in [0,1], \tag{26}$$

where $k_b^{(x)}(u)$ denotes the rescaled kernel $b^{-1}k^{(x)}(u/b)$, with

$$k^{(x)}(u) = \begin{cases} \psi_1\left(\frac{x}{b}\right) k(u) + \psi_2\left(\frac{x}{b}\right) uk(u) & x \in [0,b) \\ k(u) & x \in [b, 1-b] \\ \psi_1\left(\frac{1-x}{b}\right) k(u) - \psi_2\left(\frac{1-x}{b}\right) uk(u) & x \in (1-b, 1]. \end{cases} \tag{27}$$

For $s \in [-1,1]$, the coefficients $\psi_1(s)$, $\psi_2(s)$ are determined by

$$\psi_1(s)\int_{-1}^s k(u)\,\mathrm{d}u + \psi_2(s)\int_{-1}^s uk(u)\,\mathrm{d}u = 1$$
$$\psi_1(s)\int_{-1}^s uk(u)\,\mathrm{d}u + \psi_2(s)\int_{-1}^s u^2k(u)\,\mathrm{d}u = 0.$$

As a result, the boundary corrected kernel satisfies

$$\int_{-1}^{x/b} k^{(x)}(u)\,\mathrm{d}u = 1 \qquad \text{and} \qquad \int_{-1}^{x/b} uk^{(x)}(u)\,\mathrm{d}u = 0. \tag{28}$$

Moreover, ψ_1 and ψ_2 are continuously differentiable (in particular they are bounded). We aim at showing that in this case, Theorem 3.1 holds for the L_p-error on the whole support, i.e., with $J_n(p)$ instead of $J_n^c(p)$.

Note that boundary corrected kernel estimator coincides with the standard kernel estimator on $[b, 1-b]$. Hence the behavior of the L_p-error on $[b, 1-b]$ will be the same. We just have to deal with the boundary regions $[0, b]$ and $[1-b, 1]$.

Proposition 3.6. *Assume that (A1)-(A3) hold and let $\hat{\lambda}_n^s$ be defined in (26). Let k satisfy (1) and suppose $p \geq 1$ and $nb \to \infty$. Then*

$$b^{-1/2}(nb)^{p/2} \int_0^b \left|\hat{\lambda}_n^s(t) - \lambda(t)\right|^p \mathrm{d}\mu(t) \xrightarrow{\mathbb{P}} 0.$$

The previous result also holds if we consider the integral on $(1-b, 1]$ instead of $[0, b)$.

The proof can be found in Appendix A.

Corollary 3.7. *Assume that (A1)-(A3) hold and let $J_n(p)$ be defined in (16). Let k satisfy (1) and suppose $p \geq 1$ and $nb \to \infty$. Then*

i) if $nb^5 \to 0$, then it holds

$$(b\sigma^2(p))^{-1/2} \left\{(nb)^{p/2} J_n(p) - m_n(p)\right\} \xrightarrow{d} N(0,1);$$

ii) If $nb^5 \to C_0^2 > 0$ and B_n in Assumption (A2) is a Brownian motion, then it holds

$$(b\theta^2(p))^{-1/2} \left\{(nb)^{p/2} J_n(p) - m_n(p)\right\} \xrightarrow{d} N(0,1);$$

iii) If $nb^5 \to C_0^2 > 0$ and B_n in Assumption (A2) is a Brownian bridge, then it holds

$$(b\tilde{\theta}^2(p))^{-1/2} \left\{(nb)^{p/2} J_n(p) - m_n(p)\right\} \xrightarrow{d} N(0,1),$$

where σ^2, θ^2, $\tilde{\theta}^2$ and m_n are defined respectively in (9), (11), (12) and (8).

Proof. It follows from combining Theorem 3.1 and Proposition 3.6, together with the fact that

$$b^{-1/2} \int_{\mathbb{R}} \int_0^b \left|\sqrt{L'(t)} Dx + g_{(n)}(t)\right|^p w(t)\phi(x)\,\mathrm{d}t\,\mathrm{d}x \to 0,$$

where D and $g_{(n)}$ are defined in (7) and (4). □

4. Smoothed Grenander-type estimator

The smoothed Grenander-type estimator is defined by

$$\tilde{\lambda}_n^{SG}(t) = \int_{0 \vee (t-b)}^{1 \wedge (t+b)} k_b^{(t)}(t-u)\,\mathrm{d}\tilde{\Lambda}_n(u), \qquad \text{for } t \in [0,1], \tag{29}$$

where $\tilde{\Lambda}_n$ is the least concave majorant of Λ_n. We are interested in the asymptotic distribution of the L_p-error of this estimator:

$$I_n^{SG}(p) = \int_0^1 \left|\tilde{\lambda}_n^{SG}(t) - \lambda(t)\right|^p \,\mathrm{d}\mu(t). \tag{30}$$

We will compare the behavior of the L_p-error of $\tilde{\lambda}_n^{SG}$ with that of the regular kernel estimator $\hat{\lambda}_n^s$ from (26). Because

$$\tilde{\lambda}_n^{SG}(t) - \hat{\lambda}_n^s(t) = \int k_b^{(t)}(t-u)\,\mathrm{d}(\tilde{\Lambda}_n - \Lambda_n)(u),$$

we will make use of the behavior of $\tilde{\Lambda}_n - \Lambda_n$, which has been investigated in [38], extending similar results from [16] and [33]. The idea is to represent $\tilde{\Lambda}_n - \Lambda_n$ in terms of the mapping CM_I that maps a function $h : \mathbb{R} \to \mathbb{R}$ into the least concave majorant of h on the interval $I \subset \mathbb{R}$, or equivalently by the mapping $\mathrm{D}_h = \mathrm{CM}_I h - h$.

Let B_n be as in assumption (A2) and ξ_n a $N(0,1)$ distributed r.v. independent of B_n. Define versions W_n of Brownian motion by

$$W_n(t) = \begin{cases} B_n(t) + \xi_n t & \text{if } B_n \text{ is a Brownian bridge} \\ B_n(t) & \text{if } B_n \text{ is a Brownian motion.} \end{cases} \tag{31}$$

Define

$$\begin{aligned} A_n^E &= n^{2/3}\left(\mathrm{CM}_{[0,1]}\Lambda_n - \Lambda_n\right) = n^{2/3}\mathrm{D}_{[0,1]}\Lambda_n, \\ A_n^W &= n^{2/3}\left(\mathrm{CM}_{[0,1]}\Lambda_n^W - \Lambda_n^W\right) = n^{2/3}\mathrm{D}_{[0,1]}\Lambda_n^W. \end{aligned} \tag{32}$$

where

$$\Lambda_n^W(t) = \Lambda(t) + n^{-1/2}W_n(L(t)), \tag{33}$$

with L as in Assumption (A2). We start with the following result on the L_p-distance between $\tilde{\lambda}_n^{SG}$ and $\hat{\lambda}_n^s$. In order to use results from [38], we need that $1 \le p < \min(q, 2q-7)$, where q is from Assumption (A2). Moreover, in order to obtain suitable approximations in combination with results from [38], we require additional conditions on the rate at which $1/b$ tends to infinity. Also see Remark 4.2. For the optimal rate $b \sim n^{-1/5}$, the result in Theorem 4.1 is valid, as long as $p < 5$ and $q > 9$.

Theorem 4.1. *Assume that (A1) − (A2) hold and let μ be a finite measure on $(0,1)$. Let k satisfy (1) and let $\tilde{\lambda}_n^{SG}$ and $\hat{\lambda}_n^s$ be defined in (29) and (26), respectively. If $1 \le p < min(q, 2q-7)$ and $nb \to \infty$, such that $1/b = o\left(n^{1/3-1/q}\right)$, $1/b = o\left(n^{(q-3)/(6p)}\right)$, and $1/b = o\left(n^{1/6+1/(6p)}(\log n)^{-(1/2+1/(2p))}\right)$, then*

$$n^{2/3}\left(\int_b^{1-b}\left|\tilde{\lambda}_n^{SG}(t) - \hat{\lambda}_n^s(t)\right|^p \,\mathrm{d}\mu(t)\right)^{1/p} \xrightarrow{d} \alpha_0[\mathrm{D}_{\mathbb{R}}Z](0),$$

where $Z(t) = W(t) - t^2$, *with* W *being a two-sided Brownian motion originating from zero, and*

$$\alpha_0 = \left(\int_0^1 \left| \frac{c_1'(t)}{c_1(t)^2} \right|^p \mathrm{d}\mu(t) \right)^{1/p}, \qquad c_1(t) = \left| \frac{\lambda'(t)}{2L'(t)^2} \right|^{1/3}.$$

Proof. We write

$$n^{2/3} \left(\int_b^{1-b} \left| \tilde{\lambda}_n^{SG}(t) - \hat{\lambda}_n^s(t) \right|^p \mathrm{d}\mu(t) \right)^{1/p} = b^{-1} \left(\int_b^{1-b} |Y_n(t)|^p \, \mathrm{d}\mu(t) \right)^{1/p},$$

where

$$Y_n(t) = bn^{2/3} \left(\int_{t-b}^{t+b} k_b(t-u) \, \mathrm{d}(\tilde{\Lambda}_n - \Lambda_n)(u) \right), \qquad t \in (b, 1-b). \tag{34}$$

We first show that

$$b^{-p} \int_b^{1-b} |Y_n(t)|^p \, \mathrm{d}\mu(t) \xrightarrow{d} \alpha_0^p [D_{\mathbb{R}} Z](0)^p, \tag{35}$$

and then the result would follow from the continuous mapping theorem. Note that integration by parts yields

$$Y_n(t) = \frac{1}{b} \int_{-1}^{1} k' \left(\frac{t-v}{b} \right) A_n^E(v) \, \mathrm{d}v.$$

The proof consists of several succeeding approximations of A_n^E. For details, see Lemmas 8.2 to 8.6. First we replace A_n^E in the previous integral by A_n^W. The approximation of $Y_n(t)$ by

$$Y_n^{(1)}(t) = \frac{1}{b} \int_{-1}^{1} k' \left(\frac{t-v}{b} \right) A_n^W(v) \, \mathrm{d}v. \tag{36}$$

where A_n^W is defined in (32), is possible thanks to Assumption (A2). According to (14),

$$\begin{aligned} &\left| \int_b^{1-b} |Y_n(t)|^p \, \mathrm{d}\mu(t) - \int_b^{1-b} |Y_n^{(1)}(t)|^p \, \mathrm{d}\mu(t) \right| \\ &\leq p 2^{p-1} \int_b^{1-b} |Y_n(t) - Y_n^{(1)}(t)|^p \, \mathrm{d}\mu(t) \\ &\quad + p 2^{p-1} \left(\int_b^{1-b} |Y_n(t) - Y_n^{(1)}(t)|^p \, \mathrm{d}\mu(t) \right)^{1/p} \left(\int_b^{1-b} |Y_n^{(1)}(t)|^p \, \mathrm{d}\mu(t) \right)^{1-1/p}. \end{aligned} \tag{37}$$

According to Lemma 8.2, $b^{-p}\int_b^{1-b}|Y_n(t)-Y_n^{(1)}(t)|^p\,\mathrm{d}\mu(t) = o_P(1)$. Consequently, in view of (37), if we show that

$$b^{-p}\int_b^{1-b}|Y_n^{(1)}(t)|^p\,\mathrm{d}\mu(t)\xrightarrow{d}\alpha_0{}^p[\mathrm{D}_{\mathbb{R}}Z](0)^p, \tag{38}$$

then we obtain

$$b^{-p}\int_b^{1-b}|Y_n(t)|^p\,\mathrm{d}\mu(t) = b^{-p}\int_b^{1-b}|Y_n^{(1)}(t)|^p\,\mathrm{d}\mu(t)+o_P(1), \tag{39}$$

and (35) follows.

In order to prove (38), we replace A_n^W by $n^{2/3}\mathrm{D}_{I_{nv}}\Lambda_n^W$, i.e., we approximate $Y_n^{(1)}$ by

$$Y_n^{(2)}(t)=\frac{1}{b}\int_{t-b}^{t+b}k'\left(\frac{t-v}{b}\right)n^{2/3}[\mathrm{D}_{I_{nv}}\Lambda_n^W](v)\,\mathrm{d}v. \tag{40}$$

where $I_{nv}=[0,1]\cap[v-n^{-1/3}\log n, v+n^{-1/3}\log n]$ and Λ^W is defined in (33). From Lemma 8.3, we have that $b^{-p}\int_b^{1-b}|Y_n^{(1)}(t)-Y_n^{(2)}(t)|^p\,\mathrm{d}\mu(t) = o_P(1)$. Hence, similar to the argument that leads to (39), if we show that

$$b^{-p}\int_b^{1-b}|Y_n^{(2)}(t)|^p\,\mathrm{d}\mu(t)\xrightarrow{d}\alpha_0{}^p[D_{\mathbb{R}}Z](0)^p, \tag{41}$$

then, together with (14), it follows that

$$b^{-p}\int_b^{1-b}|Y_n^{(1)}(t)|^p\,\mathrm{d}\mu(t) = b^{-p}\int_b^{1-b}|Y_n^{(2)}(t)|^p\,\mathrm{d}\mu(t)+o_P(1).$$

Consequently, (38) is equivalent to (41).

In order to prove (41), let

$$Y_{nv}(s)=n^{1/6}\left[W_n(L(v+n^{-1/3}s))-W_n(L(v))\right]+\frac{1}{2}\lambda'(v)s^2. \tag{42}$$

Let $H_{nv}=[-n^{1/3}v, n^{1/3}(1-v)]\cap[-\log n,\log n]$ and

$$\Delta_{nv}=n^{2/3}[\mathrm{D}_{I_{nv}}\Lambda_n^W](v)-[\mathrm{D}_{H_{nv}}Y_{nv}](0).$$

We approximate $Y_n^{(2)}$ by

$$Y_n^{(3)}(t)=\frac{1}{b}\int_{t-b}^{t+b}k'\left(\frac{t-v}{b}\right)[\mathrm{D}_{H_{nv}}Y_{nv}](0)\,\mathrm{d}v. \tag{43}$$

From Lemma 8.4, we have that $b^{-p}\int_b^{1-b}|Y_n^{(2)}(t)-Y_n^{(3)}(t)|^p\,\mathrm{d}\mu(t) = o_P(1)$. Again, similar to the argument that leads to (39), if we show that

$$b^{-p}\int_b^{1-b}|Y_n^{(3)}(t)|^p\,\mathrm{d}\mu(t)\xrightarrow{d}\alpha_0{}^p[\mathrm{D}_{\mathbb{R}}Z](0)^p. \tag{44}$$

then, together with (14), it follows that

$$b^{-p}\int_b^{1-b}|Y_n^{(2)}(t)|^p\,\mathrm{d}\mu(t)=b^{-p}\int_b^{1-b}|Y_n^{(3)}(t)|^p\,\mathrm{d}\mu(t)+o_P(1),$$

which would prove (41).

We proceed with proving (44). Let W be a two sided Brownian motion originating from zero. We have that

$$n^{1/6}\left[W_n(L(v+n^{-1/3}s))-W_n(L(v))\right]\stackrel{d}{=}W\left(n^{1/3}(L(v+n^{-1/3}s)-L(v))\right)$$

as a process in s. Consequently,

$$Y_n^{(3)}(t)\stackrel{d}{=}\frac{1}{b}\int_{t-b}^{t+b}k'\left(\frac{t-v}{b}\right)[\mathrm{D}_{H_{nv}}\tilde{Y}_{nv}](0)\,\mathrm{d}v$$

where

$$\tilde{Y}_{nv}(s)=W(n^{1/3}(L(v+n^{-1/3}s)-L(v)))+\frac{1}{2}\lambda'(v)s^2. \tag{45}$$

Now define

$$Z_{nv}(s)=W(L'(v)s)+\frac{1}{2}\lambda'(v)s^2. \tag{46}$$

and $J_{nv}=[n^{1/3}(L(a_{nv})-L(v))/L'(v),n^{1/3}(L(b_{nv})-L(v))/L'(v)]$, where $a_{nv}=\max(0,v-n^{-1/3}\log n)$ and $b_{nv}=\min(1,v+n^{-1/3}\log n)$. We approximate $\tilde{Y}_{nv}$ by Z_{nv}, i.e., we approximate $Y_n^{(3)}$ by

$$Y_n^{(4)}(t)=\frac{1}{b}\int_{t-b}^{t+b}k'\left(\frac{t-v}{b}\right)[\mathrm{D}_{J_{nv}}Z_{nv}](0)\,\mathrm{d}v, \tag{47}$$

Lemma 8.5 yields $b^{-p}\int_b^{1-b}|Y_n^{(3)}(t)-Y_n^{(4)}(t)|^p\,\mathrm{d}\mu(t)=o_P(1)$. Once more, similar to the argument that leads to (39), if we show that

$$b^{-p}\int_b^{1-b}|Y_n^{(4)}(t)|^p\,\mathrm{d}\mu(t)\xrightarrow{d}{\alpha_0}^p[\mathrm{D}_{\mathbb{R}}Z](0)^p, \tag{48}$$

then, together with (14), it follows that

$$b^{-p}\int_b^{1-b}|Y_n^{(3)}(t)|^p\,\mathrm{d}\mu(t)=b^{-p}\int_b^{1-b}|Y_n^{(4)}(t)|^p\,\mathrm{d}\mu(t)+o_P(1),$$

and as a result, also (44) holds.

As a final step, we prove (48). Since $c_1(v)W\left(L'(v)c_2(v)s\right)\stackrel{d}{=}W(s)$ as a process in s, where

$$c_1(v)=\left(\frac{|\lambda'(v)|}{2L'(v)^2}\right)^{1/3},\qquad c_2(v)=\left(\frac{4L'(v)}{|\lambda'(v)|^2}\right)^{1/3} \tag{49}$$

we obtain that

$$Y_n^{(4)}(t) \stackrel{d}{=} \frac{1}{b}\int_{t-b}^{t+b} k'\left(\frac{t-v}{b}\right)\frac{1}{c_1(v)}[\mathrm{D}_{I_{nv}}Z](0)\,\mathrm{d}v$$

where $I_{nv} = c_2(v)^{-1}J_{nv}$ and $Z(t) = W(t) - t^2$. We approximate $\mathrm{D}_{I_{nv}}$ by $\mathrm{D}_{\mathbb{R}}$, i.e., we approximate $Y_n^{(4)}$ by

$$Y_n^{(5)}(t) = [\mathrm{D}_{\mathbb{R}}Z](0)\frac{1}{b}\int_{t-b}^{t+b} k'\left(\frac{t-v}{b}\right)\frac{1}{c_1(v)}\,\mathrm{d}v. \tag{50}$$

It remains to show that

$$b^{-p}\int_b^{1-b}|Y_n^{(5)}(t)|^p\,\mathrm{d}\mu(t) \xrightarrow{d} \alpha_0^p[\mathrm{D}_{\mathbb{R}}Z](0)^p, \tag{51}$$

because then, it follows that

$$b^{-p}\int_b^{1-b}|Y_n^{(4)}(t)|^p\,\mathrm{d}\mu(t) = b^{-p}\int_b^{1-b}|Y_n^{(5)}(t)|^p\,\mathrm{d}\mu(t) + o_P(1)$$

so that (48) holds. Since

$$\frac{1}{b}\int_{t-b}^{t+b} k'\left(\frac{t-v}{b}\right)\frac{1}{c_1(t)}\,\mathrm{d}v = \frac{1}{c_1(t)}\int_{-1}^{1} k'(y)\,\mathrm{d}y = 0.$$

we can write

$$\begin{aligned}\frac{1}{b}\int_{t-b}^{t+b} k'\left(\frac{t-v}{b}\right)\frac{1}{c_1(v)}\,\mathrm{d}v &= \frac{1}{b}\int_{t-b}^{t+b} k'\left(\frac{t-v}{b}\right)\left(\frac{1}{c_1(v)} - \frac{1}{c_1(t)}\right)\mathrm{d}v\\ &= \int_{-1}^{1} k'(y)\left(\frac{1}{c_1(t-by)} - \frac{1}{c_1(t)}\right)\mathrm{d}y.\end{aligned}$$

Assumptions (A1) and (A2) imply that $t \mapsto c_1(t)$ is strictly positive and continuously differentiable with bounded derivative, so by a Taylor expansion we get

$$\int_{-1}^{1} k'(y)\left(\frac{1}{c_1(t-by)} - \frac{1}{c_1(t)}\right)\mathrm{d}y = \frac{c_1'(t)}{c_1(t)^2}b\int_{-1}^{1} k'(y)\,y\,\mathrm{d}y + o(b).$$

Hence,

$$\begin{aligned}b^{-p}\int_b^{1-b}|Y_n^{(5)}(t)|^p\,\mathrm{d}\mu(t) &= [\mathrm{D}_{\mathbb{R}}Z](0)^p b^{-p}\int_b^{1-b}\left|\frac{c_1'(t)b}{c_1(t)^2}\right|^p\,\mathrm{d}\mu(t) + o_P(1)\\ &= [\mathrm{D}_{\mathbb{R}}Z](0)^p\int_0^1\left|\frac{c_1'(t)}{c_1(t)^2}\right|^p\,\mathrm{d}\mu(t) + o_P(1)\end{aligned} \tag{52}$$

which concludes the proof of (51) and finishes the proof of the theorem. □

Remark 4.2. Note that the assumption $1/b = o\left(n^{1/6+1/(6p)}(\log n)^{-(1+1/p)}\right)$ of the previous theorem puts a restriction on p, when b has the optimal rate $n^{-1/5}$. This is due to the approximation of $Y_n^{(4)}(t)$ by $Y_n^{(5)}(t)$ for $t \in (b, 1-b)$. This restriction on p can be avoided if we consider the L_p-error on the smaller interval $(b + n^{-1/3}\log n, 1 - b - n^{-1/3}\log n)$.

Remark 4.3. For $p > 1$, the boundary regions cannot be included in the CLT of Theorem 4.1. For example, for $t \in (0, b)$, it can be shown that there exists a universal constant $K > 0$, such that

$$n^{2p/3}\int_0^b \left|\tilde{\lambda}_n^{SG}(t) - \tilde{\lambda}_n^{s}(t)\right|^p \,\mathrm{d}\mu(t) > Kb^{-p+1}[\mathrm{D}_{\mathbb{R}}Z](0)^p + o_P(b^{-p+1}),$$

which is not bounded in probability for $p > 1$. The same result also holds for $t \in (1-b, 1)$.

In the special case $p = 1$, for $t \in (0, b)$ we have

$$\begin{aligned} &n^{2/3}\int_0^b \left|\tilde{\lambda}_n^{SG}(t) - \tilde{\lambda}_n^{s}(t)\right| \,\mathrm{d}\mu(t) \\ &= [\mathrm{D}_{\mathbb{R}}Z](0)\frac{1}{b}\int_0^b \left|\frac{1}{c_1(t)}\int_{-1}^{t/b}\frac{\mathrm{d}}{\mathrm{d}y}k^{(t)}(y)\,\mathrm{d}y\right| \,\mathrm{d}\mu(t) + o_P(1). \end{aligned}$$

If (A3) holds, then

$$\frac{1}{b}\int_0^b \left|\frac{1}{c_1(t)}\int_{-1}^{t/b}\frac{\mathrm{d}}{\mathrm{d}y}k^{(t)}(y)\,\mathrm{d}y\right| \mathrm{d}\mu(t) \to \frac{w(0)}{c_1(0)}\int_0^1 \left|\psi_1(y)\,k(y) + \psi_2(y)\,yk(y)\right| \,\mathrm{d}y.$$

Similarly, we can deal with the case $t \in (1-b, 1)$. It follows that

$$n^{2/3}\int_0^1 \left|\tilde{\lambda}_n^{SG}(t) - \tilde{\lambda}_n^{s}(t)\right| \,\mathrm{d}\mu(t) \xrightarrow{d} \tilde{\alpha}_0[\mathrm{D}_{\mathbb{R}}Z](0)$$

with

$$\tilde{\alpha}_0 = \alpha_0 + \left(\frac{w(0)}{c_1(0)} + \frac{w(1)}{c_1(1)}\right)\int_0^1 \left|\psi_1(y)\,k(y) + \psi_2(y)\,yk(y)\right| \,\mathrm{d}y.$$

We are now ready to formulate the CLT for the smoothed Grenander-type estimator. The result will follow from combining Corollary 3.7 with Theorem 4.1. Because we now deal with the L_p-error between $\tilde{\lambda}_n^{SG}$ and λ, the contribution of the integrals over the boundary regions $(0, 2b)$ and $(1-2b, 1)$ can be shown to be negligible. This means we no longer need the third requirement in Theorem 4.1 on the rate of $1/b$.

Theorem 4.4. *Assume that (A1) − (A3) hold and let k satisfy (1). Let I_n^{SG} be defined in (30). If $1 \le p < min(q, 2q-7)$ and $nb \to \infty$, such that $1/b = o\left(n^{1/3-1/q}\right)$ and $1/b = o\left(n^{(q-3)/(6p)}\right)$.*

i) If $nb^5 \to 0$, then

$$(b\sigma^2(p))^{-1/2}\left\{(nb)^{p/2} I_n^{SG}(p) - m_n(p)\right\} \xrightarrow{d} N(0,1);$$

ii) If $nb^5 \to C_0^2 > 0$, and B_n in assumption (A2) is a Brownian motion, then

$$(b\theta^2(p))^{-1/2}\left\{(nb)^{p/2} I_n^{SG}(p) - m_n(p)\right\} \xrightarrow{d} N(0,1);$$

iii) If $nb^5 \to C_0^2 > 0$, and B_n in assumption (A2) is a Brownian bridge, then

$$(b\tilde{\theta}^2(p))^{-1/2}\left\{(nb)^{p/2} I_n^{SG}(p) - m_n(p)\right\} \xrightarrow{d} N(0,1),$$

where I_n^{SG}, m_n, σ^2, θ^2, and $\tilde{\theta}^2$ are defined in (30), (8), (9), (11), and (12), respectively.

Proof. Define

$$\gamma^2(p) = \begin{cases} \sigma^2(p) & \text{if} \quad nb^5 \to 0 \\ \theta^2(p) & \text{if} \quad nb^5 \to C_0^2. \end{cases} \tag{53}$$

By Corollary 3.7, we already have that

$$(b\gamma^2(p))^{-1/2}\left\{(nb)^{p/2}\int_0^1 \left|\hat{\lambda}_n^s(t) - \lambda(t)\right|^p \mathrm{d}\mu(t) - m_n(p)\right\} \xrightarrow{d} N(0,1),$$

for $\hat{\lambda}_n^s$ defined in (26). Hence it is sufficient to show that

$$b^{-1/2}(nb)^{p/2}\left|\int_0^1 \left|\tilde{\lambda}_n^{SG}(t) - \lambda(t)\right|^p \mathrm{d}\mu(t) - \int_0^1 \left|\hat{\lambda}_n^s(t) - \lambda(t)\right|^p \mathrm{d}\mu(t)\right| \xrightarrow{\mathbb{P}} 0,$$

in all three cases (i)-(iii). First we show that

$$b^{-1/2}(nb)^{p/2}\left|\int_0^{2b} \left|\tilde{\lambda}_n^{SG}(t) - \lambda(t)\right|^p \mathrm{d}\mu(t) - \int_0^{2b} \left|\hat{\lambda}_n^s(t) - \lambda(t)\right|^p \mathrm{d}\mu(t)\right| \xrightarrow{\mathbb{P}} 0. \tag{54}$$

Indeed, by (14), we get

$$\begin{aligned}
&\left|\int_0^{2b} \left|\tilde{\lambda}_n^{SG}(t) - \lambda(t)\right|^p \mathrm{d}\mu(t) - \int_0^{2b} \left|\hat{\lambda}_n^s(t) - \lambda(t)\right|^p \mathrm{d}\mu(t)\right| \\
&\leq p2^{p-1}\int_0^{2b} \left|\tilde{\lambda}_n^{SG}(t) - \hat{\lambda}_n^s(t)\right|^p \mathrm{d}\mu(t) \\
&\quad + p2^{p-1}\left(\int_0^{2b} \left|\tilde{\lambda}_n^{SG}(t) - \hat{\lambda}_n^s(t)\right|^p \mathrm{d}\mu(t)\right)^{1/p}\left(\int_0^{2b} \left|\hat{\lambda}_n^s(t) - \lambda(t)\right|^p \mathrm{d}\mu(t)\right)^{1-1/p}.
\end{aligned} \tag{55}$$

Moreover, by integration by parts and the Kiefer-Wolfowitz type of result in Corollary 3.1 in [15], it follows that

$$\sup_{t\in[0,1]}\left|\tilde{\lambda}_n^{SG}(t)-\hat{\lambda}_n^s(t)\right| = \sup_{t\in[0,1]}\left|\int k_b^{(t)}(t-u)\,\mathrm{d}(\tilde{\Lambda}_n-\Lambda_n)(u)\right| \\ \leq Cb^{-1}\sup_{t\in[0,1]}|\tilde{\Lambda}_n(t)-\Lambda_n(t)| = O_P\left(b^{-1}\left(\frac{\log n}{n}\right)^{2/3}\right). \tag{56}$$

Hence

$$\int_0^{2b}\left|\tilde{\lambda}_n^{SG}(t)-\hat{\lambda}_n^s(t)\right|^p \,\mathrm{d}\mu(t) = O_P\left(b^{1-p}\left(\frac{\log n}{n}\right)^{2p/3}\right). \tag{57}$$

Together with Proposition 3.6 this implies (54). Similarly, we also have

$$b^{-1/2}(nb)^{p/2}\left|\int_{1-2b}^{1}\left|\tilde{\lambda}_n^{SG}(t)-\lambda(t)\right|^p \,\mathrm{d}\mu(t)-\int_{1-2b}^{1}\left|\hat{\lambda}_n^s(t)-\lambda(t)\right|^p \,\mathrm{d}\mu(t)\right| \xrightarrow{\mathbb{P}} 0.$$

Thus, it remains to prove

$$b^{-1/2}(nb)^{p/2}\left|\int_{2b}^{1-2b}\left|\tilde{\lambda}_n^{SG}(t)-\lambda(t)\right|^p \,\mathrm{d}\mu(t)-\int_{2b}^{1-2b}\left|\hat{\lambda}_n^s(t)-\lambda(t)\right|^p \,\mathrm{d}\mu(t)\right| \xrightarrow{\mathbb{P}} 0. \tag{58}$$

Again, from (14), we have

$$\left|\int_{2b}^{1-2b}\left|\tilde{\lambda}_n^{SG}(t)-\lambda(t)\right|^p \,\mathrm{d}\mu(t)-\int_{2b}^{1-2b}\left|\hat{\lambda}_n^s(t)-\lambda(t)\right|^p \,\mathrm{d}\mu(t)\right| \\ \leq p2^{p-1}\left\{\int_{2b}^{1-2b}\left|\tilde{\lambda}_n^{SG}(t)-\hat{\lambda}_n^s(t)\right|^p \,\mathrm{d}\mu(t)\right. \\ \left. +\left(\int_{2b}^{1-2b}\left|\tilde{\lambda}_n^{SG}(t)-\hat{\lambda}_n^s(t)\right|^p \mathrm{d}\mu(t)\right)^{1/p}\left(\int_{2b}^{1-2b}\left|\hat{\lambda}_n^s(t)-\lambda(t)\right|^p \mathrm{d}\mu(t)\right)^{1-1/p}\right\}. \tag{59}$$

Because $b^{-1}=o(n^{1/3-1/q})$ implies that

$$(2b,1-2b)\subset(b+n^{-1/3}\log n,1-b-n^{-1/3}\log n),$$

from Theorem 4.1, in particular Remark 4.2, we have

$$\int_{2b}^{1-2b}\left|\tilde{\lambda}_n^{SG}(t)-\hat{\lambda}_n^s(t)\right|^p \,\mathrm{d}\mu(t) = O_P(n^{-2p/3}) = o_P(n^{-p/2}). \tag{60}$$

Then, (58) follows immediately from (59) and the fact that, according to Theorem 3.1,

$$\int_{2b}^{1-2b}\left|\hat{\lambda}_n^s(t)-\lambda(t)\right|^p \,\mathrm{d}\mu(t) = O_P((nb)^{-p/2}).$$

This proves the theorem. □

Remark 4.5. Note that, if $b = cn^{-\alpha}$, for some $0 < \alpha < 1$, the proof is simple and short in case $\alpha < p/(3(1+p))$ because the Kiefer-Wolfowitz type of result in Corollary 3.1 in [15] is sufficient to prove (60). Indeed, from (56), it follows that

$$\int_{2b}^{1-2b} \left|\tilde{\lambda}_n^{SG}(t) - \hat{\lambda}_n^s(t)\right|^p \,\mathrm{d}\mu(t) = O_P\left(b^{-p}\left(\frac{\log n}{n}\right)^{2p/3}\right) = o_P\left(b^{1/2}\,(nb)^{-p/2}\right).$$

However, this assumption on α is quite restrictive because for example if $\alpha = 1/5$ then the theorem holds only for $p > 3/2$ (not for the L_1-loss) and if $\alpha = 1/4$ then the theorem holds only for $p > 3$.

5. Isotonized kernel estimator

The isotonized kernel estimator is defined as follows. First, we smooth the piecewise constant estimator Λ_n by means of a boundary corrected kernel function, i.e., let

$$\Lambda_n^s(t) = \int_{(t-b)\vee 0}^{(t+b)\wedge 1} k_b^{(t)}(t-u)\Lambda_n(u)\,\mathrm{d}u, \quad \text{for } t \in [0,1], \tag{61}$$

where $k_b^{(t)}(u)$ defined as in (27). Next, we define a continuous monotone estimator $\tilde{\lambda}_n^{GS}$ of λ as the left-hand slope of the least concave majorant $\widehat{\Lambda}_n^s$ of Λ_n^s on $[0,1]$. In this way we define a sort of Grenander estimator based on a smoothed naive estimator for Λ. For this reason we use the superscript GS.

We are interested in the asymptotic distribution of the L_p-error of this estimator:

$$I_n^{GS}(p) = \int_0^1 \left|\tilde{\lambda}_n^{GS}(t) - \lambda(t)\right|^p \,\mathrm{d}\mu(t).$$

It follows from Lemma 1 in [23] (in the case of a decreasing function), that $\tilde{\lambda}_n^{GS}$ is continuous and is the unique minimizer of

$$\psi(\lambda) = \frac{1}{2}\int_0^1 \left(\lambda(t) - \tilde{\lambda}_n^s(t)\right)^2 \,\mathrm{d}t$$

over all nonincreasing functions λ, where $\tilde{\lambda}_n^s(t) = \mathrm{d}\Lambda_n^s(t)/\mathrm{d}t$. This suggests $\tilde{\lambda}_n^s(t)$ as a naive estimator for $\lambda_0(t)$. Note that, for $t \in [b, 1-b]$, from integration by parts we get

$$\tilde{\lambda}_n^s(t) = \frac{1}{b^2}\int_{t-b}^{t+b} k'\left(\frac{t-u}{b}\right)\Lambda_n(u)\,\mathrm{d}u = \int_{t-b}^{t+b} k_b(t-u)\,\mathrm{d}\Lambda_n(u), \tag{62}$$

i.e., $\tilde{\lambda}_n^s$ coincides with the usual kernel estimator of λ on the interval $[b, 1-b]$.

Let $0 < \gamma < 1$. It can be shown that

$$\mathbb{P}(\tilde{\lambda}_n^s(t) = \tilde{\lambda}_n^{GS}(t) \text{ for all } t \in [b^\gamma, 1-b^\gamma]) \to 1. \tag{63}$$

See Corollary B.2 in the Appendix. Hence, their L_p-error between $\tilde{\lambda}_n^{GS}$ and $\tilde{\lambda}_n^s$ will exhibit the same behavior in the limit. Note that this holds for every $\gamma < 1$, which means that the interval we are considering is approaching $(b, 1-b)$. Consider a modified L_p-error of the isotonized kernel estimator defined by

$$I_{n,\gamma}^{GS,c}(p) = \int_{b^\gamma}^{1-b^\gamma} \left|\tilde{\lambda}_n^{GS}(t) - \lambda(t)\right|^p \,\mathrm{d}\mu(t). \tag{64}$$

We then have the following result.

Theorem 5.1. *Assume that (A1)-(A3) hold and let $I_{n,\gamma}^{GS,c}(p)$ be defined in (64). Let k satisfy (1) and let L be as in Assumption (A2). Assume $b \to 0$ and $1/b = o(n^{1/4})$ and let $1/2 < \gamma < 1$.*

i) If $nb^5 \to 0$, then

$$(b\sigma^2(p))^{-1/2}\left\{(nb)^{p/2} I_{n,\gamma}^{GS,c}(p) - m_n(p)\right\} \xrightarrow{d} N(0,1);$$

ii) If $nb^5 \to C_0^2 > 0$ and B_n in assumption (A2) is a Brownian motion, then

$$(b\theta^2(p))^{-1/2}\left\{(nb)^{p/2} I_{n,\gamma}^{GS,c}(p) - m_n(p)\right\} \xrightarrow{d} N(0,1);$$

iii) If $nb^5 \to C_0^2 > 0$ and B_n in assumption (A2) is a Brownian bridge, then

$$(b\tilde{\theta}^2(p))^{-1/2}\left\{(nb)^{p/2} I_{n,\gamma}^{GS,c}(p) - m_n(p)\right\} \xrightarrow{d} N(0,1),$$

where σ^2, θ^2, $\tilde{\theta}^2$ and m_n are defined respectively in (9), (11), (12) and (8).

Proof. It follows from Theorem 3.1 and (63). Note that the results of Theorem 3.1 do not change if we consider the interval $[b^\gamma, 1-b^\gamma]$ instead of $[b, 1-b]$ and that $b^{-1/2}|m_n^c(p) - m_n(p)| \to 0$. ☐

6. Hellinger error

In this section we investigate the global behavior of estimators by means of a weighted Hellinger distance

$$H(\hat{\lambda}_n, \lambda) = \left(\frac{1}{2}\int_0^1 \left(\sqrt{\hat{\lambda}_n(t)} - \sqrt{\lambda(t)}\right)^2 \mathrm{d}\mu(t)\right)^{1/2}, \tag{65}$$

where $\hat{\lambda}_n$ is the estimator at hand. This metric is convenient in maximum likelihood problems, which goes back to [34, 35, 3]. Consistency in Hellinger distance of shape constrained maximum likelihood estimators has been investigated in [43], [45], and [11], whereas rates on Hellinger risk measures have been obtained in [45], [29], and [28]. The first central limit theorem type of result for the Hellinger distance was presented in [39] for Grenander type estimators of a monotone function. We deal with the smooth (isotonic) estimators following the same approach.

Note that, for the Hellinger distance to be well defined we need to assume that λ takes only positive values. We follow the same line of argument as in [39]. We first establish that

$$\int_0^1 \left(\sqrt{\hat{\lambda}_n^s(t)} - \sqrt{\lambda(t)}\right)^2 \mathrm{d}\mu(t) = \int_0^1 \left(\hat{\lambda}_n^s(t) - \lambda(t)\right)^2 (4\lambda(t))^{-1} \,\mathrm{d}\mu(t) + O_P\left((nb)^{-3/2}\right),$$

which shows that the squared Hellinger loss can be approximated by a weighted squared L_2-distance. For details, see Lemma C.1 in the Appendix, which is the corresponding version of Lemma 2.1 in [39]. Hence, a central limit theorem for squared the Hellinger loss follows directly from the central limit theorem for the weighted L_2-distance (see Theorem C.2 in the Appendix, which corresponds to Theorem 3.1 in [39]). An application of the delta method will then lead to the following result.

Theorem 6.1. *Assume (A1)-(A3) hold. Let $\tilde{\lambda}_n^s$ be defined in (2), with k satisfying (1), and let H be defined in (65). Suppose that $nb \to \infty$ and that λ is strictly positive.*

i) If $nb^5 \to 0$, then

$$\left(b\frac{\tau^2(2)}{8\mu_n(2)}\right)^{-1/2} \left\{(nb)^{1/2} H(\hat{\lambda}_n^s, \lambda) - 2^{-1/2}\mu_n(2)^{1/2}\right\} \xrightarrow{d} N(0,1).$$

ii) If $nb^5 \to C_0^2 > 0$ and B_n in Assumption (A2) is a Brownian motion, then

$$\left(b\frac{\kappa^2(2)}{8\mu_n(2)}\right)^{-1/2} \left\{(nb)^{1/2} H(\hat{\lambda}_n^s, \lambda) - 2^{-1/2}\mu_n(2)^{1/2}\right\} \xrightarrow{d} N(0,1),$$

iii) If $nb^5 \to C_0^2 > 0$ and B_n in Assumption (A2) is a Brownian bridge, then

$$\left(b\frac{\tilde{\kappa}^2(2)}{8\mu_n(2)}\right)^{-1/2} \left\{(nb)^{1/2} H(\hat{\lambda}_n^s, \lambda) - 2^{-1/2}\mu_n(2)^{1/2}\right\} \xrightarrow{d} N(0,1),$$

where τ^2, κ^2, $\tilde{\kappa}^2$ and μ_n are defined as in (9), (11), (12) and (8), respectively, by replacing $w(t)$ with $w(t)(4\lambda(t))^{-1}$.

(iv) Under the conditions of Theorem 4.4, results (i)-(iii) also hold when replacing $\hat{\lambda}_n^s$ by the smoothed Grenander-type estimator $\tilde{\lambda}_n^{SG}$, defined in (29).

Proof. The proof consists of an application of the delta-method in combination with Theorem C.2 in the Appendix. According to part (i) of Theorem C.2,

$$b^{-1/2}\left(2nbH(\hat{\lambda}_n^s, \lambda) - \mu_n(2)\right) \xrightarrow{d} Z$$

where Z is a mean zero normal random variable with variance $\tau^2(2)$. Therefore, in order to obtain part (i) of Theorem 6.1, we apply the delta method with the mapping $\phi(x) = 2^{-1/2}x^{1/2}$. Parts (ii)-(iv) are obtained in the same way. □

To be complete, note that from Corollary B.2, the previous central limit theorems also hold for the isotonized kernel estimator $\tilde{\lambda}_n^{GS}$, defined in Section 5, when considering a Hellinger distance corresponding to the interval $(b^\gamma, 1-b^\gamma)$ instead of $(0,1)$ in (65).

7. Testing

In this section we investigate a possible application of the results obtained in Section 4 for testing monotonicity. For example, Theorem 4.4 could be used to construct a test for the single null hypothesis $H_0 : \lambda = \lambda_0$, for some known monotone function λ_0. Instead, we investigate a nonparametric test for monotonicity on the basis of the L_p-distance between the smoothed Grenander-type estimator and the kernel estimator, see Theorem 4.1.

The problem of testing a nonparametric null hypothesis of monotonicity has gained a lot of interest in the literature (see for example [30] for the density setting, [27], [24] for the hazard rate, [1], [4], [5],[19] for the regression function).

We consider a regression model with deterministic design points

$$Y_i = \lambda\left(\frac{i}{n}\right) + \epsilon_i, \qquad i \in \{1,\dots,n\}, \tag{66}$$

where the ϵ_i's are independent normal random variables with mean zero and variance σ^2. Such a model satisfies Assumption (A2) with $q = +\infty$,

$$\Lambda_n(t) = n^{-1}\sum_{i\le nt} Y_i$$

and $L(t) = \sigma^2 t$, for $t \in [0,1]$ (see Theorem 5 in [13]).

Assume we have a sample of n observations $Y_1,\dots,Y_n$. Let $\mathcal{D}$ be the space of decreasing functions on $[0,1]$ that satisfy (A1). We want to test $H_0\colon \lambda \in \mathcal{D}$ against $H_1\colon \lambda \notin \mathcal{D}$. Under the null hypothesis we can estimate λ by the smoothed Grenander-type estimator $\tilde{\lambda}_n^{SG}$ defined as in (29). On the other hand, under the alternative hypothesis we can estimate λ by the kernel estimator with boundary corrections $\hat{\lambda}_n^s$ defined in (26). Then, as a test statistic we take

$$T_n = n^{2/3}\left(\int_b^{1-b}\left|\tilde{\lambda}_n^{SG}(t) - \hat{\lambda}_n^s(t)\right|^2 \mathrm{d}t\right)^{1/2},$$

and at level α, we reject the null hypothesis if $T_n > c_{n,\alpha}$ for some critical value $c_{n,\alpha} > 0$.

In order to use the asymptotic quantiles of the limit distribution in Theorem 4.1, we need to estimate the constant C_0 which depends on the derivatives of λ. To avoid this, we choose to determine the critical value by a bootstrap procedure. We generate $B = 1000$ samples of size n from the model (66) with λ replaced by its estimator $\tilde{\lambda}_n^{SG}$ under the null hypothesis and independent Gaussian errors with mean zero and variance $\hat{\sigma}_n^2$. As an estimator of σ^2, we take the same as in [1]

$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n/2}\left(Y_{2i} - Y_{2i-1}\right)^2,$$

where for simplicity of notation we are assuming that n is even. The bootstrap regression model considers

$$Y_i^* = \tilde{\lambda}_n^{SG}\left(\frac{i}{n}\right) + \epsilon_i^*$$

where ϵ_i's are independent mean zero normal random variables with variance $\hat{\sigma}_n^2$.

For each of these samples we compute the estimators $\tilde{\lambda}_n^{SG,*}$, $\hat{\lambda}_n^{s,*}$ and the test statistic

$$T_{n,j}^* = n^{2/3}\left(\int_b^{1-b}\left|\tilde{\lambda}_n^{SG,*}(t) - \hat{\lambda}_n^{s,*}(t)\right|^2 \mathrm{d}t\right)^{1/2}, \qquad j = 1, \dots, B.$$

Then as a critical value, we take the 100α-th upper-percentile of the values $T_{n,1}^*, \dots, T_{n,B}^*$. Consistency of the bootstrap method follows from the next theorem. A sketch of the proof is given in Appendix D. In order to keep the notation simple, we formulate the result only for the case considered in the simulation study. However, it also holds in the general setting considered in Section 4, under appropriate conditions on p, b and q.

Theorem 7.1. *Consider observations $Y_1, \dots, Y_n$ from the regression model (66) with regression function λ that satisfies (A1) and is three times differentiable with bounded third derivative. Let k satisfy (1) and $b = cn^{-\gamma}$ for $\gamma \in (1/6, 1/5)$. Let $\tilde{\lambda}_n^{SG,*}$ and $\hat{\lambda}_n^{s,*}$ be the smoothed Grenander estimator and the kernel estimator constructed from the bootstrap sample $Y_1^*, \dots, Y_n^*$. Then*

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}^*\left(n^{2/3}\left(\int_b^{1-b}\left|\tilde{\lambda}_n^{SG,*}(t) - \hat{\lambda}_n^{s,*}(t)\right|^2 \mathrm{d}t\right)^{1/2} \le x\right) - \Psi(x)\right| \to 0,$$

in probability, as $n \to \infty$, where $\mathbb{P}^*$ is the conditional probability given the observations and Ψ is the distribution function of $\alpha_0[D_{\mathbb{R}}Z](0)$, with Z and α_0 defined as in Theorem 4.1 with $\mu(t) = t$, $p = 2$ and $L'(t) = \sigma^2$.

Remark 7.2. If $b = cn^{-1/5}$ then the previous theorem holds with Ψ replaced by Ψ^*, which is the distribution function of $\alpha_0^*[D_{\mathbb{R}}Z](0)$, for a different constant α_0^*, depending on $\tilde{\lambda}_n^{SG}$ and $(\tilde{\lambda}_n^{SG})'$ instead of λ and λ'. This is because the bound on $|\alpha_0^* - \alpha_0|$ (see (119)) will not be sufficient to make the transition from α_0^* to α_0. The problem could be solved by using a rescaled test statistic

$$\alpha_0^{-1} n^{2/3} \left(\int_b^{1-b} \left| \tilde{\lambda}_n^{SG}(t) - \hat{\lambda}_n^s(t) \right|^2 \mathrm{d}t \right)^{1/2},$$

whose limit distribution is $[\mathrm{D}_{\mathbb{R}} Z](0)$ for both the original and bootstrap version. However, we do not choose this approach because we want to avoid estimation of α_0 which, in practice, is problematic.

To investigate the performance of the test in practice, we repeat this procedure $N = 1000$ times and we count the percentage of rejections. This gives an approximation of the level (or the power) of the test if we start with a sample for which the true λ is decreasing (or non-decreasing). We investigate the performance of the test by comparing it to tests proposed in [1], [2] and in [19]. For a power comparison, [1] and [2] consider the following functions

$$\begin{aligned}
&\lambda_1(x) = -15(x-0.5)^3 \mathbb{1}_{\{x\le 0.5\}} - 0.3(x-0.5) + \exp\left(-250(x-0.25)^2\right),\\
&\lambda_2(x) = 16\sigma x, \quad \lambda_3(x) = 0.2\exp\left(-50(x-0.5)^2\right), \quad \lambda_4(x) = -0.1\cos(6\pi x),\\
&\lambda_5(x) = -0.2x + \lambda_3(x), \quad \lambda_6(x) = -0.2x + \lambda_4(x),\\
&\lambda_7(x) = -(1+x) + 0.45\exp\left(-50(x-0.5)^2\right),
\end{aligned}$$

We denote by T_B the local mean test of [2] and S_n^{reg} the test proposed in [1] on the basis of the distance between the least concave majorant of Λ_n and Λ_n. The result of the simulations for $n = 100$, $\alpha = 0.05$, $b = 0.1$, are given in Table 1. We see that, apart from the last case, all the three tests perform very well and they are comparable. However, our test behaves much better for the function λ_7, which is more difficult to detect than the others.

Table 1
Simulated power of T_n, T_B and S_n^{reg} for $n = 100$.

Function	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7
σ^2	0.01	0.01	0.01	0.01	0.004	0.006	0.01
T_n	1	1	1	1	1	1	0.99
T_B	0.99	0.99	1	0.99	0.99	0.98	0.76
S_n^{reg}	0.99	1	0.98	0.99	0.99	0.99	0.68

The second model that we consider is taken from [1] and [19], which is a regression function given by

$$\lambda_a(x) = -(1+x) + a\exp\left(-50(x-0.5)^2\right), \qquad x \in [0,1].$$

The results of the simulation, again for $n = 100$, $\alpha = 0.05$, $b = 0.1$ and various values of a and σ^2 are given in Table 2. We denote by S_n^{reg} the test of [1] and by

T_{run} the test of [19]. Note that when $a = 0$, the regression function is decreasing so H_0 is satisfied. We observe that our test rejects the null hypothesis more often than T_{run} and S_n^{reg} so the price we pay for getting higher power is higher level. As the value of a increases, the monotonicity of λ_a is perturbed. For $a = 0.25$ our test performs significantly better than the other two and, as expected, the power decreases as the variance of the errors increases. When $a = 0.45$ and σ^2 not too large, the three test have power one but, when σ^2 increases, T_n outperforms T_{run} and S_n^{reg}. We took $b = 0.1$, which seems to be a reasonable one considering that the whole interval has length one.

Table 2
Simulated power of T_n, T_{run} and S_n^{reg} for $n = 100$.

	$a = 0$			$a = 0.25$			$a = 0.45$		
σ	0.025	0.05	0.1	0.025	0.05	0.1	0.025	0.05	0.1
T_n	0.012	0.025	0.022	0.927	0.497	0.219	1	1	0.992
T_{run}	0	0	0	0.106	0.037	0.014	1	1	0.805
S_n^{reg}	0	0.002	0.013	0.404	0.053	0.007	1	1	0.683

In what follows, we investigate how the behavior of the test depends on the choice of the bandwidth and of the L_p-distance. We perform simulation studies with true level 0.05, for $p \in \{1, 2, 5, 10\}$ and $b \in \{0.05, 0.1, 0.2\}$. To check the level of the test in practice we consider a constant function $\lambda(x) = 1$, $x \in (0, 1)$, which is the limiting case of monotonicity (the least favorable assumption). In terms of power we consider the regression function λ_a with $a = 0.25$. Results, for $N = 10000$ iterations, are shown in Tables 3 and 4. We note that the various L_p-distances behave similarly and none of them is strictly better than the others. With these choices of the bandwidth the practical level of the test is higher than 0.05, while the tests T_{run}, S_n^{reg}, T_B have guaranteed level since they are calibrated against the most difficult null hypothesis (constant function). However, we gain a lot in terms of power, as illustrated before in Table 2.

Table 3
Simulated level of T_n for $n = 100$ and $\lambda(x) = 1$.

	$\sigma = 0.025$				$\sigma = 0.1$			
b	$p = 1$	$p = 2$	$p = 5$	$p = 10$	$p = 1$	$p = 2$	$p = 5$	$p = 10$
0.05	0.126	0.124	0.105	0.097	0.123	0.123	0.109	0.099
0.1	0.103	0.103	0.095	0.092	0.107	0.107	0.099	0.096
0.2	0.089	0.089	0.093	0.093	0.092	0.094	0.097	0.098

It is not the purpose of this paper to investigate methods of bandwidth selection. However, one possibility is to use a cross validation procedure. In each of the N iterations we select the optimal bandwidth by a leave-1-out cross

Table 4
Simulated power of T_n for $n = 100$ and $\lambda_a(x)$, with $a = 0.25$.

	$\sigma = 0.025$				$\sigma = 0.1$			
b	$p = 1$	$p = 2$	$p = 5$	$p = 10$	$p = 1$	$p = 2$	$p = 5$	$p = 10$
0.05	0.771	0.756	0.719	0.690	0.101	0.123	0.131	0.129
0.1	0.925	0.929	0.910	0.867	0.198	0.204	0.202	0.196
0.2	0.989	0.998	1	1	0.418	0.423	0.429	0.424

validation procedure by means of the kernel estimator ([8]). Afterwards, this bandwidth is used for estimation in the B bootstrap samples. We report the results of the simulations for the two settings considered previously ($\lambda(x) = 1$ and $\lambda_a(x)$, $a = 0.25$) and various sample sizes in Tables 5 and 6 respectively. We observe that, even with this bandwidth selection method, the simulated level of the test is higher than the nominal level but again also the power is high (compared to the other tests considered previously). The bad performance in terms of level might be due to the fact that the constant function does not satisfy the assumptions of Theorems 4.1 and 7.1.

Table 5
Simulated level of T_n using cross-validation bandwidth selection and $\lambda(x) = 1$.

	$\sigma = 0.025$				$\sigma = 0.1$			
n	$p = 1$	$p = 2$	$p = 5$	$p = 10$	$p = 1$	$p = 2$	$p = 5$	$p = 10$
50	0.154	0.157	0.155	0.150	0.158	0.163	0.156	0.154
100	0.136	0.140	0.141	0.138	0.136	0.139	0.139	0.138
200	0.116	0.121	0.121	0.120	0.126	0.130	0.126	0.123

Table 6
Simulated power of T_n using cross-validation bandwidth selection and $\lambda_a(x)$, with $a = 0.25$.

	$\sigma = 0.025$				$\sigma = 0.1$			
n	$p = 1$	$p = 2$	$p = 5$	$p = 10$	$p = 1$	$p = 2$	$p = 5$	$p = 10$
50	0.534	0.523	0.366	0.229	0.239	0.252	0.240	0.228
100	0.895	0.895	0.869	0.828	0.324	0.331	0.317	0.307
200	0.993	0.992	0.989	0.986	0.410	0.412	0.396	0.390

Actually the percentage of rejections of the null hypothesis is close to 0.05 for a regression function $\lambda(x) = -0.1x + 1$ which has a small slope but is not constant (see Table 7). This means that the test manages to recognize quite well a true null hypothesis as long as the regression function is not flat (or with a very small slope).

Table 7
Percentage of rejections for T_n using cross-validation bandwidth selection when $\lambda(x) = -0.1x + 1$.

	$\sigma = 0.025$				$\sigma = 0.1$			
n	$p = 1$	$p = 2$	$p = 5$	$p = 10$	$p = 1$	$p = 2$	$p = 5$	$p = 10$
50	0.071	0.072	0.066	0.061	0.105	0.109	0.104	0.097
100	0.064	0.065	0.058	0.055	0.078	0.079	0.075	0.072
200	0.045	0.042	0.039	0.037	0.056	0.069	0.060	0.057

8. Auxiliary results and proofs

8.1. Proofs for Section 3

Lemma 8.1. *Let $L : [0,1] \to \mathbb{R}$ be strictly positive and twice differentiable, such that $\inf_{t\in[0,1]} L'(t) > 0$ and $\sup_{t\in[0,1]} |L''(t)| < \infty$. Let $\Gamma_n^{(2)}$, $g_{(n)}$, and $m_n^c(p)$ be defined in (19), (4), and (8), respectively. Assume that (A1) and (A3) hold.*

1. *If $nb^5 \to 0$, then*

$$(b\sigma^2(p))^{-1/2} \left\{ \int_b^{1-b} \left| b^{-1/2}\Gamma_n^{(2)}(t) + g_{(n)}(t) \right|^p \mathrm{d}\mu(t) - m_n^c(p) \right\} \xrightarrow{d} N(0,1),$$

where $\sigma^2(p)$ is defined in (9).

2. *If $nb^5 \to C_0^2$, then*

$$(b\tilde{\theta}^2(p))^{-1/2} \left\{ \int_b^{1-b} \left| b^{-1/2}\Gamma_n^{(2)}(t) + g_{(n)}(t) \right|^p \mathrm{d}\mu(t) - m_n^c(p) \right\} \xrightarrow{d} N(0,1),$$

where $\tilde{\theta}^2(p)$ is defined in (12).

Proof. From the properties of the kernel function and L we have

$$\begin{aligned}\Gamma_n^{(2)}(t) &= \int k\left(\frac{t-u}{b}\right) \mathrm{d}W(L(u)) - \frac{W(L(1))}{L(1)} \int k\left(\frac{t-u}{b}\right) L'(u)\,\mathrm{d}u \\ &= \int k\left(\frac{t-u}{b}\right) \mathrm{d}W(L(u)) - b\frac{W(L(1))}{L(1)} L'(t) + O_P(b^3),\end{aligned}$$

where the O_P term is uniformly for $t \in [0,1]$. Hence, inequality (14) implies that

$$\begin{aligned}&\int_b^{1-b} \left| b^{-1/2}\Gamma_n^{(2)}(t) + g_{(n)}(t) \right|^p \mathrm{d}\mu(t) \\ &= \int_b^{1-b} \left| b^{-1/2} \int k\left(\frac{t-u}{b}\right) \mathrm{d}W(L(u)) + g_{(n)}(t) - b^{1/2}\frac{W(L(1))}{L(1)} L'(t) \right|^p \mathrm{d}\mu(t) \\ &\quad + O(b^3).\end{aligned}$$

Therefore, it is sufficient to prove a CLT for

$$\int_b^{1-b} \left| b^{-1/2} \int k\left(\frac{t-u}{b}\right) \mathrm{d}W(L(u)) + g_{(n)}(t) - b^{1/2}\frac{W(L(1))}{L(1)}L'(t) \right|^p \mathrm{d}\mu(t). \tag{67}$$

Let

$$X_{n,t} = b^{-1/2} \int k\left(\frac{t-u}{b}\right) \mathrm{d}W(L(u)) + g_{(n)}(t). \tag{68}$$

Then $X_{nt} \sim N(g_{(n)}(t), \sigma_n^2(t))$, where

$$\sigma_n^2(t) = \frac{1}{b}\int k^2\left(\frac{t-u}{b}\right) L'(u)\,\mathrm{d}u. \tag{69}$$

We can then write

$$\begin{aligned} & b^{-1/2}\left\{\int_b^{1-b} \left| b^{-1/2}\Gamma_n^{(2)}(t) + g_{(n)}(t)\right|^p \mathrm{d}\mu(t) - m_n^c(p)\right\} \\ &= b^{-1/2}\left\{\int_b^{1-b} \left| X_{n,t} - b^{1/2}\frac{W(L(1))}{L(1)}L'(t)\right|^p \mathrm{d}\mu(t) - m_n^c(p)\right\} + o(1) \\ &= b^{-1/2}\left\{\int_b^{1-b} |X_{n,t}|^p \,\mathrm{d}\mu(t) - m_n^c(p)\right\} \\ &\qquad - p\frac{W(L(1))}{L(1)}\int_b^{1-b} |X_{n,t}|^{p-1}\,\mathrm{sgn}\left\{X_{n,t}\right\} L'(t)w(t)\,\mathrm{d}t \\ &\qquad + b^{-1/2}\int_b^{1-b} O\left(bW(L(1))^2\right)\,\mathrm{d}t + o(1), \end{aligned} \tag{70}$$

where we use

$$|x|^p = |y|^p + p(x-y)|y|^{p-1}\,\mathrm{sgn}(y) + O((x-y)^2) \tag{71}$$

for the first term in the integrand on the right hand side of the first equality in (70). The third term on the right hand side of (70) converges to zero in probability, so it suffices to deal with the first two terms. To establish a central limit theorem for the first term, one can mimic the approach in [10] using a big-blocks-small-blocks procedure. See Lemmas A.1 and A.2 in the Appendix for details. It can be shown that

$$b^{-1/2}\left\{\int_b^{1-b} |X_{n,t}|^p \,\mathrm{d}\mu(t) - m_n^c(p)\right\} = b^{1/2}\sum_{i=1}^{M_3}\zeta_i + o_P(1),$$

where $\zeta_i = \sum_{j=c_i}^{d_i} \xi_j$, with $c_i = (i-1)(M_2+2)+1$ and $d_i = (i-1)(M_2+2)+M_2$, $M_2 = [(M_1-1)^\nu]$, for some $0<\nu<1$ and $M_1 = [1/b-1]$, $M_3 = [(M_1-1)/(M_2+2)]$, and

$$\xi_i = b^{-1}\int_{ib}^{ib+b}\left\{ |X_{n,t}|^p - \int_{-\infty}^{+\infty}\left|\sqrt{L'(t)}Dx + g_{(n)}(t)\right|^p \phi(x)\,\mathrm{d}x\right\} w(t)\,\mathrm{d}t.$$

The random variables ζ_i are independent and satisfy

$$b^{1/2}\sum_{i=1}^{M_3}\zeta_i \xrightarrow{d} N(0,\gamma^2(p)), \tag{72}$$

where $\gamma^2(p)$ is defined in (53).

Next, consider the second term in the right hand side of (70). We have

$$\begin{aligned}
&\mathbb{E}\left[\int_b^{1-b}|X_{n,t}|^{p-1}\operatorname{sgn}\{X_{n,t}\}\,L'(t)w(t)\,\mathrm{d}t\right]\\
&\quad=\int_b^{1-b}\int_{\mathbb{R}}\left|\sigma_n(t)x+g_{(n)}(t)\right|^{p-1}\operatorname{sgn}\left\{\sigma_n(t)x+g_{(n)}(t)\right\}\phi(x)\,\mathrm{d}x\,L'(t)w(t)\,\mathrm{d}t\\
&\quad\to\int_0^1\int_{\mathbb{R}}\left|\sqrt{L'(t)}Dx+g(t)\right|^{p-1}\operatorname{sgn}\left\{\sqrt{L'(t)}Dx+g(t)\right\}\phi(x)\,\mathrm{d}x\,L'(t)w(t)\,\mathrm{d}t,
\end{aligned}$$

where D and $\sigma_n(t)$ are defined in (7) and (69), respectively, and ϕ denotes the standard normal density. Note that

$$\frac{d}{dx}\left|\sqrt{L'(t)}Dx+g(t)\right|^p=p\left|\sqrt{L'(t)}Dx+g(t)\right|^{p-1}\operatorname{sgn}\left\{\sqrt{L'(t)}Dx+g(t)\right\}.$$

Hence, integration by parts gives

$$\int_0^1\int_{\mathbb{R}}\left|\sqrt{L'(t)}Dx+g(t)\right|^{p-1}\operatorname{sgn}\left\{\sqrt{L'(t)}Dx+g(t)\right\}\phi(x)\,\mathrm{d}x\,L'(t)w(t)\,\mathrm{d}t=\frac{\theta_1(p)}{Dp},$$

where θ_1 is defined in (13). We conclude

$$\mathbb{E}\left[\int_b^{1-b}|X_{n,t}|^{p-1}\operatorname{sgn}\{X_{n,t}\}\,L'(t)w(t)\,\mathrm{d}t\right]\to\frac{\theta_1(p)}{Dp}.$$

Moreover,

$$\begin{aligned}
&\operatorname{Var}\left(\int_b^{1-b}|X_{n,t}|^{p-1}\operatorname{sgn}\{X_{n,t}\}\,L'(t)w(t)\,\mathrm{d}t\right)\\
&=\int_b^{1-b}\int_b^{1-b}\operatorname{Covar}\left(|X_{n,t}|^{p-1}\operatorname{sgn}\{X_{n,t}\},|X_{n,s}|^{p-1}\operatorname{sgn}\{X_{n,s}\}\right)\\
&\qquad\qquad\cdot L'(t)L'(s)w(t)w(s)\,\mathrm{d}t\,\mathrm{d}s\\
&=\int_b^{1-b}\int_b^{1-b}\mathbb{1}_{\{|t-s|\le 2b\}}\operatorname{Covar}\left(|X_{n,t}|^{p-1}\operatorname{sgn}\{X_{n,t}\},|X_{n,s}|^{p-1}\operatorname{sgn}\{X_{n,s}\}\right)\\
&\qquad\qquad\cdot L'(t)L'(s)w(t)w(s)\,\mathrm{d}t\,\mathrm{d}s,
\end{aligned}$$

because for $|t-s|>2b$, $X_{n,t}$ is independent of $X_{n,s}$. As a result, using that $X_{n,t}$ has bounded moments, we obtain

$$\operatorname{Var}\left(\int_b^{1-b}|X_{n,t}|^{p-1}\operatorname{sgn}\{X_{n,t}\}\,L'(t)w(t)\,\mathrm{d}t\right)\to 0.$$

This means that

$$\int_b^{1-b} |X_{n,t}|^{p-1} \operatorname{sgn}\{X_{n,t}\} L'(t)w(t)\,\mathrm{d}t \to \frac{\theta_1(p)}{Dp},$$

in probability and

$$-p\frac{W(L(1))}{L(1)} \int_b^{1-b} |X_{n,t}|^{p-1} \operatorname{sgn}\{X_{n,t}\} L'(t)w(t)\,\mathrm{d}t = CW(L(1)) + o_P(1),$$

where

$$C = -\frac{\theta_1(p)}{DL(1)}. \tag{73}$$

Going back to (70), we conclude that

$$\begin{aligned} b^{-1/2} &\left\{ \int_b^{1-b} \left| b^{-1/2}\Gamma_n^{(2)}(t) + g_{(n)}(t) \right|^p \,\mathrm{d}\mu(t) - m_n^c(p) \right\} \\ &= b^{1/2} \sum_{i=1}^{M_3} \zeta_i + CW(L(1)) + o_P(1). \end{aligned} \tag{74}$$

In the case $nb^5 \to 0$, we have $g(t) = 0$ in the definition of $\theta_1(p)$ in (13). Hence, by the symmetry of the standard normal distribution, it follows that $\theta_1(p) = 0$ and as a result $C = 0$. According to (72) and (74), this means that

$$b^{-1/2} \left\{ \int_b^{1-b} \left| b^{-1/2}\Gamma_n^{(2)}(t) + g_{(n)}(t) \right|^p \,\mathrm{d}\mu(t) - m_n^c(p) \right\}$$

converges in distribution to a mean zero normal random variable with variance $\sigma^2(p)$.

Then, consider the case $nb^5 \to C_0^2 > 0$. Note that ζ_i depends only on the Brownian motion on the interval $[c_i b - b, c_i b + b]$. These intervals are disjoint, because $c_{i+1}b - b = d_i b + b$. We write

$$W(L(1)) = \sum_{i=1}^{M_3} [W(t_{i+1}) - W(t_i)] + W(L(1)) - W(t_{M_3}),$$

where $t_i = L(c_i b - b)$, for $i = 1, \dots, M_3$. Moreover, $W(L(1)) - W(t_{M_3}) \to 0$, in probability, since $t_{M_3} \sim L(1 + O(b)) \to L(1)$. Hence, the left hand side of (74), can be written as

$$\sum_{i=1}^{M_3} Y_i + o_P(1), \quad Y_i = b^{1/2}\zeta_i + C\left[W(t_{i+1}) - W(t_i)\right].$$

Since now we have a sum of independent random variables, we apply the Lindeberg-Feller central limit theorem. Using $\mathbb{E}[Y_i] = O(b^{5/2}M_2)$, it suffices to show that

$$\mathbb{E}\left[\left(\sum_{i=1}^{M_3} Y_i\right)^2\right] \to \tilde{\theta}^2(p) > 0, \tag{75}$$

and that the Lyapounov condition

$$\sum_{i=1}^{M_3} \mathbb{E}[Y_i^4] \left(\sum_{i=1}^{M_3} \mathbb{E}[Y_i^2] \right)^{-2} \to 0. \tag{76}$$

is satisfied. Once we have (75), condition (76) is equivalent to $\sum_{i=1}^{M_3} \mathbb{E}[Y_i^4] \to 0$. In order to prove this, we use that $\mathbb{E}[\zeta_i^4] = O(M_2^2)$, (see (93) in the proof of Lemma A.2 in the Appendix). Then, we get

$$\begin{aligned}\sum_{i=1}^{M_3} \mathbb{E}[Y_i^4] &\le O(b^2) \sum_{i=1}^{M_3} \mathbb{E}[\zeta_i^4] + O(1) \sum_{i=1}^{M_3} \mathbb{E}[(W(t_{i+1}) - W(t_i))^4] \\ &\le O(M_3 b^2 M_2^2) + O(M_3 (t_{i+1} - t_i)^2) \\ &= o(1) + O(M_3 M_2^2 b^2) = o(1).\end{aligned}$$

Because $\mathbb{E}[Y_i] = O(b^{5/2} M_2)$, for (75) we have

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{i=1}^{M_3} Y_i \right)^2 \right] &= \sum_{i=1}^{M_3} \mathbb{E}\left[Y_i^2 \right] + o(1) \\ &= b \sum_{i=1}^{M_3} \mathbb{E}\left[\zeta_i^2 \right] + C^2 \sum_{i=1}^{M_3} (t_{i+1} - t_i) \\ &\quad + 2Cb^{1/2} \sum_{i=1}^{M_3} \mathbb{E}[\zeta_i \{W(t_{i+1}) - W(t_i)\}] + o(1).\end{aligned}$$

It can be shown that $b \sum_{i=1}^{M_3} \mathbb{E}\left[\zeta_i^2\right] \to 0$, see Lemma A.2 in the Appendix for details. Moreover, $\sum_{i=1}^{M_3} (t_{i+1} - t_i) = L((M_3 - 1)(M_2 + 2)b) - L(0) = L(1) + o(1)$. Finally, since

$$\zeta_i = b^{-1} \int_{c_i b}^{d_i b} \left\{ |X_{n,t}|^p - \int_{-\infty}^{+\infty} \left| \sqrt{l(t)} D x + g_{(n)}(t) \right|^p \phi(x) \, \mathrm{d}x \right\} w(t) \, \mathrm{d}t,$$

we can write

$$2Cb^{1/2} \sum_{i=1}^{M_3} \mathbb{E}[\zeta_i \{W(t_{i+1}) - W(t_i)\}] = 2C \sum_{i=1}^{M_3} \int_{c_i b}^{d_i b} \mathbb{E}\left[|X_{n,t}|^p Z_{n,t} \right] w(t) \, \mathrm{d}t,$$

where $Z_{n,t} = b^{-1/2} \{W(t_{i+1}) - W(t_i)\}$. Note that

$$(X_{n,t}, Z_{n,t}) \sim N\left(\begin{bmatrix} g_{(n)}(t) \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_n^2(t) & \rho_n(t)\sigma_n(t)\tilde{\sigma}_n(t) \\ \rho_n(t)\sigma_n(t)\tilde{\sigma}_n(t) & \tilde{\sigma}_n^2(t)) \end{bmatrix} \right).$$

where $\sigma_n^2(t)$ is defined in (69) and

$$\tilde{\sigma}_n^2(t) = b^{-1}[L(t+b) - L(t-b)], \quad \rho_n(t) = \sigma_n(t)^{-1} \tilde{\sigma}_n(t)^{-1} b^{-1} \int k\left(\frac{t-u}{b} \right) l(u) \, \mathrm{d}u.$$

Using

$$Z_{n,t} \mid X_{n,t} = x \sim N\left(\frac{\tilde{\sigma}_n(t)}{\sigma_n(t)}\rho_n(t)(x - g_{(n)}(t)), \left(1-\rho_n^2(t)\right)\tilde{\sigma}_n^2(t)\right)$$

we obtain

$$\begin{aligned}
\mathbb{E}\left[|X_{n,t}|^p Z_{n,t}\right] &= \mathbb{E}\left[|X_{n,t}|^p\, \mathbb{E}[Z_{n,t} \mid X_{n,t}]\right] \\
&= \mathbb{E}\left[|X_{n,t}|^p \frac{\tilde{\sigma}_n(t)}{\sigma_n(t)}\rho_n(t)\left(X_{n,t} - g_{(n)}(t)\right)\right] \\
&= \frac{\tilde{\sigma}_n(t)}{\sigma_n(t)}\rho_n(t)\mathbb{E}\left[|X_{n,t}|^p\left(X_{n,t} - g_{(n)}(t)\right)\right] \\
&= \frac{\tilde{\sigma}_n(t)}{\sigma_n(t)}\rho_n(t)\int_{\mathbb{R}} |g_{(n)}(t) + \sigma_n(t)x|^p \sigma_n(t) x\phi(x)\,\mathrm{d}x \\
&= \sigma_n(t)^{-1}b^{-1}\int k\left(\frac{t-u}{b}\right) l(u)\,\mathrm{d}u \int_{\mathbb{R}} |g_{(n)}(t) + \sigma_n(t)x|^p x\phi(x)\,\mathrm{d}x.
\end{aligned}$$

Because $\sigma_n^2(t) \to D^2 l(t)$, where D is defined in (7), $g_{(n)}(t) \to g(t)$, as defined in (5), and $b^{-1}\int k\left(\frac{t-u}{b}\right) l(u)\,\mathrm{d}u \to l(t)$, we find that

$$\mathbb{E}\left[|X_{n,t}|^p Z_{n,t}\right] \to \frac{\sqrt{l(t)}}{D}\int_{\mathbb{R}} |g(t) + D\sqrt{l(t)}x|^p x\phi(x)\,\mathrm{d}x.$$

Hence

$$\begin{aligned}
\mathbb{E}\left[\left(\sum_{i=1}^{M_3} Y_i\right)^2\right] &= \theta^2(p) + C^2 L(1) \\
&\quad + \frac{2C}{D}\sum_{i=1}^{M_3}\int_{c_i b}^{d_i b}\int_{\mathbb{R}} |g(t) + \sqrt{l(t)}Dx|^p x\phi(x)\,\mathrm{d}x\,\sqrt{l(t)}w(t)\,\mathrm{d}t + o(1) \\
&= \theta^2(p) + C^2 L(1) \\
&\quad + \frac{2C}{D}\int_0^1\int_{\mathbb{R}} |g(t) + D\sqrt{l(t)}x|^p x\phi(x)\,\mathrm{d}x\,\sqrt{l(t)}w(t)\,\mathrm{d}t + o(1) \\
&= \theta^2(p) + C^2 L(1) + 2CD^{-1}\theta_1(p) + o(1) \\
&= \theta^2(p) - \frac{\theta_1^2(p)}{D^2 L(1)} + o(1),
\end{aligned}$$

applying the definitions of C and $\theta_1(p)$ in (73) and (13), respectively. It follows from the Lindeberg-Feller central limit theorem that $\sum_{i=1}^{M_3} Y_i \xrightarrow{d} N(0, \tilde{\theta}^2(p))$, where $\tilde{\theta}(p)$ is defined in (12). □

8.2. Proofs for Section 4

Lemma 8.2. *Let Y_n and $Y_n^{(1)}$ be defined in (34) and (36), respectively. Assume that (A1) − (A2) hold. If $1 \le p < min(q, 2q-7)$, $1/b = o\left(n^{1/3-1/q}\right)$ and*

$1/b = o\left(n^{(q-3)/(6p)}\right)$, *then*

$$b^{-p}\int_b^{1-b}|Y_n(t)-Y_n^{(1)}(t)|^p\,\mathrm{d}\mu(t)=o_P(1).$$

Proof. We follow the same reasoning as in the proof of Lemma 8 in [38]. Let $I_{nv}=[0,1]\cap[v-n^{-1/3}\log n, v+n^{-1/3}\log n]$ and for $J=E,W$, let

$$N_{nv}^J=\left\{[\mathrm{CM}_{[0,1]}\Lambda_n^W](s)=[\mathrm{CM}_{I_{nv}}\Lambda_n^W](s)\text{ for all }s\in I_{nv}\right\}. \tag{77}$$

Then according to Lemma 3 in [38], there exists $C>0$, independent of n,v,d, such that

$$\begin{aligned}\mathbb{P}\left((N_{nv}^W)^c\right)&=O(\mathrm{e}^{-Cd^3})\\ \mathbb{P}\left((N_{nv}^E)^c\right)&=O(n^{1-q/3}d^{-2q}+\mathrm{e}^{-Cd^3}).\end{aligned} \tag{78}$$

Let $K_{nv}=N_{nv}^E\cap N_{nv}^W$ and write

$$\begin{aligned}\mathbb{E}\left[\left|A_n^E(v)^p-A_n^W(v)\right|\right]&=\mathbb{E}\left[\left|A_n^E(v)^p-A_n^W(v)\right|\mathbb{1}_{K_{nv}^c}\right]\\ &\quad+n^{2p/3}\mathbb{E}\left[\left|[\mathrm{D}_{I_{nv}}\Lambda_n](t)^p-[\mathrm{D}_{I_{nv}}\Lambda_n^W](t)^p\right|\mathbb{1}_{K_{nv}}\right].\end{aligned}$$

From the proof of Lemma 8 in [38], using (78) with $d=\log n$, we have

$$\mathbb{E}\left[\left|A_n^E(v)^p-A_n^W(v)\right|\mathbb{1}_{K_{nv}^c}\right]=O_P(n^{1/2-q/6}(\log n)^{-q}+\mathrm{e}^{-C(\log n)^3/2}/2)$$

and

$$n^{2p/3}\mathbb{E}\left[\left|[\mathrm{D}_{I_{nv}}\Lambda_n](t)^p-[\mathrm{D}_{I_{nv}}\Lambda_n^W](t)^p\right|\mathbb{1}_{K_{nv}}\right]=O_p\left(n^{-p/3+p/q}\right).$$

It follows that

$$\begin{aligned}&b^{-p}\int_b^{1-b}|Y_n(t)-Y_n^{(1)}(t)|^p\,\mathrm{d}\mu(t)\\ &\leq Cb^{-p}\int_{-1}^1|A_n^E(t-by)-A_n^W(t-by)|^p\,\mathrm{d}y\\ &=b^{-p}O_P\left(n^{-p/3+p/q}\right)+b^{-p}O_P\left(n^{1/2-q/6}(\log n)^{-q}+\mathrm{e}^{-C(\log n)^3/2}\right).\end{aligned}$$

According to the assumptions on the order of b^{-1}, the right hand side is of order $o_P(1)$. □

Lemma 8.3. *Let $Y_n^{(1)}$ and $Y_n^{(2)}$ be defined in (36) and (40), respectively. Assume that (A1) − (A2) hold. If $b\to 0$, such that $nb\to\infty$, then*

$$b^{-p}\int_b^{1-b}|Y_n^{(1)}(t)-Y_n^{(2)}(t)|^p\,\mathrm{d}\mu(t)=o_P(1).$$

Proof. We have

$$\begin{aligned}
&\sup_{t\in(b,1-b)} \mathbb{E}\left[\left|Y_n^{(1)}(t)-Y_n^{(2)}(t)\right|^p\right] \\
&= \sup_{t\in(b,1-b)} \mathbb{E}\left[\left|\frac{1}{b}\int_{t-b}^{t+b} k'\left(\frac{t-v}{b}\right)\mathbb{1}_{(N_{nv}^W)^c}\left(A_n^W(v)-n^{2/3}[\mathrm{D}_{I_{nv}}\Lambda_n^W](v)\right)\mathrm{d}v\right|^p\right] \\
&\leq c \sup_{t\in(b,1-b)} \mathbb{E}\left[\sup_{v\in[0,1]}\left|A_n^W(v)-n^{2/3}[\mathrm{D}_{I_{nv}}\Lambda_n^W](v)\right|^p\left(\frac{1}{b}\int_{t-b}^{t+b}\mathbb{1}_{(N_{nv}^W)^c}\,\mathrm{d}v\right)^p\right],
\end{aligned}$$

where N_{nv}^W is defined in (77). Moreover, since

$$\sup_{v\in[0,1]}\left|A_n^W(v)-n^{2/3}[\mathrm{D}_{I_{nv}}\Lambda_n^W](v)\right| \leq 4n^{2/3}\left\{\Lambda(1)+n^{-1/2}\sup_{s\in[0,L(1)]}|W_n(s)|\right\},$$

from the Cauchy-Schwartz inequality we obtain

$$\begin{aligned}
&\sup_{t\in(b,1-b)} \mathbb{E}\left[\sup_{v\in[0,1]}\left|A_n^W(v)-n^{2/3}[\mathrm{D}_{I_{nv}}\Lambda_n^W](v)\right|^p\left(\frac{1}{b}\int_{t-b}^{t+b}\mathbb{1}_{(N_{nv}^W)^c}\,\mathrm{d}v\right)^p\right] \\
&\leq 4^p n^{2p/3}\mathbb{E}\left[\left\{\Lambda(1)+n^{-1/2}\sup_{s\in[0,L(1)]}|W_n(s)|\right\}^{2p}\right]^{1/2} \\
&\qquad\cdot \sup_{t\in(b,1-b)} \mathbb{E}\left[\left(\frac{1}{b}\int_{t-b}^{t+b}\mathbb{1}_{(N_{nv}^W)^c}\,\mathrm{d}v\right)^{2p}\right]^{1/2}.
\end{aligned}$$

For the last term on the right hand side, we can use Jensen's inequality:

$$\left(\frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x\right)^p \leq \frac{1}{b-a}\int_a^b f(x)^p\,\mathrm{d}x,$$

for all $a<b$, $p\geq 1$, and $f(x)\geq 0$. Because all the moments of $\sup_{s\in[0,L(1)]}|W_n(s)|$ are finite, together with (78), it follows that

$$\begin{aligned}
\sup_{t\in(b,1-b)} \mathbb{E}\left[\left|Y_n^{(1)}(t)-Y_n^{(2)}(t)\right|^p\right] &\leq Cn^{2p/3}\sup_{t\in(b,1-b)}\mathbb{E}\left[\frac{1}{b}\int_{t-b}^{t+b}\mathbb{1}_{(N_{nv}^W)^c}\,\mathrm{d}v\right]^{1/2} \\
&= O\left(n^{2p/3}\exp\left(-C(\log n)^3/2\right)\right).
\end{aligned} \tag{79}$$

Because

$$b^{-p}n^{2p/3}\exp\left(-C(\log n)^3/2\right) = (nb)^{2p/3-C(\log n)^2/2}b^{-p-2p/3+C(\log n)^2/2}\to 0,$$

this finishes the proof. ☐

Lemma 8.4. *Let $Y_n^{(2)}$ and $Y_n^{(3)}$ be defined in (40) and (43), respectively. Assume that (A1) − (A2) hold. If $1/b = o\left(n^{1/3-1/q}\right)$, then*

$$b^{-p}\int_b^{1-b}|Y_n^{(2)}(t)-Y_n^{(3)}(t)|^p\,\mathrm{d}\mu(t)=o_P(1).$$

Proof. Let $H_{nv}=[-n^{1/3}v, n^{1/3}(1-v)]\cap[-\log n,\log n]$ and

$$\Delta_{nv}=n^{2/3}[\mathrm{D}_{I_{nv}}\Lambda_n^W](v)-[\mathrm{D}_{H_{nv}}Y_{nv}](0).$$

By definition, we have

$$\int_b^{1-b}|Y_n^{(2)}(t)-Y_n^{(3)}(t)|^p\,\mathrm{d}\mu(t)=\int_b^{1-b}\left|\frac{1}{b}\int_{t-b}^{t+b}k'\left(\frac{t-v}{b}\right)\Delta_{nv}\,\mathrm{d}v\right|^p\,\mathrm{d}\mu(t).$$

Moreover, using

$$\sup_{t\in(0,1)}\mathbb{E}\left[|\Delta_{nt}|^p\right]=O\left(n^{-p/3+p/q}\right)$$

(see the proof of Lemma 6 in [38]), we obtain

$$\begin{aligned}
&\sup_{t\in(b,1-b)}\mathbb{E}\left[\left|\frac{1}{b}\int_{t-b}^{t+b}k'\left(\frac{t-v}{b}\right)\Delta_{nv}\,\mathrm{d}v\right|^p\right]\\
&\leq\sup_{u\in[-1,1]}|k'(u)|^p\sup_{t\in(b,1-b)}\mathbb{E}\left[\left|\frac{1}{b}\int_{t-b}^{t+b}\Delta_{nv}\,\mathrm{d}v\right|^p\right]\\
&\leq C\sup_{t\in(b,1-b)}\frac{1}{b}\int_{t-b}^{t+b}\mathbb{E}\left[|\Delta_{nv}|^p\right]\,\mathrm{d}v\leq 2C\sup_{v\in(0,1)}\mathbb{E}\left[|\Delta_{nv}|^p\right]\\
&=O\left(n^{-p/3+p/q}\right).
\end{aligned}\tag{80}$$

Because $1/b=o\left(n^{1/3-1/q}\right)$, this finishes the proof. □

Lemma 8.5. *Let $Y_n^{(3)}$ and $Y_n^{(4)}$ be defined in (43) and (47), respectively. Assume that (A1) − (A2) hold. If $1/b = o\left(n^{1/3-1/q}\right)$, then*

$$b^{-p}\int_b^{1-b}|Y_n^{(3)}(t)-Y_n^{(4)}(t)|^p\,\mathrm{d}\mu(t)=o_P(1).$$

Proof. Let H_{nv} be defined as in the proof of Lemma 8.4 and let

$$J_{nv}=\left[n^{1/3}\frac{L(a_{nv})-L(v)}{L'(v)},n^{1/3}\frac{L(b_{nv})-L(v)}{L'(v)}\right],$$

where $a_{nv}=\max(0,v-n^{-1/3}\log n)$ and $b_{nv}=\min(1,v+n^{-1/3}\log n)$. As in (4.31) in [33] we have

$$\sup_{v\in(0,1)}\mathbb{E}\left[\left|[\mathrm{D}_{H_{nv}}\tilde{Y}_{nv}](0)-[\mathrm{D}_{J_{nv}}Z_{nv}](0)\right|^p\right]=O(n^{-p/3}(\log n)^{3p}),\tag{81}$$

where $\tilde{Y}_{nv}$ and Z_{nv} are defined in (45) and (46). This means that,

$$\begin{aligned}
&\sup_{t\in(b,1-b)} \mathbb{E}\left[\left|Y_n^{(3)}(t)-Y_n^{(4)}(t)\right|^p\right]\\
&\leq \sup_{u\in[-1,1]}|k'(u)|^p \sup_{t\in(b,1-b)} \mathbb{E}\left[\left|\frac{1}{b}\int_{t-b}^{t+b}\left\{[\mathrm{D}_{H_{nv}}\tilde{Y}_{nv}](0)-[\mathrm{D}_{J_{nv}}Z_{nv}](0)\right\}\mathrm{d}v\right|^p\right]\\
&\leq C\sup_{t\in(b,1-b)}\frac{1}{b}\int_{t-b}^{t+b}\mathbb{E}\left[\left|[\mathrm{D}_{H_{nv}}\tilde{Y}_{nv}](0)-[\mathrm{D}_{J_{nv}}Z_{nv}](0)\right|^p\right]\mathrm{d}v\\
&\leq C\sup_{v\in(b,1-b)}\mathbb{E}\left[\left|[\mathrm{D}_{H_{nv}}\tilde{Y}_{nv}](0)-[\mathrm{D}_{J_{nv}}Z_{nv}](0)\right|^p\right]=O\left(n^{-p/3}(\log n)^{3p}\right).
\end{aligned} \tag{82}$$

Since $1/b=o\left(n^{1/3-1/q}\right)$, this finishes the proof. □

Lemma 8.6. *Let $Y_n^{(4)}$ and $Y_n^{(5)}$ be defined in (47) and (50), respectively. Assume that $(A1)-(A2)$ hold. If $nb\to\infty$, such that*

$$1/b=o(n^{1/6+1/(6p)}(\log n)^{-(1/2+1/(2p))}),$$

then

$$b^{-p}\int_b^{1-b}|Y_n^{(4)}(t)-Y_n^{(5)}(t)|^p\,\mathrm{d}\mu(t)=o_P(1).$$

Proof. We argue as in the proof of Lemma 4.4 in [33]. First, when $v\in(n^{-1/3}\log n, 1-n^{-1/3}\log n)$, there exists $M>0$, only depending λ, such that $[-M\log n, M\log n]\subset I_{nv}$, and on the interval $[-M\log n, M\log n]$ we have that $\mathrm{CM}_{[-M\log n,M\log n]}Z\leq \mathrm{CM}_{I_{nv}}Z\leq \mathrm{CM}_{\mathbb{R}}Z$. Let $N_{nM}=N(M\log n)$, where $N(d)$ is the event that $[\mathrm{CM}_{[-d,d]}Z](s)$ is equal to $[\mathrm{CM}_{\mathbb{R}}Z](s)$ for $s\in[-d/2,d/2]$. According to Lemma 1.2 in [32], it holds that

$$\mathbb{P}(N(d)^c)\leq \exp(-d^3/2^7). \tag{83}$$

For convenience, write $\delta_n=n^{-1/3}\log n$. Because $[\mathrm{CM}_{[-M\log n,M\log n]}Z](0)=[\mathrm{CM}_{I_{nv}}Z](0)=[\mathrm{CM}_{\mathbb{R}}Z](0)$ on the event N_{nM}, we have by means of Cauchy-Schwarz, we find that

$$\begin{aligned}
&\sup_{v\in(\delta_n,1-\delta_n)}\mathbb{E}\left[|[\mathrm{D}_{I_{nv}}Z](0)-[\mathrm{D}_{\mathbb{R}}Z](0)|^p\right]\\
&=\sup_{v\in(\delta_n,1-\delta_n)}\mathbb{E}\left[|[\mathrm{D}_{I_{nv}}Z](0)-[\mathrm{D}_{\mathbb{R}}Z](0)|^p\right]\mathbb{1}_{N_{nM}^c}\\
&\leq 2^p\mathbb{E}\left[\left(\sup_{s\in\mathbb{R}}|Z(s)|\right)^p\mathbb{1}_{N_{nM}^c}\right]\\
&\leq 2^p\left(\mathbb{E}\left[\left(\sup_{s\in\mathbb{R}}|Z(s)|\right)^{2p}\right]\right)^{1/2}\mathbb{P}(N_{nM}^c)^{1/2}.
\end{aligned}$$

Because $\mathbb{E}[(\sup|Z|)^{2p}] < \infty$, together with (83), we find that

$$\sup_{v\in(\delta_n,1-\delta_n)} \mathbb{E}\left[|[\mathrm{D}_{I_{nv}}Z](0) - [\mathrm{D}_{\mathbb{R}}Z](0)|^p\right] = O\left(\exp(-C(\log n)^3)\right). \tag{84}$$

Note that

$$Y_n^{(4)}(t) - Y_n^{(5)}(t) = \frac{1}{b}\int_{t-b}^{t+b} k'\left(\frac{t-v}{b}\right)\frac{1}{c_1(v)}\left([\mathrm{D}_{I_{nv}}Z](0) - [\mathrm{D}_{\mathbb{R}}Z](0)\right)\mathrm{d}v. \tag{85}$$

When $t \in (b+\delta_n, 1-b-\delta_n)$, then $v \in (t-b, t+b) \subset (\delta_n, 1-\delta_n)$, so after change of variables, it follows that

$$\begin{aligned}
&\sup_{t\in(b+\delta_n,1-b-\delta_n)} \mathbb{E}\left[\left|Y_n^{(4)}(t) - Y_n^{(5)}(t)\right|^p\right] \\
&\leq 2^p \frac{\sup_{u\in[-1,1]}|k'(u)|^p}{\inf_{v\in(0,1)} c_1(v)^p} \sup_{v\in(\delta_n,1-\delta_n)} \mathbb{E}\left[|[\mathrm{D}_{I_{nv}}Z](0) - [\mathrm{D}_{\mathbb{R}}Z](0)|^p\right] \\
&= O\left(\exp(-C(\log n)^3)\right).
\end{aligned} \tag{86}$$

Next, consider the case where $t \in (b, b+\delta_n)$. In this case we split the integral on the right hand side of (85) into an integral over $v \in (t-b, \delta_n)$ and an integral over $v \in (\delta_n, t+b)$. The latter integral can be bounded in the same way as in (86), whereas for the first integral we have

$$\begin{aligned}
&\left|\frac{1}{b}\int_{t-b}^{\delta_n} k'\left(\frac{t-v}{b}\right)\frac{1}{c_1(v)}\left([\mathrm{D}_{I_{nv}}Z](0) - [\mathrm{D}_{\mathbb{R}}Z](0)\right)\mathrm{d}v\right| \\
&\qquad \leq b^{-1}\delta_n \frac{\sup_{u\in[-1,1]}|k'(u)|}{\inf_{v\in(0,1)} c_1(v)} |[\mathrm{D}_{I_{nv}}Z](0) - [\mathrm{D}_{\mathbb{R}}Z](0)| \\
&\qquad \leq b^{-1}\delta_n \frac{\sup_{u\in[-1,1]}|k'(u)|}{\inf_{v\in(0,1)} c_1(v)} [\mathrm{D}_{\mathbb{R}}Z](0),
\end{aligned}$$

where we also use that $[\mathrm{D}_{I_{nv}}Z](0) \leq [\mathrm{D}_{\mathbb{R}}Z](0)$. Furthermore, since $[\mathrm{D}_{\mathbb{R}}Z](0)$ has bounded moments of any order, for $t \in (b, b+\delta_n)$, we obtain

$$\begin{aligned}
&\sup_{t\in(b,b+\delta_n)} \mathbb{E}\left[\left|Y_n^{(4)}(t) - Y_n^{(5)}(t)\right|^p\right] \\
&\leq b^{-p}\delta_n^p \frac{\sup_{u\in[-1,1]}|k'(u)|^p}{\inf_{v\in(0,1)} c_1(v)^p} \mathbb{E}\left[[\mathrm{D}_{\mathbb{R}}]Z](0)^p\right] + O\left(\exp(-C(\log n)^3)\right) \\
&= O_P\left(b^{-p}\delta_n^p\right) + O_P\left(\exp(-C(\log n)^3)\right).
\end{aligned} \tag{87}$$

A similar bound can be obtained for $t \in (1-b-\delta_n, 1-b)$. Putting things together yields,

$$\int_b^{1-b} \left|Y_n^{(4)}(t) - Y_n^{(5)}(t)\right|^p \mathrm{d}\mu(t) = O_P\left(\exp(-C(\log n)^3)\right) + O_P\left(b^{-p}\delta_n^{p+1}\right).$$

Because $nb \to \infty$ implies $b^{-p}\exp(-C(\log n)^3) \to 0$ and

$$1/b = o(n^{1/6+1/(6p)}(\log n)^{-(1/2+1/(2p))})$$

yields $b^{-2p}\delta_n^{p+1} \to 0$, this finishes the proof. □

Appendix A: Kernel estimator of a decreasing function

Lemma A.1. *Let* $L(t)=\int_0^t l(u)\,\mathrm{d}u$ *for a differentiable function* $l(t)$ *on* $[0,1]$ *such that* $\inf_{[0,1]} l(t)>0$ *and* $\sup_{[0,1]}|l'(t)|<\infty$*. Let* $\Gamma_n^{(1)}$ *be as in* (17)*. Assume that (A1) and (A3) hold. Then*

$$(b\gamma^2(p))^{-1/2}\left\{\int_b^{1-b}\left|b^{-1/2}\Gamma_n^{(1)}(t)+g_{(n)}(t)\right|^p\,\mathrm{d}\mu(t)-m_n^c(p)\right\}\xrightarrow{d}N(0,1),$$

where $\gamma^2(p)$*,* $g_{(n)}$ *and* $m_n^c(p)$ *are defined respectively in* (53), (4) *and* (8)*.*

Proof. With a change of variable we can write

$$\begin{aligned}
&\int_b^{1-b}\left|b^{-1/2}\Gamma_n^{(1)}(t)+g_{(n)}(t)\right|^p\,\mathrm{d}\mu(t)-m_n^c(l,p)\\
&=b\int_1^{(1-b)/b}\left\{\left|b^{-1/2}\int_{t-1}^{t+1}k(t-y)\,\mathrm{d}W(L(by))+g_{(n)}(tb)\right|^p w(tb)\right.\\
&\qquad\left.-\int_{\mathbb{R}}\left|l(tb)Dx+g_{(n)}(tb)\right|^p\phi(x)\,\mathrm{d}x\right\}\mathrm{d}t\\
&=b\left\{\sum_{i=1}^{M_1-1}\xi_i+\eta\right\},
\end{aligned}\tag{88}$$

where $M_1=[1/b-1]$,

$$\begin{aligned}
\xi_i=\int_i^{i+1}&\left\{\left|b^{-1/2}\int_{t-1}^{t+1}k(t-y)\,\mathrm{d}W(L(by))+g_{(n)}(tb)\right|^p\right.\\
&\left.-\int_{-\infty}^{+\infty}\left|l(tb)Dx+g_{(n)}(tb)\right|^p\phi(x)\,\mathrm{d}x\right\}w(tb)\,\mathrm{d}t
\end{aligned}\tag{89}$$

and

$$\begin{aligned}
\eta=\int_{M_1}^{(1-b)/b}&\left\{\left|b^{-1/2}\int_{t-1}^{t+1}k(t-y)\,\mathrm{d}W(L(by))+g_{(n)}(tb)\right|^p\right.\\
&\left.-\int_{-\infty}^{+\infty}\left|l(tb)Dx+g_{(n)}(tb)\right|^p\phi(x)\,\mathrm{d}x\right\}w(tb)\,\mathrm{d}t.
\end{aligned}$$

First, we show that η has no effect on the asymptotic distribution, i.e. is negligible. Using Jensen inequality and $(a+b)^p\leq 2^p(a^p+b^p)$ and the fact that l

and w are bounded, we obtain

$$\eta^2 \leq \int_{M_1}^{(1-b)/b} \left\{ \left| b^{-1/2} \int_{t-1}^{t+1} k(t-y)\,\mathrm{d}W(L(by)) + g_{(n)}(tb) \right|^{2p} + \left(\int_{\mathbb{R}} \left| l(tb)Dx + g_{(n)}(tb) \right|^p \phi(x)\,\mathrm{d}x \right)^2 \right\} w(tb)\,\mathrm{d}t$$
$$\leq C_1 \int_{M_1}^{(1-b)/b} \left\{ \left| b^{-1/2} \int_{t-1}^{t+1} k(t-y)\,\mathrm{d}W(L(by)) \right|^{2p} + \left| g_{(n)}(tb) \right|^{2p} \right\} \mathrm{d}t + C_2,$$

for some positive constants C_1 and C_2. On the other hand,

$$\begin{aligned}\int_{M_1}^{(1-b)/b} \left| g_{(n)}(tb) \right|^{2p} \mathrm{d}t &= (nb)^p \int_{M_1}^{(1-b)/b} \left| \lambda_{(n)}(tb) - \lambda(tb) \right|^{2p} \mathrm{d}t \\ &= (nb)^p b^{-1} \int_{M_1 b}^{1-b} \left| \lambda_{(n)}(t) - \lambda(t) \right|^{2p} \mathrm{d}t \\ &= (nb)^p b^{-1} \int_{M_1 b}^{1-b} \left| \int k(y)[\lambda(t-by) - \lambda(t)]\,\mathrm{d}y \right|^{2p} \mathrm{d}t \\ &\leq (nb)^p b^{4p} \sup_{t\in[0,1]} |\lambda''(t)|^{2p} \left| \int k(y) y^2\,\mathrm{d}y \right|^{2p}\end{aligned}$$

Hence,

$$\begin{aligned}\mathbb{E}[\eta^2] &\leq C_1 \int_{M_1}^{(1-b)/b} \mathbb{E}\left[\left| b^{-1/2} \int_{t-1}^{t+1} k(t-y)\,\mathrm{d}W(L(by)) \right|^{2p} \right] + 2C_3 (nb)^p b^{4p} + C_2 \\ &= O\left((nb)^p b^{4p} \right) = O(1).\end{aligned} \tag{90}$$

This means that $b\eta = o_P(1)$. The statement follows immediately from Lemma A.2. □

Lemma A.2. *Let $L(t) = \int_0^t l(u)\,\mathrm{d}u$ for a differentiable function $l(t)$ on $[0,1]$ such that $\inf_{[0,1]} l(t) > 0$ and $\sup_{[0,1]} |l'(t)| < \infty$. Assume that (A1) and (A3) hold. Let ξ_i, for $i = 1, \ldots, M_1 - 1$, be defined as in (89). Then we have*

$$b^{1/2} \gamma(p)^{-1} \sum_{i=1}^{M_1 - 1} \xi_i \to N(0,1),$$

where $\gamma^2(p)$ is defined in (53).

Proof. Let $\gamma \in (0,1)$ and $M_2 = [(M_1 - 1)^\gamma]$, $M_3 = [(M_1 - 1)/(M_2 + 2)]$. Define

$$\zeta_i = \sum_{j=(i-1)(M_2+2)+1}^{(i-1)(M_2+2)+M_2} \xi_j, \qquad i = 1, \ldots, M_3$$

$$\gamma_i = \xi_{iM_2+2i-1} + \xi_{iM_2+2i}, \qquad \gamma^* = \sum_{j=M_3(M_2+2)+1}^{M_1-1} \xi_j.$$

With this notation we can write

$$\sum_{i=1}^{M_1-1} \xi_i = \sum_{i=1}^{M_3} \zeta_i + \sum_{i=1}^{M_3} \gamma_i + \gamma^*$$

and we aim at showing that the first term in the right hand side of the previous equation determines the asymptotic distribution of $\sum_{i=0}^{M_1-1} \xi_i$.

Note that

$$b^{-1/2} \int_{t-1}^{t+1} k(t-y)\,\mathrm{d}W(L(by)) \sim N\left(0, \sigma_t^2\right)$$

where

$$\sigma_t^2 = \int_{t-1}^{t+1} k^2(t-y) l(by)\,\mathrm{d}y = D^2 l(bt) + O(b^2)$$

and

$$\begin{aligned}
&\mathbb{E}\left[\left|b^{-1/2} \int_{t-1}^{t+1} k(t-y)\,\mathrm{d}W(L(by)) + g_{(n)}(tb)\right|^p\right] \\
&= \int_{-\infty}^{+\infty} \left|\sigma_t x + g_{(n)}(tb)\right|^p \phi(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} \left|D\sqrt{l(tb)}x + g_{(n)}(tb)\right|^p \phi(x)\,\mathrm{d}x + O(b^2).
\end{aligned}$$

Hence, we get $\mathbb{E}[\xi_i] = O(b^2)$ and $\mathbb{E}[\gamma_i] = O(b^2)$. Furthermore, and, as we did for η, it can be seen that $\mathbb{E}[\xi_i^2] = O(1)$ and $\mathbb{E}[\gamma_i^2] = O(1)$.

Since γ_i depends only on the Brownian motion on the interval $[L(b(iM_2+2i-2)), L(b(iM_2+2i+2))]$, it follows that γ_i are independent (note that $M_2 > 2$). Moreover, γ^* is independent of γ_i, $i = 1, \ldots, M_3 - 1$ and $\mathbb{E}[\gamma^*] = O(M_2 b^2)$. In addition, since ξ_i is independent of ξ_j for $|i-j| \geq 3$, we also have $\mathbb{E}[(\gamma^*)^2] \leq CM_2$. As a result

$$\mathbb{E}\left[\left(\sum_{i=1}^{M_3} \gamma_i + \gamma^*\right)^2\right] \leq c(M_3 + M_2) = o(1/b) \tag{91}$$

because $bM_2 \to 0$ and $bM_3 \to 0$. Indeed $M_2 \leq (T/b)^\gamma$ and

$$b\left[\frac{[(1-b)/b]}{[[(1-b)/b]^\gamma] + 2}\right] \leq \frac{1-b}{[(1-b)/b]^\gamma + 1} \leq \frac{1-b}{1 + \frac{(1-2b)^\gamma}{b^\gamma}} = \frac{b^\gamma}{(1-2b)^\gamma + b^\gamma} \to 0.$$

Consequently

$$b^{1/2}\left(\sum_{i=1}^{M_3} \gamma_i + \gamma^*\right) \xrightarrow{\mathbb{P}} 0.$$

Next, since ζ_i, $i = 1, \dots, M_3$ are independent, we apply the central limit theorem to conclude that

$$b^{1/2}\gamma(p)^{-1}\sum_{i=0}^{M_3}\zeta_i \to N(0,1)$$

It suffices to show that

$$b\mathbb{E}\left[\left(\sum_{i=1}^{M_3}\zeta_i\right)^2\right] = b\sum_{i=0}^{M_3}\mathbb{E}[\zeta_i^2] \to \gamma^2(p). \tag{92}$$

and that they satisfy the Lyapunov's condition

$$\frac{\sum_i \mathbb{E}[\zeta_i^4]}{\left(\sum_i \mathbb{E}[\zeta_i^2]\right)^2} \to 0.$$

Once we have (92), the Lyapunov's condition is equivalent to $b^2\sum_i \mathbb{E}[\zeta_i^4] \to 0$. Using

$$\mathbb{E}[\zeta_i^4] = 4!\sum_{\substack{k,l,m,r\in I_i \\ k\le l\le m\le r}}\mathbb{E}[\xi_k\xi_l\xi_m\xi_r],$$

where $I_i = \{(i-1)(M_2+2)+1, \dots, (i-1)(M_2+2)+M_2\}$, the fact that

$$\mathbb{E}[\xi_k\xi_l\xi_m\xi_r] = O(b^2)^4 \qquad \text{if} \qquad l \ge k+3 \text{ or } r \ge m+3$$

and that all the moments of the ξ_i's are finite, we obtain that

$$\mathbb{E}[\zeta_i^4] = O(M_2^2), \qquad \text{(uniformly w.r.t. } i). \tag{93}$$

Consequently $b^2\sum_i \mathbb{E}[\zeta_i^4] = O(b^2M_3M_2^2) \to 0$ because $bM_2 \to 0$ and $bM_3M_2 = O(1)$. Indeed

$$bM_2M_3 \le bM_2\frac{M_1-1}{M_2+2} \le bM_1 \le 1.$$

In particular, it also follows that

$$b\sum_i \mathbb{E}[\zeta_i^2] = b\mathbb{E}\left[\left(\sum_{i=0}^{M_3}\zeta_i\right)^2\right] + bO(M_3^2M_2^2b^4) = O(bM_3M_2) = O(1). \tag{94}$$

Now we prove (92). From (88), it follows that

$$\text{Var}\left(\int_b^{1-b}\left|b^{-1/2}\Gamma_n^{(1)}(t) + g_{(n)}(t)\right|^p \,\mathrm{d}\mu(t)\right) = b^2\text{Var}\left(\sum_{i=1}^{M_1-1}\xi_i + \eta\right).$$

Moreover, since $\mathbb{E}[\xi_i] = O(b^2)$ for $i = 1, \ldots, M_1 - 1$ and $\mathbb{E}[\eta] = 0$, we get

$$\begin{aligned}
&b^{-1}\mathrm{Var}\left(\int_b^{1-b} \left|l(t)b^{-1/2}\Gamma_n^{(1)}(t) + g_{(n)}(t)\right|^p \,\mathrm{d}\mu(t)\right)\\
&= b\mathbb{E}\left[\left(\sum_{i=1}^{M_1-1} \xi_i + \eta\right)^2\right] + o(1)\\
&= b\mathbb{E}[\eta^2] + 2b\mathbb{E}\left[\left(\sum_{i=1}^{M_1-1} \xi_i\right)\eta\right] + b\mathbb{E}\left[\left(\sum_{i=1}^{M_1-1} \xi_i\right)^2\right] + o(1)
\end{aligned}$$

We have already shown in the proof of the previous lemma that $\mathbb{E}[\eta^2] = O(1)$, so the first term in the right hand side of the previous equation converges to zero. Furthermore,

$$\begin{aligned}
b\mathbb{E}\left[\left(\sum_{i=1}^{M_1-1} \xi_i\right)^2\right] &= b\mathbb{E}\left[\left(\sum_{i=1}^{M_3} \zeta_i + \sum_{i=1}^{M_3} \gamma_i + \gamma^*\right)^2\right]\\
&= b\mathbb{E}\left[\left(\sum_{i=1}^{M_3} \zeta_i\right)^2\right] + b\mathbb{E}\left[\left(\sum_{i=1}^{M_3} \gamma_i + \gamma^*\right)^2\right]\\
&\quad + b\mathbb{E}\left[\left(\sum_{i=1}^{M_3} \zeta_i\right)\left(\sum_{i=1}^{M_3} \gamma_i + \gamma^*\right)\right].
\end{aligned}$$

Now, making use of (91), (94) and the fact that, by Cauchy-Schwartz,

$$\mathbb{E}\left[\left(\sum_{i=1}^{M_3} \zeta_i\right)\left(\sum_{i=1}^{M_3} \gamma_i + \gamma^*\right)\right] \leq \mathbb{E}\left[\left(\sum_{i=1}^{M_3} \zeta_i\right)^2\right]^{1/2} \mathbb{E}\left[\left(\sum_{i=1}^{M_3} \gamma_i + \gamma^*\right)^2\right]^{1/2}$$

we obtain

$$b\mathbb{E}\left[\left(\sum_{i=1}^{M_1-1} \xi_i\right)^2\right] = b\mathbb{E}\left[\left(\sum_{i=1}^{M_3} \zeta_i\right)^2\right] + o(1).$$

Similarly,

$$\begin{aligned}
b\mathbb{E}\left[\left(\sum_{i=1}^{M_1-1} \xi_i\right)\eta\right] &= b\mathbb{E}\left[\left(\sum_{i=1}^{M_3} \zeta_i + \sum_{i=1}^{M_3} \gamma_i + \gamma^*\right)\eta\right]\\
&\leq b\mathbb{E}[\eta^2]^{1/2}\left\{\mathbb{E}\left[\left(\sum_{i=1}^{M_3} \zeta_i\right)^2\right]^{1/2} + \mathbb{E}\left[\left(\sum_{i=1}^{M_3} \gamma_i + \gamma^*\right)^2\right]^{1/2}\right\}
\end{aligned}$$

and the right hand side converges to zero. This means that

$$b\mathbb{E}\left[\left(\sum_{i=1}^{M_3} \zeta_i\right)^2\right] = b^{-1}\mathrm{Var}\left(\int_b^{1-b} \left|b^{-1/2}\Gamma_n^{(1)}(t) + g_{(n)}(t)\right|^p \,\mathrm{d}\mu(t)\right) + o(1).$$

Moreover, from Lemma A.3, it follows that

$$\begin{aligned}
&b^{-1}\mathrm{Var}\left(\int_b^{1-b}\left|b^{-1/2}\Gamma_n^{(1)}(t)+g_{(n)}(t)\right|^p\,\mathrm{d}\mu(t)\right)\\
&=\frac{1}{b}\int_b^{1-b}\int_b^{1-b}\Bigg\{\mathbb{E}\Bigg[\left|b^{-1/2}\int_{t-b}^{t+b}k\left(\frac{t-y}{b}\right)\,\mathrm{d}W(L(y))+g_{(n)}(t)\right|^p\\
&\qquad\qquad\left|b^{-1/2}\int_{u-b}^{u+b}k\left(\frac{u-y}{b}\right)\,\mathrm{d}W(L(y))+g_{(n)}(u)\right|^p\Bigg]\\
&\quad-\int_{-\infty}^{+\infty}\left|\sigma_n(t)x+g_{(n)}(t)\right|^p\phi(x)\,\mathrm{d}x\int_{-\infty}^{+\infty}\left|\sigma_n(u)y+g_{(n)}(u)\right|^p\phi(y)\,\mathrm{d}y\Bigg\}\,w(t)w(u)\,\mathrm{d}t\,\mathrm{d}u\\
&=\frac{1}{b}\int_b^{1-b}\int_b^{1-b}\int_{\mathbb{R}^2}\Big\{\left|\sigma_n(u)y+g_{(n)}(u)\right|^p\\
&\qquad\qquad\left|g_{(n)}(t)+\sigma_n(t)\rho_n(t,u)y+\sqrt{1-\rho_n^2(t,u)}\sigma_n(t)x\right|^p\\
&\quad-\left|\sigma_n(t)x+g_{(n)}(t)\right|^p\left|\sigma_n(u)y+g_{(n)}(u)\right|^p\Big\}w(t)w(u)\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t\,\mathrm{d}u\\
&=\frac{1}{b}\int_b^{1-b}\int_b^{1-b}\int_{\mathbb{R}^2}\Big\{\left|\sqrt{L'(u)}Dy+g_{(n)}(u)\right|^p\\
&\qquad\qquad\left|g_{(n)}(t)+\sigma_n(t)\rho_n(t,u)y+\sqrt{1-\rho_n^2(t,u)}\sigma_n(t)x\right|^p\\
&\quad-\left|\sqrt{L'(t)}Dx+g_{(n)}(t)\right|^p\left|\sqrt{L'(u)}Dy+g_{(n)}(u)\right|^p\Big\}w(t)w(u)\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t\,\mathrm{d}u
\end{aligned}$$

where $\rho_n(t,u)$ and $\sigma_n(t)$ are defined respectively in (96) and (95).

First we consider the case $nb^5\to 0$ and show that we can remove the $g_{(n)}$ functions from the previous integral. Indeed, since

$$\begin{aligned}
&\left|\left|\sqrt{L'(u)}Dy+g_{(n)}(u)\right|^p-|\sqrt{L'(u)}Dy|^p\right|\\
&\le p2^{p-1}|g_{(n)}(u)|^p+p2^{p-1}|\sqrt{L'(u)}Dy|^{p-1}|g_{(n)}(u)|
\end{aligned}$$

we obtain

$$\begin{aligned}
&\Bigg|A_n-\frac{1}{b}\int_b^{1-b}\int_b^{1-b}\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\sqrt{L'(u)}Dy\right|^p\\
&\quad\cdot B_n(t,u,x,y)w(t)w(u)\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t\,\mathrm{d}u\Bigg|\\
&\le\frac{c}{b}\int_b^{1-b}\int_b^{1-b}\int_{\mathbb{R}}\int_{\mathbb{R}}\left|g_{(n)}(u)\right|^p|B_n(t,u,x,y)|w(t)w(u)\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t\,\mathrm{d}u\\
&+\frac{c}{b}\int_b^{1-b}\int_b^{1-b}\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\sqrt{L'(u)}Dy\right|^{p-1}\left|g_{(n)}(u)\right|\\
&\quad\cdot|B_n(t,u,x,y)|w(t)w(u)\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t\,\mathrm{d}u,
\end{aligned}$$

where

$$A_n=b^{-1}\mathrm{Var}\left(\int_b^{1-b}\left|b^{-1/2}\Gamma_n^{(1)}(t)+g_{(n)}(t)\right|^p\,\mathrm{d}\mu(t)\right)$$

and

$$B_n(t,u,x,y) = \left| g_{(n)}(t) + \sigma_n(t)\rho_n(t,u)y + \sqrt{1-\rho_n^2(t,u)}\sigma_n(t)x \right|^p \\ - \left| \sqrt{L'(t)}Dx + g_{(n)}(t) \right|^p.$$

Note that, if $|t-u| \geq 2b$, then $\rho_n(t,u) = 0$ and the previous integrands are equal to zero. Hence, a sufficient condition for the left hand side of the previous inequality to converge to zero is to have

$$b^{-1} \int_b^{1-b} \int_b^{1-b} \mathbb{1}_{\{|t-u|<2b\}} \left| g_{(n)}(u) \right|^p \left| g_{(n)}(t) \right|^p \, \mathrm{d}u \, \mathrm{d}t \to 0.$$

and

$$b^{-1} \int_b^{1-b} \int_b^{1-b} \mathbb{1}_{\{|t-u|<2b\}} \left| g_{(n)}(u) \right|^p \, \mathrm{d}u \, \mathrm{d}t \to 0.$$

This is indeed the case because $g_n(u) = O\left((nb)^{1/2}b^2\right)$ uniformly w.r.t. u and $(nb)^{1/2}b^2 \to 0$. In the same way we can remove also the other $g_{(n)}$ functions from the integrand, i.e.

$$A_n = \frac{1}{b} \int_b^{1-b} \int_b^{1-b} \int_{\mathbb{R}} \int_{\mathbb{R}} \left| \sqrt{L'(u)}Dy \right|^p B_n'(t,u,x,y) w(t) w(u) \phi(x) \phi(y) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}t \, \mathrm{d}u + o(1)$$

where

$$B_n'(t,u,x,y) = \left| \sigma_n(t)\rho_n(t,u)y + \sqrt{1-\rho_n^2(t,u)}\sigma_n(t)x \right|^p - \left| \sqrt{L'(t)}Dx \right|^p$$

With the change of variable $t = u + sb$, we get

$$A_n = \int_b^{1-b} \int_{\substack{1-u/b \\ |s|\leq 2}}^{(1-b-u)/b} \int_{\mathbb{R}} \int_{\mathbb{R}} \left| \sqrt{L'(u)}\sqrt{L'(u+sb)}D^2 y \right|^p \\ \left\{ \left| yr(s) + \sqrt{1-r^2(s)}x \right|^p - |x|^p \right\} w(u) w(u+sb) \phi(x) \phi(y) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}s \, \mathrm{d}u + o(1),$$

where $r(s)$ is defined in (7). The continuity of the functions l and w and the dominated convergence theorem yield

$$A_n = \int_b^{1-b} \int_{|s|\leq 2} \int_{\mathbb{R}^2} \left| \sqrt{L'(u)} \right|^{2p} D^{2p} |y|^p \left\{ \left| yr(s) + \sqrt{1-r^2(s)}x \right|^p - |x|^p \right\} \\ w(u)^2 \phi(x) \phi(y) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}s \, \mathrm{d}u + o(1).$$

Then, with the change of variable $yr(s) + \sqrt{1-r^2(s)}x = z$ we can write equivalently

$$A_n = D^{2p}\int_b^{1-b}\left|\sqrt{L'(u)}\right|^{2p} w(u)^2\,\mathrm{d}u\frac{1}{2\pi}$$
$$\int_{\mathbb{R}^3}|y|^p\left\{|z|^p-\left|\frac{z-r(s)y}{\sqrt{1-r^2(s)}}\right|^p\right\}e^{-\frac{z^2+y^2-2rzy}{2(1-r^2(s))}}\frac{1}{\sqrt{1-r^2(s)}}\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}s+o(1)$$
$$=\sigma_1 D^{2p}\int_0^1\left|\sqrt{L'(u)}\right|^{2p} w(u)^2\,\mathrm{d}u+o(1)$$

where σ^1 is defined in (10).

Let us now consider the case $nb^5\to c_0^2>0$. First we show that the $g_{(n)}(u)$ functions can be replaced by $g(u)$ defined in (5). Indeed, $g_{(n)}(u)=g(u)+o((nb)^{1/2}b^2)$, where the big O term is uniform w.r.t. u and similar calculations to those of the previous case allow us to conclude that

$$A_n=\frac{1}{b}\int_b^{1-b}\int_b^{1-b}\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\sqrt{L'(u)}Dy+g(u)\right|^p B_n'(t,u,x,y)$$
$$w(t)w(u)\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t\,\mathrm{d}u+o(1)$$

where

$$B_n'(t,u,x,y)=\left|g(t)+\sqrt{L'(t)}D\left[\rho_n(t,u)y+\sqrt{1-\rho_n^2(t,u)}x\right]\right|^p$$
$$-\left|\sqrt{L'(t)}Dx+g(t)\right|^p.$$

With the change of variable $t=u+sb$, we get

$$A_n=\int_b^{1-b}\int_{\substack{(b-u)/b\\|s|\le 2}}^{(1-b-u)/b}\int_{\mathbb{R}}\int_{\mathbb{R}}\left|g(u)+\sqrt{L'(u)}Dy\right|^p$$
$$\left\{\left|g(u+sb)+\sqrt{L'(u+sb)}D[yr(s)+\sqrt{1-r^2(s)}x]\right|^p\right.$$
$$\left.-\left|g(u+sb)+\sqrt{L'(u+sb)}Dx\right|^p\right\}w(u)w(u+sb)\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\,\mathrm{d}u$$
$$+o(1).$$

Again, by the continuity of the functions l, w and g and the dominated convergence theorem we obtain that A_n converges to

$$\int_0^1\int_{\mathbb{R}^3}\left|g(u)+\sqrt{L'(u)}Dy\right|^p\left\{\left|g(u)+\sqrt{L'(u)}D[yr(s)+\sqrt{1-r^2(s)}x]\right|^p\right.$$
$$\left.-\left|g(u)+\sqrt{L'(u)}Dx\right|^p\right\}w(u)^2\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\,\mathrm{d}u,$$

which is exactly $\theta^2(p)$ defined in (11). □

Lemma A.3. *Let* $L(t)=\int_0^t l(u)\,\mathrm{d}u$ *for be a differentiable function* $l(t)$ *on* $[0,1]$ *such that* $\inf_{[0,1]} l(t)>0$ *and* $\sup_{[0,1]}|l'(t)|<\infty$*. For* $t\in[0,1]$*, define*

$$X_{n,t}=b^{-1/2}\int_{t-b}^{t+b} k\left(\frac{t-y}{b}\right)\mathrm{d}W(L(y))+g_{(n)}(t).$$

It holds

$$\mathbb{E}\left[|X_{n,t}X_{n,u}|^p\right]=\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\sigma_n(u)y+g_{(n)}(u)\right|^p \left|g_{(n)}(t)+\sigma_n(t)\rho_n(t,u)y+\sqrt{1-\rho_n^2(t,u)}\sigma_n(t)x\right|^p\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y,$$

where

$$\sigma_n^2(t)=l(t)D^2+O(b^2),\qquad \sigma_n(t,u)=b^{-1}\int k(t-y)k(u-y)l(y)\,\mathrm{d}y. \tag{95}$$

and

$$\rho_n(t,u)=\frac{\int k\left(\frac{t-y}{b}\right)k\left(\frac{u-y}{b}\right)l(y)\,\mathrm{d}y}{b\sqrt{D^2l(t)+O(b^2)}\sqrt{D^2l(u)+O(b^2)}} \tag{96}$$

Proof. First, note that

$$(X_{n,t},X_{n,u})\sim N\left(\begin{bmatrix}g_{(n)}(t)\\ g_{(n)}(u)\end{bmatrix},\begin{bmatrix}\sigma_n^2(t) & \sigma_n(t,u)\\ \sigma_n(t,u) & \sigma_n^2(u)\end{bmatrix}\right).$$

Hence, we have

$$X_{n,t}|X_{n,u}=x_2\sim N\left(g_{(n)}(t)+\frac{\sigma_n(t)}{\sigma_n(u)}\rho_n(t,u)\left(x_2-g_{(n)}(u)\right),(1-\rho_n^2(t,u))\sigma_n^2(t)\right).$$

Consequently, we obtain

$$\begin{aligned}
&\mathbb{E}\left[|X_{n,t}X_{n,u}|^p\right]=\mathbb{E}\left[\mathbb{E}\left[|X_{n,t}X_{n,u}|^p|X_{n,u}\right]\right]\\
&=\mathbb{E}\left[|X_{n,u}|^p\int_{\mathbb{R}}\left|g_{(n)}(t)+\frac{\sigma_n(t)}{\sigma_n(u)}\rho_n(t,u)\left(X_{n,u}-g_{(n)}(u)\right)\right.\right.\\
&\qquad\left.\left.+\sqrt{1-\rho_n^2(t,u)}\sigma_n(t)x\right|^p\phi(x)\,\mathrm{d}x\right]\\
&=\int_{\mathbb{R}}\left|\sigma_n(u)y+g_{(n)}(u)\right|^p\\
&\qquad\int_{\mathbb{R}}\left|g_{(n)}(t)+\sigma_n(t)\rho_n(t,u)y+\sqrt{1-\rho_n^2(t,u)}\sigma_n(t)x\right|^p\phi(x)\,\mathrm{d}x\,\phi(y)\,\mathrm{d}y\\
&=\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\sigma_n(u)y+g_{(n)}(u)\right|^p\\
&\qquad\left|g_{(n)}(t)+\sigma_n(t)\rho_n(t,u)y+\sqrt{1-\rho_n^2(t,u)}\sigma_n(t)x\right|^p\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y.
\end{aligned}$$

□

Proof of Proposition 3.3. We first prove (i). For each $t \in [0, b)$, we have

$$\begin{aligned}\tilde{\lambda}_n^s(t) - \lambda(t) &= \int_0^{t+b} k_b(t-u)\,\mathrm{d}\Lambda_n(u) - \lambda(t)\\ &= \int_0^{t+b} k_b(t-u)\,\mathrm{d}(\Lambda_n - \Lambda)(u) + \int_0^{t+b} k_b(t-u)\,\mathrm{d}\Lambda(u) - \lambda(t)\\ &= \int_0^{t+b} k_b(t-u)\,\mathrm{d}(\Lambda_n - \Lambda)(u) + \int_{-1}^{t/b} k(y)[\lambda(t-by) - \lambda(t)]\,\mathrm{d}y\\ &\quad - \lambda(t)\int_{t/b}^{1} k(y)\,\mathrm{d}y.\end{aligned}$$

Note that

$$\begin{aligned}&\left|\int_0^{t+b} k_b(t-u)\,\mathrm{d}(\Lambda_n - \Lambda)(u)\right|\\ &\quad= \frac{1}{b^2}\left|\int_0^{t+b} (\Lambda_n - \Lambda)(u) k'\left(\frac{t-u}{b}\right)\,\mathrm{d}u\right|\\ &\quad\le cb^{-1}\sup_{u \le 2b}|M_n(u) - M_n(0)|\\ &\quad\le cb^{-1}\left\{\sup_{u\le 2b}\left|M_n(u) - n^{-1/2}B_n\circ L(u)\right| + n^{-1/2}\left|B_n\circ L(u) - B_n\circ L(0)\right|\right\}\\ &\quad= O_P\left(b^{-1}n^{-1+1/q}\right) + n^{-1/2}b^{-1}\sup_{y\le cb}|B_n(y)| = O_P\left((nb)^{-1/2}\right),\end{aligned} \tag{97}$$

uniformly in $t \in [0, b]$, and that according to (21),

$$\left|\int_{-1}^{t/b} k(y)[\lambda(t-by) - \lambda(t)]\,\mathrm{d}y\right| = O(b),$$

Moreover, for $t \le b/2$,

$$\lambda(t)\int_{t/b}^{1} k(y)\,\mathrm{d}y \ge \inf_{t\in[0,1]}\lambda(t)\int_{1/2}^{1} k(y)\,\mathrm{d}y = C > 0.$$

Now, define the event

$$\mathcal{A}_n = \left\{\sup_{t\in[0,b]}\left(\left|\int_0^{t+b} k_b(t-u)\,\mathrm{d}(\Lambda_n - \Lambda)(u)\right| + \left|\int_{-1}^{t/b} k(y)[\lambda(t-by) - \lambda(t)]\,\mathrm{d}y\right|\right) \le C/2\right\}.$$

Then, $\mathbb{P}(\mathcal{A}_n) \to 1$ and on the event $\mathcal{A}_n$, $|\tilde{\lambda}_n^s(t) - \lambda(t)| \geq C/2$. Consequently we obtain

$$\mathbb{E}\left[\int_0^b \left|\tilde{\lambda}_n^s(t) - \lambda(t)\right|^p \mathrm{d}\mu(t)\right] \geq \mathbb{E}\left[\int_0^{b/2} \left|\tilde{\lambda}_n^s(t) - \lambda(t)\right|^p \mathrm{d}\mu(t)\right]$$
$$\geq \mathbb{E}\left[\mathbb{1}_{\mathcal{A}_n} \int_0^{b/2} \left|\hat{\lambda}_n^s(t) - \lambda(t)\right|^p \mathrm{d}\mu(t)\right] \geq c\mathbb{P}(\mathcal{A}_n)b, \tag{98}$$

for some $c > 0$. Hence

$$(nb)^{p/2}\mathbb{E}\left[\int_0^b \left|\tilde{\lambda}_n^s(t) - \lambda(t)\right|^p \mathrm{d}\mu(t)\right] \geq cb(nb)^{p/2}\mathbb{P}(\mathcal{A}_n) \to \infty,$$

because $b(nb)^{p/2} \geq b(nb)^{1/2} = (nb^3)^{1/2} \to \infty$.

In order to prove (ii), due to (14), we can bound

$$b^{-1/2}\left|\int_0^b (nb)^{p/2}\left|\tilde{\lambda}_n^s(t) - \lambda(t)\right|^p \mathrm{d}\mu(t) - \int_0^b \left|g_{(n)}(t)\right|^p \mathrm{d}\mu(t)\right|$$

by

$$p2^{p-1}b^{-1/2}(nb)^{p/2}\int_0^b \left|\int_0^{t+b} k_b(t-u)\,\mathrm{d}(\Lambda_n - \Lambda)(u)\right|^p \mathrm{d}\mu(t)$$
$$+ p2^{p-1}b^{-1/2}\left(\int_0^b (nb)^{p/2}\left|\int_0^{t+b} k_b(t-u)\,\mathrm{d}(\Lambda_n - \Lambda)(u)\right|^p \mathrm{d}\mu(t)\right)^{1/p}$$
$$\cdot\left(\int_0^b \left|g_{(n)}(t)\right|^p \mathrm{d}\mu(t)\right)^{1-1/p}.$$

According to (97)

$$\left|\int_0^{t+b} k_b(t-u)\,\mathrm{d}(\Lambda_n - \Lambda)(u)\right| = O_P\left((nb)^{-1/2}\right),$$

uniformly in $t \in [0, b]$. Furthermore, using (20), (21), and (22), we have

$$g_{(n)}(t) = O\left((nb)^{1/2}\right), \tag{99}$$

uniformly for $t \in [0, b]$. Hence, we obtain

$$b^{-1/2}\left|\int_0^b (nb)^{p/2}\left|\tilde{\lambda}_n^s(t) - \lambda(t)\right|^p \mathrm{d}\mu(t) - \int_0^b \left|g_{(n)}(t)\right|^p \mathrm{d}\mu(t)\right|$$
$$\leq O_P\left(b^{1/2}\right) + O_P\left(n^{(p-1)/2}b^{p/2}\right) \to 0,$$

because $n^{(p-1)/2}b^{p/2} = (bn^{1-1/p})^{p/2} \to 0$.

Next we deal with (iii). Again by means of (14), we can bound

$$b^{-1/2}\left|\int_0^b (nb)^{p/2}\left|\tilde{\lambda}_n^s(t)-\lambda(t)\right|^p \mathrm{d}\mu(t) - \int_0^b \left|Y_n(t)+g_{(n)}(t)\right|^p \mathrm{d}\mu(t)\right|$$

by

$$p2^{p-1}b^{-1/2}\left\{\int_0^b \left|(nb)^{1/2}\int_0^{t+b} k_b(t-u)\,\mathrm{d}(\Lambda_n-\Lambda-n^{-1/2}B_n\circ L)(u)\right|^p \mathrm{d}\mu(t)\right.$$
$$+\left(\int_0^b \left|(nb)^{1/2}\int_0^{t+b} k_b(t-u)\,\mathrm{d}(\Lambda_n-\Lambda-n^{-1/2}B_n\circ L)(u)\right|^p \mathrm{d}\mu(t)\right)^{1/p}$$
$$\left.\cdot\left(\int_0^b \left|Y_n(t)+g_{(n)}(t)\right|^p \mathrm{d}\mu(t)\right)^{1-1/p}\right\}$$

Note that

$$\sup_{t\in[0,b]}|Y_n(t)| = \sup_{t\in[0,b]}\left|b^{1/2}\int_0^{t+b} k_b(t-u)\,\mathrm{d}B_n(L(u))\right| = O_P(1),$$

and, as in (97),

$$\left|\int_0^{t+b} k_b(t-u)\,\mathrm{d}(\Lambda_n-\Lambda-n^{-1/2}B_n\circ L)(u)\right|$$
$$\leq \frac{1}{b}\sup_{u\leq 2b}\left|(\Lambda_n-\Lambda-n^{-1/2}B_n\circ L)(u)\right| = O_P\left(b^{-1}n^{-1+1/q}\right),$$

uniformly for $t\in[0,b]$. Together with (99), we obtain

$$b^{-1/2}\left|\int_0^b (nb)^{p/2}\left|\tilde{\lambda}_n^s(t)-\lambda(t)\right|^p \mathrm{d}\mu(t) - \int_0^b \left|Y_n(t)+g_{(n)}(t)\right|^p \mathrm{d}\mu(t)\right|$$
$$\leq O_P\left(b^{-1/2}(nb)^{p/2}bn^{-p+p/q}b^{-p}\right) + O_P\left(b^{-1/2}b(nb)^{p/2}n^{-1+1/q}b^{-1}\right)$$
$$= b^{-1/2}(nb)^{p/2}n^{-1+1/q}\left\{O_P\left((n^{-1+1/q}b^{-1})^{p-1}\right)+O_P(1)\right\}.$$

Because $n^{-1+1/q}b^{-1} = O(1)$, the term within the brackets is of order $O_P(1)$, and since $b^{p-1}n^{p-2+2/q} \to 0$, the right hand side tends to zero. This proves (25).

Then, by Jensen's inequality, we get

$$\begin{aligned}
&b^{-1}\mathrm{Var}\left(\int_0^{cb} |Y_n(t)+g_{(n)}(t)|^p \,\mathrm{d}\mu(t)\right)\\
&\quad= b^{-1}\mathbb{E}\left[\left(\int_0^{cb} |Y_n(t)+g_{(n)}(t)|^p \,\mathrm{d}\mu(t) - \int_0^{cb} \mathbb{E}\left[|Y_n(t)+g_{(n)}(t)|^p\right]\,\mathrm{d}\mu(t)\right)^2\right]\\
&\quad\geq b^{-1}\mathbb{E}\left[\left|\int_0^{cb} |Y_n(t)+g_{(n)}(t)|^p \,\mathrm{d}\mu(t) - \int_0^{cb} \mathbb{E}\left[|Y_n(t)+g_{(n)}(t)|^p\right]\,\mathrm{d}\mu(t)\right|\right]^2.
\end{aligned} \tag{100}$$

Note that $Y_n(t) \sim N(0, \sigma_n^2(t))$, where,

$$\sigma_n^2(t) = b^{-1}\int_0^{t+b} k^2\left(\frac{t-u}{b}\right) L'(u)\,\mathrm{d}u = \int_{-1}^{t/b} k^2(y)L'(t-by)\,\mathrm{d}y,$$

if B_n is a Brownian motion, and

$$\sigma_n^2(t) = \int_{-1}^{t/b} k^2(y)L'(t-by)\,\mathrm{d}y + O(b),$$

if B_n is a Brownian bridge. Now, choose $\epsilon > 0$. Then

$$\liminf_{n\to\infty} \mathbb{P}\left(\epsilon \leq Y_n(0) \leq 2\epsilon\right) > 0 \qquad \text{and} \qquad \liminf_{n\to\infty} \mathbb{P}\left(-2\epsilon \leq Y_n(0) \leq -\epsilon\right) > 0.$$

For $c > 0$, define the events

$$\begin{aligned}
\mathcal{A}_{n1} &= \left\{\epsilon/2 \leq Y_n(t) \leq 3\epsilon, \text{ for all } t \in [0, cb]\right\},\\
\mathcal{A}_{n2} &= \left\{-3\epsilon \leq Y_n(t) \leq -\epsilon/2, \text{ for all } t \in [0, cb]\right\},
\end{aligned}$$

and let

$$\mathcal{B}_n = \left\{\int_0^{cb} \mathbb{E}\left[|Y_n(t)+g_{(n)}(t)|^p\right]\,\mathrm{d}\mu(t) > \int_0^{cb} |g_{(n)}(t)|^p \,\mathrm{d}\mu(t)\right\}.$$

Then, since Y_n has continuous paths, we have

$$\liminf_{n\to\infty} \mathbb{P}(\mathcal{A}_{n1}) > 0 \qquad \text{and} \qquad \liminf_{n\to\infty} \mathbb{P}(\mathcal{A}_{n2}) > 0.$$

Moreover, $Y_n(t) > 0$ on the event $\mathcal{A}_{n1}$, and from (23), it follows that $Y_n(t) + g_{(n)}(t) < 0$, for n sufficiently large. Therefore, for n sufficiently large, we have on $\mathcal{A}_{n1}$,

$$\int_0^{cb} |Y_n(t)+g_{(n)}(t)|^p \,\mathrm{d}\mu(t) \leq \int_0^{cb} |\epsilon/2 + g_{(n)}(t)|^p \,\mathrm{d}\mu(t). \tag{101}$$

Similarly, $Y_n(t) < 0$ on the event $\mathcal{A}_{n2}$ and $Y_n(t) + g_{(n)}(t) < 0$, for large n, so that on $\mathcal{A}_{n2}$,

$$\int_0^{cb} |Y_n(t) + g_{(n)}(t)|^p \,\mathrm{d}\mu(t) \geq \int_0^{cb} | - \epsilon/2 + g_{(n)}(t)|^p \,\mathrm{d}\mu(t). \tag{102}$$

Next, write

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_0^{cb} |Y_n(t) + g_{(n)}(t)|^p \,\mathrm{d}\mu(t) - \int_0^{cb} \mathbb{E}\left[|Y_n(t) + g_{(n)}(t)|^p\right] \,\mathrm{d}\mu(t)\right|\right] \\
&\geq \mathbb{E}\left[\left|\int_0^{cb} |Y_n(t) + g_{(n)}(t)|^p \,\mathrm{d}\mu(t) - \int_0^{cb} \mathbb{E}\left[|Y_n(t) + g_{(n)}(t)|^p\right] \,\mathrm{d}\mu(t)\right| \mathbb{1}_{\mathcal{A}_{n1}}\right] \mathbb{1}_{\mathcal{B}_n} \\
&\quad + \mathbb{E}\left[\left|\int_0^{cb} |Y_n(t) + g_{(n)}(t)|^p \,\mathrm{d}\mu(t) - \int_0^{cb} \mathbb{E}\left[|Y_n(t) + g_{(n)}(t)|^p\right] \,\mathrm{d}\mu(t)\right| \mathbb{1}_{\mathcal{A}_{n2}}\right] \mathbb{1}_{\mathcal{B}_n^c}.
\end{aligned} \tag{103}$$

Consider the first term on the right hand side. Because for n large, $Y_n(t) + g_{(n)}(t) < 0$ on the event $\mathcal{A}_{n1}$, we have $|Y_n(t) + g_{(n)}(t)| \leq |g_{(n)}(t)|$. It follows that on the event $\mathcal{A}_{n1} \cap \mathcal{B}_n$:

$$\int_0^{cb} |Y_n(t)+g_{(n)}(t)|^p \,\mathrm{d}\mu(t) \leq \int_0^{cb} |g_{(n)}(t)|^p \,\mathrm{d}\mu(t) < \int_0^{cb} \mathbb{E}\left[|Y_n(t) + g_{(n)}(t)|^p\right] \,\mathrm{d}\mu(t).$$

This means that we can remove the absolute value signs in the first term on the right hand side of (103). Similarly, $Y_n(t) + g_{(n)}(t) < 0$, for n sufficiently large on the event $\mathcal{A}_{n2}$, so that on the event $\mathcal{A}_{n2} \cap \mathcal{B}_n^c$:

$$\int_0^{cb} |Y_n(t)+g_{(n)}(t)|^p \,\mathrm{d}\mu(t) \geq \int_0^{cb} |g_{(n)}(t)|^p \,\mathrm{d}\mu(t) \geq \int_0^{cb} \mathbb{E}\left[|Y_n(t) + g_{(n)}(t)|^p\right] \,\mathrm{d}\mu(t),$$

so that we can also remove the absolute value signs in the second term on the right hand side of (103). It follows that the right hand of (103) is equal to

$$\begin{aligned}
&\mathbb{E}\left[\int_0^{cb} \left(\mathbb{E}\left[|Y_n(t) + g_{(n)}(t)|^p\right] \,\mathrm{d}\mu(t) - \int_0^{cb} |Y_n(t) + g_{(n)}(t)|^p \,\mathrm{d}\mu(t)\right) \mathbb{1}_{\mathcal{A}_{n1}}\right] \mathbb{1}_{\mathcal{B}_n} \\
&+ \mathbb{E}\left[\left(\int_0^{cb} |Y_n(t) + g_{(n)}(t)|^p \,\mathrm{d}\mu(t) - \int_0^{cb} \mathbb{E}\left[|Y_n(t) + g_{(n)}(t)|^p\right] \,\mathrm{d}\mu(t)\right) \mathbb{1}_{\mathcal{A}_{n2}}\right] \mathbb{1}_{\mathcal{B}_n^c} \\
&\geq \left(\int_0^{cb} |g_{(n)}(t)|^p \,\mathrm{d}\mu(t) - \int_0^{cb} |\epsilon/2 + g_{(n)}(t)|^p \,\mathrm{d}\mu(t)\right) \mathbb{P}(\mathcal{A}_{n1}) \mathbb{1}_{\mathcal{B}_n} \\
&\quad + \left(\int_0^{cb} | - \epsilon/2 + g_{(n)}(t)|^p \,\mathrm{d}\mu(t) - \int_0^{cb} |g_{(n)}(t)|^p \,\mathrm{d}\mu(t)\right) \mathbb{P}(\mathcal{A}_{n2}) \mathbb{1}_{\mathcal{B}_n^c},
\end{aligned}$$

by using (21) and (22). Furthermore, for the first term on the right hand side

$$|g_{(n)}(t)|^p - |\epsilon/2 + g_{(n)}(t)|^p = |g_{(n)}(t)|^p \left(1 - |\epsilon_n(t) + 1|^p\right),$$

where $\epsilon_n(t) = \epsilon/(2g_{(n)}(t)) = O((nb)^{-1/2}) \to 0$, due to (20), (21) and (22), where the big-O term is uniformly for $t \in [0, b]$. This means that, for n large, $1 + \epsilon_n(t) > 0$, and by a Taylor expansion $|1 + \epsilon_n(t)|^p = 1 + p\epsilon_n(t) + O((nb)^{-1})$. It follows that

$$\begin{aligned}
&\int_0^{cb} |g_{(n)}(t)|^p \,\mathrm{d}\mu(t) - \int_0^{cb} |\epsilon/2 + g_{(n)}(t)|^p \,\mathrm{d}\mu(t) \\
&= \int_0^{cb} |g_{(n)}(t)|^p \left\{1 - |\epsilon_n(t) + 1|^p\right\} \mathrm{d}\mu(t) \\
&= -p \int_0^{cb} |g_{(n)}(t)|^p \epsilon_n(t) \,\mathrm{d}\mu(t) + cb \sup_{t\in[0,cb]} |g_{(n)}(t)|^p O((nb)^{-1}) \\
&= p(\epsilon/2) \int_0^{cb} |g_{(n)}(t)|^{p-1} \,\mathrm{d}\mu(t) + O\left(b(nb)^{(p-1)/2}\right) \\
&= O\left(b(nb)^{(p-1)/2}\right)
\end{aligned}$$

due to (99). Similarly

$$\int_0^{cb} |-\epsilon/2 + g_{(n)}(t)|^p \,\mathrm{d}\mu(t) - \int_0^{cb} |g_{(n)}(t)|^p \,\mathrm{d}\mu(t) = O\left(b(nb)^{(p-1)/2}\right).$$

Going back to (100), since $\mathbb{P}(\mathcal{A}_{n1}) \to 1$ and $\mathbb{P}(\mathcal{A}_{n2}) \to 1$, we conclude that

$$b^{-1}\mathrm{Var}\left(\int_0^{cb} |Y_n(t) + g_{(n)}(t)|^p \,\mathrm{d}\mu(t)\right) \geq b^{-1} O\left(b(nb)^{(p-1)/2}\right)^2.$$

The statement follows from the fact that $b^{-1}(nb)^{p-1}b^2 = n^{p-1}b^p \to \infty$.

Finally, one can deal in the same way with the L_p-error on the interval $(1-b, 1]$. □

Proof of Proposition 3.6. By definition we have

$$\begin{aligned}
&(nb)^{p/2} \int_0^b \left|\hat{\lambda}_n^s(t) - \lambda(t)\right|^p \mathrm{d}\mu(t) \\
&\quad = \int_0^b \left|(nb)^{1/2} \int_0^{t+b} k_b^{(t)}(t-u) \,\mathrm{d}(\Lambda_n - \Lambda)(u) + \bar{g}_{(n)}(t)\right|^p \mathrm{d}\mu(t),
\end{aligned}$$

where

$$\bar{g}_{(n)}(t) = (nb)^{1/2} \left(\int k_b^{(t)}(t-u)\lambda(u) \,\mathrm{d}u - \lambda(t)\right). \tag{104}$$

When B_n in assumption (A2) is a Brownian motion, we can argue as in the proof of Theorem 3.1. By means of (14) we can bound

$$\begin{aligned}
b^{-1/2}\Bigg|&(nb)^{p/2} \int_0^b \left|\hat{\lambda}_n^s(t) - \lambda(t)\right|^p \mathrm{d}\mu(t) \\
&- \int_0^b \left|b^{-1/2} \int_0^{t+b} k^{(t)}\left(\frac{t-u}{b}\right) \mathrm{d}B_n(L(u)) + \bar{g}_{(n)}(t)\right|^p \mathrm{d}\mu(t)\Bigg|,
\end{aligned}$$

from above by

$$
\begin{aligned}
&p2^{p-1}b^{-1/2}b^{-p/2}\int_0^b\left|\int_0^{t+b}k^{(t)}\left(\frac{t-u}{b}\right)\mathrm{d}(B_n\circ L-n^{1/2}M_n)(u)\right|^p\mathrm{d}\mu(t)\\
&+p2^{p-1}b^{-1/2}\left(b^{-p/2}\int_0^b\left|\int_0^{t+b}k^{(t)}\left(\frac{t-u}{b}\right)\mathrm{d}(B_n\circ L-n^{1/2}M_n)(u)\right|^p\mathrm{d}\mu(t)\right)^{1/p}\\
&\qquad\cdot\left(\int_0^b\left|b^{-1/2}\int_0^{t+b}k^{(t)}\left(\frac{t-u}{b}\right)\mathrm{d}B_n(L(u))+\bar{g}_{(n)}(t)\right|^p\mathrm{d}\mu(t)\right)^{1-1/p}.
\end{aligned}
\tag{105}
$$

Similar to (18),

$$
\begin{aligned}
&\sup_{t\in[0,b]}\left|\int_0^{t+b}k^{(t)}\left(\frac{t-u}{b}\right)\,\mathrm{d}(B_n\circ L-n^{1/2}M_n)(u)\right|\\
&\quad\le\left|\int_{-1}^{t/b}\left\{\psi_1\left(\frac{t}{b}\right)k(y)+\psi_2\left(\frac{t}{b}\right)yk(y)\right\}\,\mathrm{d}(B_n\circ L-n^{1/2}M_n)(t-by)\right|\\
&\quad\le C\sup_{t\in[0,1]}\left|B_n\circ L(t)-n^{1/2}M_n(t)\right|\\
&\quad=O_P(n^{-1/2+1/q}).
\end{aligned}
\tag{106}
$$

Note that here we used the boundedness of the coefficients ψ_1 and ψ_2. Similar to the proof of Theorem 3.1, the idea is to show that

$$
b^{-1/2}\int_0^b\left|b^{-1/2}\int_0^{t+b}k^{(t)}\left(\frac{t-u}{b}\right)\,\mathrm{d}B_n(L(u))+\bar{g}_{(n)}(t)\right|^p\,\mathrm{d}\mu(t)\to 0, \tag{107}
$$

in probability. We first bound the left hand side of (107) by

$$
Cb^{-1/2}\int_0^b\left\{|\bar{g}_{(n)}(t)|^p+b^{-p/2}\left|\int_0^{t+b}k^{(t)}\left(\frac{t-u}{b}\right)\,\mathrm{d}B_n(L(u))\right|^p\right\}\,\mathrm{d}\mu(t).
$$

According to (28), a Taylor expansion gives

$$
\begin{aligned}
\sup_{t\in[0,b]}|\bar{g}_{(n)}(t)|&=(nb)^{1/2}\sup_{t\in[0,b]}\left|\int_0^{t+b}k_b^{(t)}(t-u)\lambda(u)\,\mathrm{d}u-\lambda(t)\right|\\
&=(nb)^{1/2}\sup_{t\in[0,b]}\left|\int_{-1}^{t/b}k^{(t)}(y)\left[\lambda(t-by)-\lambda(t)\right]\,\mathrm{d}y\right|\\
&=(nb)^{1/2}b^2\sup_{t\in[0,b]}\left|\frac{1}{2}\int_{-1}^{t/b}k^{(t)}(y)y^2\lambda''(\xi_{t,y})\,\mathrm{d}y\right|\\
&=O_P\left((nb^5)^{1/2}\right)=O_P(1).
\end{aligned}
$$

Furthermore,

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_0^{t+b} k^{(t)}\left(\frac{t-u}{b}\right)\,\mathrm{d}B_n(L(u))\right|^p\right] \\
&= \int_{\mathbb{R}}\left(\int_0^{t+b}\left(k^{(t)}\left(\frac{t-u}{b}\right)\right)^2 L'(u)\,\mathrm{d}u\right)^{p/2} |x|^p\phi(x)\,\mathrm{d}x \\
&= b^{p/2}\int_{\mathbb{R}}\left(\int_0^{t+b}\left(k^{(t)}\left(\frac{t-u}{b}\right)\right)^2 L'(u)\,\mathrm{d}u\right)^{p/2} |x|^p\phi(x)\,\mathrm{d}x \\
&= O(b^{p/2}),
\end{aligned}$$

where ϕ denotes the standard normal density. This proves (107) for the case that B_n is a Brownian motion.

When B_n in (A2) is a Brownian bridge, then we use the representation $B_n(u) = W_n(u) - uW_n(L(1))/L(1)$, for some Brownian motion W_n. In this case, by means of (14), we can bound

$$\begin{aligned}
b^{-1/2}\Bigg|&\int_0^b\left|b^{-1/2}\int_0^{t+b} k^{(t)}\left(\frac{t-u}{b}\right)\,\mathrm{d}B_n(L(u)) + \bar{g}_{(n)}(t)\right|^p \mathrm{d}\mu(t) \\
&- \int_0^b\left|b^{-1/2}\int_0^{t+b} k^{(t)}\left(\frac{t-u}{b}\right)\,\mathrm{d}W_n(L(u)) + \bar{g}_{(n)}(t)\right|^p \mathrm{d}\mu(t)\Bigg|
\end{aligned}$$

by

$$\begin{aligned}
&p2^{p-1}b^{-1/2}\int_0^b\left|b^{-1/2}\frac{W_n(L(1)}{L(1)}\int_0^{t+b} k^{(t)}\left(\frac{t-u}{b}\right)L'(u)\,\mathrm{d}u + \bar{g}_{(n)}(t)\right|^p \mathrm{d}\mu(t) \\
&+ p2^{p-1}b^{-1/2} \\
&\quad\cdot\left(\int_0^b\left|b^{-1/2}\frac{W_n(L(1))}{L(1)}\int_0^{t+b} k^{(t)}\left(\frac{t-u}{b}\right)L'(u)\,\mathrm{d}u + \bar{g}_{(n)}(t)\right|^p \mathrm{d}\mu(t)\right)^{1/p} \\
&\quad\cdot\left(\int_0^b\left|b^{-1/2}\int_0^{t+b} k^{(t)}\left(\frac{t-u}{b}\right)\mathrm{d}W_n(L(u)) + \bar{g}_{(n)}(t)\right|^p \mathrm{d}\mu(t)\right)^{1-1/p},
\end{aligned}$$

which tends to zero in probability, due to (107). □

Appendix B: Isotonized kernel estimator

Lemma B.1. *Assume (A1)-(A2) and let $\tilde{\lambda}_n^s$ be defined in (2). Let k satisfy (1) and let $p \geq 1$. If $b \to 0$, $nb \to \infty$, and $1/b = o(n^{1/4})$, then*

$$\mathbb{P}\left(\tilde{\lambda}_n^s \text{ is decreasing on } [b, 1-b]\right) \to 1.$$

Proof. The proof is completely similar to that of Lemma A.7 in [37]. Note that condition (8) in that paper follows from our Assumption (A2) and that here λ is a decreasing function.

We use the fact that on $[b,1-b]$, $\tilde{\lambda}_n^s$ is the standard kernel estimator of λ given by (62) and we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\lambda}_n^s(t)=\int_{t-b}^{t+b}\frac{1}{b^2}k'\left(\frac{t-u}{b}\right)\mathrm{d}\left(\Lambda_n-\Lambda\right)(u)+\int_{t-b}^{t+b}\frac{1}{b^2}k'\left(\frac{t-u}{b}\right)\lambda(u)\,\mathrm{d}u. \tag{108}$$

The first term on the right hand side of (108) converges to zero because in absolute value it is bounded from above by

$$\frac{1}{b^2}\sup_{x\in[0,1]}|\Lambda_n(x)-\Lambda(x)|\sup_{y\in[-1,1]}|k''(y)|=O_p(b^{-2}n^{-1/2})=o_p(1),$$

according to Assumption (A2) and the fact that $1/b=o(n^{-1/4})$. Moreover, integration by parts gives

$$\int\frac{1}{b^2}k'\left(\frac{t-u}{b}\right)\lambda(u)\,\mathrm{d}u=\int_{-1}^{1}k(y)\lambda'(t-by)\,\mathrm{d}y.$$

Hence, the second term on the right hand side of (108) is bounded from above by a strictly negative constant because of Assumption (A1). We conclude that $\tilde{\lambda}_n^s$ is decreasing on $[b,1-b]$ with probability tending to one. ☐

Corollary B.2. *Assume (A1)-(A2) and let $\tilde{\lambda}_n^s$ and $\tilde{\lambda}_n^{GS}$ be defined in (2) and Section 5, respectively. Let k satisfy (1). Let $0<\gamma<1$ and $p\geq 1$. If $b\to 0$, $nb\to\infty$, and $1/b=o(n^{1/4})$, then*

$$\mathbb{P}\left(\tilde{\lambda}_n^s(t)=\tilde{\lambda}_n^{GS}(t)\text{ for all }t\in[b^\gamma,1-b^\gamma]\right)\to 1.$$

Proof. The proof is completely similar to that of Lemma 3.2 in [37], but now we want to extend the interval to $[b^\gamma,1-b^\gamma]$, which is not fixed but approaches the boundaries as $n\to\infty$. In this case we define the linearly extended version of Λ_n^s by

$$\hat{\Lambda}_n^*(t)=\begin{cases}\Lambda_n^s(b^\gamma)+\big(t-b^\gamma\big)\tilde{\lambda}_n^s(b^\gamma), & \text{for } t\in[0,b^\gamma),\\ \Lambda_n^s(t), & \text{for } t\in[b^\gamma,1-b^\gamma],\\ \Lambda_n^s(1-b^\gamma)+\big(t-1+b^\gamma\big)\tilde{\lambda}_n^s(1-b^\gamma), & \text{for } t\in(1-b^\gamma,1].\end{cases}$$

Choose $0<\delta<2$. It suffices to prove that, for sufficiently large n,

$$\mathbb{P}\left(\hat{\Lambda}_n^*\text{ is concave on }[0,1]\right)\geq 1-\delta/2, \tag{109}$$

and

$$\mathbb{P}\left(\hat{\Lambda}_n^*(t)\geq\Lambda_n^s(t),\ \text{for all }t\in[0,1]\right)\geq 1-\delta/2. \tag{110}$$

To prove (109), define the event

$$A_n = \left\{ \tilde{\lambda}_n^s \text{ is decreasing on } [b, 1-b] \right\}.$$

On the event A_n the curve $\hat{\Lambda}_n^*$ is concave on $[0,1]$, so

$$\mathbb{P}\left(\hat{\Lambda}_n^* \text{ is concave on } [0,1] \right) \geq \mathbb{P}(A_n),$$

and the result follows from Lemma B.1. To prove (110), we split the interval $[0,1]$ in five intervals $I_1 = [0,b)$, $I_2 = [b, b^\gamma)$, $I_3 = [b^\gamma, 1-b^\gamma]$, $I_4 = (1-b^\gamma, 1-b]$ and $I_5 = (1-b, 1]$. Then, as in Lemma 3.2 in [37], we show that

$$\mathbb{P}\left(\hat{\Lambda}_n^*(t) \geq \Lambda_n^s(t), \text{ for all } t \in I_i \right) \geq 1 - \delta/10, \quad i = 1, \dots, 5. \tag{111}$$

For $t \in I_3$, $\hat{\Lambda}_n^*(t) = \Lambda_n^s(t)$, so (111) is trivial. For $t \in I_2$, by the mean value theorem,

$$\hat{\Lambda}_n^*(t) - \Lambda_n^s(t) = \Lambda_n^s(b^\gamma) + \left(t - b^\gamma\right)\tilde{\lambda}_n^s(b^\gamma) - \Lambda_n^s(t) = (b^\gamma - t)\left[\tilde{\lambda}_n^s(\xi_t) - \tilde{\lambda}_n^s(b^\gamma)\right],$$

for some $\xi_t \in (t, b^\gamma) \subset (b, b^\gamma)$. Thus,

$$\mathbb{P}\left(\hat{\Lambda}_n^*(t) \geq \Lambda_n^s(t), \text{ for all } t \in I_2 \right) \geq \mathbb{P}(A_n) \geq 1 - \delta/10,$$

for n sufficiently large, according to Lemma B.1. The argument for I_4 is exactly the same.

Next, we consider $t \in I_1$. We have

$$\begin{aligned}
&\hat{\Lambda}_n^*(t) - \Lambda_n^s(t) \\
&\quad = \Lambda_n^s(b^\gamma) + \left(t - b^\gamma\right)\tilde{\lambda}_n^s(b^\gamma) - \Lambda_n^s(t) \\
&\quad = \left[\Lambda_n^s(b^\gamma) - \Lambda^s(b^\gamma)\right] + \left[\Lambda^s(t) - \Lambda_n^s(t)\right] + \Lambda^s(b^\gamma) - \Lambda^s(t) - \left(b^\gamma - t\right)\tilde{\lambda}_n^s(b^\gamma) \\
&\quad \geq -2 \sup_{t \in [0,1]} \left|\Lambda_n^s(t) - \Lambda^s(t)\right| + \Lambda^s(b^\gamma) - \Lambda^s(t) - (b^\gamma - t)\lambda(b^\gamma) \\
&\qquad\qquad + \left(b^\gamma - t\right)\left[\lambda(b^\gamma) - \tilde{\lambda}_n^s(b^\gamma)\right],
\end{aligned} \tag{112}$$

where Λ^s is the deterministic version of Λ_n^s,

$$\Lambda^s(t) = \int_{(t-b)\vee 0}^{(t+b)\wedge 1} k_b^{(t)}(t-u)\Lambda(u)\,\mathrm{d}u.$$

For the first term on right hand side of (112), note that

$$\begin{aligned}
\sup_{t\in[0,1]}|\Lambda_n^s(t)-\Lambda^s(t)| &= \sup_{t\in[0,1]}\left|\int_{(t-b)\vee 0}^{(t+b)\wedge 1} k_b^{(t)}(t-u)\left[\Lambda_n(u)-\Lambda(u)\right]\,\mathrm{d}u\right| \\
&= \sup_{t\in[0,1]}\left|\int k^{(t)}(y)\left[\Lambda_n(t-by)-\Lambda(t-by)\right]\,\mathrm{d}y\right| \\
&\le \sup_{t\in[0,1]}|\Lambda_n(t-by)-\Lambda(t-by)|\int\sup_{t\in[0,1]}\left|k^{(t)}(y)\right|\,\mathrm{d}y \\
&= O_P\left(n^{-1/2}\right),
\end{aligned} \tag{113}$$

due to Assumption (A2). Moreover, for the third term on right hand side of (112), for $t\in(b,1-b)$, we have

$$\begin{aligned}
\left|\lambda(t)-\tilde{\lambda}_n^s(t)\right| &\le \left|\lambda(t)-\int k_b(t-u)\lambda(u)\,\mathrm{d}u\right| + \left|\int k_b(t-u)\,\mathrm{d}(\Lambda-\Lambda_n)(u)\right| \\
&= \left|\int k(y)[\lambda(t)-\lambda(t-by)]\,\mathrm{d}y\right| + b^{-1}\left|\int k'(y)(\Lambda-\Lambda_n)(t-by)\,\mathrm{d}y\right| \\
&= O(b^2)+O_P(b^{-1}n^{-1/2}).
\end{aligned} \tag{114}$$

For the second term on right hand side of (112), for $t\in[0,b)$, we write

$$\begin{aligned}
&\Lambda^s(b^\gamma)-\Lambda^s(t)-(b^\gamma-t)\lambda(b^\gamma) \\
&\quad= \int_{b^\gamma-b}^{b^\gamma+b} k_b(b^\gamma-u)\Lambda(u)\,\mathrm{d}u - \int_0^{t+b} k_b^{(t)}(t-u)\Lambda(u)\,\mathrm{d}u - (b^\gamma-t)\lambda(b^\gamma) \\
&\quad= \int_{b^\gamma-b}^{b^\gamma+b} k_b(b^\gamma-u)[\Lambda(u)-\Lambda(b^\gamma)]\,\mathrm{d}u - \int_0^{t+b} k_b^{(t)}(t-u)[\Lambda(u)-\Lambda(t)]\,\mathrm{d}u \\
&\qquad\qquad + [\Lambda(b^\gamma)-\Lambda(t)-(b^\gamma-t)\lambda(b^\gamma)] \\
&\quad= \int_{-1}^{1} k(y)[\Lambda(b^\gamma-by)-\Lambda(b^\gamma)]\,\mathrm{d}y - \int_{-1}^{t/b} k^{(t)}(y)[\Lambda(t-by)-\Lambda(t)]\,\mathrm{d}y \\
&\qquad\qquad - \frac{1}{2}(b^\gamma-t)^2\lambda'(\xi_t) \\
&\quad\ge \int_{-1}^{1} k(y)[\Lambda(b^\gamma-by)-\Lambda(b^\gamma)]\,\mathrm{d}y - \int_{-1}^{t/b} k^{(t)}(y)[\Lambda(t-by)-\Lambda(t)]\,\mathrm{d}y \\
&\qquad\qquad - \inf_{t\in[0,1]}|\lambda'(t)|b^{1+\gamma} + \frac{1}{2}\inf_{t\in[0,1]}|\lambda'(t)|b^{2\gamma}
\end{aligned} \tag{115}$$

where $\xi_t\in(t,b^\gamma)$. Furthermore, the first two integrals on the right hand side can be written as

$$\frac{b^2}{2}\int_{-1}^{1} k(y)y^2\lambda'(\xi_{1,y})\,\mathrm{d}y - \frac{b^2}{2}\int_{-1}^{t/b} k^{(t)}(y)y^2\lambda'(\xi_{2,y})\,\mathrm{d}y$$
$$\geq -\frac{b^2}{2}\left|\int_{-1}^{1} k(y)y^2\lambda'(\xi_{1,y})\,\mathrm{d}y - \int_{-1}^{t/b} k^{(t)}(y)y^2\lambda'(\xi_{2,y})\,\mathrm{d}y\right|$$
$$\geq -\frac{b^2}{2}\left|\int_{-1}^{1} k(y)y^2\lambda'(\xi_{1,y})\,\mathrm{d}y - \int_{-1}^{t/b} k^{(t)}(y)y^2\lambda'(\xi_{2,y})\,\mathrm{d}y\right| = O(b^2),$$

with $\xi_t \in (t, b^\gamma)$, $|\xi_{1,y} - b^\gamma| \leq by$ and $|\xi_{2,y} - t| \leq by$. This means that

$$\mathbb{P}\left(\hat{\Lambda}_n^*(t) - \Lambda_n^s(t) \geq 0,\ \text{for all } x \in I_1\right) \geq \mathbb{P}\left(Y_n \leq \frac{1}{2}\inf_{t\in[0,1]}|\lambda'(t)|b^{2\gamma}\right),$$

where

$$Y_n = O_P(n^{-1/2}) + O(b^\gamma)\left\{O(b^2) + O_P(b^{-1}n^{-1/2})\right\} + O(b^2) - \inf_{t\in[0,1]}|\lambda'(t)|b^{1+\gamma}$$
$$= O_P(b^{1+\gamma}).$$

Hence, for n large enough, this probability is greater than $1 - \delta/10$, because $\gamma < 1$. □

Appendix C: CLT for the Hellinger loss

Lemma C.1. *Assume (A1)-(A3) hold. If λ is strictly positive, we have*

$$\int_0^1 \left(\sqrt{\hat{\lambda}_n^s(t)} - \sqrt{\lambda(t)}\right)^2 \mathrm{d}\mu(t) = \int_0^1 \left(\hat{\lambda}_n^s(t) - \lambda(t)\right)^2 (4\lambda(t))^{-1}\,\mathrm{d}\mu(t) + O_P\left((nb)^{-3/2}\right).$$

The previous results holds also if we replace $\hat{\lambda}_n^s$ with the smoothed Grenander-type estimator $\tilde{\lambda}_n^{SG}$.

Proof. As in the proof of Lemma 2.1 in [39] we get

$$\int_0^1 \left(\sqrt{\hat{\lambda}_n^s(t)} - \sqrt{\lambda(t)}\right)^2 \mathrm{d}\mu(t) = \int_0^1 \left(\hat{\lambda}_n^s(t) - \lambda(t)\right)^2 (4\lambda(t))^{-1}\,\mathrm{d}\mu(t) + R_n,$$

where

$$|R_n| \leq C\int_0^1 \left|\hat{\lambda}_n^s(t) - \lambda(t)\right|^3 \mathrm{d}\mu(t)$$

for some positive constant C only depending on $\lambda(0)$ and $\lambda(1)$. Then, from Corollary 3.7, it follows that $R_n = O_P\left((nb)^{-3/2}\right)$. When dealing with the smoothed Grenander-type estimator, the result follows from Theorem 4.4. □

Theorem C.2. *Assume (A1)-(A3) hold and that λ is strictly positive.*

i) If $nb^5 \to 0$, then it holds

$$(b\sigma^{2,*}(2))^{-1/2}\left\{2nbH(\hat{\lambda}_n^s,\lambda)^2 - m_n^*(2)\right\} \xrightarrow{d} N(0,1).$$

ii) If $nb^5 \to C_0^2 > 0$ and B_n in Assumption (A2) is a Brownian motion, then it holds

$$(b\theta^{2,*}(2))^{-1/2}\left\{2nbH(\hat{\lambda}_n^s,\lambda)^2 - m_n^*(2)\right\} \xrightarrow{d} N(0,1),$$

iii) If $nb^5 \to C_0^2 > 0$ and B_n in Assumption (A2) is a Brownian bridge, then it holds

$$(b\tilde{\theta}^{2,*}(2))^{-1/2}\left\{2nbH(\hat{\lambda}_n^s,\lambda)^2 - m_n^*(2)\right\} \xrightarrow{d} N(0,1),$$

where $\sigma^{2,*}$, $\theta^{2,*}$, $\tilde{\theta}^{2,*}$ and m_n^* are defined, respectively, as in (9), (11), (12) and (8) by replacing $w(t)$ with $w(t)(4\lambda(t))^{-1}$.

If $p < min(q, 2q-7)$ and $1/b = o\left(n^{(1/3-1/q)\min(q/(2p),1)}\right)$, the same results hold also when replacing $\hat{\lambda}_n^s$ by the smoothed Grenander-type estimator $\tilde{\lambda}_n^{SG}$.

Proof. According to Lemma C.1, it is sufficient to show that the results hold if we replace $2H(\hat{\lambda}_n^s,\lambda)^2$ by

$$\int_0^1 \left(\hat{\lambda}_n^s(t) - \lambda(t)\right)^2 (4\lambda(t))^{-1}\,\mathrm{d}\mu(t) = \int_0^1 \left(\hat{\lambda}_n^s(t) - \lambda(t)\right)^2 \mathrm{d}\tilde{\mu}(t),$$

where

$$d\tilde{\mu}(t) = \frac{1}{4\lambda(t)}\,\mathrm{d}\mu(t) = \frac{w(t)}{4\lambda(t)}\,\mathrm{d}t.$$

It suffices to apply Corollary 3.7 with a weight $\tilde{\mu}$ instead of μ.

For the smoothed Grenander estimator the result would follow from Theorem 4.4. □

Appendix D: Consistency of bootstrap

Let $\mathbb{P}^*(\cdot|Y_1, Y_2, \ldots)$ be the bootstrap probability measure, conditional on the observations $Y_1, Y_2, \ldots$ and let $\mathbb{P}_n^*$ be the empirical distribution of $Y_1^*, \ldots, Y_n^*$, conditional on the observations. We first have to show that bootstrap versions of assumptions (A1)-(A2) are satisfied. Note that the bootstrap versions of λ, Λ and M_n are $\tilde{\lambda}_n^{SG}$, $\Lambda^*(t) = \int_0^t \tilde{\lambda}_n^{SG}(u)\,\mathrm{d}u$, and $M_n^* = \Lambda_n^* - \Lambda^*$, respectively, where Λ_n^* is the Breslow estimator based on the bootstrap sample.

By definition, the smoothed Grenander estimator is twice continuously differentiable. Moreover, from the proofs of Lemmas 3.1 and 3.2 in [14] with $b = cn^{-\gamma}$

for $\gamma \in (1/6, 1/5)$, it can be shown that there exists an event A_n with $\mathbb{P}(A_n) \to 1$, such that, on A_n,

$$\sup_{[0,1]} \left| \tilde{\lambda}_n^{SG}(t) - \lambda(t) \right| \leq C_1 b^2 \log n,$$
$$\sup_{[0,1]} \left| (\tilde{\lambda}_n^{SG})'(t) - \lambda'(t) \right| \leq C_1 b \log n, \tag{116}$$

for some $C_1 > 0$. Furthermore, on A_n we also obtain

$$\sup_{(b,1-b)} \left| (\tilde{\lambda}_n^{SG})''(t) - \lambda''(t) \right| \leq C_1 (nb^5)^{-1/2} \log n, \tag{117}$$

by using a third order Tailor expansion for $\Lambda(t - bu)$ (instead of a second order expansion in [14]). Hence, there exists a positive constant C such that for n sufficiently large, on A_n, we have $\inf_{t\in[0,1]} |(\tilde{\lambda}_n^{SG})'(t)| > C$. This means that, conditionally on the observations, the bootstrap version of assumption (A1) holds on the event A_n. It can be proved as in Theorem 5 of [13] that, conditionally on the observations, the bootstrap version of assumption (A2) holds with a Brownian motion W_n^* with respect to $\mathbb{P}^*$, $L^*(t) = \hat{\sigma}_n^2 t$, and $q = \infty$. Furthermore, let A'_n be the event on which

$$n^{1/2} |\hat{\sigma}_n^2 - \sigma^2| \leq \log n. \tag{118}$$

Since $\sqrt{n}(\hat{\sigma}_n^2 - \sigma^2) = O_P(1)$ (it can be proved using the central limit theorem), this means $\mathbb{P}(A'_n) \to 1$, and consequently $\mathbb{P}(A_n \cap A'_n) \to 1$.

With (A1)-(A2) holding with probability tending to one, Theorem 7.1 can be proved following the same line of reasoning as in Theorem 4.1. It suffices to show, that

$$\sup_{x\in\mathbb{R}} \left| \mathbb{P}^* \left(b^{-p} \int_b^{1-b} |Y_n^*(t)|^p \, \mathrm{d}\mu(t) \leq x, A_n \cap A'_n \right) - \Psi(x) \right| \to 0,$$

in probability, as $n \to \infty$, where Ψ is the distribution function of $\alpha_0^p [\mathrm{D}_{\mathbb{R}} Z](0)^p$ and

$$Y_n^*(t) = bn^{2/3} \left(\int_{t-b}^{t+b} k_b(t-u) \, \mathrm{d}(\tilde{\Lambda}_n^* - \Lambda_n^*)(u) \right)$$

where $\tilde{\Lambda}_n^*$ is the least concave majorant of Λ_n^*. We need to prove bootstrap versions of Lemmas 8.2-8.6, which essentially relies on bootstrap versions of Lemmas 3, 5-8 in [38]. Note that this requires that $\Lambda^*(1)$ is bounded, which is satisfied on A_n.

The bootstrap version of Lemma 3 in [38] can be proved in the same way by replacing Λ_n and Λ_n^W with their bootstrap versions Λ_n^* and $\Lambda_n^{W,*} = \Lambda^* + n^{-1/2} W_n^* \circ L^*$ and restricting ourselves to the event $A_n \cap A'_n$. We can then use the fact that $(L^*)'(t)$ and $(\tilde{\lambda}_n^{SG})'(t)$ are bounded from above and below on $[0,1]$, uniformly in n. Similarly, for the bootstrap version of Lemmas 5 and 6 in [38], the proofs remain the same after replacing the quantities Y_{nt}, Λ_n^E, and Λ_n^W, with their bootstrap versions

$$Y_{nt}^*(s) = n^{1/6}\left\{W_n^*\left(L^*(t+n^{-1/3}s)\right) - W_n^*\left(L^*(t)\right)\right\} + \frac{1}{2}(\tilde{\lambda}_n^{SG})'(t)s^2,$$

$\Lambda_n^{E,*} = \Lambda_n^*$, and $\Lambda_n^{W,*}$. Furthermore, the process W_n^* has the same distribution as the process W^*, which is a standard Brownian motion with respect to $\mathbb{P}^*$.

In order to obtain a bootstrap version of Lemma 7 in [38], we keep a_{nt}, b_{nt} and H_{nt} as in [38], and replace the quantities $\tilde{Y}_{nt}$, Z_{nt} and J_{nt} by the following bootstrap versions

$$\begin{aligned} Z_{nt}^*(s) &= W^*((L^*)'(t)s) + \frac{1}{2}(\tilde{\lambda}_n^{SG})'(t)s^2, \\ \tilde{Y}_{nt}^* &= W^*\left(n^{1/3}\left(L^*\left(t+n^{-1/3}s\right) - L^*(t)\right)\right) + \frac{1}{2}(\tilde{\lambda}_n^{SG})'(t)s^2, \end{aligned}$$

and

$$J_{nt}^* = \left[\frac{n^{1/3}\left(L^*(a_{nt}) - L^*(t)\right)}{(L^*)'(t)}, \frac{n^{1/3}\left(L^*(b_{nt}) - L^*(t)\right)}{(L^*)'(t)}\right].$$

Note that, since $L^*(t) = t\hat{\sigma}_n^2$, we actually have $\tilde{Y}_{nt}^*(s) = Z_{nt}^*(s)$ and $J_{nt}^* = H_{nt}$. Hence, it is trivial that

$$\mathbb{E}^*\left[[\mathrm{D}_{H_{nt}}\tilde{Y}_{nt}^*](0)^p\right] = \mathbb{E}^*\left[[\mathrm{D}_{J_{nt}^*}Z_{nt}^*](0)^p\right].$$

If the function L was not linear, then we would follow the proof of Lemma 4.4 in [33], making use of Lemma 4.3 in [33] with a bootstrap version of the function Φ_{nt}. As in the proof of Lemma 4.4 in [33], we now have

$$c_1^*(t)Z_{nt}^*(c_2^*(t)s) \stackrel{d}{=} W^*(s) - s^2$$

where $c_1^*(t)$ and $c_2^*(t)$ are defined as in (49) with λ' and L' replaced by $(\tilde{\lambda}_n^{SG})'$ and $(L^*)'$, respectively. Moreover, since c_1^* is uniformly bounded on $A_n \cap A_n'$, it follows that

$$\mathbb{E}^*\left(\sup_{s\in\mathbb{R}}|Z_{nt}^*(s)|\right)^p < \infty,$$

where the expectation is taken with respect to $\mathbb{P}^*$. The rest of the proof continues as in [33], and for $J = E, W$, we obtain that on $A_n \cap A_n'$,

$$\mathbb{E}^*\left[A_n^{J,*}(t)^p\right] = \left(\frac{2(L^*)'(t)^2}{|(\tilde{\lambda}_n^{SG})'(t)|}\right)^{p/3}\mathbb{E}^*\left[\zeta^*(0)^p\right] + o\left(n^{-1/6}\right),$$

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

$$\mathbb{E}^*\left[A_n^{J,*}(t)^p\right] \leq \left(\frac{2(L^*)'(t)^2}{|(\tilde{\lambda}_n^{SG})'(t)|}\right)^{p/3}\mathbb{E}^*\left[\zeta^*(0)^p\right] + o\left(n^{-1/6}\right)$$

uniformly in $t \in (0,1)$, where $A_n^{J,*} = n^{2/3}(\mathrm{CM}_{[0,1]}\Lambda_n^{J,*} - \Lambda_n^{J,*})$ and $\zeta^*(s) = [\mathrm{CM}_{\mathbb{R}} Z^*](s) - Z^*(s)$, with $Z^*(s) = W^*(s) - s^2$. Finally, the bootstrap version of Lemma 8 in [38] can be proved in the same way as its original version, replacing the quantities A_n^J,with bootstrap versions $A_n^{J,*}$ for $J = E, W$.

Using these results, we perform similar succeeding approximations of $Y_n^*(t)$ by

$$Y_n^{(1),*}(t) = \frac{1}{b}\int_{-1}^{1} k'\left(\frac{t-v}{b}\right) A_n^{W,*}(v)\,\mathrm{d}v$$
$$Y_n^{(2),*}(t) = \frac{1}{b}\int_{-1}^{1} k'\left(\frac{t-v}{b}\right) [\mathrm{D}_{I_{nv}} A_n^{W,*}](v)\,\mathrm{d}v,$$

where $I_{nv} = [0,1] \cap [v - n^{-1/3}\log n, v + n^{-1/3}\log n]$, and

$$Y_n^{(3),*}(t) = \frac{1}{b}\int_{-1}^{1} k'\left(\frac{t-v}{b}\right) [\mathrm{D}_{H_{nv}} Y_{nv}^*](0)\,\mathrm{d}v,$$
$$Y_n^{(4),*}(t) = \frac{1}{b}\int_{t-b}^{t+b} k'\left(\frac{t-v}{b}\right) [\mathrm{D}_{J_{nv}^*} Z_{nv}^*](0)\,\mathrm{d}v,$$

as in Lemmas 8.2-8.5, where now $Y_n^{(3),*}(t) = Y_n^{(4),*}(t)$. Next, the approximation of $Y_n^{(4),*}(t)$ by

$$Y_n^{(5),*}(t) = [\mathrm{D}_{\mathbb{R}} Z^*](0)\frac{1}{b}\int_{t-b}^{t+b} k'\left(\frac{t-v}{b}\right)\frac{1}{c_1^*(v)}\,\mathrm{d}v,$$

can be proved exactly as in Lemma 8.6 using that c_1^* is uniformly bounded from below on the event $A_n \cap A_n'$. As in (52), we obtain

$$\begin{aligned}
b^{-p}\int_b^{1-b} |Y_n^{(5),*}(t)|^p\,\mathrm{d}\mu(t) &= [\mathrm{D}_{\mathbb{R}} Z^*](0)^p \int_b^{1-b}\left|\frac{(c_1^*)'(t)}{c_1^*(t)^2}\right|^p \mathrm{d}\mu(t) + o_{P^*}(1)\\
&= [\mathrm{D}_{\mathbb{R}} Z^*](0)^p \int_b^{1-b}\left|\frac{c_1'(t)}{c_1(t)^2}\right|^p \mathrm{d}\mu(t) + o_{P^*}(1)\\
&\quad + [\mathrm{D}_{\mathbb{R}} Z^*](0)^p \int_b^{1-b}\left\{\left|\frac{(c_1^*)'(t)}{c_1^*(t)^2}\right|^p - \left|\frac{c_1'(t)}{c_1(t)^2}\right|^p\right\}\mathrm{d}\mu(t).
\end{aligned}$$

By definition,

$$c_1'(t) = -\frac{1}{3}\lambda''(t)\left(\frac{1}{2\sigma^2|\lambda'(t)|^2}\right)^{1/3}$$

and

$$(c_1^*)'(t) = -\frac{1}{3}(\tilde{\lambda}_n^{SG})''(t)\left(\frac{1}{2\hat{\sigma}^2|(\tilde{\lambda}_n^{SG})'(t)|^2}\right)^{1/3}$$

In particular, on $A_n \cap A_n'$, $(c_1^*)'$ is uniformly bounded from above while c_1^* is uniformly bounded from above and below away from zero. Hence, through repeated use of the mean value theorem, using (116) and (117), it can be shown that on $A_n \cap A'_n$

$$\sup_{t\in(b,1-b)} \left| \left| \frac{(c_1^*)'(t)}{c_1^*(t)^2} \right|^p - \left| \frac{c_1'(t)}{c_1(t)^2} \right|^p \right| \leq K(nb^5)^{-1/2} \log n, \tag{119}$$

for some constant $K > 0$. It follows that, on $A_n \cap A'_n$,

$$\begin{aligned} b^{-p} \int_b^{1-b} |Y_n^{(5),*}(t)|^p \,\mathrm{d}\mu(t) &= [\mathrm{D}_{\mathbb{R}} Z^*](0)^p \int_b^{1-b} \left| \frac{c_1'(t)}{c_1(t)^2} \right|^p \,\mathrm{d}\mu(t) + o_{P^*}(1) \\ &= [\mathrm{D}_{\mathbb{R}} Z^*](0)^p \int_0^1 \left| \frac{c_1'(t)}{c_1(t)^2} \right|^p \,\mathrm{d}\mu(t) + o_{P^*}(1). \end{aligned}$$

Because $\mathbb{P}^*([\mathrm{D}_{\mathbb{R}} Z^*](0) \leq t) = \mathbb{P}([\mathrm{D}_{\mathbb{R}} Z](0) \leq t)$, for all $t \in \mathbb{R}$, the limit distribution for the bootstrap version is still the same as the one in (51).

References

[1] Akakpo, N., Balabdaoui, F. and Durot, C. (2014). Testing monotonicity via local least concave majorants. *Bernoulli* **20** 514–544. MR3178508

[2] Baraud, Y., Huet, S. and Laurent, B. (2005). Testing convex hypotheses on the mean of a Gaussian vector. Application to testing qualitative hypotheses on a regression function. *Ann. Statist.* **33** 214–257. MR2157802

[3] Birgé, L. and Massart, P. (1993). Rates of convergence for minimum contrast estimators. *Probability Theory and Related Fields* **97** 113–150. MR1240719

[4] Birke, M. and Dette, H. (2007). Testing strict monotonicity in nonparametric regression. *Math. Methods Statist.* **16** 110–123. MR2335093

[5] Birke, M. and Neumeyer, N. (2013). Testing monotonicity of regression functions—an empirical process approach. *Scand. J. Stat.* **40** 438–454. MR3091691

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

[7] Cheng, K. F. and Lin, P. E. (1981). Nonparametric estimation of a regression function. *Z. Wahrsch. Verw. Gebiete* **57** 223–233. MR0626817

[8] Chu, C.-K., Marron, J. S. et al. (1991). Comparison of two bandwidth selectors with dependent errors. *The Annals of Statistics* **19** 1906–1918. MR1135155

[9] Csörgő, M., Gombay, E. and Horváth, L. (1991). Central limit theorems for L_p distances of kernel estimators of densities under random censorship. *Ann. Statist.* **19** 1813–1831. MR1135150

[10] Csörgő, M. and Horváth, L. (1988). Central limit theorems for L_p-norms of density estimators. *Probab. Theory Related Fields* **80** 269–291. MR0968821

[11] Doss, C. R. and Wellner, J. A. (2016). Global rates of convergence of the MLEs of log-concave and s-concave densities. *The Annals of Statistics* **44** 954–981. MR3485950

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

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

[14] Durot, C., Groeneboom, P. and Lopuhaä, H. P. (2013). Testing equality of functions under monotonicity constraints. *J. Nonparametr. Stat.* **25** 939–970. MR3174305

[15] Durot, C. and Lopuhaä, H. P. (2014). A Kiefer-Wolfowitz type of result in a general setting, with an application to smooth monotone estimation. *Electron. J. Stat.* **8** 2479–2513. MR3285873

[16] Durot, C. and Tocquet, A.-S. (2003). On the distance between the empirical process and its concave majorant in a monotone regression framework. *Ann. Inst. H. Poincaré Probab. Statist.* **39** 217–240. MR1962134

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

[18] Friedman, J. and Tibshirani, R. (1984). The monotone smoothing of scatter plots. *Technometrics* **26** 243–350.

[19] Gijbels, I., Hall, P., Jones, M. C. and Koch, I. (2000). Tests for monotonicity of a regression mean with guaranteed level. *Biometrika* **87** 663–673. MR1789816

[20] Grenander, U. (1956). On the theory of mortality measurement. II. *Skand. Aktuarietidskr.* **39** 125–153 (1957). MR0093415

[21] Groeneboom, P. Estimating a monotone density. In *Proceedings of the Berkeley conference in honor of Jerzy Neyman and Jack Kiefer, Vol. II (Berkeley, Calif., 1983). Wadsworth Statist./Probab. Ser.* 539–555. MR0822052

[22] 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

[23] Groeneboom, P. and Jongbloed, G. (2010). Generalized continuous isotonic regression. *Statist. Probab. Lett.* **80** 248–253. MR2575453

[24] Groeneboom, P. and Jongbloed, G. (2012). Isotonic L_2-projection test for local monotonicity of a hazard. *J. Statist. Plann. Inference* **142** 1644–1658. MR2903377

[25] Groeneboom, P. and Jongbloed, G. (2013). Smooth and non-smooth estimates of a monotone hazard. In *From probability to statistics and back: high-dimensional models and processes. Inst. Math. Stat. (IMS) Collect.* **9** 174–196. Inst. Math. Statist., Beachwood, OH. MR3202633

[26] Groeneboom, P., Jongbloed, G. and Witte, B. I. (2010). Maximum smoothed likelihood estimation and smoothed maximum likelihood estimation in the current status model. *Ann. Statist.* **38** 352–387. MR2589325

[27] Hall, P. and Van Keilegom, I. (2005). Testing for monotone increasing hazard rate. *Ann. Statist.* **33** 1109–1137. MR2195630

[28] Kim, A. K. H., Guntuboyina, A. and Samworth, R. J. (2016). Adaptation in log-concave density estimation. *arXiv:1609.00861*. MR3845018

[29] Kim, A. K. H. and Samworth, R. J. (2016). Global rates of convergence in log-concave density estimation. *The Annals of Statistics* **44** 2756–2779. MR3576560

[30] Kulikov, V. N. and Lopuhaä, H. P. (2004). Testing for a monotone density using L_k-distances between the empirical distribution function and its concave majorant. *Report-028, Eurandom.* MR2274141

[31] 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

[32] Kulikov, V. N. and Lopuhaä, H. P. (2006). The limit process of the difference between the empirical distribution function and its concave majorant. *Statist. Probab. Lett.* **76** 1781–1786. MR2274141

[33] Kulikov, V. N. and Lopuhaä, H. P. (2008). Distribution of global measures of deviation between the empirical distribution function and its concave majorant. *J. Theoret. Probab.* **21** 356–377. MR2391249

[34] Le Cam, L. (1970). On the weak convergence of probability measures. *Annals of Mathematical Statistics* **41** 621–625. MR0254909

[35] LeCam, L. (1973). Convergence of estimates under dimensionality restrictions. *The Annals of Statistics* **1** 38–53. MR0334381

[36] Lopuhaä, H. P. and Musta, E. (2017). Smooth estimation of a monotone hazard and a monotone density under random censoring. *Stat. Neerl.* **71** 58–82. MR3605364

[37] Lopuhaä, H. P. and Musta, E. (2017). Isotonized smooth estimators of a monotone baseline hazard in the Cox model. *J. Statist. Plann. Inference* **191** 43–67. MR3679108

[38] Lopuhaä, H. P. and Musta, E. (2018). The distance between a naive cumulative estimator and its least concave majorant. *Statist. Prob. Letters* **139** 199–128. MR3802192

[39] Lopuhaä, H. P. and Musta, E. (2018). Central limit theorems for the L_p-error of smooth isotonic estimators. *arXiv preprint arXiv:1805.12430*.

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

[41] Marshall, A. W. and Proschan, F. (1965). Maximum likelihood estimation for distributions with monotone failure rate. *Annals of Mathematical Statistics* **36** 69–77. MR0170436

[42] Mukerjee, H. (1988). Monotone nonparameteric regression. *Ann. Statist.* **16** 741–750. MR0947574

[43] Pal, J. K., Woodroofe, M. and Meyer, M. (2007). Estimating a Polya frequency function$_2$. In *Complex datasets and inverse problems. IMS Lecture Notes Monogr. Ser.* **54** 239–249. Inst. Math. Statist., Beachwood, OH. MR2459192

[44] Ramsay, J. O. (1998). Estimating smooth monotone functions. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **60** 365–375. MR1616049

[45] Seregin, A. and Wellner, J. A. (2010). Nonparametric estimation of multivariate convex-transformed densities. *The Annals of Statistics* **38** 3751–3781. With supplementary material available online. MR2766867

[46] van der Vaart, A. W. and van der Laan, M. J. (2003). Smooth estimation of a monotone density. *Statistics* **37** 189–203. MR1986176
[47] Wright, F. T. (1982). Monotone regression estimates for grouped observations. *Ann. Statist.* **10** 278–286. MR0642739