Estimation of a $k-$monotone density, part 4: limit distribution theory and the spline connection

Fadoua Balabdaoui [1] and Jon A. Wellner [2]

University of Washington

September 8, 2004

Abstract

We study the asymptotic behavior of the Maximum Likelihood and Least Squares estimators of a k-monotone density g_0 at a fixed point x_0 when $k > 2$. In Balabdaoui and Wellner (2004a), it was proved that both estimators exist and are splines of degree $k-1$ with simple knots. These knots, which are also the jump points of the $(k-1)$-st derivative of the estimators, cluster around a point $x_0 > 0$ under the assumption that g_0 has a continuous k-th derivative in a neighborhood of x_0 and $(-1)^k g_0^{(k)}(x_0) > 0$.

If τ_n^- and τ_n^+ are two successive knots, we prove that the random "gap" $\tau_n^+ - \tau_n^-$ is $O_p\left(n^{-1/(2k+1)}\right)$ for any $k > 2$ if a conjecture about the upper bound on the error in a particular Hermite interpolation via odd-degree splines holds. Based on the order of the gap, the asymptotic distribution of the Maximum Likelihood and Least Squares estimators can be established. We find that the j-th derivative of the estimators at x_0 converges at the rate $n^{-(k-j)/(2k+1)}$ for $j = 0, \cdots, k-1$. The limiting distribution depends on an almost surely defined stochastic process H_k that stays above (below) the k-fold integral of Brownian motion plus a deterministic drift, when k is even (odd). The family of the processes H_k is studied separately in the companion manuscript Balabdaoui and Wellner (2004c).

[1] Research supported in part by National Science Foundation grant DMS-0203320

[2] Research supported in part by National Science Foundation grants DMS-0203320, and NIAID grant 2R01 AI291968-04

AMS 2000 *subject classifications.* Primary: 62G05; secondary 60G15, 62E20.

Key words and phrases. monotone, convex, k-monotone, Maximum Likelihood, Least Squares, rates of convergence, asymptotic distribution, mixture model, Hermite interpolation, splines, k-fold integral of Brownian motion.

Outline

1 Introduction

Our interest in the nonparametric estimation of a k-monotone density was first motivated by JEWELL (1982) on the nonparametric Maximum Likelihood estimator of a scale mixture of exponentials g_0,

$$g_0(x) = \int_0^\infty y^{-1} \exp(-x/y) dF_0(y), \qquad x > 0 \tag{1.1}$$

where F_0 is any distribution function on $(0, \infty)$. If $X_1, \cdots, X_n$ are i.i.d. with density g_0, JEWELL (1982) established that, $\hat{F}_n$, the Maximum Likelihood Estimator (MLE) of the mixing distribution F_0, exists, has at most n support points, and converges weakly to F_0 with probability 1. To our knowledge the rates of convergence of either the MLE of the mixing distribution F_0 or the true mixed density g_0 remain unknown. JEWELL (1982) noted that the complement of the true cumulative distribution $1 - G_0$ is the Laplace transform of the mixing distribution function F_0, and also the fact the class of scale of mixtures of Exponentials given in (1.1) can be identified as the class of *completely monotone* densities (Bernstein's theorem). By definition, a function g on $(0, \infty)$ is completely monotone if and only if g is infinitely differentiable on $(0, \infty)$ and $(-1)^k g^{(k)} \geq 0$, for $k \in \mathbb{R}$ (see e.g. SCHOENBERG (1938), WIDDER (1941), FELLER (1971), WILLIAMSON (1956), and GNEITING (1998)). If g is only differentiable up to a finite degree, then it is k-monotone if and only if $(-1)^j g^{(j)}$ is nonnegative, nonincreasing and convex for $j = 0, \cdots, k-2$ if $k \geq 2$ and simply nonnegative and nonincreasing if $k = 1$ (see e.g. WILLIAMSON (1956), LÉVY (1962), GNEITING (1999)). The class of completely monotone densities is the intersection of all the classes of k-monotone densities, $k \geq 1$ (GNEITING (1999)) and a completely monotone density can viewed then as an "∞-monotone" density. To prepare the ground for establishing the exact rate of convergence of the MLE for mixtures of exponentials (or, equivalently, completely monotone densities), it therefore seems natural to first establish asymptotic distribution theory for the MLE of a k-mononotone density.

When $k = 1$, the problem specializes to estimating a nonincreasing density g_0. In this case, the asymptotic distribution theory was established by PRAKASA RAO (1969), and revisited by GROENEBOOM (1985) and KIM AND POLLARD (1990). They showed that if x_0 is a fixed point such that $g_0'(x_0) < 0$ (and assuming that g_0' is continuous in a neighborhood of x_0), then the MLE $\hat{g}_n$ (the Grenander estimator), satisfies

$$n^{1/3} \left(\hat{g}_n(x_0) - g_0(x_0)\right) \to_d \left(\frac{1}{2} g_0(x_0) |g_0'(x_0)|\right)^{1/3} 2Z, \tag{1.2}$$

where $2Z$ is the slope at zero of the greatest convex minorant of two-sided Brownian motion $+t^2$, $t \in \mathbb{R}$. For $k = 2$, GROENEBOOM, JONGBLOED, AND WELLNER (2001B) considered both the MLE and LSE and established that if the true convex and nonincreasing density g_0 satisfies $g_0''(x_0) > 0$ (and assuming that g_0'' is continuous in a neighborhood of x_0), then

$$\begin{pmatrix} n^{2/5}\left(\bar{g}_n(x_0)-g_0(x_0)\right) \\ n^{1/5}\left(\bar{g}'_n(x_0)-g'(x_0)\right) \end{pmatrix} \to_d \begin{pmatrix} \left(\frac{1}{24}g_0^2(x_0)g_0''(x_0)\right)^{1/5} H^{(2)}(0) \\ \left(\frac{1}{24^3}g_0(x_0)g''(x_0)^3\right)^{1/5} H^{(3)}(0) \end{pmatrix}, \tag{1.3}$$

where $\bar{g}_n$ is the either the MLE or LSE and H is a random cubic spline function such that $H^{(2)}$ is convex and H stays above the integrated two-sided Brownian motion $+t^4, t \in \mathbb{R}$ and touches exactly at those points where $H^{(2)}$ changes its slope (see Groeneboom, Jongbloed, and Wellner (2001a)).

The key result that Groeneboom, Jongbloed, and Wellner (2001b) used to establish (1.3) is that $\tau_n^+ - \tau_n^- = O_p(n^{-1/5})$ as $n \to \infty$, where τ_n^- and τ_n^+ are two successive jump points of the first derivative of $\bar{g}_n$ in the neighborhood of x_0. This stochastic order was obtained using the characterizations of the estimators together with the "mid-point property" which we review in Section 2. For $k = 1$, the same property can be used to establish that $n^{-1/3}$ is the order of the gap. As a function of k, it is natural to conjecture that $n^{-1/(2k+1)}$ is the general form of the order of the gap. In the problem of nonparametric regression via splines, Mammen and van de Geer (1997) have conjectured that $n^{-1/(2k+1)}$ is the order of the distance between the knot points of their regression spline $\hat{m}$ under the assumption that the true regression curve m_0 satisfies our same working assumptions, but the question was left open (see Mammen and van de Geer (1997), page 400). In this manuscript, we refer to the problem of establishing the order of $\tau_n^+ - \tau_n^-$ as the *gap problem*.

In Section 2, we show that when $k > 2$, the gap problem is closely related to a "non-classical" Hermite interpolation problem via odd-degree splines. To put the interpolation problem encountered in the next section in context, it is useful to review briefly the related *complete Hermite interpolation problem* for odd-degree splines which is more "classical" and for which error bounds uniform in the knots are now available. Given a function $f \in C^{(k-1)}[0,1]$ and an increasing sequence $0 = t_0 < t_1 < \cdots < t_m < t_{m+1} = 1$ where $m \geq 1$ is an integer, it is well-known that there exists a unique spline, called the *complete spline* and denoted here by Cf, of degree $2k-1$ with interior knots $t_1, \cdots, t_m$ that satisfies the $2k + m$ conditions

$$\begin{cases} (Cf)(t_i) = f(t_i), \quad i = 1, \cdots, m \\ (Cf)^{(l)}(t_0) = f^{(l)}(t_0), \ (Cf)^{(l)}(t_{m+1}) = f^{(l)}(t_{m+1}), \quad l = 0, \cdots, k-1; \end{cases}$$

see Schoenberg (1963), de Boor (1974), or Nürnberger (1989), page 116, for further discussion. If $f \in C^{(2k)}[0,1]$, then there exists $c_k > 0$ such that

$$\sup_{0<t_1<\cdots<t_m<1} \|f - Cf\|_\infty \leq c_k \|f^{(2k)}\|_\infty. \tag{1.4}$$

This "uniform in knots" bound in the complete interpolation problem was first conjectured by de Boor (1973) in (1972) for $k > 4$ as a generalization that goes beyond $k = 2, 3$ and 4 for which the result was already established (see also de Boor

(1974)). However, it took more than 25 years to prove de Boor's conjecture; the proof of (1.4) is due to SHADRIN (2001). By a scaling argument, the bound (1.4) implies that, if $f \in C^{(2k)}[a,b], a < b \in \mathbb{R}$, the interpolation error in the complete Hermite interpolation problem is uniformly bounded in the knots, and that the bound is of the order of $(b-a)^{2k}$.

One key property of the complete spline interpolant Cf is that $(Cf)^{(k)}$ is the Least Squares approximation of $f^{(k)}$ when $f^{(k)} \in L_2([0,1])$; i.e., if $\mathcal{S}_k(t_1,\cdots,t_m)$ denotes the space of splines of order k (degree $k-1$) and interior knots $t_1,\cdots,t_m$, then

$$\int_0^1 \left((Cf)^{(k)} - f^{(k)}(x)\right)^2 dx = \min_{S\in\mathcal{S}_k(t_1,\cdots,t_m)} \int_0^1 \left(S(x) - f^{(k)}(x)\right)^2 dx \tag{1.5}$$

(see e.g. SCHOENBERG (1963), DE BOOR (1974), NÜRNBERGER (1989)). Consequently, if L_∞ denotes the space of bounded functions on $[0,1]$, then the properly defined map

$$\begin{array}{rcl} C^{(k)}[0,1] & \rightarrow & \mathcal{S}_k(t_1,\cdots,t_m) \\ f^{(k)} & \mapsto & (Cf)^{(k)} \end{array}$$

is the restriction of the orthoprojector, denoted here by $P_{S_k(\underline{t})}$, where $\underline{t} = (t_1,\cdots,t_m)$, from L_∞ to L_∞ with respect to the inner product $\langle g,h\rangle = \int_0^1 g(x)h(x)dx$. DE BOOR (1974) pointed out that, in order to prove the conjecture, it is enough to prove that

$$\sup_{\underline{t}} \|P_{S_k(\underline{t})}\|_\infty = \sup_{\underline{t}} \sup_g \frac{\|P_{S_k(\underline{t})}(g)\|_\infty}{\|g\|}$$

is bounded, and this was successfully achieved by SHADRIN (2001).

The Hermite interpolation problem which arises naturally in Section 2 appears to be another variation of Hermite interpolation problems via odd-degree splines, which has not yet been studied in the approximation theory or spline literature. More specifically, if f is some real-valued function in $C^{(j)}[0,1]$ for some $j \geq 2$, $0 = t_0 < t_1 < \cdots < t_{2k-4} < t_{2k-3} = 1$ is a given increasing sequence, then there exists a unique spline Hf of degree $2k-1$ and interior knots $t_1,\cdots,t_{2k-4}$ satisfying the $4k-4$ conditions

$$(Hf)(t_i) = f(t_i), \quad \text{and} \quad (Hf)'(t_i) = f'(t_i), \quad i = 0,\cdots,2k-3.$$

Note that the spline Hf matches not only the value of the function at the knots but also the value of its first derivative. Thus we should expect the interpolation error to be smaller than in the complete interpolation, and this gives hope that boundedness of the interpolation error also holds in this uncommon problem.

We conjecture then that there is a constant $c_{k,j}$ depending only on k and j such that

$$\sup_{0<t_1<\cdots<t_{2k-4}<1} \|f - Hf\|_\infty \leq c_{k,j}\ \|f^{(j)}\|_\infty, \tag{1.6}$$

where $\|\cdot\|_\infty$ is the supremum norm over $[0,1]$. *Numerical evidence suggesting that* $c_{k,2k} = 1/(2k)!$ *for* $k = 3, 4$ *and 5 will be presented.* **Jon: where? I guess I might suggest cutting this for now** Based on the conjecture, we can prove that the distance between two consecutive knots in a neighborhood of x_0 is $O_p(n^{-1/(2k+1)})$.

In Section 3, we are able to show that the asymptotic distributions in (1.2) and (1.3) takes the following general form:

$$\begin{pmatrix} n^{\frac{k}{2k+1}}(\bar{g}_n(x_0) - g_0(x_0)) \\ n^{\frac{k-1}{2k+1}}(\bar{g}_n^{(1)}(x_0) - g_0^{(1)}(x_0)) \\ \vdots \\ n^{\frac{1}{2k+1}}(\bar{g}_n^{(k-1)}(x_0) - g_0^{(k-1)}(x_0)) \end{pmatrix} \to_d \begin{pmatrix} c_0(g_0)H_k^{(k)}(0) \\ c_1(g_0)H_k^{(k+1)}(0) \\ \vdots \\ c_{k-1}(g_0)H_k^{(2k-1)}(0) \end{pmatrix}$$

where

$$c_j(g_0) = \left\{ (g_0(x_0))^{k-j} \left(\frac{(-1)^k g_0^{(k)}(x_0)}{k!} \right)^{2j+1} \right\}^{\frac{1}{2k+1}},$$

for $j = 0, \cdots, k-1$. The rate $n^{-(k-j)/(2k+1)}$ was found by Balabdaoui and Wellner (2004a) to be the asymptotic minimax lower bound for estimating $g_0^{(j)}(x_0)$, $j = 0, \cdots, k-1$ under the same working assumptions. The limiting distribution depends on the higher derivatives of H_k, an almost surely uniquely defined process that stays above (below) the $(k-1)$-fold integral of Brownian motion plus the drift $(k!/(2k)!)\, t^{2k}$, when k is even (odd), and is $(2k-2)-$convex; i.e. the $2k-2$ derivative of H_k is convex. The process H_k is studied separately in Balabdaoui and Wellner (2004c). Proving the existence of H_k relies also on our conjecture in (1.6) since the key problem, also referred to as the *gap problem*, depends on a very similar Hermite interpolation problem, except that the knots of the estimators are replaced by the points of touch between the $(k-1)$-fold integral of Brownian motion plus the drift $(k!/(2k)!)\, t^{2k}$ and H_k.

Section 4 gives a discussion of remaining unsolved problems.

For more discussion of the background and related problems, see Balabdaoui and Wellner (2004a). For a discussion of algorithms and computational issues, see Balabdaoui and Wellner (2004b).

2 The gap problem

Recall that it was assumed that g_0 is k-times continuously differentiable at x_0 and that $(-1)^k g_0^{(k)}(x_0) > 0$. Under a weaker assumption, Balabdaoui and Wellner (2004b) proved strong consistency of the $(k-1)$-st derivative of the MLE and LSE. This implies that the number of jump points of this derivative, in a small neighborhood of x_0, has to diverge to infinity almost surely as the sample size $n \to \infty$. This "clustering" phenomenon is one of the most crucial elements in studying the local asymptotics of the estimators. The jump points form then a sequence that converges to x_0 almost surely and therefore the distance between two successive jump points, for example located just before and after x_0, converges to 0 as $n \to \infty$. But it is not enough to know that the "gap" between these points converges to 0: an upper bound for this rate of convergence is needed.

In the following subsection, we describe the difficulty of establishing this result for $k > 2$. In the general case, the problem becomes more difficult than the problem in the special case $k = 2$.

2.1 Fundamental differences

Let τ_n^- and τ_n^+ be the last and first jump points of the $(k-1)$-st derivative of either the MLE or LSE, located before and after x_0 respectively. To obtain a better understanding of the gap problem, we describe the reasoning used by Groeneboom, Jongbloed, and Wellner (2001b) in order to prove that $\tau_n^+ - \tau_n^- = O_p(n^{-1/5})$ for the special case $k = 2$. Here, we restrict ourselves only to the LSE since it is a simpler case to deal with than the MLE. Recall that for $k = 2$ the characterization of the LSE, $\tilde{g}_n$, is given by

$$\tilde{H}_n(x) \begin{cases} \geq \mathbb{Y}_n(x), & x \geq 0 \\ = \mathbb{Y}_n(x), & \text{if and only if } x \text{ is a jump point of } \tilde{g}_n' \end{cases} \tag{2.1}$$

where

$$\tilde{H}_n(x) = \int_0^x (x-t)\tilde{g}_n(t)dt, \quad \text{and} \quad \mathbb{Y}_n(x) = \int_0^x (x-t)d\mathbb{G}_n(t),$$

and $\mathbb{G}_n$ is the empirical distribution function. For ease of notation, we omit writing the subscript n on the jump points, but their dependence on n should be kept in mind. On the interval $[\tau^-, \tau^+)$, the function $\tilde{g}_n'$ is constant since they are no more jump points in this interval. This implies that $\tilde{H}_n$ is polynomial of degree 3 on $[\tau^-, \tau^+)$. But, from the characterization in (2.1), it follows that

$$\tilde{H}_n(\tau^\pm) = \mathbb{Y}_n(\tau^\pm), \qquad \tilde{H}_n'(\tau^\pm) = \mathbb{Y}_n'(\tau^\pm)\,.$$

These four boundary conditions allow us to fully determine the cubic polynomial $\tilde{H}_n$ on $[\tau^-, \tau^+]$. Using the explicit expression for $\tilde{H}_n$ and evaluating it at the mid-point $\bar{\tau} = (\tau^- + \tau^+)/2$, Groeneboom, Jongbloed, and Wellner (2001b) established that

$$\tilde{H}_n(\bar{\tau}) = \frac{\mathbb{Y}_n(\tau^-) + \mathbb{Y}_n(\tau^+)}{2} - \frac{(\mathbb{G}_n(\tau^+) - \mathbb{G}_n(\tau^-))\,(\tau^+ - \tau^-)}{8}.$$

Groeneboom, Jongbloed and Wellner refer to this as the "mid-point property". By applying the first condition (the inequality condition) in (2.1), it follows that

$$\frac{\mathbb{Y}_n(\tau^-) + \mathbb{Y}_n(\tau^+)}{2} - \frac{(\mathbb{G}_n(\tau^+) - \mathbb{G}_n(\tau^-))(\tau^+ - \tau^-)}{8} \geq \mathbb{Y}_n(\bar{\tau}).$$

The inequality in the last display can be rewritten as

$$\frac{Y_0(\tau^-) + Y_0(\tau^+)}{2} - \frac{(G_0(\tau^+) - G_0(\tau^-))(\tau^+ - \tau^-)}{8} \geq \mathbb{E}_n$$

where G_0 and Y_0 are the true counterparts of $\mathbb{G}_n$ and $\mathbb{Y}_n$ respectively, and $\mathbb{E}_n$ a random error. Using empirical process theory, Groeneboom, Jongbloed, and Wellner (2001b) showed that

$$|\mathbb{E}_n| = O_p(n^{-4/5}) + o_p((\tau^+ - \tau^-)^4). \tag{2.2}$$

On the other hand, Groeneboom, Jongbloed, and Wellner (2001b) established that there exists a universal constant $C > 0$ such that

$$\begin{aligned} &\frac{Y_0(\tau^-) + Y_0(\tau^+)}{2} - \frac{(G_0(\tau^+) - G_0(\tau^-))(\tau^+ - \tau^-)}{8} \\ &= -C g_0''(x_0)(\tau^+ - \tau^-)^4 + \ o_p((\tau^+ - \tau^-)^4). \end{aligned} \tag{2.3}$$

Combining the results in (2.2) and (2.3), it follows that

$$\tau^+ - \tau^- = O_p(n^{-1/5}).$$

The problem has two main features that make the above arguments work. First of all, the polynomial $\tilde{H}_n$ can be fully determined on $[\tau^-, \tau^+]$ and therefore it can be evaluated at any point between τ^- and τ^+. Second of all, it can expressed via the empirical process $\mathbb{Y}_n$ and that enables us to "get rid of" terms depending on $\tilde{g}_n$ whose rate of convergence is still unknown at this stage. We should also add that the problem is symmetric around $\bar{\tau}$, a property that helps establishing the formula derived in (2.3).

When $k > 2$, we have established in Balabdaoui and Wellner (2004a), Proposition 2.2, that $\tilde{g}_n$ is the LSE if and only if

$$\tilde{H}_n(x) \begin{cases} \geq \mathbb{Y}_n(x), & x \geq 0 \\ = \mathbb{Y}_n(x), & \text{if and only if } x \text{ is a jump point of } \tilde{g}_n^{(k-1)} \end{cases} \tag{2.4}$$

where

$$\tilde{H}_n(x) = \int_0^x \frac{(x-t)^{k-1}}{(k-1)!} \tilde{g}_n(t) dt$$

and

$$\mathbb{Y}_n(x) = \int_0^x \frac{(x-t)^{k-1}}{(k-1)!} d\mathbb{G}_n(t).$$

If τ^-, τ^+ are two successive jump points of $\tilde{g}_n^{(k-1)}$, then the four equalities

$$\tilde{H}_n(\tau^\pm) = \mathbb{Y}_n(\tau^\pm), \quad \text{and} \quad \tilde{H}'_n(\tau^\pm) = \mathbb{Y}'_n(\tau^\pm)$$

still hold. However, these equations are not enough to determine the polynomial $\tilde{H}_n$, now of degree $2k-1$, on the interval $[\tau^-, \tau^+]$ when $k > 2$. One would need $2k$ conditions to be able to achieve this. [We would be in this situation if we had equality of the higher derivatives of $\tilde{H}_n$ and $\mathbb{Y}_n$ at τ^- and τ^+, that is

$$\tilde{H}_n^{(j)}(\tau^-) = \mathbb{Y}_n^{(j)}(\tau^-), \qquad \tilde{H}_n^{(j)}(\tau^+) = \mathbb{Y}_n^{(j)}(\tau^+) \tag{2.5}$$

for $j = 0, \cdots, k-1$, but the characterization (2.4) does not give this much.]

Thus it becomes clear that two jump points are not sufficient to determine the piecewise polynomial $\tilde{H}_n$. However, if we consider $p > 2$ jump points $\tau_0 < \cdots < \tau_{p-1}$ (all located e.g. after x_0), $\tilde{H}_n$ is a spline of degree $2k-1$ that is $(2k-2)$-times differentiable at its knot points $\tau_0, \cdots, \tau_{p-1}$. In the next subsection, we prove that if $p = 2k-2$, the spline $\tilde{H}_n$ is completely determined on $[\tau_0, \tau_{2k-3}]$ by the conditions

$$\tilde{H}_n(\tau_i) = \mathbb{Y}(\tau_i), \ \text{ and } \ \tilde{H}'_n(\tau_i) = \mathbb{Y}'(\tau_i), \qquad i = 0, \ldots, 2k-3\,. \tag{2.6}$$

This result proves to be very useful for determining the stochastic order of the distance between two successive jump points in a small neighborhood of x_0.

2.2 A Hermite interpolation problem

In the next lemma, we prove that given $2k-2$ successive jump points $\tau_0 < \cdots < \tau_{2k-3}$, of $\tilde{g}_n^{(k-1)}$, $\tilde{H}_n$ is the unique solution of the Hermite problem given by (2.6). But before that, we need the following lemma which gives a definition of B-splines.

Lemma 2.1 *Let $m \geq 1$ be an integer and $x_1 < \cdots < x_{m+1}$ be arbitrary $(m+1)$ points in $\mathbb{R}$. There exists a unique vector $(a_1, \cdots, a_{m+1}) \in \mathbb{R}^{m+1}$ such that the spline*

$$B(t) = \sum_{i=1}^{m+1} a_i (t - x_i)_+^{m-1}, \qquad t \in \mathbb{R}$$

satisfies

$$B(t) = 0, \qquad \text{if } t \leq x_1 \ \text{ or } \ t \geq x_{m+1} \tag{2.7}$$

$$B_k(t) > 0, \qquad \text{if } t \in (x_1, x_{m+1}) \tag{2.8}$$

$$\int_{x_1}^{x_{m+1}} B(t)dt = 1. \tag{2.9}$$

B is called the B-spline of degree $m-1$ with support $[x_1, x_{m+1}]$. Furthermore,

$$B(t) = [x_1, \cdots, x_{m+1}](-1)^m m(t - \cdot)_+^{m-1}, \qquad t \in \mathbb{R}; \tag{2.10}$$

thus $B(t)$ is the divided difference of order m of the function $x \mapsto (-1)^m m(t-x)_+^{m-1}$, $x \in \mathbb{R}$ with respect to the knots $x_1, \ldots, x_{m+1}$.

Proof. See e.g. NÜRNBERGER (1989), Theorems 2.2 and 2.9, pages 96 and 99. ■

Remark 2.1 *Note that for any a and b in $\mathbb{R}$, we have*

$$(b-a)^{m-1} = (b-a)_+^{m-1} + (-1)^{m-1}(a-b)_+^{m-1}.$$

On the other hand, we can write

$$\begin{aligned}\sum_{i=1}^{m+1} a_i(t-x_i)^{m-1} &= \sum_{i=1}^{m+1} a_i \sum_{l=0}^{m-1} \binom{m-1}{l} x_i^l \; t^{m-1-l} \\ &= \sum_{l=0}^{m-1} \binom{m-1}{l} \left(\sum_{i=1}^{m+1} a_i x_i^l \right) t^{m-1-l} = 0, \quad \text{for } t \in \mathbb{R},\end{aligned}$$

where the last equality follows from the identities in (2.4) of Theorem 2.2 in NÜRNBERGER (1989)*. Therefore, B can also be given by*

$$B(t) = (-1)^m \sum_{i=1}^{m+1} a_i(x_i - t)_+^{m-1} \qquad t \in \mathbb{R},$$

or equivalently

$$B(t) = [x_1, \cdots, x_{m+1}]m(\cdot - t)_+^{m-1}. \tag{2.11}$$

The latter form will be used in the rest of this chapter.

Lemma 2.2 *Let $k \geq 2$. The function $\tilde{H}_n$ characterized by (2.4) is a spline of degree $2k-1$. Moreover, given any $2k-2$ successive jump points of $\tilde{H}_n^{(2k-1)}$, $\tau_0 < \cdots < \tau_{2k-3}$, the $(2k-1)$-th spline $\tilde{H}_n$ is uniquely determined on $[\tau_0, \tau_{2k-3}]$ by the values of the process $\mathbb{Y}_n$ and of its derivative $\mathbb{Y}_n'$ at $\tau_0, \cdots, \tau_{2k-3}$. Furthermore, for any arbitrary points $\tau_{-(2k-1)} < \cdots < \tau_{-1}$ to the left of τ_0 and $\tau_{2k-2} < \cdots < \tau_{4k-4}$ to the right of τ_{2k-3}, there exist coefficients $\alpha_{-(2k-1)}, \cdots, \alpha_{2k-4}$ depending on $\mathbb{Y}_n(\tau_i)$ and $\mathbb{Y}_n'(\tau_i)$, $i = 0, \cdots, 2k-3$, such that the spline $\tilde{H}_n$ can be written as*

$$\tilde{H}_n(t) = \sum_{i=-(2k-1)}^{2k-4} \alpha_i B_i(t), \tag{2.12}$$

for all $t \in [\tau_0, \tau_{2k-3}]$ where, for $i = -(2k-1), \cdots, 2k-4$, B_i is the B-spline of degree $2k-1$ corresponding to the set of knots $\{\tau_i, \cdots, \tau_{i+2k}\}$.

Proof. We know that for any jump point τ of $\tilde{H}_n^{(2k-1)}$, we have

$$\tilde{H}_n(\tau) = \mathbb{Y}_n(\tau) \quad \text{and} \quad \tilde{H}_n'(\tau) = \mathbb{Y}_n'(\tau).$$

This can viewed as a *Hermite interpolation problem* if we consider that the *interpolated function* is the process $\mathbb{Y}_n$ and that the *interpolating spline* is $\tilde{H}_n$ (see e.g. NÜRNBERGER (1989), Definition 3.6, pages 108 and 109).

Now let $p = 2k-2$ and consider successive $2k-2$ jump points $\tau_0 < \cdots < \tau_{2k-3}$. We denote $\tau_0 = x_0 = a$, $\tau_{2k-3} = x_{2k-3} = b$ and $\tau_1 = x_1, \cdots, \tau_{2k-4} = x_{2k-4}$. Also, for $i = 1, \cdots, 4k-4$, consider the points t_i such that $t_1 = t_2 = x_0$, $t_3 = t_4 = x_1, \ldots,$ $t_{4k-5} = t_{4k-4} = x_{2k-3}$. Using this notation, we see that the $(2k-1)-th$ spline $\tilde{H}_n$ satisfies

$$\tilde{H}_n(t_i) = \mathbb{Y}_n(t_i) \quad \text{and} \quad \tilde{H}_n'(t_i) = \mathbb{Y}_n'(t_i) \tag{2.13}$$

for all $i = 1, \cdots, 4k-4$. Furthermore, we can check that for all $i = 1, \cdots, 2k-4$, we have

$$t_i < x_i < t_{i+2k}.$$

Indeed, for a given $i = 1, \cdots, 2k-4$, we know that $x_i = t_{2i+1} = t_{2i+2}$ and it is easy to see that

$$t_i < t_{2i+1} = t_{2i+2} < t_{i+2k}.$$

Therefore, by Theorem 3.7 in NÜRNBERGER (1989), page 109, the *Hermite interpolation problem* defined in (2.13) has a unique solution in $S_{2k-1}(x_1, \cdots, x_{2k-4})$, the space of splines of degree $2k-1$ that are $(2k-2)$-times continuously differentiable at the knots $x_1, \cdots, x_{2k-4}$ (or, see DEVORE AND LORENTZ (1993), Theorem 9.2, page 162). Notice that in Nürnberger's notation (see NÜRNBERGER (1989)), the parameters $p-2$ and $2k-1$ play the role of k and m respectively. Also, note that the integer $p = 2k-2$ was chosen here so that the number of equations $(2p)$ and the dimension of the space $S_{2k-1}(x_1, \cdots, x_p)$ $(\dim(S_{2k-1}(x_1, \cdots, x_p)) = p-2+2k)$ are equal. It follows that we can find $\alpha_{-(2k-1)}, \cdots, \alpha_{2k-4}$ such that

$$\tilde{H}_n(t) = \sum_{i=-(2k-1)}^{2k-4} \alpha_i B_i(t)$$

for all $t \in [a,b] \equiv [\tau_0, \tau_{2k-3}]$, where $\underline{\alpha}^t = (\alpha_{-(2k-1)}, \cdots, \alpha_{2k-4})^t$ is the unique solution of the linear system

$$M\alpha \equiv \begin{pmatrix} B_{-(2k-1)}(\tau_0) & \cdots & B_{2k-4}(\tau_0) \\ (B_{-(2k-1)})'(\tau_0) & \cdots & (B_{2k-4})'(\tau_0) \\ \vdots & \vdots & \vdots \\ B_{-(2k-1)}(\tau_{2k-3}) & \cdots & B_{2k-4}(\tau_{2k-3}) \\ (B_{-(2k-1)})'(\tau_{2k-3}) & \cdots & (B_{2k-4})'(\tau_{2k-3}) \end{pmatrix} \underline{\alpha} = \begin{pmatrix} \mathbb{Y}_n(\tau_0) \\ \mathbb{Y}_n'(\tau_0) \\ \vdots \\ \mathbb{Y}_n(\tau_{2k-3}) \\ \mathbb{Y}_n'(\tau_{2k-3}) \end{pmatrix} \tag{2.14}$$

and $B_i, i = -(2k-1), \cdots, 2k-4$, are $(4k-4)$ linearly independent B-splines of degree $2k-1$ and knots $\tau_i, \cdots < \tau_{i+2k}$. ∎

In the following lemma, we prove a preparatory result that will be used later for deriving the stochastic order of the distance between the jump points.

Lemma 2.3 *Let $\bar{\tau} \in \cup_{i=0}^{2k-4}(\tau_i, \tau_{i+1})$. If $e_k(t)$ denotes the error at t of the Hermite interpolation of the function $y^{2k}/(2k)!$ at the points $\tau_0, \cdots, \tau_{2k-3}$, then*

$$-g_0^{(k)}(\bar{\tau})e_k(\bar{\tau}) \leq \mathbb{E}_n + \mathbb{R}_n$$

where $\mathbb{E}_n$ defined in (2.16) is a random error and $\mathbb{R}_n$ defined in (2.18) is a remainder that both depend on the knots $\tau_0, \cdots, \tau_{2k-3}$ and the point $\bar{\tau}$.

Proof. In this proof, we use the explicit B-spline representation of $\tilde{H}_n$ that was introduced in the previous lemma. Let $A = (a_{ij})_{ij}$ and $B = (b_{ij})_{ij}$ be the $(4k-4) \times (k-1)$ sub-matrices obtained by extracting the odd and even columns of the inverse of the matrix M given in (2.14). We can write,

$$\tilde{H}_n(t) = \sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} (a_{ij}\mathbb{Y}_n(\tau_j) + b_{ij}\mathbb{Y}_n'(\tau_j)) \right) B_i(t)$$

for all $t \in [\tau_0, \tau_{2k-3}]$. Fix $t = \bar{\tau} \in \cup_{i=0}^{2k-4}(\tau_i, \tau_{i+1})$. From the inequality condition in the characterization of the LSE, it follows that

$$\sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} (a_{ij}\mathbb{Y}_n(\tau_j) + b_{ij}\mathbb{Y}_n'(\tau_j)) \right) B_i(\bar{\tau}) \geq \mathbb{Y}_n(\bar{\tau})$$

or equivalently

$$\sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} (a_{ij}Y_0(\tau_j) + b_{ij}Y_0'(\tau_j)) \right) B_i(\bar{\tau}) - Y_0(\bar{\tau}) \geq -\mathbb{E}_n \tag{2.15}$$

where Y_0 is the k-fold integral of the true density g_0 and $\mathbb{E}_n$ is given by

$$\mathbb{E}_n = \sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} (a_{ij}(\mathbb{Y}_n - Y_0)(\tau_j) + b_{ij}(\mathbb{Y}_n' - Y_0')(\tau_j)) \right) B_i(\bar{\tau}) \\ + Y_0(\bar{\tau}) - \mathbb{Y}_n(\bar{\tau}). \tag{2.16}$$

Based on the working assumptions, the function Y_0 is $(2k)$-times continuously differentiable in a small neighborhood of x_0. Using Taylor expansion of $Y_0(\tau_j)$ and

$Y_0'(\tau_j)$ around $\bar\tau$ up to the orders $2k$ and $2k-1$ respectively, the inequality in (2.15) can be rewritten as

$$\begin{aligned}
&\left(\sum_{i=-(2k-1)}^{2k-4}\left\{\sum_{j=0}^{2k-3} a_{ij}\right\} B_i(\bar\tau)-1\right) Y_0(\bar\tau)\\
&+\left(\sum_{i=-(2k-1)}^{2k-4}\left\{\sum_{j=0}^{2k-3} a_{ij}(\tau_j-\bar\tau)+b_{ij}\right\} B_i(\bar\tau)\right) Y_0'(\bar\tau)\\
&\vdots\\
&+\left(\sum_{i=-(2k-1)}^{2k-4}\left\{\sum_{j=0}^{2k-3} a_{ij}\frac{(\tau_j-\bar\tau)^{2k}}{(2k)!}+b_{ij}\frac{(\tau_j-\bar\tau)^{2k-1}}{(2k-1)!}\right\} B_i(\bar\tau)\right) Y_0^{(2k)}(\bar\tau)\\
&+\mathbb{R}_n\\
&\geq -\mathbb{E}_n
\end{aligned} \tag{2.17}$$

where $\mathbb{R}_n$ is the remainder of the Taylor expansion and can be given in the integral form

$$\begin{aligned}
\mathbb{R}_n \;=\; &\sum_{i=-(2k-1)}^{2k-4}\left(\sum_{j=0}^{2k-3} a_{ij}\int_{\bar\tau}^{\tau_j}\frac{(\tau_j-t)^{2k-1}}{(2k)!}(g_0^{(k)}(t)-g_0^{(k)}(x_0))dt\right.\\
&\left.+\,b_{ij}\int_{\bar\tau}^{\tau_j}\frac{(\tau_j-t)^{2k-2}}{(2k-2)!}(g_0^{(k)}(t)-g_0^{(k)}(x_0))dt\right)B_i(\bar\tau).
\end{aligned} \tag{2.18}$$

The remainder $\mathbb{R}_n$ can be viewed as the error of Hermite interpolation at the point $\bar\tau$ where

$$x\mapsto\int_{\bar\tau}^{x}\frac{(x-t)^{2k-1}}{(2k-1)!}(g_0^{(k)}(t)-g_0^{(k)}(x_0))dt$$

is the function being interpolated. The order of $\mathbb{R}_n$ will be determined in the next subsection. Note that

$$\begin{aligned}
\sum_{i=-(2k-1)}^{2k-4}\left(\sum_{j=0}^{2k-3} a_{ij}\right) B_i(\bar\tau)-1 &= 0 \\
\sum_{i=-(2k-1)}^{2k-4}\left(\sum_{j=0}^{2k-3} a_{ij}(\tau_j-\bar\tau)+b_{ij}\right) B_i(\bar\tau) &= 0\\
&\vdots\\
\sum_{i=-(2k-1)}^{2k-4}\left(\sum_{j=0}^{2k-3} a_{ij}\frac{(\tau_j-\bar\tau)^{2k-1}}{(2k-1)!}+b_{ij}\frac{(\tau_j-\bar\tau)^{2k-2}}{(2k-2)!}\right) B_i(\bar\tau) &= 0.
\end{aligned} \tag{2.19}$$

Indeed, since the space of splines of degree $2k-1$ and with simple knots $\tau_0, \cdots, \tau_{2k-3}$ includes all the polynomials of degree $\leq 2k-1$, the solution of the Hermite problem when the interpolated function is a polynomial of degree $\leq 2k-1$ is the polynomial itself. Therefore, if we consider $P_0(t) = 1, P_1(t) = t-\bar{\tau}, \cdots, P_{2k-1}(t) = (t-\bar{\tau})^{2k-1}/(2k-1)!$, the previous terms are identically zero since they are exactly equal to $P_j(\bar{\tau}) = 0$, $j = 0, \cdots, 2k-1$.

Now

$$\sum_{i=-(2k-1)}^{2k-4} \sum_{j=0}^{2k-3} \left(a_{ij} \frac{(\tau_j - \bar{\tau})^{2k}}{(2k)!} + b_{ij} \frac{(\tau_j - \bar{\tau})^{2k-1}}{(2k-1)!} \right) B_i(\bar{\tau})$$

can be recognized as the Hermite interpolation error at the point $\bar{\tau}$ when $(y-\bar{\tau})^{2k}/(2k)!$ is the function being interpolated at the knots $\tau_0, \cdots, \tau_{2k-3}$. But this error is equal to $e_k(\bar{\tau})$. Indeed, using the binomial identity, we can write

$$\begin{aligned}
&\sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} a_{ij} \frac{(\tau_j - \bar{\tau})^{2k}}{(2k)!} + b_{ij} \frac{(\tau_j - \bar{\tau})^{2k-1}}{(2k-1)!} \right) B_i(\bar{\tau}) \\
&= \sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} a_{ij} \frac{(\tau_j)^{2k}}{(2k)!} + b_{ij} \frac{(\tau_j)^{2k-1}}{(2k-1)!} \right) B_i(\bar{\tau}) \\
&+ \sum_{r=1}^{2k-1} \left(\sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} a_{ij} \binom{2k}{r} \frac{(\tau_j)^{2k-r}}{(2k)!} + b_{ij} \binom{2k-1}{r} \frac{(\tau_j)^{2k-1-r}}{(2k-1)!} \right) B_i(\bar{\tau}) \right) (-1)^r \bar{\tau}^r \\
&+ \left(\sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} a_{ij} \right) B_i(\bar{\tau}) \right) \frac{\bar{\tau}^{2k}}{(2k)!}.
\end{aligned}$$

Using the identity

$$\binom{2k-1}{r} = \frac{2k-r}{2k} \binom{2k}{r}$$

for all $r \in \{0, \cdots, 2k\}$, it follows that

$$\begin{aligned}
&\sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} a_{ij} \binom{2k}{r} \frac{(\tau_j)^{2k-r}}{(2k)!} + b_{ij} \binom{2k-1}{r-1} \frac{(\tau_j)^{2k-1-r}}{(2k-1)!} \right) B_i(\bar{\tau}) \\
&= \binom{2k}{r} \sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} a_{ij} \frac{(\tau_j)^{2k-r}}{(2k)!} + b_{ij} (2k-r) \frac{(\tau_j)^{2k-1-r}}{(2k)!} \right) B_i(\bar{\tau}) \\
&= \binom{2k}{r} \frac{\bar{\tau}^{2k-r}}{(2k)!}
\end{aligned}$$

since for all $t \in [\tau_0, \tau_{2k-3}]$ and $1 \leq r \leq 2k-1$

$$\sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} a_{ij}(\tau_j)^{2k-r} + b_{ij}(2k-r)(\tau_j)^{2k-1-r} \right) B_i(t) = t^{2k-r}.$$

Therefore,

$$\begin{aligned}
&\sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} a_{ij} \frac{(\tau_j - \bar{\tau})^{2k}}{(2k)!} + b_{ij} \frac{(\tau_j - \bar{\tau})^{2k-1}}{(2k-1)!} \right) B_i(\bar{\tau}) \\
&= \sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} a_{ij} \frac{(\tau_j)^{2k}}{(2k)!} + b_{ij} \frac{(\tau_j)^{2k-1}}{(2k-1)!} \right) B_i(\bar{\tau}) \\
&\quad + \left(\sum_{r=1}^{2k} (-1)^r \binom{2k}{r} \right) \frac{\bar{\tau}^{2k}}{(2k)!} \\
&= \sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} a_{ij} \frac{(\tau_j)^{2k}}{(2k)!} + b_{ij} \frac{(\tau_j)^{2k-1}}{(2k-1)!} \right) B_i(\bar{\tau}) - \frac{\bar{\tau}^{2k}}{(2k)!} \\
&= e_k(\bar{\tau})
\end{aligned}$$

since $\sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} a_{ij} \right) B_i(\bar{\tau}) = 1$ and $\sum_{r=0}^{2k} (-1)^r \binom{2k}{r} = 0$. We conclude that the inequality in (2.17) can be rewritten as stated in the lemma. ■

2.3 The order of the gap

In this subsection, we show that the gap problem can be reduced to a conjecture concerning the structure of the error bound for a certain Hermite interpolation problem (with uniformity in the knots). We restrict here ourselves to the LSE. For the MLE, the proof follows the same steps except that the notation is more cumbersome.

The error $e_k(t)$ defined in the previous lemma can be recognized as a monospline of degree $2k$ with $2k-2$ simple knots $\tau_0, \cdots, \tau_{2k-3}$. For a definition of monosplines, see e.g. Michelli (1972), Bojanov, Hakopian and Sahakian (1993), Nürnberger (1989), page 194 or DeVore and Lorentz (1993), page 136. Our bound on the random error $\mathbb{E}_n$ will be based on the following conjecture:

Conjecture 2.1 *Let $a = x_0 < x_1 < \cdots < x_{2k-3} = b$ be $2k-2$ arbitrary points and $1 \leq r \leq 2k-1$. Suppose that f that is a function that is r-times differentiable on $[a, b]$ except for a finite number of points. If Hf denotes the unique interpolating spline of degree $2k-1$ that solves the Hermite problem:*

$$Hf(x_j) = f(x_j), \text{ and } (Hf)'(x_j) = f'(x_j)$$

for $j = 0, \cdots, 2k-3$*, then there exists a constant* $C > 0$ *(depending only on* k*) such that*

$$\sup_{t \in [a,b]} |Hf(t) - f(t)| \leq C\omega(f^{(r)}; b-a)\,(b-a)^r$$

where $\omega(f^{(r)}; \cdot)$ *is the modulus of continuity of* $f^{(r)}$ *on* $[a,b]$*:*

$$\omega(h; \delta) = \sup\{|h(t_2) - h(t_1)| : \ t_1, t_2 \in [a,b], \ |t_2 - t_1| \leq \delta\}.$$

The above conjecture remains to be proved. In the case of quasi-interpolation, a similar result is available, and was proved by de Boor and Fix (1973); see e.g. Nürnberger (1989), page 189. However, we believe that such a result should also be true for our Hermite interpolation problem. Although the literature seems to be more concerned with the approximation error of other types of interpolating splines, we believe that there is no reason that our spline fails to satisfy a similar property especially that it tries to "recover" better the original function f by interpolating its tangent at the knots as well. Also, it should be mentioned that it is known that, given an interval $[a,b]$, the minimal deviation of a function f from the space of splines $S_m(x_1, \cdots, x_p)$ satisfies

$$d_\infty(f, S_m(x_1, \cdots, x_p)) \leq K\delta^r \omega(f^{(r)}; \delta)$$

if $f^{(r)} \in C[a,b]$ for some $r \in \{0, \cdots, m\}$, where $K > 0$ is a universal constant that depends only on r and $\delta = \max_{0 \leq i \leq p} |x_{i+1} - x_i|$ with $x_0 = a$ and $x_{p+1} = b$ (see e.g. Nürnberger (1989), Theorem 4.27, page 159).

Now we will derive an upper bound for the random error $\mathbb{E}_n$ based on Conjecture 2.1.

Lemma 2.4 *If Conjecture 2.1 holds, then the random error* $\mathbb{E}_n$ *satisfies*

$$|\mathbb{E}_n| = O_p(n^{-k/(2k+1)}) + o_p((\tau_{2k-3} - \tau_0)^{2k}).$$

Proof. Let f be the function given by

$$f(t) = \sum_{i=-(2k-1)}^{2k-3} \left(\sum_{j=0}^{2k-4} (a_{ij} \frac{(\tau_j - t)^{k-1}}{(k-1)!} + b_{ij} \frac{(\tau_j - t)^{k-2}}{(k-2)!}) 1_{[\tau_j, \bar{\tau}]}(t) \right) B_i(\bar{\tau}),$$

where $[\tau_j, \bar{\tau}] \equiv [\bar{\tau}, \tau_j]$ if $\tau_j > \bar{\tau}$. Then, the error $\mathbb{E}_n$ can be rewritten as

$$\mathbb{E}_n = \int_0^\infty f(t) d(\mathbb{G}_n(t) - G_0(t)). \tag{2.20}$$

Indeed, we found in the previous subsection that $\mathbb{E}_n$ is given by

$$\mathbb{E}_n = \sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} (a_{ij}(\mathbb{Y}_n - Y_0)(\tau_j) + b_{ij}(\mathbb{Y}'_n - Y'_0)(\tau_j)) \right) B_i(\bar{\tau}) + Y_0(\bar{\tau}) - \mathbb{Y}_n(\bar{\tau}).$$

Let us denote $\mathbb{D}_n = \mathbb{Y}_n - Y_0$. The error $\mathbb{E}_n$ can be rewritten as

$$\mathbb{E}_n = \sum_{i=-(2k-1)}^{2k-4} (\sum_{j=0}^{2k-3} (a_{ij}\mathbb{D}_n(\tau_j) + b_{ij}\mathbb{D}'_n(\tau_j))B_i(\bar{\tau}) - \mathbb{D}_n(\bar{\tau}).$$

Now for arbitrary x and y, we can write

$$\mathbb{D}_n(y) = \mathbb{D}_n(x) + (y-x)\mathbb{D}'_n(x) + \cdots + \int_x^y \frac{(y-t)^{k-1}}{(k-1)!} d(\mathbb{G}_n(t) - G_0(t))$$

and similarly

$$\mathbb{D}'_n(y) = \mathbb{D}'_n(x) + (y-x)\mathbb{D}''_n(x) + \cdots + \int_x^y \frac{(y-t)^{k-2}}{(k-2)!} d(\mathbb{G}_n(t) - G_0(t)).$$

Taking $x = \bar{\tau}$ and $y = \tau_j$ for $j = 0, \cdots, 2k-3$ and using the identities in (2.19) up to the order $(k-2)$, it follows that

$$\begin{aligned}
\mathbb{E}_n &= \sum_{i=-(2k-1)}^{2k-4} \left(\sum_{j=0}^{2k-3} \int_{\bar{\tau}}^{\tau_j} (a_{ij} \frac{(\tau_j - t)^{k-1}}{(k-1)!} + b_{ij} \frac{(\tau_j - t)^{k-2}}{(k-2)!}) d(\mathbb{G}_n(t) - G_0(t)) \right) B_i(\bar{\tau}) \\
&= \sum_{i=-(2k-1)}^{2k-4} \Big(\sum_{j=0}^{2k-3} \int_0^\infty (a_{ij} \frac{(\tau_j - t)^{k-1}}{(k-1)!} + b_{ij} \frac{(\tau_j - t)^{k-2}}{(k-2)!}) 1_{[\bar{\tau},\tau_j]}(t) d(\mathbb{G}_n(t) - G_0(t)) \Big) \\
&\qquad B_i(\bar{\tau}) \\
&= \int_0^\infty \Bigg[\sum_{i=-(2k-1)}^{2k-4} \Big(\sum_{j=0}^{2k-3} (a_{ij} \frac{(\tau_j - t)^{k-1}}{(k-1)!} + b_{ij} \frac{(\tau_j - t)^{k-2}}{(k-2)!}) 1_{[\bar{\tau},\tau_j]}(t) \Big) \\
&\qquad B_i(\bar{\tau}) \Bigg] d(\mathbb{G}_n(t) - G_0(t))
\end{aligned}$$

which is the form claimed in (2.20).

Even if the function f is formally integrated on $(0, \infty)$, it is clear that we can assume that f is compactly supported on $[\tau_0, \tau_{2k-3}]$. For a fixed $t \in [\tau_0, \tau_{2k-3}]$, there are two possibilities: $t < \bar{\tau}$ or $t \geq \bar{\tau}$. Suppose without loss of generality that $t \geq \bar{\tau}$. Then, $f(t)$ which can be also given by

$$f(t) = \sum_{i=-(2k-1)}^{2k-4} \left\{ \sum_{j=0}^{2k-3} \left(a_{ij} \frac{(\tau_j - t)^{k-1}}{(k-1)!} + b_{ij} \frac{(\tau_j - t)^{k-2}}{(k-2)!} \right) \right\} 1_{[\tau_j \geq t]} B_i(\bar{\tau})$$

$$= \sum_{i=-(2k-1)}^{2k-4} \left\{ \sum_{j=0}^{2k-3} \left(a_{ij} g_t(\tau_j) + b_{ij} g'_t(\tau_j) \right) \right\} B_i(\bar{\tau})$$

with

$$g_t(x) = \frac{(x-t)^{k-1}}{(k-1)!} 1_{[x \geq t]},$$

is nothing but the error at the point $\bar{\tau}$ of the Hermite interpolation of g_t at the points $\tau_0, \cdots, \tau_{2k-3}$. Note that g_t is a spline of degree $k-1$ that is $(k-1)$-times differentiable except at its unique knot t. If Conjecture 2.1 holds, there exists $C > 0$, such that

$$|f(t)| \leq C\omega(g_t^{(k-1)}, \tau_{2k-3} - \tau_0)(\tau_{2k-3} - \tau_0)^{k-1}.$$

But

$$\omega(g_t^{(k-1)}, \tau_{2k-3} - \tau_0) \leq 1.$$

Therefore, it follows that

$$\sup_{t \in [\tau_0, \tau_{2k-3}]} |f(t)| \leq C(\tau_{2k-3} - \tau_0)^{k-1}. \tag{2.21}$$

Since the function $f(t)$ depends on the knots $\tau_0, \cdots, \tau_{2k-3}$ and the point $\bar{\tau}$ (which is a fixed point in $\cup_{j=0}^{2k-4}(\tau_j, \tau_{j+1})$, it can be viewed as an element of the class

$$\mathcal{F}_{x,\underline{r}} = \left\{ f_{x,y_1,\cdots,y_{2k-2}} : x \leq y_1 \leq x + r_1, \cdots, y_{2k-3} \leq y_{2k-2} \leq y_{2k-3} + r_{2k-2} \right\}$$

where $x > 0$ and $\underline{r} = (r_1, \cdots, r_{2k-2}) : r_j > 0, \ j = 1, \cdots, 2k-2$ is a fixed $(2k-2)$-vector. To make the link between the members of the class $\mathcal{F}_{x,\underline{r}}$ and the function $f(t)$, the latter can be written as

$$f(t) = f_{\tau_0, \tau_1, \cdots, \bar{\tau}, \cdots, \tau_{2k-3}}(t), \quad t \in [\tau_0, \tau_{2k-3}].$$

In this case, $x = \tau_0$, $y_1 = \tau_1$, $y_{2k-2} = \tau_{2k-3}$ and $\{y_1, \cdots, y_{2k-2}\} = \{\tau_1, \cdots, \tau_{2k-3}\} \cup \{\bar{\tau}\}$.

Let Q be an arbitrary measure on $(0, \infty)$. The collection $\mathcal{F}_{x,\underline{r}}$ admits a finite covering number with respect to $L_2(Q)$. In fact, any element $f_{x,y_1,\cdots,y_{2k-2}} \in \mathcal{F}_{x,\underline{r}}$ is $(k-2)$-times differentiable on $[x, y_{2k-2}]$. Therefore, for every $\epsilon > 0$, the collection $\mathcal{F}_{x,\underline{r}}$ admits a finite bracketing number that is bounded by $(K/\epsilon)^{1/(k-2)}$, for some $0 < K < \infty$. More specifically, there exists a constant $K > 0$ depending only on k and $R = r_1 + \cdots + r_{2k-2}$ (an upper bound for the length of the interval $[x, y_{2k-2}]$) such that

$$\log N_{[]}(\epsilon, \mathcal{F}_{x,\underline{r}}, L_2(Q)) \leq K \left(\frac{1}{\epsilon} \right)^{\frac{1}{k-2}} \tag{2.22}$$

(see e.g. van der Vaart and Wellner (1996), Corollary 2.7.2, page 157). It follows that

$$\int_0^1 \sqrt{1+\log N_{[]}(\epsilon, \mathcal{F}_{x,\underline{r}}, L_2(G_0))} d\epsilon < \infty.$$

On the other hand, using Conjecture 2.1, we have

$$|f_{x,y_1,\cdots,y_{2k-2}}(t)| \leq C(y_{2k-2}-x)^{k-1} 1_{[x,y_{2k-2}]}(t)$$

(compare with the bound in (2.21)) and hence the function $F_{x,R}$ given by

$$F_{x,R}(t) = CR^{k-1} 1_{[x,x+R]}(t).$$

is an envelope for the class $\mathcal{F}_{x,\underline{r}}$. On the other hand, if x belongs to a small neighborhood $[x_0-\delta, x_0+\delta]$ for some small $\delta > 0$, then we can find some constant $M > 0$ depending only on δ, R and $g_0(x_0)$ such that $0 < \sup_{t \in [x_0-\delta, x_0+\delta+R]} g_0(t) < M$. Therefore,

$$EF^2_{x,R}(X_1) = C^2 R^{2(k-1)} \int_x^{x+R} g_0(x) dx \leq C^2 M R^{2k-1}.$$

By van der Vaart and Wellner (1996), Theorem 2.14.2, page 240, it follows that

$$E\left\{\left(\sup_{f_{x,y_1,\cdots,y_{2k-2}} \in \mathcal{F}_{x,\underline{r}}} \left|(\mathbb{G}_n - G_0)(f_{x,y_1,\cdots,y_{2k-2}})\right|\right)^2\right\} \leq \frac{K'}{n} EF^2_{x,R}(X_1) = O(n^{-1}R^{2k-1}) \tag{2.23}$$

for some constant $K' > 0$ depending only on x_0, δ and R. j We denote

$$(\mathbb{P}_n - P_0)(f_{x,y_1,\cdots,y_{2k-2}}) = (\mathbb{G}_n - G_0)(f_{x,y_1,\cdots,y_{2k-2}})$$

where $f_{x,y_1,\cdots,y_{2k-2}}$ is an element in $\mathcal{F}_{x,\underline{r}}$ and define M_n as

$$\begin{aligned} M_n &= \inf\Big\{D > 0 : \left|(\mathbb{P}_n - P_0)(f_{x,y_1,\cdots,y_{2k-3},y})\right| \leq \epsilon(y-x)^{2k} \\ &\qquad + n^{-2k/(2k+1)} D, \ \text{ for all } \ y \in [x, x+R]\Big\}. \end{aligned}$$

and $M_n = \infty$ if no $D > 0$ satisfies the required inequality. For $1 \leq j \leq \lfloor Rn^{1/(2k+1)} \rfloor = j_n$, we have

$$\begin{aligned} &P(M_n > m) \\ &\leq \sum_{1 \leq j \leq j_n} P\Big\{(j-1)n^{-1/(2k+1)} \leq y - x \leq jn^{-1/(2k+1)}, \end{aligned}$$

$$
\begin{aligned}
& \qquad\qquad \left|(\mathbb{P}_n - P_0)(f_{x,y_1,\cdots,y_{2k-3},y})\right| > \epsilon(y-x)^{2k} + n^{-2k/(2k+1)}m \Big\} \\
= & \sum_{1\le j\le j_n} P\Big\{(j-1)n^{-1/(2k+1)} \le y-x \le jn^{-1/(2k+1)}, \\
& \qquad\qquad \left|(\mathbb{P}_n - P_0)(f_{x,y_1,\cdots,y_{2k-3},y})\right| > \epsilon(y-x)^{2k} + n^{-2k/(2k+1)}m \Big\} \\
\le & \sum_{1\le j\le j_n} P\Big\{(j-1)n^{-1/(2k+1)} \le y-x \le jn^{-1/(2k+1)}, \\
& \qquad\qquad n^{2k/(2k+1)}\left|(\mathbb{P}_n - P_0)(f_{x,y_1,\cdots,y_{2k-3},y})\right| > \epsilon(j-1)^{2k} + m \Big\} \\
\le & \sum_{1\le j\le j_n} n^{4k/(2k+1)} \frac{E\left\{\left(\sup_{y:0\le y-x<jn^{-1/(2k+1)}} \left|(\mathbb{P}_n - P_0)(f_{x,y_1,\cdots,y_{2k-3},y})\right|\right)^2\right\}}{\left(\epsilon(j-1)^{2k}+m\right)^2} \\
= & \sum_{1\le j\le j_n} n^{4k/(2k+1)} \frac{E\left\{\left(\sup_{f_{x,y_1,\cdots,y_{2k-3},y}\in\mathcal{F}_{x,jn^{-1/(2k+1)}}} \left|(\mathbb{P}_n - P_0)(f_{x,y_1,\cdots,y_{2k-3},y})\right|\right)^2\right\}}{\left(\epsilon(j-1)^{2k}+m\right)^2} \\
\le & \; C \sum_{1\le j\le j_n} n^{4k/(2k+1)} n^{-1} n^{-(2k-1)/(2k+1)} \frac{j^{2k-1}}{\left(\epsilon(j-1)^{2k}+m\right)^2} \\
= & \; C \sum_{1\le j\le j_n} \frac{j^{2k-1}}{\left(\epsilon(j-1)^{2k}+m\right)^2} \\
\le & \; C \sum_{j=1}^{\infty} \frac{j^{2k-1}}{\left(\epsilon(j-1)^{2k}+m\right)^2} \searrow 0 \text{ as } m \nearrow \infty
\end{aligned}
$$

where $C > 0$ is a constant that is independent of $x \in [x_0 - \delta, x_0 + \delta]$. Therefore, $M_n = O_p(1)$ and hence it follows that

$$
\left|(\mathbb{P}_n - P_0)(f_{x,y_1,\cdots,y_{2k-3},y})\right| \le \epsilon(y-x)^{2k} + O_p(n^{-2k/(2k+1)})
$$

which holds for *all* $f_{x,y_1,\cdots,y_{2k-3},y} \in \mathcal{F}_{x,\underline{r}}$ and x in some small neighborhood $[x_0-\delta, x_0+\delta]$ of x_0. It follows that

$$
|\mathbb{E}_n| = o_p((\tau_{2k-3}-\tau_0)^{2k}) + O_p(n^{-2k/(2k+1)}).
$$

■

To show that $\tau_{2k-3} - \tau_0 = O_p(n^{-1/(2k+1)})$, we need the following result:

Lemma 2.5 *The error $e_k(t)$ has no other zeros than $\tau_0, \cdots, \tau_{2k-3}$ in $[\tau_0, \tau_{2k-3}]$.*

Proof. The result follows from Proposition 1 of MICHELLI (1972); see also DE BOOR (2004). ∎

Recall that $e_k(t)$ is a monospline of degree $2k$ with $2k-2$ simple knots $\tau_0, \cdots, \tau_{2k-3}$. Furthermore, by construction, these knots are also double zeros; i.e. $e_k(\tau_j) = e'_k(\tau_j) = 0$ for $j = 0, \cdots, 2k-3$. Now, we state two preparatory lemmas that will help determine the sign of the error $e_k(t)$ at any point $t \in \cup_{j=0}^{2k-4}(\tau_j, \tau_{j+1})$.

Lemma 2.6 *Let $k \geq 2$ be an integer. The monospline M_k of degree $2k$ with simple knots $\xi_0 = -k + 3/2, \xi_1 = -k + 5/2, \cdots, \xi_{2k-4} = k + 1/2, \xi_{2k-3} = k - 3/2$ and such that $M_k(\xi_j) = M'_k(\xi_j) = 0$ for $j = 0, \cdots, 2k-3$ has a constant sign: +1 (-1) if k is odd (even).*

Proof. Let $\mathcal{B}_{2k}$ be the Bernoulli monospline of degree $2k$. The function $\mathcal{B}_{2k}(t-1/2) - \mathcal{B}_{2k}(0)$ is equal to the error of the Hermite interpolation of $t^{2k}/(2k)!$ at the equispaced knots $\xi_0, \cdots, \xi_{2k-3}$. By uniqueness, it follows that

$$M_k(t) = \mathcal{B}_{2k}(t-1/2) - \mathcal{B}_{2k}(0)$$

for all $t \in [-k+3/2, k-3/2]$. The Bernoulli monospline $\mathcal{B}_{2k}$ is the 1-periodic extension of the Bernoulli polynomial p_{2k} of degree $2k$ which takes extreme values at 0 when considered as a function on $[0,1]$. It follows that M_k is of one sign on $[-k+3/2, k-3/2]$. Furthermore, $p_{2k}(1/2) < p_{2k}(0)$ if k is even and $p_{2k}(1/2) > p_{2k}(0)$. Therefore, M_k is nonpositive if k is even and nonnegative if k is odd. ∎

Lemma 2.7 *If $t \in \cup_{j=0}^{2k-4}(\tau_j, \tau_{j+1})$, then*

$$(-1)^{k-1} e_k(t) > 0;$$

i.e., $e_k(t)$ is nonnegative (nonpositive) if k is odd (even).

Proof. Let $\bar{\tau}$ be a fixed point in $\cup_{j=0}^{2k-4}(\tau_j, \tau_{j+1})$. We can assume without loss of generality that $\bar{\tau} \in (\tau_0, \tau_1)$. There exists $\lambda \in (0,1)$ such that $\bar{\tau} = \lambda\tau_0 + (1-\lambda)\tau_1$. Consider now the function

$$(\tau_0, \cdots, \tau_{2k-3}) \mapsto \frac{e_k(\bar{\tau}) + |e_k(\bar{\tau})|}{2e_k(\bar{\tau})}.$$

Note that it is possible to divide by $e_k(\bar{\tau})$ since $e_k(\bar{\tau}) \neq 0$ as $\bar{\tau}$ is different from the knots. It is easy to see that the function is continuous in $\tau_0, \cdots, \tau_{2k-3}$. Furthermore, it can only take two possible values, 0 or 1, and therefore has to be constant. But, when the knots are equally distant, we know from Lemma 2.6 that the constant is 1 (0) if k is odd (even). It follows that $(-1)^{k-1} e_k(\bar{\tau}) > 0$. ∎

We can now state the main result of this section:

Lemma 2.8 *Let $k \geq 2$. If $g_0 \in \mathcal{D}_k$ satisfies $g_0^{(k)}(x_0) \neq 0$ and Conjecture 2.1 holds, then $\tau_{2k-3} - \tau_0 = O_p(n^{-1/(2k+1)})$.*

Proof. Let $j_0 \in \{0, \cdots, 2k-4\}$ be such that $[\tau_{j_0}, \tau_{j_0+1}]$ be the largest knot interval; i.e., $\tau_{j_0+1} - \tau_{j_0} = \max_{0 \leq j \leq 2k-4}(\tau_{j+1} - \tau_j)$. Let $a = \tau_0$, $b = \tau_{2k-3}$.

By Conjecture 2.1, there exists a constant $C > 0$ depending only on k such that

$$|\mathbb{R}_n| \leq C \sup_{t \in [\tau_0, \tau_{2k-3}]} |g_0^{(k)}(t) - g_0^{(k)}(x_0)| \, (b-a)^{2k}$$

using the fact that if f is $\in C^{2k}[a,b]$, then

$$\omega(f^{(2k-1)}, b-a) \leq \sup_{t \in [a,b]} |f^{(2k)}(t)| \, (b-a).$$

Therefore, it follows that

$$|\mathbb{R}_n| \leq C \sup_{t \in [\tau_0, \tau_{2k-3}]} |g_0^{(k)}(t) - g_0^{(k)}(x_0)|(\tau_{2k-3} - \tau_0)^{2k} = o_p((\tau_{2k-3} - \tau_0)^{2k}).$$

Using the result of Lemma 2.3 and since the bounds on $\mathbb{R}_n$ and $\mathbb{E}_n$ (see Lemma 2.4) are independent of the choice of $\bar{\tau}$ in $\cup_{j=0}^{2k-4}(\tau_j, \tau_{j+1})$, it follows that

$$\sup_{\bar{\tau} \in (\tau_{j_0}, \tau_{j_0+1})} (-1)^{k-1} e_k(\bar{\tau}) \leq O_p(n^{-2k/(2k+1)}) + o_p((\tau_{2k-3} - \tau_0)^{2k}).$$

Now, on the interval $[\tau_{j_0}, \tau_{j_0+1}]$, the Hermite interpolation spline is a polynomial of degree $2k-1$. On the other hand, the best uniform approximation of the function t^{2k} on $[\tau_{j_0}, \tau_{j_0+1}]$ from the space of polynomials of degree $\leq 2k-1$ is given by the polynomial

$$t \mapsto t^{2k} - \left(\frac{\tau_{j_0+1} - \tau_{j_0}}{2}\right)^{2k} \frac{1}{2^{2k-1}} T_{2k}\left(\frac{2t - (\tau_{j_0} + \tau_{j_0+1})}{\tau_{j_0+1} - \tau_{j_0}}\right), \tag{2.24}$$

where T_{2k} is the Chebyshev polynomial of degree $2k$ (defined on $[-1,1]$), see, e.g., Nürnberger (1989), Theorem 3.23, page 46 or DeVore and Lorentz (1993), Theorem 6.1, page 75. It follows that

$$\begin{aligned} (-1)^{k-1} e_k(\bar{\tau}) &\geq \left\| \frac{T_{2k}}{2^{4k-1}(2k)!} \right\|_\infty (\tau_{j_0+1} - \tau_{j_0})^{2k} \\ &= \frac{1}{2^{4k-1}(2k)!} (\tau_{j_0+1} - \tau_{j_0})^{2k} \end{aligned} \tag{2.25}$$

since $\|T_{2k}\|_\infty = 1$. But,

$$\tau_{2k-3} - \tau_0 = \sum_{j=0}^{2k-4} (\tau_{j+1} - \tau_j) \leq (2k-3)(\tau_{j_0+1} - \tau_{j_0}).$$

It follows that

$$(-1)^{k-1} e_k(\bar{\tau}) \geq \frac{1}{(2k-3)^{2k} 2^{4k-1} (2k)!} (\tau_{2k-3} - \tau_0)^{2k}.$$

Combining the results obtained above, we conclude that

$$\frac{(-1)^k g_0^{(k)}(x_0)}{(2k-3)^{2k} 2^{4k-1} (2k)!} (\tau_{2k-3} - \tau_0)^{2k} \leq O_p(n^{-2k/(2k+1)}) + o_p((\tau_{2k-3} - \tau_0)^{2k})$$

which implies that $\tau_{2k-3} - \tau_0 = O_p(n^{-1/(2k+1)})$. ■

3 The asymptotic distribution

The prepare for a statement of the main result, we first recall the following theorem from Balabdaoui and Wellner (2004c) giving existence of the processes H_k.

Theorem 3.1 *For all $k \geq 1$, let Y_k denote the stochastic process defined by*

$$Y_k(t) = \begin{cases} \int_0^t \frac{(t-s)^{k-1}}{(k-1)!} dW(s) + \frac{(-1)^k k!}{(2k)!} t^{2k}, & t \geq 0 \\ \int_t^0 \frac{(t-s)^{k-1}}{(k-1)!} dW(s) + \frac{(-1)^k k!}{(2k)!} t^{2k}, & t < 0. \end{cases}$$

If Conjecture 2.1 holds (also see the discussion in Balabdaoui and Wellner (2004c)*), then there exists an almost surely uniquely defined stochastic process H_k characterized by the three following conditions:*

(i) The process H_k stays everywhere above the process Y_k:

$$H_k(t) \geq Y_k(t), \quad t \in \mathbb{R}.$$

(ii) $(-1)^k H_k$ is $2k$-convex; i.e. $(-1)^k H_k^{(2k-2)}$ exists and convex.

(iii) The process H_k satisfies

$$\int_{-\infty}^{\infty} (H_k(t) - Y_k(t))\, dH_k^{(2k-1)}(t) = 0.$$

(iv) If k is even, $\lim_{|t|\to\infty}(H_k^{(2j)}(t) - Y_k^{(2j)}(t)) = 0$ for $j = 0, \cdots, (k-2)/2$; if k is odd, $\lim_{t\to\infty}(H_k(t) - Y_k(t)) = 0$ and $\lim_{|t|\to\infty}(H_k^{(2j+1)}(t) - Y_k^{(2j+1)}(t)) = 0$ for $j = 0, \cdots, (k-3)/2$.

This is Theorem 2.1 in Balabdaoui and Wellner (2004c). Now we are able to state the main result of this paper:

Theorem 3.2 *Let $x_0 > 0$ and g_0 be a k-monotone density such that g_0 is k-times differentiable at x_0 with $(-1)^k g_0^{(k)}(x_0) > 0$ and assume that $g_0^{(k)}$ is continuous in a neighborhood of x_0. Let $\bar{g}_n$ denote either the LSE, $\tilde{g}_n$ or the MLE $\hat{g}_n$ and let $\bar{F}_n$ be the corresponding mixing measure. If Conjecture 2.1 holds, then*

$$\begin{pmatrix} n^{\frac{k}{2k+1}}(\bar{g}_n(x_0) - g_0(x_0)) \\ n^{\frac{k-1}{2k+1}}(\bar{g}_n^{(1)}(x_0) - g_0^{(1)}(x_0)) \\ \vdots \\ n^{\frac{1}{2k+1}}(\bar{g}_n^{(k-1)}(x_0) - g_0^{(k-1)}(x_0)) \end{pmatrix} \to_d \begin{pmatrix} c_0(g_0) H_k^{(k)}(0) \\ c_1(g_0) H_k^{(k+1)}(0) \\ \vdots \\ c_{k-1}(g_0) H_k^{(2k-1)}(0) \end{pmatrix}$$

and

$$n^{\frac{1}{2k+1}}(\bar{F}_n(x_0) - F(x_0)) \to_d \frac{(-1)^k x_0^k}{k!} c_{k-1}(g_0) H_k^{(2k-1)}(0)$$

where

$$c_j(g_0) = \left\{ (g_0(x_0))^{k-j} \left(\frac{(-1)^k g_0^{(k)}(x_0)}{k!} \right)^{2j+1} \right\}^{\frac{1}{2k+1}},$$

for $j = 0, \cdots, k-1$.

To prove Theorem 3.2 we first use the result of the previous section to derive rates of convergence of $\bar{g}_n^{(j)}$, $j = 0, \cdots, k-1$ at a fixed point $x_0 > 0$.

Consider the event $J_n = J_n^{(1)} \cap J_n^{(2)}$ where $J_n^{(i)}$, $i = 1, 2$, are defined by

$$\begin{aligned} J_n^{(1)} &\equiv J_n^{(1)}(x_0, k, M) \\ &= \{\text{there exist } (k+1) \text{ jump points } \tau_{n,1}, \cdots, \tau_{n,k+1} \\ &\qquad \text{(not necessarily successive) satisfying} \\ &\qquad x_0 - n^{-1/(2k+1)} \le \tau_{n,1} < \cdots < \tau_{n,k+1} \le x_0 + M n^{-1/(2k+1)} \\ &\qquad k n^{-1/(2k+1)} \le \tau_{n,k+1} - \tau_{n,1} \le M n^{-1/(2k+1)} \Big\}, \end{aligned}$$

and

$$J_n^{(2)} \equiv J_n^{(2)}(j, k, c_j) = \left\{ \inf_{t \in [\tau_{n,1}, \tau_{n,k+1}]} \left| \bar{g}_n^{(j)}(t) - g_0^{(j)}(t) \right| \le c_j n^{-(k-j)/(2k+1)} \right\}.$$

Proposition 3.1 *Suppose that $(-1)^k g_0^{(k)}(x_0) > 0$ and $g_0^{(k)}$ is continuous in a neighborhood of x_0. Let $\bar{g}_n$ be either the MLE $\hat{g}_n$ or the LSE $\tilde{g}_n$ and let $0 \le j \le k-1$. Suppose also that $\int_0^\infty y^{-1/2} dG_0(y) < \infty$ holds. Then, if Conjecture 2.1 holds, for any $\epsilon > 0$, there exists $M > 0$ and $c_j > 0$ such that $P(J_n) > 1 - \epsilon$ for all sufficiently large n.*

Proof. Fix $\epsilon > 0$. We will consider first the LSE and we will start with $j = 0$. Fix $\epsilon > 0$. For ease of notation, we will write the jump points of $\tilde{g}_n^{(k-1)}$ without the subscript n. Let τ_1 be the first jump point of $\tilde{g}_n^{(k-1)}$ after $x_0 - n^{-1/(2k+1)}$, τ_2 the first jump point after $\tau_1 + n^{-1/(2k+1)}$, ..., τ_{k+1} the first jump point after $\tau_k + n^{-1/(2k+1)}$. By Lemma 2.8, there exists $M > 0$ such that

$$0 \le \tau_{k+1} - \tau_1 \le M n^{-1/(2k+1)}$$

with probability $> 1-\epsilon$. Note that by construction $\tau_{k+1} - \tau_1 \ge k n^{-1/(2k+1)}$. Fix $c > 0$ and consider the event

$$\inf_{t \in [\tau_1, \tau_{k+1}]} |\tilde{g}_n(t) - g_0(t)| > c n^{-k/(2k+1)}. \tag{3.1}$$

On this set and for any nonnegative function g on $[\tau_1, \tau_{k+1}]$, we have

$$\left| \int_{\tau_1}^{\tau_{k+1}} (\tilde{g}_n(t) - g_0(t))\, g(t) dt \right| \ge c n^{-k/(2k+1)} \int_{\tau_n^-}^{\tau_n^+} g(t) dt. \tag{3.2}$$

Now, let B be the B-spline of degree $k-1$ and with support $[x_1, x_{k+1}]$. Recall from (2.11) in Section 5 that B can be given by

$$B(t) = [\tau_1, \cdots, \tau_{k+1}] k \ (\cdot - t)_+^{k-1}$$

where $[x_1, \cdots, x_m] g$ denotes the divided difference of degree m with respect to the points $x_1, \cdots, x_m$. After some algebra, we find that B can be given by

$$B(t) = (-1)^k k \left(\frac{(t-\tau_1)_+^{k-1}}{\prod_{j \ne 1} (\tau_j - \tau_1)} + \cdots + \frac{(t-\tau_k)_+^{k-1}}{\prod_{j \ne k} (\tau_j - \tau_k)} \right).$$

for all $t \in [\tau_1, \tau_{k+1}]$.

Let $|\eta| > 0$ and consider the perturbation function

$$p(t) = \prod_{1 \le i < j \le k+1} (\tau_j - \tau_i) \times B(t).$$

It is easy to check that for $|\eta|$ small enough, the perturbed function

$$\tilde{g}_{\eta,n}(t) = \tilde{g}_n(t) + \eta p(t)$$

is k-monotone on $(0, \infty)$. Indeed, p was chosen so that it satisfies $p^{(j)}(\tau_1) = p^{(j)}(\tau_{k+1}) = 0$ for $0 \le j \le k-2$, which guarantees that the perturbed function $\tilde{g}_{\eta,n}$ belongs to $C^{k-2}(0, \infty)$. For $0 \le j \le k-3$, the properties of strict convexity and monotonicity of $(-1)^j \tilde{g}_n^{(j)}$ on $(0, \infty)$ are preserved by $\tilde{g}_{\eta,n}^{(j)}$ as long as $|\eta|$ is small enough. For $k-2$, $(-1)^{k-2} \tilde{g}_n^{(k-2)}$ is piecewise linear and hence not strictly convex on $(0, \infty)$. Since p is a spline of degree $k-1$, the function $(-1)^{k-2}\tilde{g}_{\eta,n}^{(k-2)}$ is also piecewise linear and one can check that it is nonincreasing and convex for very small values of η. It follows that

$$\lim_{\eta\to 0}\frac{Q_n(\tilde{g}_{\eta,n})-Q_n(\tilde{g}_n)}{\eta}=0\,.$$

This implies that

$$\int_{\tau_1}^{\tau_{k+1}} p(t)d(\tilde{G}_n-\mathbb{G}_n)(t)=0.$$

The previous equality can be rewritten as

$$\int_{\tau_1}^{\tau_{k+1}} p(t)\left(\tilde{g}_n(t)-g_0(t)\right)dt=\int_{\tau_1}^{\tau_{k+1}} p(t)d(\mathbb{G}_n(t)-G_0(t)).$$

Taking $g\equiv p$ in (3.2), we obtain

$$\begin{aligned}
\left|\int_{\tau_1}^{\tau_{k+1}} p(t)d(\mathbb{G}_n(t)-G_0(t))\right| &\geq cn^{-k/(2k+1)}\int_{\tau_1}^{\tau_{k+1}} p(t)dt \\
&= cn^{-k/(2k+1)}\prod_{1\leq i<j\leq k+1}(\tau_j-\tau_i) \qquad (3.3)\\
&\geq cn^{-k/(2k+1)}\left(n^{-1/(2k+1)}\right)^{k(k+1)/2} \qquad (3.4)\\
&= cn^{-(3+k)k/(2(2k+1))}
\end{aligned}$$

where in (3.3), we used the fact that B-splines integrate to 1, whereas in (3.4) we used the facts that there are $k(k+1)/2$ terms in the product $\prod_{1\leq i<j\leq k+1}(\tau_j-\tau_i)$ and that $\tau_j-\tau_i\geq n^{-1/(2k+1)}, 1\leq i<j\leq k+1$.

Let $0<x<y_1<\cdots<y_{k-1}<y$ be $(k+1)$ points in $(0,\infty)$ and consider the function $f_{x,y_1,\cdots,y_{k-1},y}$ defined by

$$f_{x,y_1,\cdots,y_{k-1},y_k}(t)=(-1)^k k\prod_{0\leq i<j\leq k}(y_j-y_i)\left(\frac{(y_0-t)_+^{k-1}}{\prod_{j\neq 0}(y_j-y_0)}+\cdots+\frac{(y_{k-1}-t)_+^{k-1}}{\prod_{j\neq k-1}(y_j-y_{k-1})}\right)$$

where $y_0=x$. Let $\underline{r}=(r_1,\cdots,r_k)$, $r_i>0$ for $i=1,\cdots,k$, be a fixed k-vector and consider the collection of functions

$$\mathcal{F}_{x,\underline{r}}=\left\{f_{x,y_1,\cdots,y_{k-1},y_k}\;:\;x<y_1\leq x+r_1,\cdots,y_{k-1}<y_k\leq y_{k-1}+r_k\right\}.$$

For a fixed $x>0$ and $\underline{r}$, the collection $\mathcal{F}_{x,\underline{r}}$ has a finite covering number with respect to $L_2(Q)$ where Q is an arbitrary probability measure. In fact, denote

$$\alpha_j=(-1)^k k\frac{\prod_{0\leq l<l'\leq k}(y_{l'}-y_l)}{\prod_{j'\neq j}(y_{j'}-y_j)}$$

and consider the collections of functions

$$\mathcal{F}_{x,R_j} = \left\{ t \mapsto \alpha_j (y_j - t)_+^{k-1} 1_{[x,y_k]}(t), x \le y_j \le x + R_j, x \le y_k \le x + R \right\}$$

where $R_j = r_1 + \cdots + r_j$ for $j = 1, \cdots, k$ and $R = R_k$. By VAN DER VAART AND WELLNER (1996), Lemmas 2.6.16 and 2.6.18, pages 146 and 147, the collections $\mathcal{F}_{x,R_j}, j = 1, \cdots, k-1$ are VC-subgraph classes. Furthermore, the function

$$F_{x,R}(t) = kR^{k(k-1)/2}(x-t)_+^{k-1} 1_{[x,x+R]}(t)$$

is a common envelope for these classes. To see that, notice that for $j = 0, \cdots, k$, the product $\prod_{j' \neq j}(y_{j'} - y_j)$ contains k terms and hence α_j is a product of $k(k+1)/2 - k = k(k-1)/2$ that are at most R distant from one another. It follows that

$$\alpha_j \le kR^{k(k-1)/2}, \quad \text{for } j = 0, \cdots, k.$$

For an arbitrary probability measure Q, we have

$$\|F_{x,R}\|_{Q,2}^2 = k^2 R^{k(k-1)} \int_x^{x+R} (t-x)^{2k-2} dQ(t) \le k^2 R^{k(k+1)-2}$$

which is independent of Q. By Theorem 2.6.7 in VAN DER VAART AND WELLNER (1996), there exist a universal constant $K > 0$, two constants $D_j > 0$ and $V_j > 0$ that depend only on x, R_j and R such that the $\epsilon\|F_{x,R}\|_{Q,2}^2$-covering number of $\mathcal{F}_{x,R_j}$ with respect to $L_2(Q)$ is given by

$$N\left(\epsilon\|F_{x,R}\|_{Q,2}^2, \mathcal{F}_{x,R_j}, L_2(Q)\right) \le KD_j \left(\frac{1}{\epsilon}\right)^{V_j}.$$

It follows that the collection $\mathcal{F}_{x,\underline{r}}$ admits a finite ϵ-covering number with respect to $L_2(Q)$. Furthermore, it is easy to see that the function $k \times F_{x,R}$ is an envelope for this collection. Therefore, there exist a universal constant $K > 0$, $D > 0$ and $V > 0$ depending only on x and $\underline{R}_j$, $j = 1, \cdots, k$ such that

$$N\left(\epsilon\|F_{x,R}\|_{Q,2}^2, \mathcal{F}_{x,\underline{r}}, L_2(Q)\right) \le KD \left(\frac{1}{\epsilon}\right)^{V}$$

and therfore

$$\sup_Q \int_0^1 \sqrt{1 + \log(N\left(\epsilon\|F_{x,R}\|_{Q,2}^2, \mathcal{F}_{x,\underline{r}}, L_2(Q)\right)} d\epsilon < \infty.$$

On the other hand, if x is in a small neighborhood $[x_0 - \delta, x_0 + \delta]$ for some small $\delta > 0$, there exists some constant $C > 0$ depending only on δ, R and $g_0(x_0)$ such that $0 < g_0 < C$ on $[x, x+R]$ for all $x \in [x_0 - \delta, x_0 + \delta]$. It follows that

$$\begin{aligned} EF_{x,R}^2(X_1) &\le k^2 R^{k(k-1)} \int_x^{x+R} (t-x)^{2k-2} g_0(x) dx \\ &\le \frac{k^2 C}{2k-1} R^{k(k-1)} R^{2k-1} = \frac{k^2 C}{2k-1} R^{k(k+1)-1}. \end{aligned}$$

Therefore, by the Theorem 2.14.1 in van der Vaart and Wellner (1996), we have

$$E\left\{\left(\sup_{f_{x,y_1,\cdots,y_k}\in\mathcal{F}_{x,\underline{r}}}\left|(\mathbb{G}_n-G_0)(f_{x,y_1,\cdots,y_k})\right|\right)^2\right\} \leq \frac{K'}{n}EF_{x,R}^2(X_1)=O(n^{-1}R^{k(k+1)+1}), \qquad (3.5)$$

for some constant K' depending only on x_0, δ and R.

We denote

$$(\mathbb{P}_n-P_0)(f_{x,y_1,\cdots,y_{k-1},y}) = (\mathbb{G}_n-G_0)(f_{x,y_1,\cdots,y_{k-1},y})$$

where $f_{x,y_1,\cdots,y_{k-1},y}\in\mathcal{F}_{x,R}$ and define M_n as

$$M_n = \inf\left\{D>0:\left|(\mathbb{P}_n-P_0)(f_{x,y_1,\cdots,y_{k-1},y})\right|\leq\epsilon(y-x)^{(3k+1)k/2} + n^{-(3+k)k/(2(2k+1))}D, \quad \text{for all} \quad y\in[x,x+R]\right\};$$

note that M_n is possibly equal to infinity if no $D>0$ satisfies the required inequality. Let $n>N$. For $1\leq j\leq\lfloor Rn^{1/(2k+1)}\rfloor=j_n$, we have

$$\begin{aligned}
&P(M_n>m)\\
&\leq \sum_{1\leq j\leq j_n}P\Big\{\exists\, y:\ (j-1)n^{-1/(2k+1)}\leq y-x\leq jn^{-1/(2k+1)},\\
&\qquad \left|(\mathbb{P}_n-P_0)(f_{x,y_1,\cdots,y_{k-1},y})\right|>\epsilon(y-x)^{(3+k)k/2}+\ n^{-(3+k)k/(2(2k+1))}\ m\Big\}\\
&\leq \sum_{1\leq j\leq j_n}P\Big\{\exists\, y:\ 0\leq y-x\leq jn^{-1/(2k+1)},\\
&\qquad n^{(3+k)k/(2(2k+1))}\left|(\mathbb{P}_n-P_0)(f_{x,y_1,\cdots,y_{k-1},y})\right|>\epsilon(j-1)^{(3+k)k/2}+\ m\Big\}\\
&\leq \sum_{1\leq j\leq j_n}n^{(3+k)k/(2k+1)}\frac{E\left\{\left(\sup_{y:0\leq y-x<jn^{-1/(2k+1)}}\left|(\mathbb{P}_n-P_0)(f_{x,y_1,\cdots,y_{k-1},y})\right|\right)^2\right\}}{\left(\epsilon(j-1)^{(3+k)k/2}+m\right)^2}\\
&= \sum_{1\leq j\leq j_n}n^{(3+k)k/(2k+1)}\frac{E\left\{\left(\sup_{f_{x,y_1,\cdots,y_{k-1},y}\in\mathcal{F}_{x,jn^{-1/(2k+1)}}}\left|(\mathbb{P}_n-P_0)(f_{x,y_1,\cdots,y_{k-1},y})\right|\right)^2\right\}}{\left(\epsilon(j-1)^{(3+k)k/2}+m\right)^2}\\
&\leq C\sum_{1\leq j\leq j_n}n^{(3+k)k/(2k+1)}n^{-1}n^{-(k(k+1)-1)/(2k+1)}\frac{j^{k(k+1)-1}}{\left(\epsilon(j-1)^{(3+k)k/2}+m\right)^2}
\end{aligned}$$

$$
\begin{aligned}
&= C \sum_{1\le j\le j_n} \frac{j^{k(k+1)-1}}{\left(\epsilon(j-1)^{(3+k)k/2}+m\right)^2} \\
&\le C\sum_{j=1}^{\infty} \frac{j^{k(k+1)-1}}{\left(\epsilon(j-1)^{(3+k)k/2}+m\right)^2}, \ \searrow 0 \ \text{ as } m\to\infty,
\end{aligned}
$$

where $C > 0$ is a constant independent of $x \in [x_0 - \delta, x_0 + \delta]$. Therefore, $M_n = O_p(1)$ and hence

$$
\left|(\mathbb{P}_n - P_0)(f_{x,y_1,\cdots,y_{k-1},y})\right| \le \epsilon(y-x)^{(3+k)k/2} + O_p\left(n^{-(3+k)k/(2(2k+1))}\right)
$$

uniformly in x, y. It follows that

$$
\left|\int_{\tau_1}^{\tau_{k+1}} p(t)d(\mathbb{G}_n - G_0)(t)\right| = O_p\left(n^{-(3+k)k/(2(2k+1))}\right)
$$

and we can choose $c_0 = c$ to be large enough so that the probability of the event (3.1) is arbitrarily small. This proves the result for $j = 0$.

Now let $1 \le j \le k-1$. This time we will need $(k+1+j)$ jump points $\tau_1 < \cdots < \tau_{k+1+j}$. As for $j = 0$, τ_1 is taken to be the first jump point of $\tilde{g}_n^{(k-1)}$ after $x_0 - n^{-1/(2k+1)}$, τ_2 the first jump point after $\tau_1 + n^{-1/(2k+1)}$ and so on. Notice that the existence of at least $k+1+j$ jump points is guaranteed by the fact that $g_0^{(k)}(x_0) \ne 0$ which implies that with probability 1, the number of jump points tends to infinity with increasing sample size n. Consider the function

$$
q_j(t) = \prod_{1\le i<j\le k+j+1} (\tau_j - \tau_i) \times B_j(t)
$$

where B_j is the B-spline of degree $k+j-1$ with support $[\tau_1, \tau_{k+1+j}]$; i.e.,

$$
B_j(t) \;=\; (-1)^{k+j}(k+j)\left(\frac{(\tau_1 - t)_+^{k+j-1}}{\prod_{j\ne 1}(\tau_j - \tau_1)} + \cdots + \frac{(\tau_{k+j} - t)_+^{k+j-1}}{\prod_{j\ne k+j}(\tau_j - \tau_{k+j})}\right).
$$

It is easy to check that $p_j = q_j^{(j)}$ is a valid perturbation function (it is a spline of degree $k-1$) since for $|\eta|$ small enough, the function

$$
\tilde{g}_{\eta,n,j} = \tilde{g}_n + \eta p_j
$$

is k-monotone. It follows that

$$
\lim_{\eta\to 0} \frac{Q_n(\tilde{g}_{\eta,n,j}) - Q_n(\tilde{g}_n)}{\eta} = 0
$$

which implies that

$$
\int_{\tau_1}^{\tau_{k+1+j}} p_j(t)(\tilde{g}_n(t) - g_0(t))dt = \int_{\tau_1}^{\tau_{k+1+j}} p_j(t)d(\mathbb{G}_n(t) - G_0(t))dt
$$

By successive integrations by parts and using the fact that $q_j^{(i)}(\tau_1) = q_j^{(i)}(\tau_{k+1+j}) = 0$ for $i = 0, \cdots, k+j-2$, we obtain

$$\int_{\tau_1}^{\tau_{k+1+j}} (-1)^j q_j(t)(\tilde{g}_n^{(j)}(t) - g_0^{(j)}(t))dt = \int_{\tau_1}^{\tau_{k+1+j}} p_j(t)d(\mathbb{G}_n(t) - G_0(t))dt.$$

Therefore, if we assume that there exists $c > 0$ such that

$$\inf_{t\in[\tau_1,\tau_{k+1+j}]} \left|\tilde{g}_n^{(j)}(t) - g_0^{(j)}(t)\right| > c\, n^{-(k-j)/(2k+1)} \tag{3.6}$$

then

$$\begin{aligned}
&\left|\int_{\tau_1}^{\tau_{k+1+j}} p_j(t)d(\mathbb{G}_n(t) - G_0(t))dt\right| \\
&\geq \; c\, n^{-(k-j)/(2k+1)} \int_{\tau_1}^{\tau_{k+1+j}} q_j(t)dt \\
&\geq \; c\,(k+j)\, n^{-(k-j)/(2k+1)} \left(n^{-1/(2k+1)}\right)^{(k+1+j)(k+2+j)/2} \\
&= \; c\,(k+j)\, n^{-((2(k-j)+(k+j)(k+j+1))/(2(2k+1))} \\
&= \; c\,(k+j)\, n^{-(3k-j+(k+j)^2)/(2(2k+1))}.
\end{aligned}$$

Using similar empirical process arguments as in the proof for $j = 0$ it can be shown that

$$\left|\int_{\tau_1}^{\tau_{k+1+j}} p_j(t)d(\mathbb{G}_n(t) - G_0(t))dt\right| = O_p\left(n^{-(3k-j+(k+j)^2)/(2(2k+1))}\right)$$

and the result for $1 \leq j \leq k-1$ follows. For the MLE, the result can be proved similarly by using the same perturbation functions and also consistency of the MLE. ∎

Proposition 3.2 *Let $x_0 > 0$ and g_0 a k-monotone density such that $(-1)^k g_0^{(k)}(x_0) > 0$. Let $\bar{g}_n$ denote either the MLE $\hat{g}_n$ or the LSE $\tilde{g}_n$. If Conjecture 2.1 holds, then for each $M > 0$ we have,*

$$\sup_{|t|\leq M} \left|\bar{g}_n^{(k-1)}(x_0 + n^{-1/(2k+1)}t) - g_0^{(k-1)}(x_0)\right| = O_p(n^{-1/(2k+1)}) \tag{3.7}$$

and

$$\sup_{|t|\leq M} \left|\bar{g}_n^{(j)}(x_0 + n^{-1/(2k+1)}t) - \sum_{i=j}^{k-1} \frac{n^{-(i-j)/(2k+1)} g_0^{(i)}(x_0)}{(i-j)!} t^{i-j}\right| = O_p(n^{-(k-j)/(2k+1)}) \tag{3.8}$$

for $j = 0, \cdots, k-2$.

Proof. To prove (3.8), we will use induction starting from the highest order of differentiation $k-1$. The techniques used here are very much analogous to the ones used in the case $k=2$ in GROENEBOOM, JONGBLOED, AND WELLNER (2001B). But this was possible mainly because of the result established in the previous lemma.

We begin by establishing (3.7). Let $M > 0$ and $0 < \epsilon < 1$. We consider two sequences of $(k+1)$ jump points $\tau_{1,1}, \cdots, \tau_{k+1,1}$ and $\tau_{1,2}, \cdots, \tau_{k+1,2}$ as described in the previous theorem, where $\tau_{1,1}$ is the first jump point of $\bar{g}_n^{(k-1)}$ after $x_0 + Mn^{-1/(2k+1)}$ and $\tau_{1,2}$ is the first jump after $\tau_{k+1,1} + n^{-1/(2k+1)}$. Similarly, we define two other sequences $\tau_{1,-1} \cdots, \tau_{k+1,-1}$ and $\tau_{1,-2}, \cdots, \tau_{k+1,-2}$ to the left of x_0. By the previous theorem, we can find $c > 0$ so that,

$$\inf_{t \in [\tau_{1,i}, \tau_{k+1,i}]} |\bar{g}_n^{(k-2)}(t) - g_0^{(k-2)}(t)| < cn^{-2/(2k+1)}$$

for $i = -2, -1, 1, 2$ with probability greater than $1-\epsilon$. Let ξ_1 and ξ_2 be the minimizer of $|\bar{g}_n^{(k-2)} - g_0^{(k-2)}|$ on $[\tau_{1,1}, \tau_{k+1,1}]$ and $[\tau_{1,2}, \tau_{k+1,2}]$ respectively. Define ξ_{-1} and ξ_{-2} similarly to the left of x_0. For all $t \in [x_0 - Mn^{-1/(2k+1)}, x_0 + Mn^{-1/(2k+1)}]$, we have with probability greater than $1-\epsilon$

$$\begin{aligned}
(-1)^{k-2}\bar{g}_n^{(k-1)}(t-) &\leq (-1)^{k-2}\bar{g}_n^{(k-1)}(t+) \\
&\leq \frac{(-1)^{k-2}\bar{g}_n^{(k-2)}(\xi_2) - (-1)^{k-2}\bar{g}_n^{(k-2)}(\xi_1)}{\xi_2 - \xi_1} \\
&\leq \frac{(-1)^{k-2}g_0^{(k-2)}(\xi_2) - (-1)^{k-2}g_0^{(k-2)}(\xi_1) + 2cn^{-2/(2k+1)}}{\xi_2 - \xi_1} \\
&\leq (-1)^{k-2}g_0^{(k-1)}(\xi_2) + 2cn^{-1/(2k+1)}
\end{aligned}$$

since $\xi_2 - \xi_1 \geq n^{-1/(2k+1)}$. Similarly, with probability greater than $1-\epsilon$, we have that

$$(-1)^{k-2}\bar{g}_n^{(k-1)}(t+) \geq (-1)^{k-2}\bar{g}_n^{(k-1)}(t-) \geq (-1)^{k-2}g_0^{(k-1)}(\xi_{-2}) - 2cn^{-1/(2k+1)}.$$

Now, using the fact that $\xi_{\pm 2} = x_0 + O_p(n^{-1/(2k+1)})$ and differentiability of $g_0^{(k-1)}$ at the point x_0, we obtain (3.7).

Using similar arguments in the proof of Lemma 4.4 in GROENEBOOM, JONGBLOED, AND WELLNER (2001B), we can show (3.8) for $j = k-2$ which specializes to

$$\sup_{|t| \leq M} \left| \bar{g}_n^{(k-2)}(x_0 + n^{-1/(2k+1)}t) - g_0^{(k-2)}(x_0) - n^{-1/(2k+1)}tg_0^{(k-1)}(x_0) \right| = O_p(n^{-2/(2k+1)})$$

for all $M > 0$. Indeed, since the jump points $\tau_{j,i}, j = 1, \cdots, k+1, i = -2, -1, 1, 2$ are at distance from x_0 that is $O_p(n^{-1/(2k+1)})$, we can find with probability exceeding $1-\epsilon$, $K > M$ such that ξ_1 and ξ_2 are in $[x_0 + Mn^{-1/(2k+1)}, x_0 + Kn^{-1/(2k+1)}]$, ξ_{-2} and ξ_{-1} in $[x_0 - Kn^{-1/(2k+1)}, x_0 - Mn^{-1/(2k+1)}]$. But we know that, with probability greater than $1-\epsilon$, we can find $c > 0$ such that

$$|\bar{g}_n^{(k-2)}(\xi_{\pm 1}) - g_0^{(k-2)}(\xi_{\pm 1})| \leq cn^{-2/(2k+1)}.$$

Also, with probability greater than $1-\epsilon$, we can find $c'>0$ such that

$$\sup_{t\in[x_0-Kn^{-1/(2k+1)},x_0+Kn^{-1/(2k+1)}]}\left|\bar{g}_n^{(k-1)}(t)-g_0^{(k-1)}(x_0)\right|\le c'n^{-1/(2k+1)}.$$

Hence, with probability greater than $1-3\epsilon$, we have for any $t\in[x_0-Mn^{-1/(2k+1)},x_0+Mn^{-1/(2k+1)}]$

$$\begin{aligned}
&(-1)^{k-2}\bar{g}_n^{(k-2)}(t)\\
&\quad\ge (-1)^{k-2}\bar{g}_n^{(k-2)}(\xi_1)+(-1)^{k-2}\bar{g}_n^{(k-1)}(\xi_1)(t-\xi_1)\\
&\quad\ge (-1)^{k-2}g_0^{(k-2)}(\xi_1)-cn^{-2/(2k+1)}+((-1)^{k-2}g_0^{(k-1)}(x_0)+c'n^{-1/(2k+1)})(t-\xi_1)\\
&\quad\ge (-1)^{k-2}g_0^{(k-2)}(x_0)+(\xi_1-x_0)(-1)^{k-2}g_0^{(k-1)}(x_0)+(t-\xi_1)(-1)^{k-2}g_0^{(k-1)}(x_0)\\
&\qquad -cn^{-2/(2k+1)}-c'n^{-1/(2k+1)}(\xi_1-t) \qquad\qquad (3.9)\\
&\quad\ge (-1)^{k-2}g_0^{(k-2)}(x_0)+(t-x_0)(-1)^{k-2}g_0^{(k-1)}(x_0)-(c+2Kc')n^{-2/(2k+1)}.
\end{aligned}$$

where in (3.9), we used convexity of $(-1)^{k-2}g_0^{(k-2)}$ "from below". On the other hand, using convexity of $(-1)^{k-2}g_0^{(k-2)}$ but this time "from above", we have

$$\begin{aligned}
&(-1)^{k-2}\bar{g}_n^{(k-2)}(t)\\
&\le (-1)^{k-2}\bar{g}_n^{(k-2)}(\xi_{-1})+\frac{(-1)^{k-2}\bar{g}_n^{(k-2)}(\xi_1)-(-1)^{k-2}\bar{g}_n^{(k-2)}(\xi_{-1})}{\xi_1-\xi_{-1}}(t-\xi_{-1})\\
&\le (-1)^{k-2}\bar{g}_0^{(k-2)}(\xi_{-1})+cn^{-2/(2k+1)}\\
&\quad+\frac{(-1)^{k-2}g_0^{(k-2)}(\xi_1)-(-1)^{k-2}g_0^{(k-2)}(\xi_{-1})+2cn^{-2/(2k+1)}}{\xi_1-\xi_{-1}}(t-\xi_{-1})\\
&\le (-1)^{k-2}g_0^{(k-2)}(x_0)+(\xi_{-1}-x_0)(-1)^{k-2}g_0^{(k-2)}(x_0)+\frac{1}{2}(\xi_{-1}-x_0)^2(-1)^{k-2}g_0^{(k)}(\nu)\\
&\quad+(-1)^{k-2}g_0^{(k-1)}(\xi_1)(t-\xi_{-1})+2cn^{-2/(2k+1)}\frac{(t-\xi_{-1})}{\xi_1-\xi_{-1}}\\
&\le (-1)^{k-2}g_0^{(k-2)}(x_0)+(\xi_{-1}-x_0)(-1)^{k-2}g_0^{(k-2)}(x_0)+\frac{1}{2}(\xi_{-1}-x_0)^2(-1)^{k-2}g_0^{(k)}(\nu)\\
&\quad+\left((-1)^{k-2}g_0^{(k-1)}(x_0)+c'n^{-1/(2k+1)}\right)(t-\xi_{-1})+2cn^{-2/(2k+1)}\frac{(t-\xi_{-1})}{\xi_1-\xi_{-1}}\\
&\le (-1)^{k-2}g_0^{(k-2)}(x_0)+(t-x_0)(-1)^{k-2}g_0^{(k-1)}(x_0)+\left(\frac{D_1}{2}+2c+2Kc'\right)n^{-2/(2k+1)}
\end{aligned}$$

where $\nu\in(\xi_{-1},x_0)$, $D_1=\sup_{x\in[x_0-\delta,x_0+\delta]}|g_0^{(k)}(x)|$ and $[x_0-\delta,x_0+\delta]$ can be taken to be the largest neighborhood where $g_0^{(k)}$ exists and is continuous. In all the previous calculations, n is taken sufficiently large so that $[x_0-Kn^{-1/(2k+1)},x_0+Kn^{-1/(2k+1)}]\subseteq[x_0-\delta,x_0+\delta]$. We conclude that (3.8) holds for $j=k-2$.

Now, suppose that (3.8) is true for all $j' > j-1$; i.e., for all $M > 0$

$$\sup_{|t|<M} \left| \bar{g}_n^{(j')}(x_0 + n^{-1/(2k+1)}t) - \sum_{i=j'}^{k-1} \frac{n^{-(i-j')/(2k+1)} g_0^{(i)}(x_0)}{(i-j')!} t^{i-j'} \right| = O_p(n^{-(k-j')/(2k+1)}).$$

We are going to prove (3.8) for $j-1$. We assume without loss of generality that k and $j-1$ are even. In what follows, $\xi_{\pm 1}$ denotes the same numbers introduced before but this time there are associated with $\bar{g}_n^{(j-1)}$; i.e., for any $0 < \epsilon < 1$, there exist $c > 0$ and $K > M$ such that

$$|\bar{g}_n^{(j-1)}(\xi_{\pm 1}) - g_0^{(j-1)}(\xi_{\pm 1})| \le c n^{-(k-j+1)/(2k+1)}$$

with probability greater than $1-\epsilon$ and where $\xi_1 \in [x_0 + Mn^{-1/(2k+1)}, x_0 + Kn^{-1/(2k+1)}]$ and $\xi_{-1} \in [x_0 - Kn^{-1/(2k+1)}, x_0 - Mn^{-1/(2k+1)}]$.

Now, using the induction assumption, we know that we can find $c' > 0$ such that, with probability greater than $1-\epsilon$,

$$\begin{aligned} -c' n^{-(k-j')/(2k+1)} &\le \bar{g}_n^{(j')}(x_0 + n^{-1/(2k+1)}t) - \sum_{i=j'}^{k-1} \frac{n^{-(i-j')/(2k+1)} g_0^{(i)}(x_0)}{(i-j')!} t^{i-j'} \\ &\le c' n^{-(k-j')/(2k+1)} \end{aligned} \tag{3.10}$$

for all $|t| \le M$ and $j' > j-1$.

Using convexity of $\bar{g}_n^{(j-1)}$ "from below", we have for all $|t - x_0| \le Mn^{-1/(2k+1)}$ with probability greater than $1-2\epsilon$,

$$\begin{aligned} &\bar{g}_n^{(j-1)}(t) \\ &\ge \bar{g}_n^{(j-1)}(\xi_1) + \bar{g}_n^{(j)}(\xi_1)(t-\xi_1) + \cdots + \frac{1}{(k-j)!} \bar{g}_n^{(k-1)}(\xi_1)(t-\xi_1)^{k-j} \\ &\ge g_0^{(j-1)}(\xi_1) - c n^{-(k-j+1)/(2k+1)} + \left(\sum_{i=j}^{k-1} \frac{g_0^{(i)}(x_0)}{(i-j)!} (\xi_1 - x_0)^{i-j} (t-\xi_1) \right) \\ &\quad + \left(\sum_{i=j+1}^{k-1} \frac{g_0^{(i)}(x_0)}{(i-j-1)!} (\xi_1 - x_0)^{i-j-1} \right) \frac{(t-\xi_1)^2}{2!} + \cdots + g_0^{(k-1)}(x_0) \frac{(t-\xi_1)^{k-j}}{(k-j)!} \\ &\quad + c' n^{-(k-j)/(2k+1)}(t-\xi_1) - c' n^{-(k-j-1)/(2k+1)} \frac{(t-\xi_1)^2}{2!} \\ &\quad + \cdots - c' n^{-1/(2k+1)} \frac{(t-\xi_1)^{k-j}}{(k-j)!}. \end{aligned} \tag{3.11}$$

Using Taylor expansion of $g_0^{(j-1)}(\xi_1)$ around $g_0^{(j-1)}(x_0)$, we can write

$$g_0^{(j-1)}(\xi_1) = g_0^{(j-1)}(x_0) + g_0^{(j)}(x_0)(\xi_1 - x_0) + \cdots + \frac{g_0^{(k-1)}(x_0)}{(k-j)!}(\xi_1 - x_0)^{k-j}$$

$$+ \frac{g_0^{(k)}(\nu)}{(k-j+1)!}(\xi_1 - x_0)^{k-j+1}$$

where $\nu \in (x_0, \xi_1)$. Using this expansion and the fact that

$$|t - \xi_1| \leq K n^{-1/(2k+1)},$$

the right side of (3.11) can be bounded below by

$$\begin{aligned}
& \sum_{i=j-1}^{k-1} \frac{g_0^{(i)}(x_0)}{(i-j+1)!}(\xi_1 - x_0)^{i-j+1} + \sum_{i=j}^{k-1} \frac{g_0^{(i)}(x_0)}{(i-j)!}(\xi_1 - x_0)^{i-j}(t - \xi_1) \\
& + \sum_{i=j+1}^{k-1} \frac{g_0^{(i)}(x_0)}{(i-j-1)!}(\xi_1 - x_0)^{i-j-1}\frac{(t-\xi_1)^2}{2!} + \cdots + g_0^{(k-1)}(x_0)\frac{(t-\xi_1)^{k-j}}{(k-j)!} \\
& - \left(c + c' \sum_{p=1}^{k-j} \frac{K^p}{p!}\right) n^{-(k-j+1)/(2k+1)} + \frac{g_0^{(k)}(\nu)}{(k-j+1)!}(\xi_1 - x_0)^{k-j+1} \\
= \; & g_0^{(j-1)}(x_0) + g_0^{(j)}(x_0)(t - x_0) \\
& + \frac{g_0^{(j+1)}(x_0)}{2!}\left((\xi_1 - x_0)^2 + 2(\xi_1 - x_0)(t - \xi_1) + (t - \xi_1)^2\right) \\
& + \cdots + \frac{g_0^{(k-1)}(x_0)}{(k-j)!} \sum_{p=0}^{k-j} \frac{(k-j)!}{(k-j-p)!p!}(\xi_1 - x_0)^{k-j-p}(t-\xi_1)^p \\
& - \left(c + c' \sum_{p=1}^{k-j} \frac{K^p}{p!}\right) n^{-(k-j+1)/(2k+1)} + \frac{g_0^{(k)}(\nu)}{(k-j+1)!}(\xi_1 - x_0)^{k-j+1} \\
= \; & g_0^{(j-1)}(x_0) + g_0^{(j)}(x_0)(t - x_0) + \cdots + \frac{g^{(k-1)}(x_0)}{(k-j)!}(t - x_0)^{k-j} \\
& - \left(c + c' \sum_{p=1}^{k-j} \frac{K^p}{p!}\right) n^{-(k-j+1)/(2k+1)} - \frac{D_1 K^{k-j+1}}{(k-j+1)!} n^{-(k-j+1)/(2k+1)}
\end{aligned}$$

since $0 \leq \xi_1 - x_0 \leq K n^{-1/(2k+1)}$.

Now, we use convexity of $\bar{g}_n^{(j-1)}$ "from above". We first need to establish a useful inequality. Since $\bar{g}_n^{(k-2)}$ is convex, we have for all $t' \in [x_0 - M n^{-1/(2k+1)}, x_0 + M n^{-1/(2k+1)}]$ and

$$\bar{g}_n^{(k-2)}(t') \leq \bar{g}_n^{(k-2)}(\xi_{-1}) + \frac{\bar{g}_n^{(k-2)}(\xi_1) - \bar{g}_n^{(k-2)}(\xi_{-1})}{\xi_{n,1} - \xi_{-1}}(t' - \xi_{-1}).$$

By successive integrations of the last inequality between ξ_{-1} and t, we obtain that

$$\bar{g}_n^{(j-1)}(t) - \bar{g}_n^{(j-1)}(\xi_{-1}) \quad \leq \quad \bar{g}_n^{(j)}(\xi_{-1})(t - \xi_1) + \bar{g}_n^{(j+1)}(\xi_{-1})\frac{(t - \xi_{-1})^2}{2!}$$

$$+\cdots+\frac{\bar{g}_n^{(k-2)}(\xi_1)-\bar{g}_n^{(k-2)}(\xi_{-1})}{\xi_1-\xi_{-1}}\frac{(t-\xi_{-1})^{k-j}}{(k-j)!}.$$

It follows that with probability greater than $1-2\epsilon$, we have

$$\begin{aligned}
&\bar{g}_n^{(j-1)}(t)\\
&\le \bar{g}_n^{(j-1)}(\xi_{-1})+\bar{g}_n^{(j)}(\xi_{-1})(t-\xi_{-1})+\bar{g}_n^{(j+1)}(\xi_{-1})\frac{(t-\xi_{-1})^2}{2!}\\
&\quad+\cdots+\frac{g_0^{(k-2)}(\xi_1)-g_0^{(k-2)}(\xi_{-1})+2cn^{-2/(2k+1)}}{\xi_1-\xi_{-1}}\frac{(t-\xi_{-1})^{k-j}}{(k-j)!}\\
&\le g_0^{(j-1)}(\xi_{-1})+cn^{-(k-j+1)/(2k+1)}\\
&\quad+\left(\sum_{i=j}^{k-1}\frac{g_0^{(i)}(x_0)}{(i-j)!}(\xi_{-1}-x_0)^{i-j}+c'n^{-(k-j)/(2k+1)}\right)(t-\xi_{-1})\\
&\quad+\left(\sum_{i=j+1}^{k-1}\frac{g_0^{(i)}(x_0)}{(i-j-1)!}(\xi_{-1}-x_0)^{i-j-1}+c'n^{-(k-j-1)/(2k+1)}\right)\frac{(t-\xi_{-1})^2}{2!}\\
&\quad+\cdots+\left(g_0^{(k-1)}(\xi_1)+\frac{c}{K}n^{-1/(2k+1)}\right)\frac{(t-\xi_{-1})^{k-j}}{(k-j)!}\\
&\le \sum_{i=j-1}^{k-1}\frac{g_0^{(i)}(x_0)}{(i-j+1)!}(\xi_{-1}-x_0)^{i-j+1}+\frac{g^{(k)}(\nu)}{k!}(\xi_{-1}-x_0)^{k-j+1}\\
&\quad+\left(\sum_{i=j}^{k-1}\frac{g_0^{(i)}(x_0)}{(i-j)!}(\xi_{-1}-x_0)^{i-j}\right)(t-\xi_{-1})\\
&\quad+\cdots+\left(\sum_{i=j+1}^{k-1}\frac{g_0^{(i)}(x_0)}{(i-j-1)!}(\xi_{-1}-x_0)^{i-j-1}\right)\frac{(t-\xi_{-1})^2}{2!}\\
&\quad+\left(g_0^{(k-1)}(x_0)+cn^{-1/(2k+1)}\right)\frac{(t-\xi_{-1})^{k-j}}{(k-j)!}\\
&\quad+\left(c(1+K^{k-j})+c'\sum_{p=1}^{k-j}\frac{K^p}{p!}+\frac{D_1K^{k-j+1}}{k!}\right)n^{-(k-j+1)/(2k+1)}\\
&= g_0^{(j-1)}(x_0)+g_0^{(j)}(x_0)(t-x_0)+\cdots+g_0^{(k-j)}(x_0)\frac{(t-x_0)^{k-j}}{(k-j)!}+K'n^{-(k-j+1)/(2k+1)}
\end{aligned}$$

with $K'=c(1+K^{k-j})+c'\sum_{p=1}^{k-j}\frac{K^p}{p!}+\frac{D_1K^{k-j+1}}{k!}$. It follows that (3.8) holds for $j-1$. ■

Recall that the characterization of the LSE $\tilde{g}_n$ involved the processes $\mathbb{Y}_n$ and $\tilde{H}_n$

defined by

$$\mathbb{Y}_n(x) = \int_0^x \int_0^{t_{k-1}} \cdots \int_0^{t_2} \mathbb{G}_n(t_1) dt_1 dt_2 \cdots dt_{k-1}, \qquad x \geq 0,$$

and

$$\tilde{H}_n(x) = \int_0^x \int_0^{t_k} \cdots \int_0^{t_2} \tilde{g}_n(t_1) dt_1 dt_2 \cdots dt_k. \qquad x \geq 0,$$

Since we are interested in estimating the true density or its l-th derivative ($l \leq k-1$) at a point $x_0 > 0$, we need to define a local version of these processes. We define the local $\mathbb{Y}_n$ and $\tilde{H}_n$-processes respectively by

$$\begin{aligned}
\mathbb{Y}_n^{loc}(t) &= n^{\frac{2k}{2k+1}} \int_{x_0}^{x_0 + tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_2} \\
&\qquad \left\{ \mathbb{G}_n(v_1) - \mathbb{G}_n(x_0) - \int_{x_0}^{v_1} \sum_{j=0}^{k-1} \frac{(u-x_0)^j}{j!} g_0^{(j)}(x_0) du \right\} \Pi_{i=1}^{k-1} dv_i,
\end{aligned}$$

and

$$\begin{aligned}
\tilde{H}_n^{loc}(t) &= n^{\frac{2k}{2k+1}} \int_{x_0}^{x_0 + tn^{-1/(2k+1)}} \int_{x_0}^{v_k} \cdots \int_{x_0}^{v_2} \\
&\qquad \left\{ \tilde{g}_n(v_1) - \sum_{j=0}^{k-1} \frac{(v_1 - x_0)^j}{j!} g_0^{(j)}(x_0) \right\} dv_1 \cdots dv_k \\
&\qquad + \tilde{A}_{(k-1)n} t^{k-1} + \tilde{A}_{(k-2)n} t^{k-2} + \cdots + \tilde{A}_{1n} t + \tilde{A}_{0n},
\end{aligned}$$

where

$$\begin{aligned}
\tilde{A}_{(k-1)n} &= \frac{n^{(k+1)/(2k+1)}}{(k-1)!} \left(\tilde{H}_n^{(k-1)}(x_0) - \mathbb{Y}_n^{(k-1)}(x_0) \right) = \frac{n^{(k+1)/(2k+1)}}{(k-1)!} \left(\tilde{G}_n(x_0) - \mathbb{G}_n(x_0) \right) \\
\tilde{A}_{(k-2)n} &= \frac{n^{(k+2)/(2k+1)}}{(k-2)!} \left(\tilde{H}_n^{(k-2)}(x_0) - \mathbb{Y}_n^{(k-2)}(x_0) \right) \\
&\vdots \\
\tilde{A}_{1n} &= n^{(2k-1)/(2k+1)} \left(\tilde{H}_n'(x_0) - \mathbb{Y}_n'(x_0) \right) \\
\tilde{A}_{0n} &= n^{2k/(2k+1)} \left(\tilde{H}_n(x_0) - \mathbb{Y}_n(x_0) \right),
\end{aligned}$$

and $\tilde{G}_n(x) = \int_0^x \tilde{g}_n(y) dy$.

Example 3.1 $k = 3$

$$\begin{aligned}\mathbb{Y}_n^{loc}(t) &= n^{6/7}\int_{x_0}^{x_0+tn^{-1/7}}\int_{x_0}^{w}\Big\{\mathbb{G}_n(v)-\mathbb{G}_n(x_0)\\ &\quad -\int_{x_0}^{v}\Big(g_0(x_0)+(u-x_0)g_0'(x_0)+\frac{1}{2}(u-x_0)^2g_0''(x_0)\Big)du\Big\}dvdw,\end{aligned}$$

and

$$\begin{aligned}\tilde{H}_n^{loc}(t) &= n^{6/7}\int_{x_0}^{x_0+tn^{-1/7}}\int_{x_0}^{w}\int_{x_0}^{v}\Big\{\tilde{g}_n(u)-g_0(x_0)-(u-x_0)g_0'(u)\\ &\quad -\frac{1}{2}(u-x_0)^2g_0''(x_0)\Big\}dudvdw+\tilde{A}_{2n}t^2+\tilde{A}_{1n}t+\tilde{A}_{0n}\end{aligned}$$

where

$$\begin{aligned}\tilde{A}_{2n} &= \frac{n^{4/7}}{2}\Big(\tilde{G}_n(x_0)-\mathbb{G}_n(x_0)\Big),\\ \tilde{A}_{1n} &= n^{5/7}\Big(\tilde{H}_n'(x_0)-\mathbb{Y}_n'(x_0)\Big),\end{aligned}$$

and

$$\tilde{A}_{0n} = n^{6/7}\Big(\tilde{H}_n(x_0)-\mathbb{Y}_n(x_0)\Big).$$

In the following lemma, we will give the asymptotic distribution of the local process $\mathbb{Y}_n^{loc}$ in terms of the $(k-1)$-fold integral of two-sided Brownian motion, $g_0(x_0)$, and $g_0^{(k)}(x_0)$ assuming that the true density g_0 is k-differentiable at x_0 and continuous in an open neighborhood around x_0.

Lemma 3.1 *Let x_0 be a point where g_0 is k-differentiable and $g_0^{(k)}$ is continuous at x_0. Then*

$$\mathbb{Y}_n^{loc}(t)\Rightarrow\begin{cases}\sqrt{g_0(x_0)}\int_0^t\int_0^{s_{k-1}}\cdots\int_0^{s_2}W(s_1)ds_1\cdots ds_{k-1}+\frac{1}{2k!}t^{2k}g_0^{(k)}(x_0),\ t\geq 0\\ \sqrt{g_0(x_0)}\int_t^0\int_{s_{k-1}}^0\cdots\int_{s_2}^0W(s_1)ds_1\cdots ds_{k-1}+\frac{1}{2k!}t^{2k}g_0^{(k)}(x_0),\ t<0\end{cases}$$

in $D[-K,K]$ for every $K>0$ and where W is standard Brownian motion starting at 0.

Proof. Fix $K>0$. We will prove the lemma for $t\geq 0$ and similar arguments can be used for $t\in[-K,0)$. We have

$$\mathbb{Y}_n^{loc}(t) = n^{2k/(2k+1)}\int_{x_0}^{x_0+tn^{-1/(2k+1)}}\int_{x_0}^{v_{k-1}}\cdots\int_{x_0}^{v_2}\Big\{\mathbb{G}_n(v_1)-\mathbb{G}_n(x_0)$$

$$
\begin{aligned}
& - \int_{x_0}^{v_1} \left(g_0(x_0) + (u - x_0) g_0'(x_0) + \cdots + \frac{1}{(k-1)!} (u - x_0)^{k-1} g_0^{(k-1)}(x_0) \right) du \Big\} \\
& \qquad dv_1 dv_2 \cdots dv_{k-1} \\
= \; & A_n + B_n,
\end{aligned}
$$

where

$$
\begin{aligned}
A_n \; = \; & n^{2k/(2k+1)} \int_{x_0}^{x_0 + tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_2} \\
& \Big\{ \mathbb{G}_n(v_1) - \mathbb{G}_n(x_0) - (G_0(v_1) - G_0(x_0)) \Big\} dv_1 dv_2 \cdots dv_{k-1},
\end{aligned}
$$

and

$$
\begin{aligned}
B_n \; = \; & n^{2k/(2k+1)} \int_{x_0}^{x_0 + tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_2} \\
& \Big\{ G_0(v_1) - G_0(x_0) - \int_{x_0}^{v_1} \Big(g_0(x_0) + (u - x_0) g_0'(x_0) \\
& + \; \cdots + \frac{1}{(k-1)!} (u - x_0)^{k-1} g_0^{(k-1)}(x_0) \Big) du \Big\} dv_1 dv_2 \cdots dv_{k-1}.
\end{aligned}
$$

But, with $\mathbb{U}_n$ denoting $\sqrt{n}(\Gamma_n - I)$, $\Gamma_n(t) = n^{-1} \sum_{i=1}^n 1_{[\xi_i \le t]}$ where $\xi_1, \cdots, \xi_n$ are i.i.d. $U(0,1)$ random variables, we have

$$
\begin{aligned}
A_n \; \stackrel{d}{=} \; & n^{2k/(2k+1) - 1/2} \int_{x_0}^{x_0 + tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_2} \Big(\mathbb{U}_n(G_0(v_1)) - \mathbb{U}_n(G_0(x_0)) \Big) \Big) \\
& dv_1 dv_2 \cdots dv_{k-1} \\
= \; & n^{\frac{2k-1}{2(2k+1)}} \int_{x_0}^{x_0 + tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_2} \Big(\mathbb{U}_n(G_0(v_1)) - \mathbb{U}_n(G_0(x_0)) \Big) \Big) \\
& dv_1 dv_2 \cdots dv_{k-1},
\end{aligned}
$$

and using Taylor expansion of $G_0(v_1)$ in the neighborhood of x_0,

$$
\begin{aligned}
B_n \; = \; & n^{\frac{2k}{2k+1}} \int_{x_0}^{x_0 + tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_2} \frac{(v_1 - x_0)^{k+1}}{(k+1)!} \Big(g_0^{(k)}(v_1^*) - g_0^{(k)}(x_0) \Big) \prod_{i=1}^{k-1} dv_i \\
& + n^{\frac{2k}{2k+1}} \int_{x_0}^{x_0 + tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_2} \frac{(v_1 - x_0)^{k+1}}{(k+1)!} g_0^{(k)}(x_0) \prod_{i=1}^{k-1} dv_i \\
= \; & B_{n1} + B_{n2},
\end{aligned}
$$

where $|v_1^* - x_0| \le |v_1 - x_0|$. Now,

$$\begin{aligned}
B_{n2} &= n^{\frac{2k}{2k+1}} \frac{1}{(k+1)!} g_0^{(k)}(x_0) \int_{x_0}^{x_0+tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_3} \frac{1}{k+2}(v_2 - x_0)^{k+2} dv_2 \cdots dv_{k-1} \\
&= n^{\frac{2k}{2k+1}} \frac{1}{(k+3)!} g_0^{(k)}(x_0) \int_{x_0}^{x_0+tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_4} (v_3 - x_0)^{k+3} dv_4 \cdots dv_{k-1} \\
&\vdots \\
&= n^{\frac{2k}{2k+1}} \frac{1}{(2k-1)!} g_0^{(k)}(x_0) \int_{x_0}^{x_0+tn^{-1/(2k+1)}} (v_{k-1} - x_0)^{2k-1} dv_{k-1} \\
&= n^{\frac{2k}{2k+1}} g_0^{(k)}(x_0) \frac{1}{(2k)!} \left(\frac{t}{n^{1/2k+1}} \right)^{2k} \\
&= \frac{1}{(2k)!} g_0^{(k)}(x_0) t^{2k}.
\end{aligned}$$

Furthermore, by continuity of $g_0^{(k)}$ at x_0, we deduce that $B_{n1}(t) = o(1)$ uniformly in $0 \le t \le K$ and hence

$$B_n \to \frac{1}{(2k)!} g_0^{(k)}(x_0) t^{2k}, \tag{3.12}$$

as $n \to \infty$ uniformly in $0 \le t \le K$.

Using the identity

$$\mathbb{U}(G_0(v)) - \mathbb{U}(G_0(x_0)) \stackrel{d}{=} W(G_0(v)) - W(G_0(x_0)) - (G_0(v) - G_0(x_0))W(1),$$

where W is two-sided Brownian motion process, we have

$$\begin{aligned}
A_n &\stackrel{d}{=} n^{\frac{2k-1}{2(2k+1)}} \int_{x_0}^{x_0+tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_2} \\
&\qquad \Big(\mathbb{U}_n(v_1) - \mathbb{U}(v_1) - (\mathbb{U}_n(x_0) - \mathbb{U}(x_0) \Big) dv_1 \cdots dv_{k-1} \\
&\quad + n^{\frac{2k-1}{2(2k+1)}} \int_{x_0}^{x_0+tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_2} \Big(W(G_0(v)) - W(G_0(x_0)) \Big) \\
&\quad - W(1) n^{\frac{2k-1}{2(2k+1)}} \int_{x_0}^{x_0+tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_2} \\
&\qquad (G_0(v_1) - G_0(x_0)) dv_1 \cdots dv_{k-1} \\
&= A_{n1} + A_{n2} + A_{n3}.
\end{aligned}$$

But,

$$
\begin{aligned}
A_{n1} &\leq 2n^{\frac{2k-1}{2(2k+1)}}\|\mathbb{U}_n-\mathbb{U}\|_\infty \int_{x_0}^{x_0+tn^{-1/(2k+1)}}\int_{x_0}^{v_{k-1}}\cdots\int_{x_0}^{v_2} dv_1\cdots dv_{k-1}\\
&= 2n^{\frac{2k-1}{2(2k+1)}}\|\mathbb{U}_n-\mathbb{U}\|_\infty \int_{x_0}^{x_0+tn^{-1/(2k+1)}}\int_{x_0}^{v_{k-1}}\cdots\int_{x_0}^{v_3}(v_2-x_0)dv_2\cdots dv_{k-1}\\
&= 2n^{\frac{2k-1}{2(2k+1)}}\|\mathbb{U}_n-\mathbb{U}\|_\infty \int_{x_0}^{x_0+tn^{-1/(2k+1)}}\int_{x_0}^{v_{k-1}}\cdots\int_{x_0}^{v_4}\frac{1}{2}(v_3-x_0)^2dv_3\\
&\vdots\\
&= 2n^{\frac{2k-1}{2(2k+1)}}\|\mathbb{U}_n-\mathbb{U}\|_\infty \frac{1}{(k-2)!}\int_{x_0}^{x_0+tn^{-1/(2k+1)}}(v_{k-1}-x_0)^{k-2}dv_{k-1}\\
&= 2n^{\frac{2k-1}{2(2k+1)}}\|\mathbb{U}_n-\mathbb{U}\|_\infty \frac{1}{(k-1)!}\left(\frac{t}{n^{1/(2k+1)}}\right)^{k-1}\\
&= 2t^{k-1}n^{\frac{1/2}{2k+1}}O\left(\frac{\log(n)^2}{n^{1/2}}\right)\\
&= 2t^{k-1}O\left(\frac{\log(n)^2}{n^{k/(2k+1)}}\right) \qquad (3.13)
\end{aligned}
$$

since $\|\mathbb{U}_n-\mathbb{U}\|_\infty = O\left(n^{-1/2}(\log(n))^2\right)$ via Komlós, Major and Tusnády (1975); see e.g. Shorack and Wellner (1986), page 494.

On the other hand, using the fact that g_0 is nonincreasing, we have

$$
\begin{aligned}
A_{n3} &\leq |W(1)|g_0(x_0)n^{\frac{2k-1}{2(2k+1)}}\int_{x_0}^{x_0+tn^{-1/(2k+1)}}\int_{x_0}^{v_{k-1}}\cdots\int_{x_0}^{v_2}(v_1-x_0)dv_1\cdots dv_{k-1}\\
&= |W(1)|g_0(x_0)n^{\frac{2k-1}{2(2k+1)}}\int_{x_0}^{x_0+tn^{-1/(2k+1)}}\int_{x_0}^{v_{k-1}}\cdots\int_{x_0}^{v_3}\frac{1}{2}(v_1-x_0)^2dv_2\\
&\vdots\\
&= |W(1)|g_0(x_0)n^{\frac{2k-1}{2(2k+1)}}\frac{1}{(k-1)!}\int_{0}^{x_0+tn^{-1/(2k+1)}}(v_{k-1}-x_0)^{k-1}dv_{k-1}\\
&= |W(1)|g_0(x_0)n^{\frac{2k-1}{2(2k+1)}}\frac{1}{k!}\left(\frac{t}{n^{1/(2k+1)}}\right)^{k}\\
&= |W(1)|g_0(x_0)t^kn^{-\frac{1}{2(2k+1)}} \to_p 0, \qquad (3.14)
\end{aligned}
$$

as $n\to\infty$ uniformly in $0\leq t\leq K$.

Finally, using the change of variables $s_j = n^{1/(2k+1)}(v_j-x_0)$ for $j=1,\ldots,k-1$, we have

$$
\begin{aligned}
A_{n2} &= n^{\frac{2k-1}{2(2k+1)}} \int_{x_0}^{x_0+tn^{\frac{-1}{2k+1}}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_2} \Big(W(G_0(v_1)) - W(G_0(x_0)) \Big) dv_1 \cdots dv_{k-1} \\
&= n^{\frac{2k-1}{2(2k+1)}} n^{-\frac{(k-1)}{(2k+1)}} \int_0^t \int_0^{s_{k-1}} \cdots \int_0^{s_2} \Big(W(G_0(n^{\frac{-1}{2k+1}} s_1 + x_0)) - W(G_0(x_0)) \Big) \\
&\qquad ds_1 \cdots ds_{k-1} \\
&\stackrel{d}{=} n^{\frac{1}{2(2k+1)}} \int_0^t \int_0^{s_{k-1}} \cdots \int_0^{v_2} W\Big(G_0(n^{\frac{-1}{2k+1}} s_1 + x_0) - G_0(x_0) \Big) ds_1 \cdots ds_{k-1} \\
&\stackrel{d}{=} \int_0^t \int_0^{s_{k-1}} \cdots \int_0^{s_2} W\Big(n^{\frac{1}{2k+1}} (G_0(n^{\frac{-1}{2k+1}} s_1 + x_0) - G_0(x_0)) \Big) ds_1 \cdots ds_{k-1} \\
&\to \int_0^t \int_0^{s_{k-1}} \cdots \int_0^{s_2} W(s_1 g_0(x_0)) ds_1 \cdots ds_{k-1} \quad \text{as } n \to \infty \\
&\stackrel{d}{=} \sqrt{g_0(x_0)} \int_0^t \int_0^{s_{k-1}} \cdots \int_0^{s_2} W(s_1) ds_1 \cdots ds_{k-1}. \qquad (3.15)
\end{aligned}
$$

Therefore, combining (3.12), (3.13), (3.14) and (3.15) yields

$$
\mathbb{Y}_n^{loc}(t) \Rightarrow \sqrt{g_0(x_0)} \int_0^t \int_0^{s_{k-1}} \cdots \int_0^{s_2} W(s_1) ds_1 \cdots ds_{k-1} + \frac{1}{(2k)!} t^{2k} g_0^{(k)}(x_0)
$$

for $0 \le t \le K$. A similar argument for $-K \le t < 0$ yields the conclusion. ■

We will now rescale this limiting process to obtain a "canonical" version. In the case of $k = 2$, Groeneboom, Jongbloed and Wellner (Groeneboom, Jongbloed, and Wellner (2001b)) chose the "canonical process" to be

$$
Y(t) = \int_0^t W(y) dy + t^4,
$$

and one can establish a link between estimating a non-decreasing convex density and the following Gaussian problem:

$$
dX(t) = f_0(t) dt + dW(t) \qquad (3.16)
$$

where f_0 is convex. Integrating (3.16) twice and choosing $f_0(t) = 12t^2$, we have

$$
\int_0^t X(y) dy = \int_0^t W(y) dy + t^4 = Y(t).
$$

Similarly, one can establish a link between the k-monotone density estimation problem and the Gaussian problem:

$$
dX(t) = f_0(t) dt + dW(t)
$$

where $(-1)^k f_0$ has a convex $(k-2)$-th derivative. If we choose $f_0(t) = t^k$ and integrate the previous stochastic differential equation $k-1$ times, we get

$$
\begin{aligned}
X(t) &= \frac{1}{k+1} t^{k+1} + W(t) \\
X_1(t) &= \int_0^t X(s) ds = \frac{1}{(k+1)(k+2)} t^{k+2} + \int_0^t W(s) ds \\
X_2(t) &= \int_0^t \int_0^{s_2} X(s_1) ds_1 ds_2 = \frac{k!}{(k+3)!} t^{k+3} + \int_0^t \int_0^{s_2} W(s_1) ds_1 ds_2 \\
&\vdots \\
X_{k-1}(t) &= \frac{k!}{(2k)!} t^{2k} + \int_0^t \int_0^{s_{k-1}} \cdots \int_0^{s_2} W(s_1) ds_1 ds_2 \cdots ds_{k-1} \stackrel{d}{=} Y_k(t).
\end{aligned}
$$

Here we will rescale the limiting process so that we obtain the "canonical process"

$$
Y_k(t) = \int_0^t \int_0^{s_{k-1}} \cdots \int_0^{s_2} W(s_1) ds_1 ds_2 \cdots ds_{k-1} + (-1)^k \frac{k!}{(2k)!} t^{2k}, \quad t \geq 0.
$$

Let us denote by σ and a, the multiplicative term $\sqrt{g_0(x_0)}$ and $(-1)^k g_0^{(k)}(x_0)/k!$, the leading coefficient of the drift term in the limiting process

$$
Y_{a,\sigma}(t) = \sqrt{g_0(x_0)} \int_0^t \int_0^{s_{k-1}} \cdots \int_0^{s_2} W(s_1) ds_1 \cdots ds_{k-1} + \frac{(-1)^k}{k!} g_0^{(k)}(x_0) (-1)^k \frac{k!}{(2k)!} t^{2k}
$$

respectively. In the following, we are going to find constants r_1 and r_2 such that

$$
r_1 Y_{a,\sigma}(r_2 t) \stackrel{d}{=} Y_k(t).
$$

We have,

$$
\begin{aligned}
Y_{a,\sigma}(t) &= a(-1)^k \frac{k!}{(2k)!} t^{2k} + \sigma \int_0^t \int_0^{s_{k-1}} \cdots \int_0^{s_2} W(s_1) ds_1 \cdots ds_{k-1} \\
&\stackrel{d}{=} a(-1)^k \frac{k!}{(2k)!} t^{2k} + \alpha^{-1/2} \sigma \int_0^t \int_0^{s_{k-1}} \cdots \int_0^{s_2} W(\alpha s_1) ds_1 \cdots ds_{k-1} \\
&\stackrel{d}{=} a(-1)^k \frac{k!}{(2k)!} t^{2k} + \alpha^{-1/2} \sigma \int_0^t \int_0^{s_{k-1}} \cdots \int_0^{\alpha s_2} \frac{1}{\alpha} W(s_1) ds_1 \cdots ds_{k-1} \\
&\stackrel{d}{=} a(-1)^k \frac{k!}{(2k)!} t^{2k} + \alpha^{-1/2} \sigma \int_0^t \int_0^{s_{k-1}} \cdots \int_0^{\alpha s_3} \int_0^{s_2} \frac{1}{\alpha^2} W(s_1) ds_1 \cdots ds_{k-1} \\
&\vdots \\
&\stackrel{d}{=} a(-1)^k \frac{k!}{(2k)!} t^{2k} + \alpha^{-1/2} \sigma \int_0^{\alpha t} \int_0^{s_{k-1}} \cdots \int_0^{s_2} W(s_1) \frac{1}{\alpha^{k-1}} ds_1 \cdots ds_{k-1} \\
&\stackrel{d}{=} a(-1)^k \frac{k!}{(2k)!} t^{2k} + \alpha^{1/2-k} \sigma \int_0^{\alpha t} \int_0^{s_{k-1}} \cdots \int_0^{s_2} W(s_1) ds_1 \cdots ds_{k-1}.
\end{aligned}
$$

Therefore,

$$r_1 Y_{a,\sigma}(r_2 t) \stackrel{d}{=} a(-1)^k \frac{k!}{(2k)!} r_1 (r_2 t)^{2k} + r_1 \alpha^{1/2-k} \sigma \int_0^{r_2 \alpha t} \int_0^{s_{k-1}} \cdots \int_0^{s_2} W(s_1) ds_1 \cdots ds_{k-1},$$

and

$$\begin{cases} a r_1 r_2^{2k} = 1, \\ r_1 \alpha^{1/2-k} \sigma = 1, \\ r_2 \alpha = 1. \end{cases}$$

Solving the previous system of equations yields

$$\alpha = \left(\frac{a}{\sigma}\right)^{2/(2k+1)}$$

and therefore

$$r_1 = \frac{1}{\sqrt{g_0(x_0)}} \left(\frac{(-1)^k g_0^{(k)}(x_0)}{k! \sqrt{g_0(x_0)}} \right)^{(2k-1)/(2k+1)} \quad \text{and} \tag{3.17}$$

$$r_2 = \left(\frac{\sqrt{g_0(x_0)}}{\frac{(-1)^k g_0^{(k)}(x_0)}{k!}} \right)^{2/(2k+1)}. \tag{3.18}$$

Thus,

$$\begin{aligned} Y_{a,\sigma}(t) &\stackrel{d}{=} \frac{1}{r_1} Y_k \left(\frac{t}{r_2} \right) \\ &= \sqrt{g_0(x_0)} \left(\frac{k! \sqrt{g_0(x_0)}}{(-1)^k g_0^{(k)}(x_0)} \right)^{(2k-1)/(2k+1)} Y_k \left(\left(\frac{k! \sqrt{g_0(x_0)}}{(-1)^k g_0^{(k)}(x_0)} \right)^{-2/(2k+1)} t \right). \end{aligned}$$

Note that (3.17) specializes to A.9 in Groeneboom, Jongbloed, and Wellner (2001a), page 1651 when $k = 2$.

Let us now have a closer look at the difference of the two local processes $\mathbb{Y}_n^{loc}$ and $\tilde{H}_n^{loc}$. The asymptotic behavior of this difference, as we will show later, will have a crucial role in establishing the asymptotic theory of the LSE.

We have,

$$\begin{aligned} &\tilde{H}_n^{loc}(t) - \mathbb{Y}_n^{loc}(t) \\ &= n^{\frac{2k}{2k+1}} \int_{x_0}^{x_0 + tn^{-\frac{1}{2k+1}}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_2} \Big\{ \Big((\tilde{G}_n(v_1) - \tilde{G}_n(x_0)) - (\mathbb{G}_n(v_1) - \mathbb{G}_n(x_0)) \Big) \\ &\qquad dv_1 \cdots dv_{k-1} \Big\} + \tilde{A}_{(k-1)n} t^{k-1} + \cdots + \tilde{A}_{1n} t + \tilde{A}_{0n} \\ &= n^{\frac{2k}{2k+1}} \int_{x_0}^{x_0 + tn^{-\frac{1}{2k+1}}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_2} \Big(\tilde{G}_n(v_1) - \mathbb{G}_n(v_1) \Big) dv_1 \cdots dv_{k-1} \end{aligned}$$

$$
\begin{aligned}
& - \frac{n^{(k+1)/(2k+1)}}{(k-1)!}\Big(\tilde{G}_n(x_0) - \mathbb{G}_n(x_0)\Big)t^{k-1} + \tilde{A}_{(k-1)n}t^{k-1} + \cdots + \tilde{A}_{1n}t + \tilde{A}_{0n} \\
= \;& n^{\frac{2k}{2k+1}} \int_{x_0}^{x_0+tn^{-\frac{1}{2k+1}}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_2} \Big(\tilde{G}_n(v_1) - \mathbb{G}_n(v_1)\Big) dv_1 \cdots dv_{k-1} \\
& - \tilde{A}_{(k-1)n}t^{k-1} + \tilde{A}_{(k-1)n}t^{k-1} + \cdots + \tilde{A}_{1n}t + \tilde{A}_{0n} \\
= \;& n^{\frac{2k}{2k+1}} \int_{x_0}^{x_0+tn^{-\frac{1}{2k+1}}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_2} \Big(\tilde{G}_n(v_1) - \mathbb{G}_n(v_1)\Big) dv_1 \cdots dv_{k-1} \\
& + \tilde{A}_{(k-2)n}t^{k-2} + \cdots + \tilde{A}_{1n}t + \tilde{A}_{0n} \\
= \;& n^{\frac{2k}{2k+1}} \int_{x_0}^{x_0+tn^{-\frac{1}{2k+1}}} \int_{x_0}^{v_{k-1}} \cdots \int_{0}^{v_2} \Big(\tilde{G}_n(v_1) - \mathbb{G}_n(v_1)\Big) dv_1 \cdots dv_{k-1} \\
& - n^{\frac{2k}{2k+1}} \int_{x_0}^{x_0+tn^{\frac{-1}{2k+1}}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_3} dv_2 \cdots dv_{k-1} \times \int_0^{x_0} \Big(\tilde{G}_n(v_1) - \mathbb{G}_n(v_1)\Big) dv_1 \\
& + \tilde{A}_{(k-2)n}t^{k-2} + \cdots + \tilde{A}_{1n}t + \tilde{A}_{0n} \\
= \;& n^{\frac{2k}{2k+1}} \int_{x_0}^{x_0+tn^{\frac{-1}{2k+1}}} \int_{x_0}^{v_{k-1}} \cdots \int_{0}^{v_2} \Big(\tilde{G}_n(v_1) - \mathbb{G}_n(v_1)\Big) dv_1 \cdots dv_{k-1} \\
& - n^{(k+2)/(2k+1)} \frac{t^{k-2}}{(k-2)!} \times \int_0^{x_0} \Big(\tilde{G}_n(v_1) - \mathbb{G}_n(v_1)\Big) dv_1 + \tilde{A}_{(k-2)n}t^{k-2} \\
& + \tilde{A}_{(k-3)n}t^{k-3} + \cdots + \tilde{A}_{1n}t + \tilde{A}_{0n} \\
= \;& n^{\frac{2k}{2k+1}} \int_{x_0}^{x_0+tn^{\frac{-1}{2k+1}}} \int_{x_0}^{v_{k-1}} \cdots \int_{0}^{v_2} \Big(\tilde{G}_n(v_1) - \mathbb{G}_n(v_1)\Big) dv_1 \cdots dv_{k-1} \\
& - \tilde{A}_{(k-2)n}t^{k-2} + \tilde{A}_{(k-2)n}t^{k-2} + \cdots + \tilde{A}_{1n}t + \tilde{A}_{0n} \\
= \;& n^{\frac{2k}{2k+1}} \int_{x_0}^{x_0+tn^{\frac{-1}{2k+1}}} \int_{x_0}^{v_{k-1}} \cdots \int_{0}^{v_2} \Big(\tilde{G}_n(v_1) - \mathbb{G}_n(v_1)\Big) dv_1 \cdots dv_{k-1} \\
& + \tilde{A}_{(k-3)n}t^{k-2} + \cdots + \tilde{A}_{1n}t + \tilde{A}_{0n} \\
\vdots & \\
= \;& n^{\frac{2k}{2k+1}} \left(\tilde{H}_n(x_0 + tn^{\frac{-1}{2k+1}}) - \mathbb{Y}_n(x_0 + tn^{\frac{-1}{2k+1}})\right) \geq 0,
\end{aligned}
$$

by the first Fenchel condition satisfied by the LSE.

A natural thing to do is to rescale the processes $\mathbb{Y}_n^{loc}(t)$ and $\tilde{H}_n^{loc}(t)$ so that the rescaled $\mathbb{Y}_n^{loc}(t)$ converges to the process Y_k we defined already. Since the scaling of $\mathbb{Y}_n^{loc}(t)$ will be exactly the same as the one we used for Y_k, we define $\tilde{H}_n^l$ as

$$
\tilde{H}_n^l(t) = r_1 \tilde{H}_n^{loc}(r_2 t)
$$

where

$$r_1 = \frac{1}{\sqrt{g_0(x_0)}} \left(\frac{(-1)^k g_0^{(k)}(x_0)}{\sqrt{g_0(x_0)} k!} \right)^{(2k-1)/(2k+1)}, \quad r_2 = \left(\frac{(-1)^k g_0^{(k)}(x_0)}{\sqrt{g_0(x_0)} k!} \right)^{-2/(2k+1)}.$$

Now, we can write

$$\begin{aligned}
(\tilde{H}_n^l)^{(k)}(0) &= r_1 r_2^k (\tilde{H}_n^{loc})^{(k)}(0) = n^{k/(2k+1)} c_k(g_0)(\tilde{g}_n(x_0) - g_0(x_0)) \\
(\tilde{H}_n^l)^{(k+1)}(0) &= r_1 r_2^{k+1} (\tilde{H}_n^{loc})^{(k+1)}(0) = n^{(k-1)/(2k+1)} c_{k-1}(g_0)(\tilde{g}_n'(x_0) - g_0'(x_0)) \\
(\tilde{H}_n^l)^{(k+2)}(0) &= r_1 r_2^{k+2} (\tilde{H}_n^{loc})^{(k+2)}(0) = n^{(k-2)/(2k+1)} c_{k-2}(g_0)(\tilde{g}_n''(x_0) - g_0''(x_0)) \\
&\vdots \\
(\tilde{H}_n^l)^{(2k-1)}(0) &= r_1 r_2^{2k-1} (\tilde{H}_n^{loc})^{(2k-1)}(0) = n^{1/(2k+1)} c_1(g_0)(\tilde{g}_n^{(k-1)}(x_0) - g_0^{(k-1)}(x_0)).
\end{aligned}$$

Now, let us consider the MLE $\hat{g}_n$. Recall that the characterization of this estimator involves the process $\widehat{H}_n$ given by

$$\widehat{H}_n(t) = \int_0^t \frac{(t-u)^{k-1}}{\hat{g}_n(u)} d\mathbb{G}_n(t), \quad \text{for all} \quad t \geq 0$$

and that

$$\widehat{H}_n(t) \begin{cases} \leq \frac{t^k}{k}, & t \geq 0 \\ = \frac{t^k}{k}, & \text{when } \ t \ \text{ is a jump point of } \ \hat{g}_n^{(k-1)} \end{cases}$$

is a necessary and sufficient condition for $\hat{g}_n$ to be the MLE. Note that $\widehat{H}_n$ and $\hat{H}_n$ defined in Balabdaoui and Wellner (2004a), Lemma 2.6 are different: $\widehat{H}_n(t) = (t^k/k)\hat{H}_n(t)$ for $t \geq 0$.

We define the local processes $\widehat{Y}_n^{loc}$ and $\widehat{H}_n^{loc}$ as

$$\begin{aligned}
\widehat{Y}_n^{loc}(t) &= n^{2k/(2k+1)} g_0(x_0) \int_{x_0}^{x_0 + tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_1} \frac{g_0(v) - \sum_{j=1}^{k-1} \frac{(v-x_0)^j}{j!} g_0^{(j)}(x_0)}{\hat{g}_n(v)} \\
&\qquad dv dv_1 \cdots dv_{k-1} \\
&\quad + n^{2k/(2k+1)} g_0(x_0) \int_{x_0}^{x_0 + tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_1} \frac{1}{\hat{g}_n(v)} d(\mathbb{G}_n - G_0)(v) \\
&\qquad dv_1 \cdots dv_{k-1}
\end{aligned}$$

and

$$\begin{aligned}
\widehat{H}_n^{loc}(t) &= n^{2k/(2k+1)} g_0(x_0) \int_{x_0}^{x_0 + tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_1} \frac{\hat{g}_n(v) - \sum_{j=1}^{k-1} \frac{(v-x_0)^j}{j!} g_0^{(j)}(x_0)}{\hat{g}_n(v)} \\
&\qquad dv dv_1 \cdots dv_{k-1} + \widehat{A}_{(k-1)n} t^{k-1} + \cdots + \widehat{A}_{0n}
\end{aligned}$$

where for $0 \le j \le k-1$

$$\widehat{A}_{jn} = -\frac{n^{(2k-j)/(2k+1)}}{(k-1)!j!} g_0(x_0) \left(\widehat{H}_n^{(j)}(x_0) - \frac{(k-1)!}{(k-j)!} x_0^{k-j} \right).$$

With this particular choice of $\widehat{A}_{jn}$, $0 \le j \le k-1$, we have

$$\begin{aligned}
&\widehat{H}_n^{loc}(t) - \widehat{Y}_n^{loc}(t) \\
&= n^{2k/(2k+1)} g_0(x_0) \int_{x_0}^{x_0+tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_1} \frac{\hat{g}_n(v) - g_0(v)}{\hat{g}_n(v)} dv dv_1 \cdots dv_{k-1} \\
&\quad - n^{2k/(2k+1)} g_0(x_0) \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_1} \frac{1}{\hat{g}_n(v)} d(\mathbb{G}_n - G_0)(v) dv_1 \cdots dv_{k-1} \\
&\quad + \widehat{A}_{(k-1)n} t^{k-1} + \cdots + \widehat{A}_{0n} \\
&= n^{2k/(2k+1)} g_0(x_0) \left(\frac{t^k}{k!} n^{-k/(2k+1)} - \int_{x_0}^{x_0+n^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_1} \frac{1}{\hat{g}_n(v)} d\mathbb{G}_n(v) \prod_{i=1}^{k-1} dv_i \right) \\
&\quad + \widehat{A}_{(k-1)n} t^{k-1} + \cdots + \widehat{A}_{0n}.
\end{aligned}$$

But notice that for any $t \ge 0$

$$\int_0^t \frac{1}{\hat{g}_n(u)} d\mathbb{G}_n(u) = \frac{1}{(k-1)!} \widehat{H}_n^{(k-1)}(t).$$

It follows that

$$\begin{aligned}
&\int_{x_0}^{x_0+tn^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_1} \frac{1}{\hat{g}_n(v)} d\mathbb{G}_n(v) dv_1 \cdots dv_{k-1} \\
&= \frac{1}{(k-1)!} \int_{x_0}^{x_0+n^{-1/(2k+1)}} \int_{x_0}^{v_{k-1}} \cdots \int_{x_0}^{v_1} \left(\widehat{H}_n^{(k-1)}(v_1) - \widehat{H}_n^{(k-1)}(x_0) \right) dv_1 \cdots dv_{k-1} \\
&= \frac{1}{(k-1)!} \left(\widehat{H}_n(x_0 + tn^{-1/(2k+1)}) - \sum_{j=0}^{k-1} \frac{t^j n^{-j/(2k+1)}}{j!} \widehat{H}_n^{(j)}(x_0) \right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\widehat{H}_n^{loc}(t) - \widehat{Y}_n^{loc}(t) \\
&= n^{2k/(2k+1)} g_0(x_0) \left\{ -\frac{\widehat{H}_n(x_0 + tn^{-1/(2k+1)})}{(k-1)!} + \frac{t^k}{k!} n^{-k/(2k+1)} + \sum_{j=0}^{k-1} \frac{t^j n^{-j/(2k+1)}}{(k-1)!j!} \widehat{H}_n^{(j)}(x_0) \right\} \\
&\quad + \widehat{A}_{(k-1)n} t^{k-1} + \cdots + \widehat{A}_{0n} \\
&= n^{2k/(2k+1)} \frac{g_0(x_0)}{k!} \Big\{ -k\widehat{H}_n(x_0 + tn^{-1/(2k+1)}) + t^k n^{-k/(2k+1)}
\end{aligned}$$

$$+\sum_{j=0}^{k-1}\frac{t^j n^{-j/(2k+1)}}{j!}k\left(\widehat{H}_n^{(j)}(x_0)-\frac{1}{k}\frac{k!}{(k-j)!}x_0^{k-j}\right)+\sum_{j=0}^{k-1}\frac{k!}{j!(k-j)!}t^j n^{-j/(2k+1)}x_0^{k-j}\Bigg\}$$
$$+\widehat{A}_{(k-1)n}t^{k-1}+\cdots+\widehat{A}_{0n}$$
$$= n^{2k/(2k+1)}\frac{g_0(x_0)}{k!}\Big\{-k\widehat{H}_n(x_0+tn^{-1/(2k+1)})+(x_0+tn^{-1/(2k+1)})^k\Big\}$$

by replacing the coefficients $\widehat{A}_{jn}$, $0\le j\le k-1$ by their expressions. It follows that

$$\widehat{H}_n^{loc}(t)-\widehat{Y}_n^{loc}(t)=n^{2k/(2k+1)}\frac{g_0(x_0)}{(k-1)!}\left(\frac{1}{k}(x_0+tn^{-1/(2k+1)})^k-\widehat{H}_n(x_0+tn^{-1/(2k+1)})\right)\ge 0.$$

As for the LSE, we define $\widehat{Y}_n^l$ and $\widehat{H}_n^l$ by

$$\widehat{Y}_n^l(t)=r_1\widehat{Y}_n^{loc}(r_2 t)$$

and

$$\widehat{H}_n^l(t)=r_1\widehat{H}_n^{loc}(r_2 t).$$

Lemma 3.2 *Let $K>0$. Then*

$$\widehat{Y}_n\Rightarrow Y_k$$

in $D[-K,K]$.

Proof. We apply the same arguments in the proof of Lemma 3.1 in the case of the LSE. ■

Now, let $\bar{H}_n^l$ denote either $\tilde{H}_n^l$ or $\widehat{H}_n^l$. Recall that

$$\tilde{A}_{jn}=\frac{n^{(2k-j)/(2k+1)}}{j!}\left(\tilde{H}_n^{(j)}(x_0)-\mathbb{Y}_n^{(j)}(x_0)\right)$$

and

$$\widehat{A}_{jn}=-\frac{n^{(2k-j)/(2k+1)}}{(k-1)!j!}g_0(x_0)\left(\widehat{H}_n^{(j)}(x_0)-\frac{(k-1)!}{(k-j)!}x_0^{k-j}\right).$$

To show that the derivatives of $\bar{H}_n^l$ are tight, we need the following lemma.

Lemma 3.3 *For all $j\in\{0,\ldots,k-1\}$, let $\bar{A}_{jn}$ denote either $\tilde{A}_{jn}$ or $\widehat{A}_{jn}$. If Conjecture 2.1 holds, then*

$$\bar{A}_{jn}=O_p(1). \tag{3.19}$$

Proof. We will show the lemma only for the LSE as the arguments are very similar for the MLE. Let $j \in \{0, \ldots, k-1\}$ and denote $\tilde{\Delta}_n(x) = \tilde{H}_n(x) - \mathbb{Y}_n(x)$ for all $x \geq 0$. We will start by proving (3.19) for $j = k-1$ and $k-2$ and then use induction for $2 \leq j \leq k-3$. Proving (3.19) for $j = k-1$ would have been sufficient but we wanted to show it for $j = k-2$ to give a better idea about how the proof works.

Now consider k successive jump points, $\tau_1, \cdots, \tau_k$, of $\tilde{g}_n^{(k-1)}$ where τ_1 is the first jump after x_0. By the mean value theorem, there exist $\tau_1^{(1)} \in (\tau_1, \tau_2)$, $\tau_2^{(1)} \in (\tau_2, \tau_3)$, $\ldots$, $\tau_{k-1}^{(1)} \in (\tau_{k-1}, \tau_k)$ such that $\tilde{\Delta}_n'(\tau_i^{(1)}) = 0$ for $1 \leq i \leq k-1$. Also, by the same theorem there exist $\tau_1^{(2)} \in (\tau_1^{(1)}, \tau_2^{(1)}), \ldots, \tau_{k-2}^{(2)} \in (\tau_{k-2}^{(1)}, \tau_{k-1}^{(1)})$ such that $\tilde{\Delta}_n''(\tau_i^{(2)}) = 0$ for $1 \leq i \leq k-2$. It is easy to see that we can carry on this reasoning up to the $(k-1)$-st level of differentiation and so there exists $\tau^{(k-1)}$ such that

$$\tilde{\Delta}_n^{(k-1)}(\tau^{(k-1)}) = 0.$$

Denote $\tau = \tau^{(k-1)}$. We can write

$$\tilde{\Delta}_n^{(k-1)}(x_0) = \tilde{\Delta}_n^{(k-1)}(x_0) - \tilde{\Delta}_n^{(k-1)}(\tau).$$

But since

$$\tilde{\Delta}_n^{(k-1)}(x) \quad = \quad \int_0^x d(\tilde{G}_n(t) - \mathbb{G}_n(t)), \quad \text{for } x \geq 0,$$

we can write,

$$\begin{aligned}
|\tilde{\Delta}_n^{(k-1)}(x_0)| &= \left| \int_{x_0}^{\tau} d(\tilde{G}_n(t) - \mathbb{G}_n(t)) \right| \\
&\leq \left| \int_{x_0}^{\tau} d(\tilde{G}_n(t) - G_0(t)) \right| + \left| \int_{x_0}^{\tau} d(\mathbb{G}_n(t) - G_0(t)) \right| \\
&= \left| \int_{x_0}^{\tau} (\tilde{g}_n(t) - g_0(t)) dt \right| + \left| \int_{x_0}^{\tau} d(\mathbb{G}_n(t) - G_0(t)) \right| \\
&\leq \int_{x_0}^{\tau} |\tilde{g}_n(t) - g_0(t)|\, dt + \left| \int_{x_0}^{\tau} d(\mathbb{G}_n(t) - G_0(t)) \right|.
\end{aligned}$$

Fix $0 < \epsilon < 1$. By Lemma 2.8 and Proposition 3.2, we can find $M > 0$ and $c > 0$ such that with probability greater than $1 - \epsilon$

$$x_0 \leq \tau \leq x_0 + Mn^{-1/(2k+1)}$$

and

$$\left| \tilde{g}_n(t) - g_0(x_0) - g_0'(x_0)(t - x_0) - \cdots - \frac{g_0^{(k-1)}(x_0)}{(k-1)!}(t - x_0)^{k-1} \right| \leq cn^{-k/(2k+1)}$$

for $x_0 - Mn^{-1/(2k+1)} \le t \le x_0 + Mn^{-1/(2k+1)}$. On the other hand, using Taylor expansion, we can find $d > 0$ that

$$\begin{aligned} \left| g_0(t) - g_0(x) + g_0'(x_0)(t-x_0) - \cdots - \frac{g_0^{(k-1)}(x_0)}{(k-1)!}(t-x_0)^{k-1} \right| &\le d\,(t-x_0)^k \\ &\le c' n^{-k/(2k+1)} \end{aligned}$$

for $x_0 - Mn^{-1/(2k+1)} \le t \le x_0 + Mn^{-1/(2k+1)}$ and where $c' = dM^k$. It follows that

$$\begin{aligned} \int_{x_0}^{\tau} |\tilde{g}_n(t) - g_0(t)|\, dt &\le (c+c')n^{-k/(2k+1)} \int_{x_0}^{\tau} dt \\ &= (c+c')n^{-k/(2k+1)} \times (\tau - x_0) \\ &\le (c+c')Mn^{-(k+1)/(2k+1)}. \end{aligned}$$

To finish off the proof, we only need to check that

$$\left| \int_{x_0}^{\tau} d(\mathbb{G}_n(t) - G_0(t)) \right| = O_p(n^{-(k+1)/(2k+1)}).$$

But this can be shown using similar arguments to those in the proof of Proposition 3.1. Indeed,

$$\int_{x_0}^{\tau} d(\mathbb{G}_n(t) - G_0(t)) = \int_0^{\infty} 1_{[x_0,\tau]}(t) d(\mathbb{G}_n(t) - G_0(t))$$

is an empirical process indexed by the point $\tau \in [x_0, x_0 + Mn^{-1/(2k+1)}]$.

Consider now the empirical process

$$U_n(y,z) = \int_0^{\infty} 1_{[y,z]}(t) d(\mathbb{G}_n(t) - G_0(t))$$

for $0 < y \le z$ and the class of functions

$$\mathcal{F}_{y,R} = \left\{ f_{y,z} : f_{y,z}(t) = 1_{[y,z]}(t), y \le z \le y + R \right\}$$

for a fixed $y > 0$ and $R > 0$. One can prove that there exist, $\delta > 0$ and $R > 0$ such that

$$|U_n(y,z)| \le \epsilon (z-y)^{k+1} + O_p(n^{-(k+1)/(2k+1)})$$

for all $|y - x_0| \le \delta$, $z \in [y, y+R]$ and for all $\epsilon > 0$. It follows that

$$\begin{aligned} \left| \int_{x_0}^{\tau} d(\mathbb{G}_n(t) - G_0(t)) \right| &= o_p\left((\tau - x_0)^{k+1} \right) + O_p(n^{-(k+1)/(2k+1)}) \\ &= O_p((n^{-(k+1)/(2k+1)}) \end{aligned}$$

and the result follows for $j = k-1$. Note that we obtain the same result if we replace x_0 by any x in an neighborhood of x_0 of the form $]x_0 - Mn^{-1/(2k+1)}, x_0 + Mn^{-1/(2k+1)}]$, for some constant $K > 0$; i.e., we can find $K > 0$ indenpendent of x such that

$$\left|\tilde{\Delta}_n^{(k-1)}(x)\right| \leq Kn^{-(k+2)/(2k+1)}$$

with large probability.

Now, let $j = k-2$. We have,

$$\tilde{\Delta}_n^{(k-2)}(x_0) = \int_0^{x_0} (x_0 - t)d(\tilde{G}_n(t) - \mathbb{G}_n(t)).$$

Let τ be a zero of $\tilde{\Delta}_n^{(k-2)}$ (we can find such a zero the same way as we did for $\tilde{\Delta}_n^{(k-1)}$). We can write

$$\begin{aligned}
\tilde{\Delta}_n^{(k-2)}(x_0) &= \tilde{\Delta}_n^{(k-2)}(x_0) - \tilde{\Delta}_n^{(k-2)}(\tau) \\
&= \int_0^{x_0} (x_0 - t)d(\tilde{G}_n(t) - \mathbb{G}_n(t)) - \int_0^{\tau} (\tau - t)d(\tilde{G}_n(t) - \mathbb{G}_n(t)) \\
&= -\int_{x_0}^{\tau} (x_0 - t)d(\tilde{G}_n(t) - \mathbb{G}_n(t)) - (\tau - x_0)\int_0^{\tau} d(\tilde{G}_n(t) - \mathbb{G}_n(t)) \\
&= -\int_{x_0}^{\tau} (x_0 - t)d(\tilde{G}_n(t) - \mathbb{G}_n(t)) - (\tau - x_0)\tilde{\Delta}_n^{(k-1)}(\tau).
\end{aligned}$$

Let $M > 0$ be such that $x_0 \leq \tau \leq x_0 + Mn^{-1/(2k+1)}$. By the previous result, there exists $c > 0$ such that

$$\left|(\tau - x_0)\tilde{\Delta}_n^{(k-1)}(\tau)\right| \leq cn^{-2/(2k+1)}$$

with large probability.
Now

$$\left|\int_{x_0}^{\tau} (x_0 - t)d(\tilde{G}_n(t) - \mathbb{G}_n(t))\right| \leq \int_{x_0}^{\tau} (t - x_0)|\tilde{g}_n(t) - g_0(t)|dt + \left|\int_{x_0}^{\tau} (t - x_0)d(\mathbb{G}_n(t) - G_0(t))\right|.$$

We can find $d > 0$ such that

$$\left|\tilde{g}_n(t) - g_0(x_0) - g_0'(x_0)(t - x_0) - \cdots - \frac{g_0^{(k-1)}(x_0)}{(k-1)!}(t - x_0)^{k-1}\right| \leq dn^{-k/(2k+1)}$$

and

$$\left|g_0(t) - g_0(x_0) - g_0'(x_0)(t - x_0) - \cdots - \frac{g_0^{(k-1)}(x_0)}{(k-1)!}(t - x_0)^{k-1}\right| \leq dn^{-k/(2k+1)}$$

for all $t \in [x_0 - Mn^{-1/(2k+1)}, x_0 + Mn^{-1/(2k+1)}]$ with large probability. It follows that

$$\begin{aligned}\int_{x_0}^{\tau} (t - x_0)|\tilde{g}_n(t) - g_0(t)|dt &\leq 2d\, n^{-k/(2k+1)} \int_{x_0}^{\tau} (t - x_0)dt \\ &= d\, n^{-k/(2k+1)}(\tau - x_0)^2 \\ &\leq 4dM^2\, n^{-(k+2)/(2k+1)}.\end{aligned}$$

with large probability. Finally, using again empirical processes arguments, we can show that

$$\left| \int_{x_0}^{\tau} (t - x_0)(\mathbb{G}_n(t) - G_0(t)) \right| = O_p(n^{-(k+2)/(2k+1)})$$

and the result follows for $j = k - 2$. The same result holds if we replace x_0 by any $x \in [x_0 - Mn^{-1/(2k+1)}, n^{-1/(2k+1)}, x_0 + Mn^{-1/(2k+1)}]$, for some $M > 0$; i.e., we can find $K > 0$ indenpendent of x such that

$$\left| \tilde{\Delta}_n^{(k-2)}(x) \right| \leq Kn^{-(k+2)/(2k+1)}$$

with large probability.

Now let $0 \leq j \leq k - 3$ and fix $\epsilon > 0$. Suppose that for all $j' > j$ and $M > 0$, there exists $c > 0$ such that for all $z \in [x_0 - Mn^{-1/(2k+1)}, x_0 + Mn^{-1/(2k+1)}]$,

$$(k - 1 - j')!|\tilde{\Delta}_n^{(j')}(z)| \leq cn^{-(2k-j')/(2k+1)}.$$

with probability greater than $1 - \epsilon$. We can write,

$$\begin{aligned}&(k-1-j)!\tilde{\Delta}_n^{(j)}(y) \\ &= \int_0^y (y-t)^{k-1-j} d(\tilde{G}_n(t) - \mathbb{G}_n(t)) \\ &= \int_0^y ((y-x) + (x-t))^{k-1-j}\, d(\tilde{G}_n(t) - \mathbb{G}_n(t)) \\ &= \sum_{l=0}^{k-1-j} \binom{k-1-j}{l} (y-x)^l \int_0^y (x-t)^{k-1-j-l} d(\tilde{G}_n(t) - \mathbb{G}_n(t)) \\ &= \sum_{l=1}^{k-1-j} \binom{k-1-j}{l} (y-x)^l \int_0^y (x-t)^{k-1-j-l} d(\tilde{G}_n(t) - \mathbb{G}_n(t)) \\ &\quad + \int_0^y (x-t)^{k-1-j} d(\tilde{G}_n(t) - \mathbb{G}_n(t)) \\ &= \sum_{l=1}^{k-1-j} \binom{k-1-j}{l} (y-x)^l \tilde{\Delta}_n^{(j+l)}(y) + \tilde{\Delta}_n^{(j)}(x) + \int_x^y (x-t)^{k-1-j} d(\tilde{G}_n(t) - \mathbb{G}_n(t))\end{aligned}$$

Take x to be a zero of $\tilde{\Delta}_n^{(j)}$ (such zero can be constructed using the mean value theorem as we did for $j = k-2$ and $j = k-1$). Thus there exists $M > 0$ such that $x_0 - Mn^{-1/(2k+1)} \le x \le x_0 + Mn^{-1/(2k+1)}$. Now by applying the induction hypothesis, there exists $c > 0$ such that we have for all $y \in [x_0 - Mn^{-1/(2k+1)}, x_0 + Mn^{-1/(2k+1)}]$, we have

$$\begin{aligned}\left|(k-1-j)!\tilde{\Delta}_n^{(j)}(y)\right| &\le c \sum_{l=1}^{k-1-j} \binom{k-1-j}{l} |y-x|^l n^{-(2k-(j+l))/(2k+1)} \\ &\quad + \left|\int_x^y (x-t)^{k-1-j} d(\tilde{G}_n(t) - \mathbb{G}_n(t))\right|.\end{aligned}$$

But,

$$\sum_{l=1}^{k-1-j} \binom{k-1-j}{l} |y-x|^l n^{-(2k-(j+l))/(2k+1)} \le \left(\sum_{l=1}^{k-1-j} \binom{k-1-j}{l} (2M)^l\right) n^{-(2k-j)/(2k+1)}$$

and

$$\left|\int_x^y (x-t)^{k-1-j} d(\tilde{G}_n(t) - \mathbb{G}_n(t))\right| = O_p(n^{-(2k-j)/(2k+1)})$$

by using empirical processes arguments. Therefore, the result holds for j and hence for all $j = 0, \cdots, k-1$. ■

Lemma 3.4 *Let $0 \le j \le 2k-1$ and $c > 0$. Let $\bar{H}_n^l$ denote either $\widehat{H}_n^l$ or $\tilde{H}_n^l$. Then*

$$(\bar{H}_n^l)^{(j)} \Rightarrow H_k^{(j)}$$

in $D[-c,c]$ for $j = 0, \cdots, 2k-1$ and where H_k is the stochastic process defined in Theorem 3.1.

Proof. The arguments are very similar to the ones used in Groeneboom, Jongbloed and Wellner (Groeneboom, Jongbloed, and Wellner (2001b)). We prove the lemma for $\tilde{H}_n^l$ as the arguments are similar for $\widehat{H}_n^l$. Let $c > 0$. On $[-c,c]$, define the vector-valued stochastic process

$$Z_n(t) = \left(\tilde{H}_n^l(t), \cdots, (\tilde{H}_n^l)^{(2k-2)}(t), \mathbb{Y}_n^l(t), \cdots, (\mathbb{Y}_n^l)^{(k-2)}(t), (\tilde{H}_n^l)^{(2k-1)}(t), (\mathbb{Y}_n^l)^{(k-1)}(t)\right).$$

This stochastic process belongs to the space

$$E_k[-c,c] = (C[-c,c])^{3k-2} \times (D[-c,c])^2$$

where $C[-c,c]$ and $D[-c,c]$ are respectively the space of continuous and right-continuous functions on $[-c,c]$. We endow the space $E_k[-c,c]$ with the product topology induced by the uniform topology on $C[-c,c]$ and the Skorohod topology on $D[-c,c]$.

By Lemma 3.3, we know that $(\tilde{H}_n^l)^{(j)}$ is tight in $C[-c,c]$ for $j=0,\cdots,2k-2$. It follows from the same lemma together with the monotonicity of $(\tilde{H}_n^l)^{(2k-1)}$ that the latter is tight in $D[-c,c]$. On the other hand, since the processes $\left(\mathbb{Y}_n^l,\cdots,(\mathbb{Y}_n^l)^{(k-2)}\right)$ and $(\mathbb{Y}_n^l)^{(k-1)}$ converge weakly, they are tight in $(C[-c,c])^{k-1}$ and $D[-c,c]$ respectively. Now, for a fixed $\epsilon>0$, there exists an $M>0$ such that with probability greater than $1-\epsilon$, the process Z_n belongs to $E_{k,M}[-c,c]$ where $E_{k,M}=(C_M[-c,c])^{3k-2}\times(D_M[-c,c])^2$, and $C_M[-c,c]$ and $D_M[-c,c]$ are respectively the subset of functions in $C[-c,c]$ and the subset of monotone functions in $D[-c,c]$ that are bounded by M. Since the subspace $E_{k,M}[-c,c]$ is compact, we can extract from any arbitrary sequence $\{Z_{n'}\}$ a further subsequence $\{Z_{n''}\}$ that is weakly converging to some process

$$Z_0=\left(H_0,\cdots,H_0^{(2k-1)},Y_0,\cdots,Y_0^{(k-2)},H_0^{(2k-1)},Y_0^{(k-1)}\right) \tag{3.20}$$

in $E_k[-c,c]$ and where $Y_0=Y_k$.

Now, consider the functions ϕ_1 and $\phi_2 : E_k[-c,c]\mapsto\mathbb{R}$ defined by

$$\phi_1(z_1,\cdots,z_{3k})=\inf_{t\in[-c,c]}(z_1(t)-z_{2k}(t))\wedge 0$$

and

$$\phi_2(z_1,\cdots,z_{3k})=\int_{-c}^{c}(z_1(t)-z_{2k}(t))dz_{3k-1}(t).$$

It is easy to check that the functions ϕ_1 and ϕ_2 are both continuous. By the continuous mapping theorem, it follows that $\phi_1(Z_0)=\phi_2(Z_0)=0$ since $\phi_1(Z_{n''})=\phi_2(Z_{n''})=0$ and therefore,

$$H_0(t)\geq Y_k(t),$$

for all $t\in[-c,c]$ and

$$\int_{-c}^{c}(H_0(t)-Y_k(t))dH_0^{(2k-1)}(t)=0.$$

It is easy to see check that $(-1)^kH_0^{(2k-2)}$ is convex. Since $c>0$ is arbitrary, we see that H_0 satisfies conditions (i) and (iii) of Theorem 3.1. Furthermore, outside the interval $[-c,c]$ we can take $\tilde{H}_n^l$ and $\mathbb{Y}_n^l$ to be identically 0. With this choice, the condition (iv) of Theorem 3.1 is satisfied. By uniqueness of the process H_k, it follows that $H_0=H_k$. Since the limit is the same for any subsequence $\{Z_{n_l}\}$, we conclude that the sequence $\{Z_n\}$ converges weakly to

$$Z_k=\left(H_k,\cdots,H_k^{(2k-1)},Y_k,\cdots,Y_k^{(k-2)},H_k^{(2k-1)},Y_k^{(k-1)}\right)$$

and in particular $Z_n(0) \to_d Z_k(0)$ and $(\tilde{H}_n^l)^{(j)}(0) \to_d H_k^{(j)}(0)$ for $j = 0, \cdots, 2k-1$. ■

Proof of Theorem 3.2. For the direct problems, we apply Lemma 3.4 at $t = 0$ together with the fact that for $j = 0, \cdots, k-1$,

$$(\tilde{H}_n^l)^{k+j}(0) = c_j(g_0) n^{(k-j)/(2k+1)} (\tilde{g}_n(x_0) - g_0(x_0))$$

and

$$(\widehat{H}_n^l)^{k+j}(0) - c_j(g_0) n^{(k-j)/(2k+1)} (\hat{g}_n(x_0) - g_0(x_0)) \to_p 0 \quad \text{as } n \to \infty$$

which follow from the respective definitions of $\tilde{H}_n^l$ and $\widehat{H}_n^l$, and also strong consistency of the MLE (for $\widehat{H}_n^l$). For the inverse problem, the claim follows from Lemma 3.4 and the inverse formula in Balabdaoui and Wellner (2004a), Lemma 2.2. ■

Acknowledgements: We gratefully acknowledge helpful conversations with Carl de Boor, Nira Dyn, Tilmann Gneiting, and Piet Groeneboom.

References

Balabdaoui, F. and Wellner, J. A. (2004a). Estimation of a k-monotone density, part 1: characterizations, consistency, and minimax lower bounds. Technical Report 459, Department of Statistics, University of Washington. Available at: *http://www.stat.washington.edu/www/research/reports/2004/*.

Balabdaoui, F. and Wellner, J. A. (2004b). Estimation of a k-monotone density, part 2: algorithms for computation and numerical results. Technical Report 460, Department of Statistics, University of Washington. Available at: *http://www.stat.washington.edu/www/research/reports/2004/*.

Balabdaoui, F. and Wellner, J. A. (2004c). Estimation of a k-monotone density, part 3: limiting Gaussian versions of the problem; invelopes and envelopes. Technical Report 461, Department of Statistics, University of Washington. Available at: *http://www.stat.washington.edu/www/research/reports/2004/*.

Bojanov, B. D., Hakopian, H. A. and Sahakian, A. A. (1993). *Spline Functions and Multivariate Interpolations*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

de Boor, C. (1973). The quasi-interpolant as a tool in elementary polynomial spline theory. *In Approximation Theory* (Austin, TX, 1973), 269-276. Academic Press, New York.

de Boor, C. (1974). Bounding the error in spline interpolation. *SIAM Rev.* **16**, 531 - 544.

de Boor, C. (2004). *http://www.cs.wisc.edu/ deboor/toast/pages09.html.*

de Boor, C. and Fix G. J. (1973). Spline approximation by quasi-interpolants. *J. Approx. Theory* **8**, 19 - 45.

DeVore, R. A. and Lorentz, G. G. (1993). *Constructive Approximation*. Springer-Verlag, Berlin.

Feller, W. (1971) *An Introduction to Probability Theory and Its Applications*. Vol. 2, 2nd ed. Wiley, New York.

Gneiting, T. (1998). On the Bernstein-Hausdorff-Widder conditions for completely monotone functions. *Exposition. Math.* **16**, 181 - 183.

Gneiting, T. (1999). Radial positive definite functions generated by Euclid's hat. *J. Multivariate Analysis* **69**, 88 - 119.

Groeneboom, P. (1985). Estimating a monotone density. *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer,* Vol. II. Lucien M. LeCam and Richard A. Olshen eds. Wadsworth, New York. 529 - 555.

Groeneboom, P., Jongbloed, G., and Wellner, J. A. (2001a). A canonical process for estimation of convex functions: The "invelope" of integrated Brownian motion $+t^4$. *Ann. Statist.* **29**, 1620 - 1652.

Groeneboom, P., Jongbloed, G., and Wellner, J. A. (2001b). Estimation of convex functions: characterizations and asymptotic theory. *Ann. Statist.* **29**, 1653 - 1698.

Jewell, N. P. (1982). Mixtures of exponential distributions. *Ann. Statist.* **10**, 479 - 484.

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

Komlós, J., Major, P., and Tusnády, G. (1975). An approximation of partial sums of independent rv's and the sample disttribution function. *Z. Wahrsch. verw. Geb.* **32**, 111 - 131.

Lévy, P. (1962). Extensions d'un théorème de D. Dugué et M. Girault. *Z. Wahrscheinlichkeitstheorie* **1**, 159 - 173.

Mammen, E. and van de Geer, S. (1997). Locally adaptive regression splines. *Ann. Statist.* **19**, 387 - 413.

Michelli, C. (1972). The fundamental theorem of algebra for monosplines with multiplicities. In *Linear Operators and Approximation*, 419 - 430. Birkhäuser, Basel.

Nürnberger, G. (1989). *Approximation by Spline Functions*. Springer-Verlag, New York.

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

Schoenberg, I. J. (1938). Metric spaces and completely monotone functions. *Ann. of Math.* **39**, 811 - 841.

Schoenberg, I. J. (1963). On interpolation by spline functions and its minimal property. In *Proceedings at the Conference held at the Mathematical Research*

Institute at Oberwolfach, August 4-10, 1963. Birkhäuser Verlag Basel, 109-128.

Shadrin, A. Yu. (2001). The L_∞-norm of the L_2-spline projector is bounded independently of the knot sequence: A proof a de Boor's conjecture. *Acta Math.* **187**, 59-137.

Shorack, G. and Wellner J. A. (1986). *Empirical Processes with Applications to Statistics*. John Wiley and Sons, New York.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer-Verlag, New York.

Widder, D. V. (1941). *The Laplace Transform.* Princeton University Press, Princeton.

Williamson, R. E. (1956). Multiply monotone functions and their Laplace transforms. *Duke Math. J.* **23**, 189 - 207.