CentER

Discussion Paper

No. 2008–24

THE SHORTH PLOT

By John H.J. Einmahl, Maria Gantner, Günther Sawitzki

February 2008

ISSN 0924-7815

The shorth plot*

John H.J. Einmahl Maria Gantner Günther Sawitzki

February 19, 2008

Abstract: The shorth plot is a tool to investigate probability mass concentration. It is a graphical representation of the length of the shorth, the shortest interval covering a certain fraction of the distribution, localized by forcing the intervals considered to contain a given point x. It is easy to compute, avoids bandwidth selection problems and allows scanning for local as well as for global features of the probability distribution. We prove functional central limit theorems for the empirical shorth plot. The good rate of convergence of the empirical shorth plot makes it useful already for moderate sample size.

JEL codes: C13, C14.
Key words: Data analysis, distribution diagnostics, functional central limit theorem, probability mass concentration.

1 Introduction

Using exploratory diagnostics is one of the first steps in data analysis. Here graphical displays are essential tools. For detecting specific features, specialized displays may be available. For example, if there is a model distribution F to be compared with, from a mathematical point of view the empirical distribution F_n is a key instrument, and its

*John Einmahl is Professor of Statistics, Department of Econometrics & OR, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands (E-mail: *j.h.j.einmahl@uvt.nl*). Maria Gantner is Ph.D. Candidate of Statistics, Department of Econometrics & OR, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands (E-mail: *m.gantner@uvt.nl*). Günther Sawitzki is Academic Director, StatLab, Institute for Applied Mathematics, Im Neuenheimer Feld 294, D 69120 Heidelberg, Germany (E-mail: *gs@statlab.uni-heidelberg.de*).

graphical representations by means of PP-plots

$$x \mapsto (F(x), F_n(x))$$

or QQ-plots

$$\alpha \mapsto \left(F^{-1}(\alpha), F_n^{-1}(\alpha)\right)$$

are tools of first choice. If we consider the overall scale and location, box & whisker plots are a valuable tool. The limitation of box & whisker plots is that they give a global view which ignores any local structure. In particular, they are not an appropriate tool if it comes to analyze the modality of a distribution. More specialized tools are needed in this case, such as the silhouette and the excess density plot, both tools being introduced in Müller and Sawitzki (1991).

While we have some instruments for specific tasks, the situation is not satisfactory if it comes to general purpose tools. PP-plots and QQ-plots need considerable training to be used as diagnostic tools, as they do not highlight qualitative features.

Focussing on the density in contrast to the distribution function leads to density estimators and their visual representations, such as histograms and kernel density plots. These, however, introduce another complexity, such as the choice of cut points or bandwidth choice. The qualitative features revealed or suggested by density estimation based methods may critically depend on bandwidth choice. Moreover, estimating a density is a more specific task than understanding the shape of a density. Density estimation based methods are prone to pay for these initial steps in terms of slow convergence or large fluctuation, or disputable choices of smoothing.

We will use the length of the shorth to analyze the qualitative shape of a distribution. Originally, the shorth is the shortest interval containing half of the distribution; more generally, the α-shorth is the the shortest interval containing fraction α of the distribution. The shorth was introduced in the Princeton robustness study as a candidate for a robust location estimator, using the mean of a shorth as an estimator for a mode, see Andrews et al. (1972). As a location estimator, it performs poorly; it has an asymptotic rate of only $n^{-1/3}$, with non-trivial limiting distribution, see Andrews at al. (1972), p. 50, or Shorack and Wellner (1986), p. 767. Moreover, the shorth interval is not well defined, since there may be several competing intervals. The length of the shorth however is a functional which is easy to estimate and it gives a graphical representation which is easy to interpret. As pointed out in Grübel (1988), the length of the shorth has a convergence rate of $n^{-\frac{1}{2}}$ with a Gaussian limit. The critical conditions for Gaussianity are that the shorth interval is sufficiently pronounced, essentially this means that the shorth interval must not be in a flat part of the density, see Section 3.3 in Grübel (1988). In Einmahl and Mason (1992) it was shown that the good convergence of $n^{-\frac{1}{2}}$ is retained under much weaker conditions, including flat-part densities, but that the limit can be non-Gaussian. We extend the definition of the length of the shorth to supply localization. We will vary the coverage, and hence allow for multi-scale analysis. Thus the global estimator is extended into a tool for local and global diagnostics.

The paper is organized as follows. In the remainder of this section we present the definition and elementary properties of the localized length of the shorth. In Section 2 we define the shorth plot, the central object of this paper. Asymptotic results for the empirical shorth plot are presented in Section 3. It will be shown that the rate of convergence of the localized empirical length of the shorth to the theoretical length is $n^{-\frac{1}{2}}$, uniformly in α and the point of localization. In Section 4 we study some real data examples. The paper is completed by a discussion section and a section containing the proofs of the results from Section 3.

In order to be more explicit we specify our setup and notation. Let $X_1, \ldots, X_n$, $n \geq 1$, be independent random variables with common distribution function F. Let P be the probability measure corresponding to F. Let $\mathbb{I} = \{[a, b] : -\infty < a < b < \infty\}$ be the class of closed intervals and let $\mathbb{I}_x = \{[a, b] : -\infty < a < b < \infty,\ x \in [a, b]\}$ be the class of closed intervals that contain $x \in \mathbb{R}$. Define the empirical measure P_n on the Borel sets $\mathbb{B}$ on $\mathbb{R}$ by

$$P_n(B) = \frac{1}{n} \sum_{i=1}^{n} 1_B(X_i), \qquad B \in \mathbb{B},$$

where 1_B denotes the indicator function. Let $|\cdot|$ denote Lebesgue measure.

Definition 1 *The length of the shorth at point $x \in \mathbb{R}$ for coverage level $\alpha \in (0, 1)$ is*

$$S_\alpha(x) = \inf\{|I| : P(I) \geq \alpha, I \in \mathbb{I}_x\}.$$

We get the length of the shorth as originally defined by taking $\inf_x S_{0.5}(x)$. The definition in terms of a theoretical probability P has an immediate empirical counterpart, the empirical length of the shorth

$$S_{n,\alpha}(x) = \inf\{|I| : P_n(I) \geq \alpha, I \in \mathbb{I}_x\}.$$

To get a picture of the optimization problem behind the length of the shorth, we consider the bivariate function

$$(a,b) \longmapsto \big(|I|\,, P\,(I)\big) \qquad \text{with } I = [a,b], \text{ where } a < b.$$

This is defined on the half space $\{(a,b) : a < b\}$ above the diagonal. The level curves

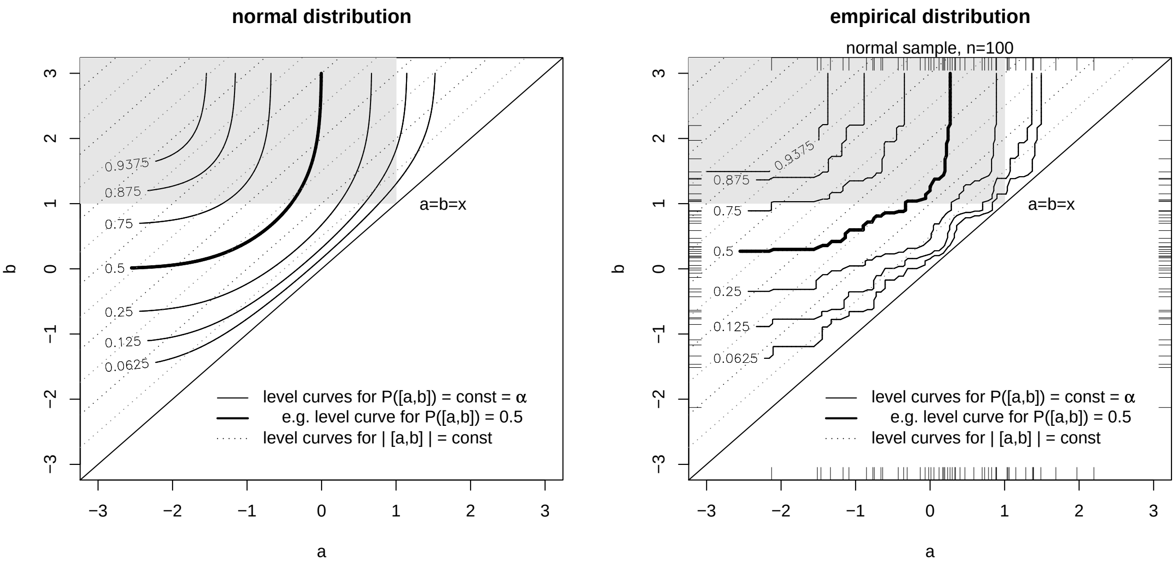

Figure 1: The length of the shorth as an optimization problem: minimize $|[a,b]|$ under the restriction $P([a,b]) \geq \alpha$. Localizing at x restricts the optimization to the quadrant top left of (x,x).

of $|I|$ are parallel to the diagonal. The level curves of $P(I)$ depend on the distribution. The α-shorth minimizes $|I|$ in the area above the level curve at level α, i.e. $P(I) \geq \alpha$. Going to the empirical version replaces the level curves of $P(I)$ by those of of $P_n(I)$. The theoretical curves for the Gaussian distribution and for a Gaussian sample are shown in Figure 1. Localizing the α-shorth at a point x restricts optimization to the (grey) top left quadrant anchored at (x,x).

Let the distribution function F be absolutely continuous with density f. Assume there exist $-\infty \leq x_* < x^* \leq \infty$ such that $f(x) > 0$ on $\mathcal{S} = (x_*, x^*)$ and $f(x) = 0$ outside $\mathcal{S}$; also assume that f is uniformly continuous on $\mathcal{S}$. As a consequence we have that F is strictly increasing on $\mathcal{S}$ and that f is bounded.

We have the following elementary properties concerning $S_\alpha(x)$.

- *Minimizing intervals:* For every α and x, there exists an interval I with length $S_\alpha(x)$

such that $x \in I$ and $P(I) = \alpha$.
• *Continuity:* For all α, $|S_\alpha(x) - S_\alpha(y)| \le |x - y|$. Moreover, the function

$$(x, \alpha) \mapsto S_\alpha(x)$$

is continuous as a function of two variables.
• *Monotonicity:* For all x,

$$\alpha \mapsto S_\alpha(x)$$

is strictly increasing in α.
• *Invariance:* For all α,

$$x \mapsto S_\alpha(x)$$

is invariant under shift transformations and equivariant under scale transformations, that is when we apply a transformation $u' = cu + d$ (for some constants $c > 0, d$), then the new $S'_\alpha(x')$ satisfies

$$S'_\alpha(x') = cS_\alpha(x),$$

with $x' = cx + d$.

Denote the j-th order statistic by $X_{(j)}$; $X_{(0)} = -\infty, X_{(n+1)} = \infty$. For computing the empirical length of the shorth, observe that $S_{n,\alpha}(x)$ can be interpolated from $S_{n,\alpha}(X_{(j)})$ and $S_{n,\alpha}(X_{(j+1)})$ where j is such that $X_{(j)} \le x < X_{(j+1)}$. Therefore we can focus on computing $S_{n,\alpha}(X_i)$. Write $k_\alpha = \lceil n\alpha \rceil - 1$, with $\lceil \cdot \rceil$ the ceiling function. Then we simply have

$$S_{n,\alpha}(X_i) = \min\{X_{(j+k_\alpha)} - X_{(j)} :\ 1 \le j \le i \le j + k_\alpha \le n\}.$$

Using a stepwise algorithm, a further reduction of complexity is possible since we can easily relate $S_{n,\alpha}(X_i)$ to $S_{n,\alpha}(X_{i-1})$. This yields an algorithm with linear complexity in n.

2 The Shorth Plot

Definition 2 (Sawitzki, 1994) *The shorth plot is the graph of the function*

$$x \mapsto S_\alpha(x),\ x \in \mathbb{R}$$

for (all or) a selection of coverages α.

The empirical shorth plot is the graph of

$$x \mapsto S_{n,\alpha}(x),\ x \in \mathbb{R}.$$

Mass concentration now can be represented by the graphs of $x \mapsto S_\alpha(x)$ and $x \mapsto S_{n,\alpha}(x)$, see Figure 2. A *small* length of the shorth signals probability mass concentration, whereas

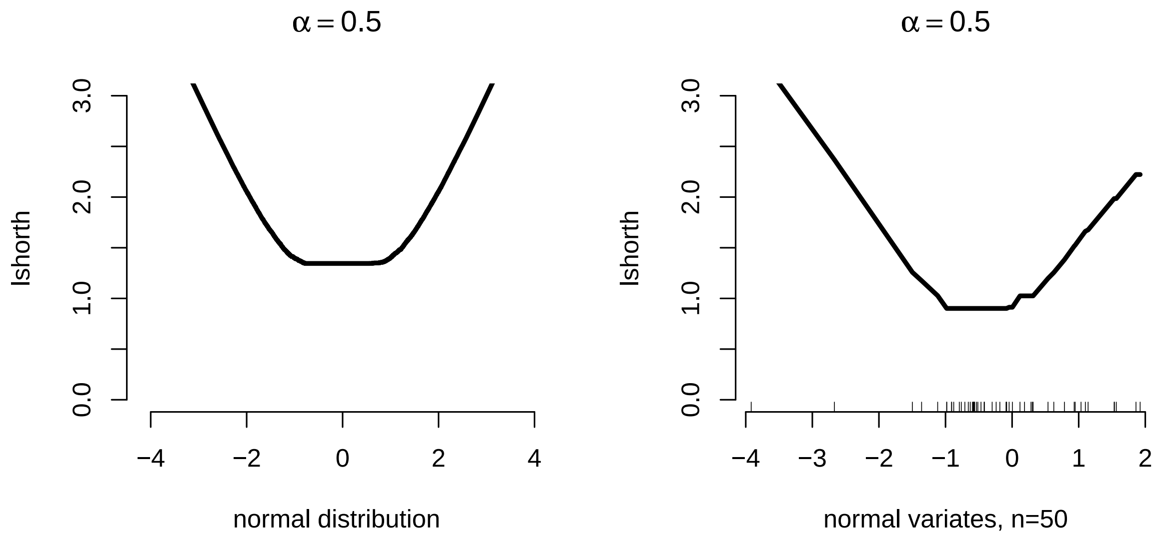

Figure 2: Short plot and empirical shorth plot for a sample of 50 standard normal random variables for $\alpha = 0.5$. Note that different scales are used.

large values of the density indicate mass concentration. To make the interpretation of the shorth plot easier, we will in the sequel use a downward orientation of the vertical axis so that it is aligned with the density plot.

Figure 3 shows the shorth plots for a uniform, a normal and a log-normal distribution for sample sizes 50 and 200 and the theoretical ones. Varying the coverage level α gives a good impression of the mass concentration. Small levels give information about the local behavior, in particular near modes. Higher levels give information about skewness of the overall distribution shape. The high coverage levels show the range of the distribution. A "dyadic" scale for α, e.g., 0.125, 0.25, 0.5, 0.75, 0.875 is a recommended choice. The Monotonicity property (Section 1) allows the multiple scales to be displayed simultaneously without overlaps, thus giving a multi-resolution image of the distribution.

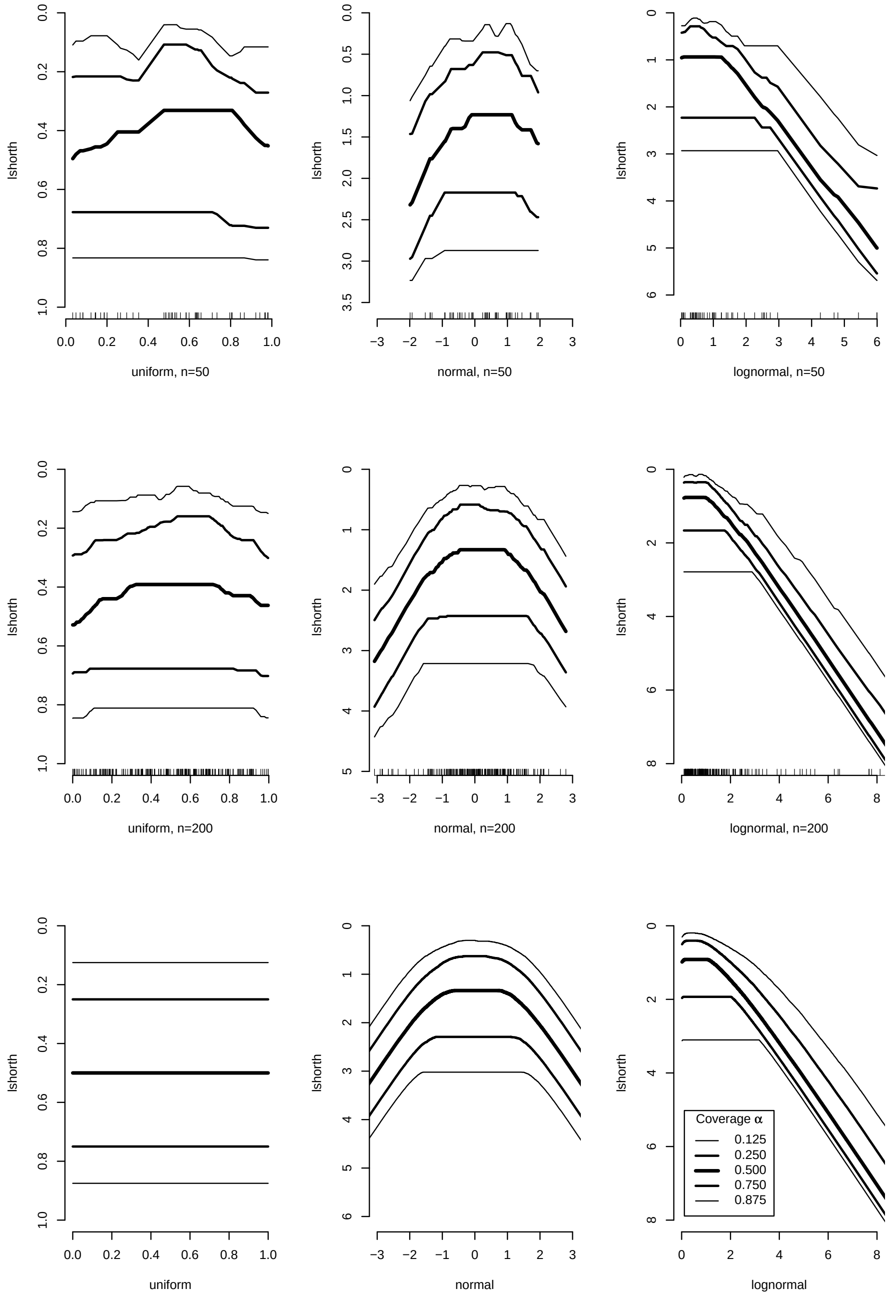

Figure 3: Shorth plots for a uniform, a normal, and a log-normal distribution for sample sizes 50 and 200 and the theoretical ones, for various coverage levels α. Note that different scales are used.

3 Asymptotic Results

In this section we consider the asymptotic behavior of the empirical shorth plot; recall the notation and assumptions of Section 1. For asymptotic analysis, it is more convenient to view S as a process in α, and therefore we write in this section and the proofs section, without confusion, $S_x(\alpha)$ instead of $S_\alpha(x)$, and so on. Let $\mathbb{I}^* = \mathbb{I} \cup \{\mathbb{R}, \emptyset\}$. Define for each $n \geq 1$ the empirical process indexed by intervals to be

$$U_n(I) = n^{\frac{1}{2}}\{P_n(I) - P(I)\}, \quad I \in \mathbb{I}^*.$$

Introduce the pseudometric d_0 defined on $\mathbb{B}$ by

$$d_0(B_1, B_2) = P(B_1 \triangle B_2), \quad \text{for } B_1, B_2 \in \mathbb{B},$$

with $B_1 \triangle B_2 = (B_1 \setminus B_2) \cup (B_2 \setminus B_1)$. Let B_P be a bounded, mean zero Gaussian process indexed by $\mathbb{I}^*$, uniformly continuous in d_0, with covariance function $P(A_1 \cap A_2) - P(A_1)P(A_2)$, $A_1, A_2 \in \mathbb{I}^*$. Then, by the functional central limit theorem and the Skorohod representation theorem, there exist $\tilde{B}_P \overset{d}{=} B_P$ and a sequence $\tilde{U}_n \overset{d}{=} U_n$ such that

$$\text{(1)} \qquad \sup\{|\tilde{U}_n(I) - \tilde{B}_P(I)| : I \in \mathbb{I}^*\} \to 0 \quad \text{a.s. as } n \to \infty.$$

Henceforth we will drop the tildes from the notation.

We will need the following assumption:

(A) *There exist $x_1, x_2 \in [x_*, x^*]$, $x_1 \leq x_2$, such that f is strictly increasing on $(x_*, x_1]$, constant on $[x_1, x_2]$, and strictly decreasing on $[x_2, x^*)$; also $\lim_{x \downarrow x_*} f(x) = \lim_{x \uparrow x^*} f(x)$.*

We introduce some more notation. Write $g_x = 1/S'_x$; see Chapter 6, Lemma 3 for the existence of g_x. Set $T_{x,0} = \{\emptyset\}$, $T_{x,1} = \{\mathbb{R}\}$, $x \in \mathbb{R}$ and

$$T_{x,\alpha} = \{I \in \mathbb{I}_x : |I| = S_x(\alpha),\ P(I) = \alpha\} \quad \text{for } 0 < \alpha < 1,\ x \in \mathbb{R}.$$

For any $x \in \mathbb{R}$ let

$$B_x(\alpha) = \sup\{B_P(I) : I \in T_{x,\alpha}\}, \quad 0 \leq \alpha \leq 1.$$

Consider the *shorth plot process*:

$$Q_{n,x}(\alpha) = g_x(\alpha) n^{\frac{1}{2}} \left(S_{n,x}(\alpha) - S_x(\alpha)\right), \quad 0 < \alpha < 1,\ x \in \mathbb{R}.$$

Theorem 1 *Under the assumptions on F of Section 1 and assumption (A) we have for all $0 < \eta < 1/2$, on the probability space of (1),*

$$\sup_{x \in \mathbb{R}} \sup_{\eta \leq \alpha \leq 1-\eta} |Q_{n,x}(\alpha) + B_x(\alpha)| \to 0 \quad a.s. \quad as \;\; n \to \infty. \tag{2}$$

We note that if $T_{x,\alpha}$ contains at least two sets I_1 and I_2 with $P(I_1 \vartriangle I_2) > 0$, then $B_x(\alpha)$ is not a normal random variable and $\mathbb{E}B_x(\alpha) > 0$. Since B_P is bounded, $B_{\cdot}(\cdot)$ is a bounded process on $\mathbb{R} \times [0,1]$ with $B_x(0) = B_x(1) = 0$ almost surely. It will be shown (Section 6) that $(B_x)_{x \in \mathbb{R}}$ is a collection of uniformly equicontinuous functions on $[0,1]$.

We will need two additional assumptions in order to extend the convergence on $[\eta, 1-\eta]$ in (2) to convergence on the entire interval $(0,1)$. The first one is the classical Csörgő and Révész (1978) condition.

(B) *If $\lim_{x \downarrow x_*} f(x) = 0$, then f' exists on $\mathcal{S}$ and for some $0 < M < \infty$*

$$\sup_{x \in \mathcal{S}} F(x)\,(1 - F(x))\,\frac{|f'(x)|}{f^2(x)} < M.$$

For the second assumption, let $I_\alpha = [a,b]$ be a shortest interval such that $P(I_\alpha) = \alpha$, $\alpha \in (0,1)$; note that $f(a) = f(b)$. If $\lim_{x \downarrow x_*} f(x) = 0$, then for large enough α, I_α is unique. For such an α define $\lambda_\alpha = F(a)/(1-\alpha)$.

(C) *If $\lim_{x \downarrow x_*} f(x) = 0$, then $0 < \lim_{\alpha \uparrow 1} \lambda_\alpha < 1$.*

Theorem 2 *In addition to the assumptions of Theorem 1, suppose (B) and (C) hold. Then on the probability space of (1),*

$$\sup_{x \in \mathbb{R}} \sup_{0 < \alpha < 1} |Q_{n,x}(\alpha) + B_x(\alpha)| \xrightarrow{\mathbb{P}} 0 \quad as \;\; n \to \infty.$$

It readily follows that for every interval $I \in \mathbb{I}$, there exists an $x \in \mathbb{R}$ such that $I \in T_{x,P(I)}$. This implies that

$$\sup_{x \in \mathbb{R}} \sup_{0 \leq \alpha \leq 1} |B_x(\alpha)| \;=\; \sup_{I \in \mathbb{I}} |B_P(I)| \stackrel{d}{=} \sup_{0 \leq \alpha, \beta \leq 1} |B(\alpha) - B(\beta)|, \tag{3}$$

with B a standard Brownian bridge. The right-hand side of (3) is the limiting distribution of the Kuiper statistic, see e.g. Shorack and Wellner (1986), p. 144.

4 Examples

While the theory presented above covers the general unimodal situation, the following examples focus on the use of the shorth plot as an exploratory tool in multimodal situations.

4.1 Old Faithful Geyser

As a first example, we use the eruption durations of the Old Faithful geyser. The data are just one component of a bivariate time series data set. Looking at a one dimensional marginal distribution ignores the process structure. However, these data have been used repeatedly to illustrate smoothing algorithms like kernel density estimators (Figure 4, left) and we reuse it to illustrate our approach (Figure 4, right). This is a good natured data set showing two distinct nodes with sizeable observation counts, and some overall skewness. The high coverage levels of the shorth plot ($\alpha > 50\%$) just show the overall range of the

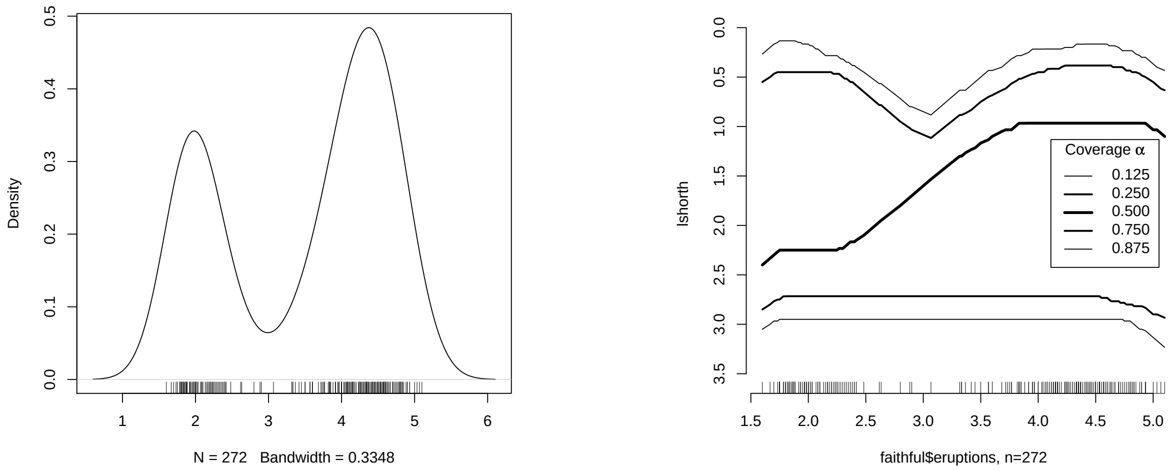

Figure 4: (left) Eruption durations of the Old Faithful geyser: density estimation (right) Eruption durations of the Old Faithful geyser: shorth plot.

data. The 50% level indicates a pronounced skewness. The small levels reveal that we have two modes, with a comparable coverage range. The multi-scale property of the shorth plot allows to combine these aspects in one picture.

4.2 Melbourne Temperature Data

In Hyndman et al. (1996) the bifurcation to bimodality in the Melbourne temperature data set is pointed out. We use an extended version of the data set (Melbourne temperature data 1955-2007, provided by the Bureau of Meteorology, Victorian Climate Services Centre, Melbourne) and analyze the day by day difference in temperature at 15.00h, conditioned on today's temperature. The shorth plot view is in Figure 5. We indeed clearly see the bimodality (and some skewness) when conditioning on high temperatures and the unimodality when conditioning on the lower temperatures.

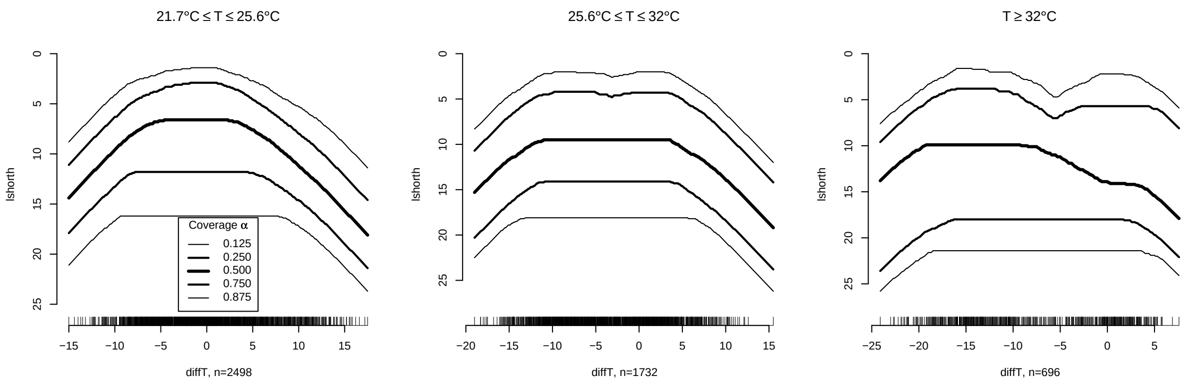

Figure 5: Melbourne day by day temperature difference at 15:00h conditioned at today's temperature.

5 Discussion

The α-shorth is a well-defined concept. Its length is studied in detail in Grübel (1988). The shorth can be extended to higher dimensions by replacing the class of intervals by a class of sets (e.g., all ellipsoids) and length by "volume"; see Einmahl and Mason (1992) for the asymptotic behavior of these minimal volumes. In higher dimensions, however, there is no canonical class of sets, like the intervals in dimension one. It is an open question whether the shorth plot can be carried over to a regression context.

The shorth plot was introduced in Sawitzki (1994), but no further analysis or theory was provided. A closely related idea is the balloonogram in Tukey and Tukey (1981). Their multivariate procedure reduces in dimension one to considering the shortest interval, centered at a data point, that contains a certain number of data. In contrast to the balloonogram, the shorth plot avoids centering, thus reducing random fluctuation. No theory is provided, however, and also only one coverage level is used at a time.

The shorth plot is based on the concept of mass concentration, an idea which is shared with the excess density plot and the silhouette plot (Müller and Sawitzki, 1991). Excess density and silhouette plots are designed to detect the modes of a density. They use a global approach: there is no localization in x, like in the shorth plot. In Hyndman (1996), so-called highest density regions boxplots are introduced. These boxplots use mass concentration in a regression context.

Kernel density estimators with varying bandwidths are widely studied and somewhat related to our approach. The coverage α of the shorth plot bears some similarity with the bandwidth chosen for kernel estimation. The SiZer (Chaudhuri and Marron, 1999) is a kernel-based approach which studies simultaneously a wide range of bandwidths. Another approach that combines kernel estimation explicitly with detecting modes is that of the mode trees (Minnotte and Scott, 1993). Here the mode locations are plotted against the bandwidth of the density estimator with those modes. Mass concentration is a local concept, but not, like a density, an infinitesimal concept. Therefore the shorth plot avoids the smoothing step and can be based directly on the empirical measure.

6 Proofs

The proof of Theorem 1 is based on a number of lemmas and a proposition, which we will state and prove below. In Lemmas 1 and 2 we present certain extensions of the conditions (C_6) and (C_8) in Einmahl and Mason (1992), respectively.

Lemma 1 *For every $\varepsilon > 0$, whenever $x \in \mathbb{R}$, $0 \le \alpha_1, \alpha_2 \le 1$ with $|\alpha_1 - \alpha_2| < \varepsilon$ and $I_1 \in T_{x,\alpha_1}$, there is an $I_2 \in T_{x,\alpha_2}$ with $d_0(I_1, I_2) < \varepsilon$.*

Proof: Write $\beta = P([x_1, x_2])$. Observe that $T_{x,\alpha}$ contains infinitely many intervals if and only if $x \in (x_1, x_2)$ and $\alpha < \beta$, otherwise it contains exactly one interval. From this it follows that for $\alpha_1 < \alpha_2$, an I_2 as in the lemma can be found with $I_2 \supset I_1$. Similarly, for $\alpha_1 > \alpha_2$, we can take $I_2 \subset I_1$. □

Lemma 2 *For every $\varepsilon > 0$ there exists a $\delta > 0$ such that whenever $I \in \mathbb{I}_x$, $x \in \mathbb{R}$, satisfies $0 < \alpha - \delta < P(I) < \alpha < 1$ and $|I| < S_x(\alpha)$, there is an $I' \in T_{x,S_x^{-1}(|I|)}$ such that $d_0(I, I') < \varepsilon$.*

Proof: The proof can be given along the same lines as in Example 2 in Einmahl and Mason (1992). We will omit details. $\square$

Let J be an open or closed interval and consider a function $f : J \to \mathbb{R}$; let $\delta > 0$. The modulus of continuity of f is defined by

$$\omega(f, \delta) = \sup\{|f(u) - f(v)| : u, v \in J, |u - v| \leq \delta\}.$$

Lemma 3 *For all $x \in \mathbb{R}$, g_x exists on $(0, 1)$ and is positive. Moreover, $(g_x)_{x\in\mathbb{R}}$ is a collection of uniformly equicontinuous functions on $(0, 1)$, i.e.*

$$\lim_{\delta \downarrow 0} \sup_{x\in\mathbb{R}} \omega(g_x, \delta) = 0 \ \ a.s.$$

In addition, for any $0 < \varepsilon < \frac{1}{2}$,

$$\inf_{x\in\mathbb{R}} \inf_{\varepsilon \leq \alpha \leq 1-\varepsilon} g_x(\alpha) > 0.$$

Proof: Let $I_{x,\alpha} \in T_{x,\alpha}$. If x is on the boundary of $I_{x,\alpha}$, then $g_x(\alpha) = f(y)$, where y is the other endpoint of the interval $I_{x,\alpha}$. If x is not on the boundary of $I_{x,\alpha} =: [y_1, y_2]$, then, since $\lim_{x\downarrow x_*} f(x) = \lim_{x\uparrow x_*} f(x)$, $g_x(\alpha) = f(y_1) = f(y_2)$. Hence we obtain

$$\sup_{x\in\mathbb{R}} \sup_{\substack{|\alpha-\beta|\leq\delta \\ 0<\alpha,\beta<1}} |g_x(\alpha) - g_x(\beta)| \leq \sup_{\substack{|\alpha'-\beta'|\leq\delta \\ 0<\alpha',\beta'<1}} \left| f(F^{-1}(\alpha')) - f(F^{-1}(\beta')) \right|.$$

Now the uniform continuity of $f \circ F^{-1}$ on $(0, 1)$ yields the uniform equicontinuity of $(g_x)_{x\in\mathbb{R}}$.

Let $0 < \varepsilon < \frac{1}{2}$ and $a < b$ such that $P([a, b]) = 1 - \varepsilon$ and $f(a) = f(b) > 0$. Then it follows that

$$\inf_{x\in\mathbb{R}} \inf_{\varepsilon \leq \alpha \leq 1-\varepsilon} g_x(\alpha) \geq f(a). \qquad \square$$

Set

$$\bar{P}_{n,x}(\alpha) = \sup\{P_n(I) : \ |I| \leq S_x(\alpha), \ I \in \mathbb{I}_x\}, \quad 0 < \alpha < 1, \ \ x \in \mathbb{R},$$

$\bar{P}_{n,x}(0) = 0$ and $\bar{P}_{n,x}(1) = 1$. Consider the process

$$\bar{U}_{n,x}(\alpha) = n^{\frac{1}{2}}(\bar{P}_{n,x}(\alpha) - \alpha), \quad 0 \leq \alpha \leq 1, \ \ x \in \mathbb{R}.$$

This is the right place to describe the main steps in the proofs of the theorems. First we present in Proposition 1 the appropriate convergence result for the "uniformized" process $\bar{U}_{n,x}$. Next we "invert" this statement to get a similar convergence result for a uniformized quantile-type process (Corollary 1), and finally we obtain our theorems by "stretching out" this process in the vertical direction.

Proposition 1 *On the probability space of (1),*

$$\sup_{x\in\mathbb{R}}\ \sup_{0\le\alpha\le1}|\bar{U}_{n,x}(\alpha)-B_x(\alpha)|\to 0 \quad a.s.\ \ as\ \ n\to\infty. \tag{4}$$

Proof: The proof follows closely the lines of that in Proposition 3.1 in Einmahl and Mason (1992), but now the supremum over x has also to be taken into account. Clearly it is sufficient to show that $\sup_{x\in\mathbb{R}}\sup_{0<\alpha<1}|\bar{U}_{n,x}(\alpha)-B_x(\alpha)|\to 0$ almost surely as $n\to\infty$. First we show that

$$\limsup_{n\to\infty}\sup_{x\in\mathbb{R}}\sup_{0<\alpha<1}\left(B_x(\alpha)-\bar{U}_{n,x}(\alpha)\right)\le 0 \quad \text{a.s.} \tag{5}$$

For any $0<\alpha<1$ and $x\in\mathbb{R}$,

$$\begin{aligned}
B_x(\alpha)-\bar{U}_{n,x}(\alpha) &\le \sup\{B_P(I): I\in T_{x,\alpha}\}\\
&\quad -n^{\frac12}\left(\sup\{P_n(I):|I|\le S_x(\alpha),\ P(I)=\alpha,\ I\in\mathbb{I}_x\}-\alpha\right)\\
&= \sup\{B_P(I): I\in T_{x,\alpha}\}-n^{\frac12}\sup\{P_n(I)-P(I): I\in T_{x,\alpha}\}\\
&\le \sup\{B_P(I)-U_n(I): I\in\mathbb{I}\}.
\end{aligned}$$

Now (5) follows from (1).

It remains to show that

$$\limsup_{n\to\infty}\sup_{x\in\mathbb{R}}\sup_{0<\alpha<1}\left(\bar{U}_{n,x}(\alpha)-B_x(\alpha)\right)\le 0 \quad a.s.$$

For any $0<\alpha<1$ and $x\in\mathbb{R}$ we have

$$\begin{aligned}
&\bar{U}_{n,x}(\alpha)-B_x(\alpha)\\
&\quad\le \left\{n^{\frac12}\left(\sup\{P_n(I):|I|\le S_x(\alpha),\ \alpha-n^{-\frac14}<P(I)\le\alpha,\ I\in\mathbb{I}_x\}-\alpha\right)-B_x(\alpha)\right\}\\
&\quad\quad\vee\left\{n^{\frac12}\left(\sup\{P_n(I):P(I)\le\alpha-n^{-\frac14},\ I\in\mathbb{I}_x\}-\alpha\right)-B_x(\alpha)\right\}.
\end{aligned}\tag{6}$$

The second term in the right-hand side of (6) is bounded from above by

$$\begin{aligned}
&\sup\left\{n^{\frac12}\left(P_n(I)-P(I)-n^{-\frac14}\right): I\in\mathbb{I}_x\right\}-B_x(\alpha)\\
&\quad\le \sup\{|U_n(I)-B_P(I)|: I\in\mathbb{I}_x\}+\sup\{|B_P(I)|: I\in\mathbb{I}_x\}\\
&\quad\quad+\sup\{|B_P(I')|: I'\in T_{x,\alpha}\}-n^{\frac14}\\
&\quad\le \sup\{|U_n(I)-B_P(I)|: I\in\mathbb{I}\}+2\sup\{|B_P(I)|: I\in\mathbb{I}\}-n^{\frac14},
\end{aligned}$$

which, by (1) and the boundedness of B_P, converges almost surely to $-\infty$, as $n \to \infty$. Next consider the first term in the right-hand side of (6). For any $0 < \alpha < 1$ and $x \in \mathbb{R}$, this term is equal to

$$\begin{aligned}
&\sup\left\{n^{\frac{1}{2}}(P_n(I) - \alpha) : |I| \leq S_x(\alpha),\ \alpha - n^{-\frac{1}{4}} < P(I) \leq \alpha,\ I \in \mathbb{I}_x\right\} - B_x(\alpha) \\
&\quad \leq \ \sup\{|U_n(I) - B_P(I)| : I \in \mathbb{I}_x\} \\
&\qquad\qquad + \left\{\sup\left\{B_P(I) : |I| \leq S_x(\alpha),\ \alpha - n^{-\frac{1}{4}} < P(I) \leq \alpha,\ I \in \mathbb{I}_x\right\} - B_x(\alpha)\right\}.
\end{aligned}$$

The first term tends to zero, uniformly in $x \in \mathbb{R}$, almost surely as $n \to \infty$ because of $\mathbb{I}_x \subset \mathbb{I}$ and (1), so the proof of (4) will be complete if we show

$$\begin{aligned}
&\sup_{x\in\mathbb{R}} \sup_{0<\alpha<1} \Big\{ \sup\left\{B_P(I) : |I| \leq S_x(\alpha),\ \alpha - n^{-\frac{1}{4}} < P(I) \leq \alpha,\ I \in \mathbb{I}_x\right\} \\
&\qquad\qquad - \sup\left\{B_P(I') : I' \in T_{x,\alpha}\right\} \Big\} \to 0 \ \text{ as } \ n \to \infty. \qquad (7)
\end{aligned}$$

By Lemma 2 combined with Lemma 1, and uniform continuity of B_P for any $\eta > 0$, we have for all large n

$$\begin{aligned}
&\sup_{x\in\mathbb{R}} \sup_{0<\alpha<1} \Big| \sup\left\{B_P(I) : |I| \leq S_x(\alpha),\ \alpha - n^{-\frac{1}{4}} < P(I) \leq \alpha,\ I \in \mathbb{I}_x\right\} \\
&\qquad\qquad - \ \sup\ \left\{B_P(I') : I' \in T_{x,\alpha}\right\} \Big| \ \leq \ \eta.
\end{aligned}$$

Since $\eta > 0$ is arbitrary, this implies (7). □

For $s : [0, 1] \to \mathbb{R}$, write $\|s\| = \sup\{|s(\alpha)| : 0 \leq \alpha \leq 1\}$. If $x \in \mathfrak{X}$ (some index set) and $s_x : [0, 1] \to \mathbb{R}$, write $|||s_x\ ||| = \sup\{|s_x(\alpha)| :\ 0 \leq \alpha \leq 1,\ x \in \mathfrak{X}\}$. Let I denote the identity function.

Lemma 4 *Let Γ be a nondecreasing function on $[0, 1]$ with $\Gamma(0) = 0$ and $\Gamma(1) = 1$. Define $\Gamma^{-1}(\alpha) = \inf\{\beta : \Gamma(\beta) \geq \alpha\}$, $0 \leq \alpha \leq 1$. Then*

$$\|\Gamma + \Gamma^{-1} - 2I\| \leq \omega(\Gamma - I,\ \|\Gamma - I\|).$$

Proof: Write $S = [-\|\Gamma - I\|,\ \|\Gamma - I\|]$. We have

$$\begin{aligned}
\Gamma^{-1}(\alpha) - \alpha &= \inf\{\beta - \alpha : \Gamma(\beta) \geq \alpha\} \\
&= \inf\{\beta - \alpha \in S : \Gamma(\beta) \geq \alpha\}. \qquad (8)
\end{aligned}$$

The second equality in (8) follows, since for $\beta-\alpha > \|\Gamma-I\|$ we have $\alpha-\beta < -\|\Gamma-I\| \leq \Gamma(\beta) - \beta$ and hence $\alpha < \Gamma(\beta)$; for $\beta - \alpha < -\|\Gamma - I\|$, we have $\alpha - \beta > \|\Gamma - I\|$ and therefore $\alpha - \beta \leq \Gamma(\beta) - \beta$ or $\alpha \leq \Gamma(\beta)$ is impossible. Thus

$$\begin{aligned}
\Gamma^{-1}(\alpha) - \alpha &= \inf\{\beta - \alpha \in S : \beta - \alpha \geq \alpha - \Gamma(\alpha) + \beta - \Gamma(\beta) - (\alpha - \Gamma(\alpha))\} \\
&\geq \inf\{\beta - \alpha \in S : \beta - \alpha \geq \alpha - \Gamma(\alpha) - \omega(\Gamma - I, \|\Gamma - I\|)\} \\
&\geq \alpha - \Gamma(\alpha) - \omega(\Gamma - I, \|\Gamma - I\|).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\Gamma^{-1}(\alpha) - \alpha &\leq \inf\{\beta - \alpha \in S : \beta - \alpha \geq \alpha - \Gamma(\alpha) + \omega(\Gamma - I, \|\Gamma - I\|)\} \\
&\leq \alpha - \Gamma(\alpha) + \omega(\Gamma - I, \|\Gamma - I\|),
\end{aligned}$$

and hence

$$-\omega(\Gamma - I, \|\Gamma - I\|) \;\leq\; \Gamma^{-1}(\alpha) - 2\alpha + \Gamma(\alpha) \;\leq\; \omega(\Gamma - I, \|\Gamma - I\|). \qquad \square$$

We will use this lemma to establish a generalization to a collection of functions of the well-known lemma in Vervaat (1972).

Lemma 5 *Let $\Gamma_{n,x}$ be a collection of nondecreasing functions on $[0,1]$ indexed by $n \in \mathbb{N}$ and $x \in \mathfrak{X}$. Assume for all n and x, $\Gamma_{n,x}(0) = 0$ and $\Gamma_{n,x}(1) = 1$. Moreover, let b_x, $x \in \mathfrak{X}$, be a collection of uniformly bounded ($\sup_{x\in\mathfrak{X}} \sup_{0\leq\alpha\leq1} |b_x(\alpha)| < \infty$) and uniformly equicontinuous functions on $[0,1]$. Finally let $(m_n)_{n\in\mathbb{N}}$ be a sequence of positive numbers tending to infinity.*
If, as $n \to \infty$,

$$\sup_{x\in\mathfrak{X}} \sup_{0\leq\alpha\leq1} |m_n(\Gamma_{n,x}(\alpha) - \alpha) - b_x(\alpha)| \to 0,$$

then

$$\sup_{x\in\mathfrak{X}} \sup_{0\leq\alpha\leq1} |m_n(\Gamma_{n,x}^{-1}(\alpha) - \alpha) + b_x(\alpha)| \to 0.$$

Proof: Write $D_n = \sup_{x\in\mathfrak{X}} \sup_{0\le\alpha\le 1} |m_n(\Gamma_{n,x}(\alpha)-\alpha) - b_x(\alpha)|$. From Lemma 4 we have

$$\begin{aligned}
&|||m_n(\Gamma_{n,x}^{-1} - I) \ + \ b_x \ ||| \\
&\quad \le |||m_n(\Gamma_{n,x}^{-1} - I + \Gamma_{n,x} - I) \ ||| + \ ||| - \left[m_n(\Gamma_{n,x} - I) - b_x\right] \ ||| \\
&\quad \le \ m_n \sup_{x\in\mathfrak{X}} \omega(\Gamma_{n,x} - I, \|\Gamma_{n,x} - I\|) \ + \ D_n \\
&\quad \le \ \sup_{x\in\mathfrak{X}} \left\{ 2 \sup_{0\le\alpha\le 1} |m_n(\Gamma_{n,x}(\alpha)-\alpha) - b_x(\alpha)| \ + \sup_{|\beta-\alpha|\le\|\Gamma_{n,x}-I\|} |b_x(\beta) - b_x(\alpha)| \right\} \ + \ D_n \\
&\quad \le \ \sup_{x\in\mathfrak{X}} \omega(b_x, \|\Gamma_{n,x} - I\|) \ + \ 3D_n \\
&\quad \le \sup_{x\in\mathfrak{X}} \omega \left(b_x, \frac{D_n + \ |||b_x \ |||}{m_n} \right) + 3D_n \to 0 \ \text{ as } \ n \to \infty \ . \qquad \square
\end{aligned}$$

Define

$$V_{n,x}(\beta) = \inf\{\alpha : \bar{P}_{n,x}(\alpha) \ge \beta, \ 0 \le \alpha \le 1\}, \quad 0 \le \beta \le 1, \ \ x \in \mathbb{R},$$

$$\bar{Q}_{n,x}(\alpha) = n^{\frac{1}{2}}(V_{n,x}(\alpha) - \alpha), \quad 0 \le \alpha \le 1, \ \ x \in \mathbb{R}.$$

It is immediate from Lemma 1 and the continuity of B_P, that $(B_x)_{x\in\mathbb{R}}$ is a collection of uniformly equicontinuous functions on $[0,1]$. Hence combining Lemma 5 and Proposition 1, we obtain the following important result.

Corollary 1 *On the probability space of (1), as $n \to \infty$,*

$$\sup_{x\in\mathbb{R}} \sup_{0\le\alpha\le 1} |\bar{Q}_{n,x}(\alpha) + B_x(\alpha)| \to 0 \ \ a.s.$$

and hence

$$\sup_{x\in\mathbb{R}} \sup_{0\le\alpha\le 1} |V_{n,x}(\alpha) - \alpha| \to 0 \ \ a.s.$$

Define $S_x(0) = \lim_{\alpha\downarrow 0} S_x(\alpha)$. Similar to Lemma 3.1 in Einmahl and Mason (1992) we can show:

Lemma 6 *With probability 1, for all $0 < \alpha < 1$ and $x \in \mathbb{R}$,*

$$S_{n,x}(\alpha) = S_x(V_{n,x}(\alpha)).$$

Proof of Theorem 1: For each $\eta \leq \alpha \leq 1-\eta$ and $x \in \mathbb{R}$ we get by Lemma 6 and the mean value theorem, that almost surely

$$
\begin{aligned}
Q_{n,x}(\alpha) + B_x(\alpha) &= g_x(\alpha)\, n^{\frac{1}{2}} \; (S_x(V_{n,x}(\alpha)) - S_x(\alpha)) + B_x(\alpha) \\
&= \frac{g_x(\alpha)}{g_x(\theta_{n,x})}\, \bar{Q}_{n,x}(\alpha) + B_x(\alpha) \\
&= \frac{g_x(\alpha)}{g_x(\theta_{n,x})} \big(\bar{Q}_{n,x}(\alpha) + B_x(\alpha)\big) - \frac{g_x(\alpha) - g_x(\theta_{n,x})}{g_x(\theta_{n,x})} B_x(\alpha), \qquad (9)
\end{aligned}
$$

where $\theta_{n,x}$ lies between α and $V_{n,x}(\alpha)$. Assertion (2) follows from Corollary 1 and Lemma 3. $\square$

For the proof of Theorem 2 we need three more auxiliary results.

Fact 1 *[Lemma 1 in Csörgő and Révész (1978)] Under the assumptions of Theorem 2, in particular condition (B), we have*

$$
\frac{f\,(F^{-1}(\alpha))}{f\,(F^{-1}(\beta))} \leq \left[\frac{\beta \vee \alpha}{\beta \wedge \alpha}\, \frac{1-(\beta \wedge \alpha)}{1-(\beta \vee \alpha)} \right]^M \quad \textit{for all } 0 < \alpha, \beta < 1.
$$

Fact 2 *[Lemma 3.2 in Einmahl and Mason (1992)] Let $(Y_{n,k})_{n \geq 1,\, k \geq 1}$ be a double sequence of random variables such that for each $n, k \in \mathbb{N}$, $Y_{n,k}$ is Binomial$(n, 2^{-k})$. Then*

$$
Y_n := \sup_{k \in \mathbb{N}} \; n^{-1}\, 2^k\, Y_{n,k} = O_{\mathbb{P}}(1) \quad \textit{as} \quad n \to \infty.
$$

Lemma 7 *On the probability space of (1),*

$$
\sup_{x \in \mathbb{R}} \; \sup_{0<\alpha<1} \frac{1-\alpha}{1-V_{n,x}(\alpha)} = O_{\mathbb{P}}(1).
$$

Proof: For $k \in \mathbb{N}$, $x \in \mathbb{R}$ choose $I_{k,x} \in T_{x,1-2^{-k}}$ and for $1-2^{-k} \leq \alpha < 1-2^{-k-1}$ set $I_{\alpha,x} = I_{k,x}$. Following the proof of Lemma 3.3 in Einmahl and Mason (1992), we can now show that

$$
\sup_{x \in \mathbb{R}} \; \sup_{0<\alpha<1} \frac{1-\alpha}{1-V_{n,x}(\alpha)} \; \leq \; 2 \vee \left\{ \sup_{x \in \mathbb{R}} \; \sup_{k \geq 1} 2^{k+1} (1 - P_n(I_{k,x})) \right\}. \qquad (10)
$$

Write $W_k := [F^{-1}(2^{-k}), F^{-1}(1-2^{-k})]$. Then $I_{k,x} \supset W_k$ holds for every $x \in \mathbb{R}$. Hence the second term of the right-hand side of (10) is bounded from above by

$$
\sup_{k \geq 1} 2^{k+1} (1 - P_n(W_k)). \qquad (11)
$$

Since $n(1 - P_n(W_k))$ is Binomial$(n, 2^{-(k-1)})$, Fact 2 yields that the expression in (11) is $O_{\mathbb{P}}(1)$. $\square$

Proof of Theorem 2: Theorem 1 states that for all $0 < \eta < 1/2$

$$\sup_{x\in\mathbb{R}} \sup_{\eta\le\alpha\le 1-\eta} |Q_{n,x}(\alpha) + B_x(\alpha)| \to 0 \quad \text{a.s. as } n \to \infty. \tag{12}$$

If $\lim_{x\downarrow x_*} f(x) > 0$, then $\inf_{x\in\mathbb{R}} \inf_{0<\alpha<1} g_x(\alpha) > 0$ and Theorem 2 holds using the same argument as in the proof of Theorem 1.

So in the sequel we assume $\lim_{x\downarrow x_*} f(x) = 0$. Because of (12) we only need to consider the supremum on the region where $x \in \mathbb{R}$ and $\alpha < \eta$ and on the region $x \in \mathbb{R}$ and $\alpha > 1 - \eta$. We have from (9):

$$|Q_{n,x}(\alpha) + B_x(\alpha)| \le \left| \frac{g_x(\alpha)}{g_x(\theta_{n,x})} \left(\bar{Q}_{n,x}(\alpha) + B_x(\alpha)\right) \right| + \left| \left(\frac{g_x(\alpha)}{g_x(\theta_{n,x})} - 1 \right) B_x(\alpha) \right|.$$

Therefore it follows from a routine argument, using Corollary 1 and the equicontinuity of $(B_x)_{x\in\mathbb{R}}$ in conjunction with $B_x(0) = B_x(1) = 0$ for all $x \in \mathbb{R}$ almost surely, that it is sufficient to show that for small enough $\eta > 0$, $g_x(\alpha)/g_x(\theta_{n,x})$ is bounded in probability uniformly over both regions.

First we will show

$$\sup_{\alpha<\eta} \sup_{x\in\mathbb{R}} \frac{g_x(\alpha)}{g_x(\theta_{n,x})} = O_{\mathbb{P}}(1). \tag{13}$$

Let $\delta \in (0, 1/2)$ be small. The region over which the supremum is taken will be split up into three regions depending on x, namely $\alpha < \eta$ and $F(x) < \delta$, $\delta \le F(x) \le 1 - \delta$, $F(x) > 1 - \delta$, respectively. For the middle region we have for small enough η, because of Corollary 1, that almost surely for large n, $g_x(\theta_{n,x})$ is bounded away from 0; see proof of Lemma 3. Since $g_x(\alpha) \le \sup_{y\in\mathbb{R}} f(y) = f(x_1)$, we hence have

$$\sup_{\alpha<\eta} \sup_{\delta\le F(x)\le 1-\delta} \frac{g_x(\alpha)}{g_x(\theta_{n,x})} = O_{\mathbb{P}}(1).$$

In order to complete the proof of (13) we need to consider the regions $\alpha < \eta$ and $F(x) < \delta$, $F(x) > 1 - \delta$, respectively. Because of symmetry we will restrict ourselves to the region $\alpha < \eta$ and $F(x) < \delta$. Note that from the proof of Lemma 3 it follows that g_x is nondecreasing for $\alpha \le |F(x_1) - F(x)|$ and nonincreasing for $\alpha \ge |F(x_1) - F(x)|$.

Therefore for $\alpha \geq V_{n,x}(\alpha)$ (when $\alpha \leq V_{n,x}(\alpha)$ we use 1 as an upper bound), almost surely for large n

$$\begin{aligned}
\sup_{\alpha<\eta} \sup_{F(x)<\delta} \frac{g_x(\alpha)}{g_x(\theta_{n,x})} &\leq \sup_{\alpha<\eta} \sup_{F(x)<\delta} \frac{g_x(\alpha)}{g_x(V_{n,x}(\alpha))} \\
&\leq \sup_{\alpha<\eta} \sup_{F(x)\leq\alpha} \frac{g_x(\alpha)}{g_x(V_{n,x}(\alpha))} + \sup_{\alpha<\eta} \sup_{\alpha<F(x)<\delta} \frac{g_x(\alpha)}{g_x(V_{n,x}(\alpha))} . \qquad (14)
\end{aligned}$$

For x and α both small enough we have $g_x(\alpha) = f\left(F^{-1}\left(F(x)+\alpha\right)\right)$. Hence the second term in the right-hand side of (14) is equal to

$$\sup_{\alpha<\eta} \sup_{\alpha<F(x)<\delta} \frac{f\left(F^{-1}\left(F(x)+\alpha\right)\right)}{f\left(F^{-1}\left(F(x)+V_{n,x}(\alpha)\right)\right)} \leq \sup_{F(x)<\delta} \frac{f\left(F^{-1}\left(2F(x)\right)\right)}{f\left(F^{-1}\left(F(x)\right)\right)} , \qquad (15)$$

because $f \circ F^{-1}$ is increasing on $(0, F(x_1))$. It is immediate from Fact 1 that the right-hand side of (15) is bounded.

Similarly, the first term in the right-hand side of (14) is bounded from above by

$$\begin{aligned}
&\sup_{\alpha<\eta} \sup_{F(x)\leq\alpha} \frac{f\left(F^{-1}\left(2\alpha\right)\right)}{f\left(F^{-1}\left(F(x)+V_{n,x}(\alpha)\right)\right)} \\
&\leq \sup_{\alpha<\eta} \sup_{F(x)\leq\alpha} \left(\frac{2\alpha}{F(x)+V_{n,x}(\alpha)} \cdot \frac{1-F(x)-V_{n,x}(\alpha)}{1-2\alpha} \right)^M . \qquad (16)
\end{aligned}$$

Because the second factor in the right-hand side of (16) is clearly bounded in probability, we need to show that

$$\sup_{\alpha<\eta} \sup_{F(x)\leq\alpha} \frac{\alpha}{F(x)+V_{n,x}(\alpha)} = O_{\mathbb{P}}(1).$$

The proof of this is based on the following crucial inequality: with probability 1

$$V_{n,x}(\alpha) \geq F(X_{(\lceil n\alpha \rceil)}) - F(x) \quad \text{for all } x \in \mathbb{R} \text{ and } 0 < \alpha < 1. \qquad (17)$$

When proving this inequality, we assume $F(X_{(\lceil n\alpha \rceil)}) > F(x)$, otherwise there is nothing to prove. Using Lemma 6 and the Monotonicity property (Section 1), we see that we need to show

$$S_{n,x}(\alpha) \geq S_x(F(X_{(\lceil n\alpha \rceil)}) - F(x)).$$

From the definitions of S_x and $S_{n,x}$ we obtain

$$\begin{aligned}
S_x(F(X_{(\lceil n\alpha \rceil)}) - F(x)) &= \inf\{b-a : F(b)-F(a) \geq F(X_{(\lceil n\alpha \rceil)}) - F(x), x \in [a,b]\} \\
&\leq X_{(\lceil n\alpha \rceil)} - x \leq S_{n,x}(\alpha),
\end{aligned}$$

and hence (17). Hence, almost surely,

$$\sup_{\alpha<\eta}\sup_{F(x)\leq\alpha}\frac{\alpha}{F(x)+V_{n,x}(\alpha)}\leq\sup_{\alpha<\eta}\frac{\alpha}{F(X_{(\lceil n\alpha\rceil)})}\leq\sup_{0<\alpha<1}\frac{\alpha}{F(X_{(\lceil n\alpha\rceil)})}\ . \tag{18}$$

The denominator on the right is equal in distribution to the empirical quantile function of a sample of n independent uniform-$(0,1)$ variables. Hence it is well-known that the expression on the right in (18) is bounded in probability, see, e.g., Shorack and Wellner (1986), p. 419. This proves (13).

The proof of Theorem 2 is complete if we show that

$$\sup_{\alpha>1-\eta}\sup_{x\in\mathbb{R}}\frac{g_x(\alpha)}{g_x(\theta_{n,x})}=O_{\mathbb{P}}(1). \tag{19}$$

For symmetry reasons we can restrict x to $(-\infty,F^{-1}(1/2)]$. For large enough α, λ_α is defined and we can write g_x as follows:

$$g_x(\alpha)=\begin{cases} f(F^{-1}(F(x)+\alpha)) & \text{for } x<F^{-1}((1-\alpha)\lambda_\alpha)\\ f(F^{-1}((1-\alpha)\lambda_\alpha)) & \text{for } F^{-1}((1-\alpha)\lambda_\alpha)\leq x\leq F^{-1}(1/2)\ . \end{cases}$$

For small enough η and $\alpha\leq V_{n,x}(\alpha)$ (again, when $\alpha\geq V_{n,x}(\alpha)$ we can use 1 as an upper bound) we obtain

$$\begin{aligned}\sup_{\alpha>1-\eta}\sup_{F(x)\leq\frac12}\frac{g_x(\alpha)}{g_x(\theta_{n,x})}\quad &\leq \sup_{\alpha>1-\eta}\sup_{F(x)<(1-V_{n,x}(\alpha))\lambda_{V_{n,x}(\alpha)}}\frac{g_x(\alpha)}{g_x(V_{n,x}(\alpha))}\\ &\vee\sup_{\alpha>1-\eta}\sup_{(1-V_{n,x}(\alpha))\lambda_{V_{n,x}(\alpha)}\leq F(x)<(1-\alpha)\lambda_\alpha}\frac{g_x(\alpha)}{g_x(V_{n,x}(\alpha))}\\ &\vee\sup_{\alpha>1-\eta}\sup_{(1-\alpha)\lambda_\alpha\leq F(x)\leq\frac12}\frac{g_x(\alpha)}{g_x(V_{n,x}(\alpha))}\ .\end{aligned} \tag{20}$$

The last term in the right-hand side of (20) can be bounded from above by applying Fact 1:

$$\begin{aligned}\sup_{\alpha>1-\eta}\sup_{(1-\alpha)\lambda_\alpha\leq F(x)\leq\frac12}\frac{g_x(\alpha)}{g_x(V_{n,x}(\alpha))} &= \sup_{\alpha>1-\eta}\sup_{(1-\alpha)\lambda_\alpha\leq F(x)\leq\frac12}\frac{f\left(F^{-1}\left((1-\alpha)\lambda_\alpha\right)\right)}{f\left(F^{-1}\left(\left(1-V_{n,x}(\alpha)\right)\lambda_{V_{n,x}(\alpha)}\right)\right)}\\ &\leq\sup_{\alpha>1-\eta}\sup_{x\in\mathbb{R}}\left(\frac{(1-\alpha)\lambda_\alpha}{\left(1-V_{n,x}(\alpha)\right)\lambda_{V_{n,x}(\alpha)}}\ \frac{1-\left(1-V_{n,x}(\alpha)\right)\lambda_{V_{n,x}(\alpha)}}{1-(1-\alpha)\lambda_\alpha}\right)^M,\end{aligned}$$

which is easily seen to be $O_{\mathbb{P}}(1)$, using Lemma 7 and condition (C).

The first term in the right-hand side of (20) is equal to

$$\sup_{\alpha>1-\eta}\sup_{F(x)<(1-V_{n,x}(\alpha))\lambda_{V_{n,x}(\alpha)}}\frac{f\left(F^{-1}\left(F(x)+\alpha\right)\right)}{f\left(F^{-1}\left(F(x)+V_{n,x}(\alpha)\right)\right)},$$

which is, since f is decreasing for large values and because of Fact 1, bounded from above by

$$(21)\quad \sup_{\alpha>1-\eta} \sup_{F(x)<(1-V_{n,x}(\alpha))\lambda_{V_{n,x}(\alpha)}} \frac{f\left(F^{-1}\left(F(x)+\alpha\right)\right)}{f\left(F^{-1}\left(\left(1-V_{n,x}(\alpha)\right)\lambda_{V_{n,x}(\alpha)}+V_{n,x}(\alpha)\right)\right)}$$
$$\leq \sup_{\alpha>1-\eta} \sup_{x\in\mathbb{R}} \left(\frac{\left(1-V_{n,x}(\alpha)\right)\lambda_{V_{n,x}(\alpha)}+V_{n,x}(\alpha)}{\alpha} \frac{1-\alpha}{\left(1-V_{n,x}(\alpha)\right)\left(1-\lambda_{V_{n,x}(\alpha)}\right)}\right)^M .$$

Again, Lemma 7 and condition (C) yield that this term is $O_{\mathbb{P}}(1)$.

The middle term in the right-hand side of (20), rewritten as (using $f(F^{-1}((1-\beta)\lambda_\beta)) = f(F^{-1}((1-\beta)\lambda_\beta+\beta))$)

$$\sup_{\alpha>1-\eta} \sup_{(1-V_{n,x}(\alpha))\lambda_{V_{n,x}(\alpha)}\leq F(x)<(1-\alpha)\lambda_\alpha} \frac{f\left(F^{-1}\left(F(x)+\alpha\right)\right)}{f\left(F^{-1}\left(\left(1-V_{n,x}(\alpha)\right)\lambda_{V_{n,x}(\alpha)}+V_{n,x}(\alpha)\right)\right)} ,$$

is bounded by the right-hand side of (21). This completes the proof of (19). □

References

Andrews, D.F., Bickel, P.J., Hampel, F.R., Huber, P.J., Rogers, W.H., and Tukey, J.W. (1972), *Robust Estimation of Location: Survey and Advances*, Princeton Univ. Press.

Chaudhuri, P., and Marron, J.S. (1999), "SiZer for Exploration of Structures in Curves," *Journal of the American Statistical Association*, 94, 807–823.

Csörgő, M. and Révész, P. (1978), "Strong Approximations of the Quantile Process," *The Annals of Statistics*, 6, 882–894.

Einmahl, J.H.J., and Mason, D.M. (1992), "Generalized Quantile Processes," *The Annals of Statistics*, 20, 1062–1078.

Grübel, R. (1988), "The Length of the Shorth," *The Annals of Statistics*, 16, 619–628.

Hyndman, R.J. (1996), "Computing and Graphing Highest Density Regions," *The American Statistician*, 50, 120–126.

Hyndman, R.J., Bashtannyk, D.M., and Grunwald, G.K. (1996), "Estimating and Visualizing Conditional Densities," *Journal of Computational and Graphical Statistics* 5, 315–336.

Minnotte, M.C., and Scott, D.W. (1993), "The Mode Tree: A Tool for Visualization of Nonparametric Density Features," *Journal of Computational and Graphical Statistics*, 2, 51–68.

Müller, D.W., and Sawitzki, G. (1991), "Excess Mass Estimates and Tests for Multimodality," *Journal of the American Statistical Association*, 86, 738–746.

Sawitzki, G. (1994), Diagnostic Plots for One-Dimensional Data, In: R.O. Peter Dirschedl, ed., "Computational Statistics, 25th Conference on Statistical Computing at Schloss Reisensburg.", Physica-Verlag/Springer, Heidelberg, 237–258.

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

Tukey, P.A., and Tukey, J.W. (1981), "Data-Driven View Selection; Agglomeration and Sharpening," in *Interpreting Multivariate Data* ed. V. Barnett, New York: Wiley.

Vervaat, W. (1972), "Functional Central Limit Theorems for Processes with Positive Drift and their Inverses," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 23, 245–253.