

A NEW APPROACH TO TESTS AND CONFIDENCE BANDS FOR DISTRIBUTION FUNCTIONS

BY LUTZ DÜMBGEN^{1,a} AND JON A. WELLNER^{2,b}

¹Department of Mathematics and Statistics, University of Bern, ^aduembgen@stat.unibe.ch

²Department of Statistics, University of Washington, ^bjaw@stat.washington.edu

We introduce new goodness-of-fit tests and corresponding confidence bands for distribution functions. They are inspired by multiscale methods of testing and based on refined laws of the iterated logarithm for the normalized uniform empirical process $\mathbb{U}_n(t)/\sqrt{t(1-t)}$ and its natural limiting process, the normalized Brownian bridge process $\mathbb{U}(t)/\sqrt{t(1-t)}$. The new tests and confidence bands refine the procedures of Berk and Jones (1979) and Owen (1995). Roughly speaking, the high power and accuracy of the latter methods in the tail regions of distributions are essentially preserved while gaining considerably in the central region. The goodness-of-fit tests perform well in signal detection problems involving sparsity, as in Ingster (1997), Donoho and Jin (2004) and Jager and Wellner (2007), but also under contiguous alternatives. Our analysis of the confidence bands sheds new light on the influence of the underlying ϕ -divergences.

1. Introduction and motivations.

1.1. *Some well-known facts.* Let \mathbb{F}_n be the empirical distribution function of independent random variables X_1, X_2, \dots, X_n with unknown distribution function F on the real line. The main topic of the present paper is to construct a confidence band $(A_{n,\alpha}, B_{n,\alpha})$ for F with given confidence level $1 - \alpha \in (0, 1)$. That is, $A_{n,\alpha} = A_{n,\alpha}(\cdot, (X_i)_{i=1}^n)$ and $B_{n,\alpha} = B_{n,\alpha}(\cdot, (X_i)_{i=1}^n)$ are data-driven functions on the real line such that for any true distribution function F ,

$$(1.1) \quad P_F(A_{n,\alpha} \leq F \leq B_{n,\alpha} \text{ on } \mathbb{R}) \geq 1 - \alpha.$$

Let us recall some well-known facts about \mathbb{F}_n (cf. [30, 31]). The stochastic process $(\mathbb{F}_n(x))_{x \in \mathbb{R}}$ has the same distribution as $(\mathbb{G}_n(F(x)))_{x \in \mathbb{R}}$, where \mathbb{G}_n is the empirical distribution of independent random variables $\xi_1, \xi_2, \dots, \xi_n$ with uniform distribution on $[0, 1]$. This enables the well-known Kolmogorov–Smirnov confidence bands: let

$$\mathbb{U}_n(t) := \sqrt{n}(\mathbb{G}_n(t) - t),$$

and let $\kappa_{n,\alpha}^{\text{KS}}$ be the $(1 - \alpha)$ -quantile of $\|\mathbb{U}_n\|_\infty := \sup_{t \in [0,1]} |\mathbb{U}_n(t)|$. Then the confidence band $(A_{n,\alpha}^{\text{KS}}, B_{n,\alpha}^{\text{KS}})$ with $A_{n,\alpha}^{\text{KS}} := \max(\mathbb{F}_n - n^{-1/2}\kappa_{n,\alpha}^{\text{KS}}, 0)$ and $B_{n,\alpha}^{\text{KS}} := \min(\mathbb{F}_n + n^{-1/2}\kappa_{n,\alpha}^{\text{KS}}, 1)$ satisfies (1.1) with equality if F is continuous. Since \mathbb{U}_n converges in distribution in $\ell^\infty([0, 1])$ to standard Brownian bridge \mathbb{U} , $\kappa_{n,\alpha}^{\text{KS}}$ converges to the $(1 - \alpha)$ -quantile $\kappa_\alpha^{\text{KS}}$ of $\|\mathbb{U}\|_\infty$. In particular, the width $B_{n,\alpha}^{\text{KS}} - A_{n,\alpha}^{\text{KS}}$ of the Kolmogorov–Smirnov band is bounded uniformly by $2n^{-1/2}\kappa_{n,\alpha}^{\text{KS}} = O(n^{-1/2})$. (Throughout this paper, asymptotic statements refer to $n \rightarrow \infty$, unless stated otherwise.) On the other hand, it is well known that Kolmogorov–Smirnov confidence bands give little or no information in the tails of the distribution F ; see, for example, [22], [19] and [20], Chapter 14, for a useful summary.

Received November 2021; revised October 2022.

MSC2020 subject classifications. Primary 60E10, 60F10; secondary 62D99.

Key words and phrases. Confidence band, goodness-of-fit, law of the iterated logarithm, limit distribution, multiscale test statistics.

1.2. *Confidence bands by inversion of tests.* In general, confidence bands can be obtained by inverting goodness-of-fit tests. For a given continuous distribution function F_0 , let $T_n(F_0) = T_n(F_0, (X_i)_{i=1}^n)$ be some test statistic for the null hypothesis that $F \equiv F_0$. Suppose that for any test level $\alpha \in (0, 1)$, the $(1 - \alpha)$ -quantile $\kappa_{n,\alpha}$ of $T_n(F_0)$ under the null hypothesis does not depend on F_0 . Then a $(1 - \alpha)$ -confidence band $(A_{n,\alpha}, B_{n,\alpha})$ for a continuous distribution function F is given by

$$A_{n,\alpha}(x) := \inf\{F(x) : T_n(F) \leq \kappa_{n,\alpha}\}, \quad B_{n,\alpha}(x) := \sup\{F(x) : T_n(F) \leq \kappa_{n,\alpha}\}.$$

Depending on the specific choice of T_n , these functions $A_{n,\alpha}$ and $B_{n,\alpha}$ can be computed explicitly, and the constraint (1.1) is even satisfied for arbitrary, possibly noncontinuous distribution functions F ; see Section S.6 for further details.

Since $(A_{n,\alpha}^{KS}, B_{n,\alpha}^{KS})$ corresponds to $T_n^{KS}(F_0) := \sqrt{n}\|\mathbb{F}_n - F_0\|_\infty$, one possibility to enhance precision in the tails is to consider weighted supremum norms such as

$$(1.2) \quad T_n(F_0) := \sup_{x: 0 < F_0(x) < 1} \frac{\sqrt{n}|\mathbb{F}_n - F_0|(x)}{w(F_0)}$$

or

$$(1.3) \quad T_n(F_0) := \sup_{x \in [X_{n:1}, X_{n:n}]} \frac{\sqrt{n}|\mathbb{F}_n - F_0|(x)}{w(\mathbb{F}_n)},$$

where $X_{n:1} \leq X_{n:2} \leq \dots \leq X_{n:n}$ are the order statistics of X_1, X_2, \dots, X_n . Here, $w : (0, 1) \rightarrow (0, \infty)$ is some continuous weight function such that $w(1 - t) = w(t)$ for $0 < t < 1$ and $w(t) \rightarrow 0$ as $t \rightarrow 0$. Specific proposals include

$$w(t) := \sqrt{t(1 - t)h(t)},$$

where $h \equiv 1$; see [17] and [11], or $h(t) \rightarrow \infty$ sufficiently fast as $t \rightarrow 0$, see [24] or [5]. Specifically, Stepanova and Pavlenko [32] propose to construct confidence bands with the test statistic (1.3) and $h(t) := \log \log(1/[t(1 - t)])$. The latter choice is motivated by the law of the iterated logarithm (LIL) for the Brownian bridge process \mathbb{U} , stating that

$$(1.4) \quad \limsup_{t \searrow 0} \frac{\mathbb{U}(t)}{\sqrt{2t \log \log(1/t)}} = \limsup_{t \nearrow 1} \frac{\mathbb{U}(t)}{\sqrt{2(1 - t) \log \log(1/(1 - t))}} = 1$$

almost surely.

1.3. *The tests of Berk and Jones and Owen’s bands.* Another goodness-of-fit test, proposed by Berk and Jones [3], uses the test statistic

$$(1.5) \quad T_n^{BJ}(F_0) := n \sup_{x: 0 < F_0(x) < 1} K(\mathbb{F}_n(x), F_0(x)),$$

where

$$K(u, t) := u \log\left(\frac{u}{t}\right) + (1 - u) \log\left(\frac{1 - u}{1 - t}\right)$$

for $u \in [0, 1]$ and $t \in (0, 1)$. Note that $K(u, t)$ is the Kullback–Leibler divergence between the Bernoulli(u) and Bernoulli(t) distributions. Owen [25] proposed and analyzed confidence bands for F based on this test statistic. As noted by [18], the test statistic $T_n^{BJ}(F_0)$ can be embedded into a general family of test statistics $T_{n,s}^{BJ}(F_0)$, $s \in \mathbb{R}$. Let

$$(1.6) \quad T_{n,s}^{BJ}(F_0) := \begin{cases} \sup_{x: 0 < F_0(x) < 1} nK_s(\mathbb{F}_n(x), F_0(x)) & \text{if } s > 0, \\ \sup_{x \in [X_{n:1}, X_{n:n}]} nK_s(\mathbb{F}_n(x), F_0(x)) & \text{if } s \leq 0, \end{cases}$$

with the following divergence function K_s : for $t, u \in (0, 1)$:

$$(1.7) \quad K_s(u, t) = \begin{cases} (t(u/t)^s + (1-t)[(1-u)/(1-t)]^s - 1)/[s(s-1)] & s \neq 0, 1, \\ u \log(u/t) + (1-u) \log[(1-u)/(1-t)] & s = 1, \\ t \log(t/u) + (1-t) \log[(1-t)/(1-u)] & s = 0. \end{cases}$$

(An alternative representation of K_s is given in (3.5).) Moreover, for fixed $t \in (0, 1)$ and $u \in \{0, 1\}$, the limit $K(u, t) := \lim_{u' \rightarrow u} K_s(u', t)$ equals ∞ if $s \leq 0$ and exists in $(0, \infty)$ otherwise. A detailed discussion of these divergences is given in Section S.3 of the online supplement A. At present, it suffices to note that for any fixed $t \in (0, 1)$, $K_s(u, t)$ is strictly convex in u with unique minimum 0 at $u = t$ and second derivative $[t(1-t)]^{-1}$ there. Interesting special cases are $K = K_1$, $K_{1/2}(u, t) = 4(1 - \sqrt{ut} - \sqrt{(1-u)(1-t)})$ and

$$K_2(u, t) = \frac{(u-t)^2}{2t(1-t)}, \quad K_{-1}(u, t) = \frac{(u-t)^2}{2u(1-u)}.$$

Consequently, if $w(t) := \sqrt{t(1-t)}$, then the test statistic $T_{n,2}^{BJ}(F_0)$ coincides with 0.5 times the square of $T_n(F_0)$ in (1.2), and $T_{n,-1}^{BJ}(F_0)$ equals 0.5 times the square of (1.3). As shown by [18], for any $s \in [-1, 2]$, the null distribution of $T_{n,s}^{BJ}(F_0)$ has the same asymptotic behavior, and the corresponding $(1 - \alpha)$ -quantiles $\kappa_{n,s,\alpha}^{BJ}$ satisfy

$$(1.8) \quad \kappa_{n,s,\alpha}^{BJ} = \log \log n + 2^{-1} \log \log \log n + O(1).$$

From this, one can deduce that the resulting confidence band $(A_{n,s,\alpha}^{BJO}, B_{n,s,\alpha}^{BJO})$ for F satisfies

$$B_{n,s,\alpha}^{BJO}(x) - A_{n,s,\alpha}^{BJO}(x) \leq 2\sqrt{2\gamma_n \mathbb{F}_n(1 - \mathbb{F}_n)(x)} + 4\gamma_n,$$

where $\gamma_n := n^{-1} \kappa_{n,s,\alpha}^{BJ} = (1 + o(1))n^{-1} \log \log n$; see Lemma S.12 in Section S.3. Hence, the band $(A_{n,s,\alpha}^{BJO}, B_{n,s,\alpha}^{BJO})$ is substantially more accurate than $(A_{n,\alpha}^{KS}, B_{n,\alpha}^{KS})$ in the tail regions. But in the central region, that is, when $\mathbb{F}_n(x)$ is bounded away from 0 and 1, they are of width $O(n^{-1/2}(\log \log n)^{1/2})$ rather than $O(n^{-1/2})$.

1.4. *Goals revisited.* The goal of Berk and Jones [3] was to find goodness-of-fit tests with optimal Bahadur efficiencies. They interpret their test statistic $T_n^{BJ}(F_0)$ also as a union-intersection test statistic, where $nK(\mathbb{F}_n(x), F_0(x))$ is the negative likelihood ratio statistic for the null hypothesis that $F(x) = F_0(x)$, based on the binomial distribution of $n\mathbb{F}_n(x)$. The union-intersection and related paradigms for the present goodness-of-fit testing problem have been treated in more generality by [14].

In view of the previous considerations, the confidence band $(A_{n,\alpha}^{SP}, B_{n,\alpha}^{SP})$ of [32], based on the test statistic

$$(1.9) \quad T_n^{SP}(F_0) := \sup_{x \in [X_{n:1}, X_{n:n}]} \frac{\sqrt{n}|\mathbb{F}_n - F_0|}{\sqrt{\mathbb{F}_n(1 - \mathbb{F}_n)h(\mathbb{F}_n)}}(x)$$

with $h(t) := \log \log(1/[t(1-t)])$, provides a trade-off between tail behavior and behavior in the center of the distribution. Previous proposals for the same purpose include [21] and [27]. But we shall demonstrate later that with purely multiplicative correction factors as in (1.9), the tail regions are asymptotically underemphasized in comparison with the new methods presented here.

1.5. *Our new test statistics and confidence bands.* To obtain a better compromise between the Kolmogorov–Smirnov and Berk–Jones tests, we propose a refined adjustment of $\mathbb{F}_n(x)$ involving a pointwise standardization together with a pointwise additive correction, where the latter takes into account whether x is in the center or in the tails of F_0 or \mathbb{F}_n . Only after standardization and additive correction, we take a supremum over x . This approach of pointwise standardization plus additive correction before taking a supremum has been developed in the context of multiscale testing and has proved quite successful there; see, for example, [7], [8], [29] and [28]. In the present setting, pointwise standardization means that we consider $nK_s(\mathbb{F}_n(x), F_0(x))$, which behaves asymptotically like $\mathbb{U}(F_0(x))^2/[2F_0(x)(1 - F_0(x))]$ under the null hypothesis, that is, a squared standard Gaussian random variable times 0.5. To identify an appropriate additive correction term, we utilize a refinement of the LIL (1.4), based on Kolmogorov’s upper class test (cf. [12], or [16], Chapter 1.8). For $t \in (0, 1)$, define

$$C(t) := \log \log \frac{e}{4t(1-t)} = \log(1 - \log(1 - (2t - 1)^2)) \geq 0,$$

$$D(t) := \log(1 + C(t)^2) \in [0, \min\{C(t), C(t)^2\}].$$

Then, for any fixed $\nu > 3/4$,

$$(1.10) \quad T_\nu := \sup_{t \in (0,1)} \left(\frac{\mathbb{U}(t)^2}{2t(1-t)} - C_\nu(t) \right) < \infty$$

almost surely, where $C_\nu := C + \nu D$. Note that $C(t) = C(1 - t)$, $D(t) = D(1 - t)$, and as $t \searrow 0$,

$$C(t) = \log \log(1/t) + O((\log(1/t))^{-1}),$$

$$D(t) = 2 \log \log \log(1/t) + O((\log \log(1/t))^{-1}).$$

This indicates why (1.10) follows from Kolmogorov’s test (see Section S.1), and shows the connection between (1.10) and (1.4). On $(0, 1/2]$, both functions C and D are decreasing with $C(1/2) = D(1/2) = 0$ and

$$\lim_{t \rightarrow 1/2} \frac{C(t)}{(2t - 1)^2} = \lim_{t \rightarrow 1/2} \frac{D(t)}{(2t - 1)^4} = 1.$$

Consequently, we propose the following test statistics:

$$(1.11) \quad T_{n,s,\nu}(F_0) := \begin{cases} \sup_{x: 0 < F_0(x) < 1} [nK_s(\mathbb{F}_n(x), F_0(x)) - C_\nu(\mathbb{F}_n(x), F_0(x))] & \text{if } s > 0, \\ \sup_{x \in [X_{n:1}, X_{n:n}]} [nK_s(\mathbb{F}_n(x), F_0(x)) - C_\nu(\mathbb{F}_n(x), F_0(x))] & \text{if } s \leq 0, \end{cases}$$

where for $t, u \in [0, 1]$,

$$C_\nu(u, t) := \min_{\min(u,t) \leq v \leq \max(u,t)} C_\nu(v) = \begin{cases} C_\nu(\min(u, t)) & \text{if } \min(u, t) > 1/2, \\ C_\nu(\max(u, t)) & \text{if } \max(u, t) < 1/2, \\ 0 & \text{else,} \end{cases}$$

with $C(0), C(1), D(0), D(1) := \infty$. As seen later, using this bivariate version $C_\nu(\mathbb{F}_n(x), F_0(x))$ instead of $C_\nu(F_0(x))$ or $C_\nu(\mathbb{F}_n(x))$ has computational advantages and increases power. The additive correction term $C_\nu(\mathbb{F}_n(x), F_0(x))$ is large only if x is far in the tails of \mathbb{F}_n and of F_0 .

The remainder of this paper is organized as follows:

- In Section 2, we show that under the null hypothesis, the test statistics $T_{n,s,v}(F_0)$ in (1.11) converge in distribution to T_v in (1.10) for any fixed value of $s \in \mathbb{R}$.
- Section 3 discusses statistical implications of this finding. As explained in Section 3.1, goodness-of-fit tests based on $T_{n,s,v}(F_0)$ have desirable asymptotic power. In particular, they are shown to attain a detection boundary of Ingster [15] for Gaussian mixture models. Moreover, even under contiguous alternatives they have nontrivial asymptotic power as opposed to goodness-of-fit tests based on $T_{n,s}^{BJ}$ in (1.6).
- In Section 3.2, we analyze the confidence bands $(A_{n,s,v,\alpha}, B_{n,s,v,\alpha})$ resulting from inversion of the tests $T_{n,s,v}(\cdot)$. It will be shown that these bands have similar accuracy as those of Owen [25] and the bands $(A_{n,s,\alpha}^{BJO}, B_{n,s,\alpha}^{BJO})$ based on $T_{n,s}^{BJ}(\cdot)$ in the tail regions while achieving the usual root- n consistency everywhere. In addition, we compare our bands with the confidence bands of [32], confirming our claim that a purely multiplicative adjustment of $\mathbb{F}_n - F_0$ is necessarily suboptimal in the tail regions.
- Our results for the confidence bands elucidate the impact of the parameter s on these bands for large sample sizes. These considerations are based on new inequalities and expansions for the divergences K_s , which are of independent interest.

All proofs and auxiliary results are deferred to Sections 4, 5 and an online supplement A. References to the latter start with ‘S.’ or ‘(S.’. Essential ingredients for the proofs in Section 4 are tools and techniques of Csörgő et al. [5]. A first version of this paper used a different, more self-contained approach, which is probably of independent interest and outlined in Section S.2. This also includes an alternative proof of (1.10).

2. Limit distributions under the null hypothesis. Recall the uniform empirical process \mathbb{G}_n mentioned in the Introduction. Under the null hypothesis that $F \equiv F_0$, the test statistic $T_{n,s,v}(F_0)$ has the same distribution as

$$(2.1) \quad T_{n,s,v} := \begin{cases} \sup_{t \in (0,1)} [nK_s(\mathbb{G}_n(t), t) - C_v(\mathbb{G}_n(t), t)] & \text{if } s > 0, \\ \sup_{t \in [\xi_{n:1}, \xi_{n:n}]} [nK_s(\mathbb{G}_n(t), t) - C_v(\mathbb{G}_n(t), t)] & \text{if } s \leq 0, \end{cases}$$

where $\xi_{n:1} < \dots < \xi_{n:n}$ are the order statistics of the uniform sample ξ_1, \dots, ξ_n . In particular, the $(1 - \alpha)$ -quantile of $T_{n,s,v}(F_0)$ under the null hypothesis coincides with the $(1 - \alpha)$ -quantile $\kappa_{n,s,v,\alpha}$ of $T_{n,s,v}$. Here is our main result for $T_{n,s,v}$ and $\kappa_{n,s,v,\alpha}$.

THEOREM 2.1. For all $v > 3/4$ and $s \in \mathbb{R}$,

$$T_{n,s,v} \rightarrow_d T_v.$$

Moreover, $\kappa_{n,s,v,\alpha} \rightarrow \kappa_{v,\alpha} > 0$ for any fixed test level $\alpha \in (0, 1)$, where $\kappa_{v,\alpha}$ is the $(1 - \alpha)$ -quantile of T_v .

A key step along the way to proving Theorem 2.1 will be to consider the case $s = 2$ and prove the following theorem for the uniform empirical process $\mathbb{U}_n = \sqrt{n}(\mathbb{G}_n - I)$, where I denotes the distribution function of the uniform distribution on $[0, 1]$.

THEOREM 2.2. For all $v > 3/4$,

$$\tilde{T}_{n,v} := \sup_{t \in (0,1)} \left(\frac{\mathbb{U}_n(t)^2}{2t(1-t)} - C_v(t) \right) \rightarrow_d T_v.$$

REMARK 2.3 (The impact of s and the definition of $T_{n,s,v}$). Note that the parameter s could be an arbitrary real number. However, numerical experiments indicate that the convergence to the asymptotic distribution is very slow if, say, $s < -0.5$ or $s > 1.5$. More precisely, Monte Carlo experiments show that for parameters $s \notin [-0.5, 1.5]$ the test statistics $T_{n,s,v}$ are mainly influenced by just a few very small or very large order statistics. Moreover, if $s \in (0, 0.5]$, one should redefine $T_{n,s,v}$ as a supremum over $[\xi_{n:1}, \xi_{n:n})$ rather than $(0, 1)$. As shown in our proof of Theorem 2.1, this modification does not alter the asymptotic distribution, but for realistic sample sizes n , taking the supremum over the full set $(0, 1)$ for small parameters $s > 0$ leads to distributions, which are mainly influenced by $\xi_{n:1}$.

Tables S.1 and S.2 provide exact critical values $\kappa_{n,s,v,\alpha}$ for various sample sizes n , $s \in \{j/10 : -10 \leq j \leq 20\}$, $v = 1$ and $\alpha = 0.5, 0.1, 0.05, 0.01$.

Similar discrepancies between asymptotic theory and finite sample behavior can be observed for the Berk–Jones quantiles $\kappa_{n,s,\alpha}^{BJ}$ if $s \notin [-0.5, 1.5]$; see Tables S.3 and S.4.

3. Statistical implications.

3.1. *Goodness-of-fit tests.* As explained in the Introduction, we can reject the null hypothesis that F is a given continuous distribution function F_0 at level α if the test statistic $T_{n,s,v}(F_0)$, defined in (1.11), exceeds the $(1 - \alpha)$ -quantile $\kappa_{n,s,v,\alpha}$ of $T_{n,s,v}$. The test statistics $T_{n,s,v}$ and $T_{n,s,v}(F_0)$ can be represented as the maximum of at most $2n$ terms: with $u_{n,i} := i/n$, the statistic $T_{n,s,v}$ equals

$$\max_{1 \leq i \leq n} \max \{ nK_s(u_{n,i-1}, \xi_{n:i}) - C_v(u_{n,i-1}, \xi_{n:i}), nK_s(u_{n,i}, \xi_{n:i}) - C_v(u_{n,i}, \xi_{n:i}) \}$$

if $s > 0$, and

$$\max_{1 \leq i < n} \max \{ nK_s(u_{n,i}, \xi_{n:i}) - C_v(u_{n,i}, \xi_{n:i}), nK_s(u_{n,i}, \xi_{n:i+1}) - C_v(u_{n,i}, \xi_{n:i+1}) \}$$

if $s \leq 0$. The statistic $T_{n,s,v}(F_0)$ can be represented analogously with $F_0(X_{n:i})$ in place of $\xi_{n:i}$. These formulae follow from the fact that for fixed $u \in (0, 1)$, the function $t \mapsto nK_s(u, t) - C_v(u, t)$ is continuous on $(0, 1)$, increasing on $[u, 1)$ and decreasing on $(0, u]$. For $K_s(u, t) = K_{1-s}(t, u)$ is convex in t with minimum at $t = u$ (see (S.12) in Section S.3), and $C_v(u, t)$ is increasing in $t \in (0, u]$ and decreasing in $t \in [u, 1)$. If $s > 0$, these monotonicities are also true for $u \in \{0, 1\}$, precisely,

$$C_v(0, t) = C_v(\min(t, 1/2)) \quad \text{and} \quad K_s(0, t) = \begin{cases} -\log(1 - t) & \text{if } s = 1, \\ ((1 - t)^{1-s} - 1)/(s(s - 1)) & \text{if } s \neq 1, \end{cases}$$

while $C_v(1, t) = C_v(0, 1 - t)$ and $K_s(1, t) = K_s(0, 1 - t)$.

3.1.1. *Noncontiguous alternatives.* Now suppose that the true distribution function of the observations X_i is a continuous distribution function F_n such that $\{x \in \mathbb{R} : 0 < F_n(x) < 1\} \subset \{x \in \mathbb{R} : 0 < F_0(x) < 1\}$. A first question is: under what conditions on the sequence $(F_n)_n$ does our goodness-of-fit test have asymptotic power one for any fixed test level $\alpha \in (0, 1)$. Since $\kappa_{n,s,v,\alpha} \rightarrow \kappa_{v,\alpha} < \infty$, this goal is equivalent to

$$(3.1) \quad P_{F_n}(T_{n,s,v}(F_0) > \kappa) \rightarrow 1 \quad \text{for any fixed } \kappa > 0.$$

To verify this property, the following function $\Delta_n : \mathbb{R} \rightarrow [0, \infty)$ plays a key role:

$$\Delta_n := \frac{\sqrt{n}|F_n - F_0|}{\min\{H_n(F_n), H_n(F_0)\}} \quad \text{with } H_n(t) := \sqrt{(1 + C(t))t(1 - t)} + \frac{1 + C(t)}{\sqrt{n}}$$

for $t \in [0, 1]$ with the conventions $C(t) := \infty$ and $C(t)t(1 - t) := 0$ for $t \in \{0, 1\}$.

THEOREM 3.1. *Suppose that the sequence $(F_n)_n$ satisfies the condition*

$$(3.2) \quad \sup_{x \in \mathbb{R}} \Delta_n(x) \rightarrow \infty.$$

Then (3.1) holds true for any $s \in [-1, 2]$.

It follows immediately from this theorem that (3.1) is satisfied whenever $F_n \equiv F_*$ for all sample sizes n , where $F_* \neq F_0$.

As a litmus test for our procedures and Theorem 3.1, we consider a testing problem studied in detail by [15]. The null hypothesis is given by $F_0 = \Phi$, the standard Gaussian distribution function, whereas

$$F_n(x) := (1 - \epsilon_n)\Phi(x) + \epsilon_n\Phi(x - \mu_n),$$

for certain numbers $\epsilon_n \in (0, 1)$ and $\mu_n > 0$. By means of Theorem 3.1, one can derive the following result.

COROLLARY 3.2.

(a) *Suppose that $\epsilon_n = n^{-\beta+o(1)}$ for some fixed $\beta \in (1/2, 1)$. Furthermore, let $\mu_n = \sqrt{2r \log n}$ for some $r \in (0, 1)$. Then (3.1) is satisfied for any $s \in [-1, 2]$ if*

$$r > \begin{cases} \beta - 1/2 & \text{if } \beta \in (1/2, 3/4], \\ (1 - \sqrt{1 - \beta})^2 & \text{if } \beta \in [3/4, 1). \end{cases}$$

(b) *Suppose that $\epsilon_n = n^{-1/2+o(1)}$ such that $\pi_n := \sqrt{n}\epsilon_n \rightarrow 0$. Then (3.1) is satisfied for any $s \in [-1, 2]$ if $\mu_n = \sqrt{2\lambda \log(1/\pi_n)}$ for some $\lambda > 1$.*

As explained by [15], any goodness-of-fit test at fixed level $\alpha \in (0, 1)$ has trivial asymptotic power α whenever $\epsilon_n = n^{-\beta}$ for some $\beta \in (1/2, 1)$ and $\mu_n = \sqrt{2r \log n}$ with

$$r < \begin{cases} \beta - 1/2 & \text{if } \beta \in (1/2, 3/4], \\ (1 - \sqrt{1 - \beta})^2 & \text{if } \beta \in [3/4, 1). \end{cases}$$

Thus, part (a) of the previous corollary shows that our new family of tests achieves this detection boundary, as do the goodness-of-fit tests of [6], [18] and [14].

A connection between parts (a) and (b) of Corollary 3.2 can be seen as follows: let $\epsilon_n = n^{-\beta}$ for some fixed $\beta \in (1/2, 3/4]$, and $\mu_n = \sqrt{2r \log(n)}$ for some $r > \beta - 1/2$. Then $r = \lambda(\beta - 1/2)$ for some $\lambda > 1$, and with $\pi_n = \sqrt{n}\epsilon_n = n^{1/2-\beta}$, we may write $\sqrt{2r \log(n)} = \sqrt{2\lambda \log(1/\pi_n)}$.

3.1.2. *Contiguous alternatives.* Suppose that the distribution functions F_0 and F_n have densities f_0 and f_n , respectively, with respect to some continuous measure Λ on \mathbb{R} such that, for some function a ,

$$(3.3) \quad \sqrt{n}(f_n^{1/2} - f_0^{1/2}) \rightarrow 2^{-1}af_0^{1/2} \quad \text{in } L_2(\Lambda).$$

Then it follows easily that $a \in L_2(F_0)$, $\int adF_0 = 0$ and

$$\sqrt{n}(F_n - F_0)(t) \rightarrow A(t) := \int_{-\infty}^t adF_0 \quad \text{uniformly in } t \in \mathbb{R}.$$

Furthermore, since $\int_{-\infty}^t adF_0 = \int_{\mathbb{R}} (1_{[x \leq t]} - F_0(t))a(x) dF_0(x)$, the Cauchy–Schwarz inequality yields that

$$(3.4) \quad |A(t)| \leq \sqrt{F_0(t)(1 - F_0(t))} \|a\|_{L_2(F_0)}.$$

LEMMA 3.3 (Power of “tail-dominated” tests under contiguous alternatives). *Let (φ_n) be a sequence of tests with the following two properties:*

(i) *For a fixed level $\alpha \in (0, 1)$,*

$$E_{F_0} \varphi_n(X_1, \dots, X_n) \rightarrow \alpha.$$

(ii) *For any fixed $0 < \rho < 1/2$ and $x_\rho := F_0^{-1}(\rho)$, $y_\rho := F_0^{-1}(1 - \rho)$, there exists a test $\varphi_{n,\rho}$ depending only on $(\mathbb{F}_n(x))_{x \notin [x_\rho, y_\rho]}$ such that*

$$P_{F_0}(\varphi_n \neq \varphi_{n,\rho}) \rightarrow 0.$$

Then under assumption (3.3),

$$\limsup_{n \rightarrow \infty} E_{F_n} \varphi_n(X_1, \dots, X_n) \leq \alpha.$$

Note that the Berk–Jones tests with $T_{n,s}^{\text{BJ}}(F_0)$ satisfy the assumptions of Lemma 3.3, if tuned to have asymptotic level α . For all of them involve a test statistic of the type,

$$T_n(F_0) = \sup_{x \in \mathbb{R}} \Gamma_n(\mathbb{F}_n(x)),$$

with a function $\Gamma_n : \mathbb{R} \rightarrow [0, \infty]$ such that under the null hypothesis,

$$\sup_{x \in \mathbb{R}} \Gamma_n(\mathbb{F}_n(x)) \rightarrow_p \infty,$$

but for any $0 < \rho < 1/2$,

$$\sup_{x \in [x_\rho, y_\rho]} \Gamma_n(\mathbb{F}_n(x)) = O_p(1).$$

Hence, $T_n(F_0)$ equals

$$T_n^{(\rho)}(F_0) := \sup_{x \notin [x_\rho, y_\rho]} \Gamma_n(\mathbb{F}_n(x))$$

with asymptotic probability one. Thus, we may replace the test statistic $T_n(F_0)$ with $T_n^{(\rho)}(F_0)$ while keeping the critical value.

By way of contrast, the goodness-of-fit test based on $T_{n,s,v}(F_0)$ has nontrivial asymptotic power in the present setting.

THEOREM 3.4 (Power of new tests under contiguous alternatives). *In the setting (3.3), the test statistic $T_{n,s,v}(F_0)$ converges in distribution to*

$$T_v(A) := \sup_{t \in (0,1)} \left(\frac{(\mathbb{U}(t) + A(F_0^{-1}(t)))^2}{2t(1-t)} - C_v(t) \right).$$

In particular,

$$P_{F_n}[T_{n,s,v}(F_0) \geq \kappa_{n,s,v,\alpha}] \rightarrow P[T_v(A) \geq \kappa_{v,\alpha}] \geq \alpha.$$

Concerning the impact of A ,

$$P[T_v(A) \geq \kappa_{v,\alpha}] \rightarrow 1 \quad \text{as} \quad \sup_{t \in (0,1)} \left(\frac{|A(F_0^{-1}(t))|}{\sqrt{2t(1-t)}} - \sqrt{C(t)} \right) \rightarrow \infty.$$

3.2. *Confidence bands.* The confidence bands of [25], defined in terms of $K = K_1$, may be generalized to arbitrary fixed $s \in [-1, 2]$, but we restrict our attention to $s \in (0, 2]$, because for $s \leq 0$ and a large range of sample sizes n , the resulting bands would focus mainly on small regions in the tails and be rather wide elsewhere. With confidence $1 - \alpha$ we may claim that $\sup_{x: 0 < F(x) < 1} nK_s(\mathbb{F}_n(x), F(x))$ does not exceed the $(1 - \alpha)$ -quantile $\kappa_{n,s,\alpha}^{BJ}$ of $\sup_{t \in (0,1)} nK_s(\mathbb{G}_n(t), t)$. As explained in Section S.6, inverting the inequality $nK_s(\mathbb{F}_n(x), F(x)) \leq \kappa_{n,s,\alpha}^{BJ}$ for fixed x with respect to $F(x)$ reveals that for $0 \leq i \leq n$ and $X_{n:i} \leq x < X_{n:i+1}$,

$$F(x) \in [A_{n,s,\alpha}^{BJO}(x), B_{n,s,\alpha}^{BJO}(x)] = [a_{n,s,\alpha,i}^{BJO}, b_{n,s,\alpha,i}^{BJO}],$$

where $a_{n,s,\alpha,i}^{BJO} \leq u_{n,i} \leq b_{n,s,\alpha,i}^{BJO}$ are given by $a_{n,s,\alpha,0}^{BJO} := 0, b_{n,s,\alpha,n}^{BJO} := 1$ and for $0 \leq i < n$,

$$b_{n,s,\alpha,i}^{BJO} := \max\{t \in (u_{n,i}, 1] : nK_s(u_{n,i}, t) \leq \kappa_{n,s,\alpha}^{BJ}\},$$

$$a_{n,s,\alpha,n-i}^{BJO} := 1 - b_{n,s,\alpha,i}^{BJO}.$$

Thus, computing the confidence band $(A_{n,s,\alpha}^{BJO}, B_{n,s,\alpha}^{BJO})$ boils down to determining the $2(n + 1)$ numbers $a_{n,s,\alpha,i}^{BJO}$ and $b_{n,s,\alpha,i}^{BJO}, 0 \leq i \leq n$.

Our new method is analogous: with confidence $1 - \alpha$, for $0 \leq i \leq n$ and $X_{n:i} \leq x < X_{n:i+1}$, the value $F(x)$ is contained in

$$[A_{n,s,v,\alpha}(x), B_{n,s,v,\alpha}(x)] = [a_{n,s,v,\alpha,i}, b_{n,s,v,\alpha,i}],$$

where $a_{n,s,v,\alpha,0} := 0, b_{n,s,v,\alpha,n} := 1$ and for $0 \leq i < n$,

$$b_{n,s,v,\alpha,i} := \max\{t \in (u_{n,i}, 1] : nK(u_{n,i}, t) - C_v(u_{n,i}, t) \leq \kappa_{n,s,v,\alpha}\},$$

$$a_{n,s,v,\alpha,n-i} := 1 - b_{n,s,v,\alpha,i}.$$

To understand the asymptotic performance of these confidence bands properly, we need auxiliary functions $a_s, b_s : [0, \infty) \rightarrow [0, \infty)$. Note first that for any $s \in [-1, 2]$, $K_s(u, t)$ in (1.7) may be represented as

$$(3.5) \quad K_s(u, t) = t\phi_s(u/t) + (1 - t)\phi_s[(1 - u)/(1 - t)],$$

where

$$(3.6) \quad \phi_s(x) = \begin{cases} (x^s - sx + s - 1)/[s(s - 1)] & s \neq 0, 1, \\ x \log x - x + 1 & s = 1, \\ x - 1 - \log x & s = 0, \end{cases}$$

for $x \in (0, \infty)$, and $\phi_s(0) := \lim_{x \searrow 0} \phi_s(x)$ equals $1/s^+$. If u and t are close to 0, one may approximate $K_s(u, t)$ by

$$H_s(u, t) := t\phi_s(u/t).$$

The properties of $H_s : [0, \infty) \times (0, \infty) \rightarrow [0, \infty]$ are treated in Lemma S.13. In particular, it is shown that

$$a_s(x) := \begin{cases} 0 & \text{if } x = 0, \\ \inf\{y \in (0, x) : H_s(x, y) \leq 1\} & \text{else,} \end{cases}$$

$$b_s(x) := \begin{cases} s^+ & \text{if } x = 0, \\ \max\{y > x : H_s(x, y) \leq 1\} & \text{else,} \end{cases}$$

define continuous functions $a_s, b_s : [0, \infty) \rightarrow [0, \infty)$, where a_s is convex with $a_s(0) = 0 = a_s'(0)$, $a_s(x) = 0$ if and only if $x \leq (1 - s)^+$, and b_s is concave. Moreover, $a_s(x) = x - \sqrt{2x} +$

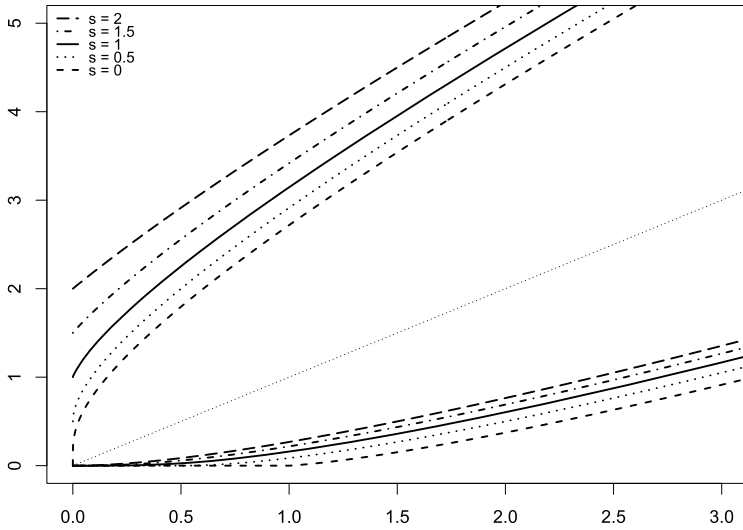


FIG. 1. The auxiliary functions a_s (below diagonal), b_s (above diagonal) for $s \in \{0, 0.5, 1, 1.5, 2\}$.

$O(1)$ and $b_s(x) = x + \sqrt{2x} + O(1)$ as $x \rightarrow \infty$. Finally, for fixed $x > 0$, $a_s(x)$ and $b_s(x)$ are nondecreasing in $s \in [-1, 2]$ with $a_s(x) < x < b_s(x)$. Figure 1 depicts these functions a_s, b_s on the interval $[0, 3]$ for $s \in \{0, 0.5, 1, 1.5, 2\}$.

Our first result shows that the confidence bands $(A_{n,s,\alpha}^{\text{BJO}}, B_{n,s,\alpha}^{\text{BJO}})$ and $(A_{n,s,v,\alpha}, B_{n,s,v,\alpha})$ are asymptotically equivalent in the tail regions, that is, for $\mathbb{F}_n(x)$ close to zero or close to one. Moreover, the test level α is asymptotically irrelevant there, but the parameter s does play a role when $\min\{\mathbb{F}_n(x), 1 - \mathbb{F}_n(x)\} \leq O(n^{-1} \log \log n)$.

THEOREM 3.5. *Let $\gamma_n := n^{-1} \log \log n$. For any fixed $s \in (0, 2]$, $v > 3/4$ and $\delta \in (0, 1)$,*

$$\left. \begin{array}{l} u_{n,i} - a_{n,s,\alpha,i}^{\text{BJO}} \\ u_{n,i} - a_{n,s,v,\alpha,i} \\ b_{n,s,\alpha,n-i}^{\text{BJO}} - u_{n,n-i} \\ b_{n,s,v,\alpha,n-i} - u_{n,n-i} \end{array} \right\} = \gamma_n (i / \log \log n - a_s(i / \log \log n))(1 + o(1))$$

and

$$\left. \begin{array}{l} b_{n,s,\alpha,i}^{\text{BJO}} - u_{n,i} \\ b_{n,s,v,\alpha,i} - u_{n,i} \\ u_{n,n-i} - a_{n,s,\alpha,n-i}^{\text{BJO}} \\ u_{n,n-i} - a_{n,s,v,\alpha,n-i} \end{array} \right\} = \gamma_n (b_s(i / \log \log n) - i / \log \log n)(1 + o(1)),$$

uniformly in $i \in \{0, 1, \dots, n\} \cap [0, n^\delta]$.

REMARK 3.6 (Choice of s). Concerning the choice of s , Theorem 3.5 shows that smaller (resp., larger) values of s lead to better upper (resp., lower) and worse lower (resp., upper) bounds for $F(x)$ in the left tail and better lower (resp., upper) and worse upper (resp., lower) bounds for $F(x)$ in the right tail. The choice $s = 1$ seems to be a good compromise; see also the numerical examples later.

The next result shows that in the central region, the parameter s is asymptotically irrelevant, and the width of the band $(A_{n,s,v,\alpha}, B_{n,s,v,\alpha})$ is of smaller order than the width of $(A_{n,s,\alpha}^{\text{BJO}}, B_{n,s,\alpha}^{\text{BJO}})$.

THEOREM 3.7. For any fixed $s \in (0, 2]$, $\nu > 3/4$ and $\delta \in (0, 1)$,

$$\left. \begin{aligned} u_{n,i} - a_{n,s,\alpha,i}^{\text{BJO}} \\ b_{n,s,\alpha,i}^{\text{BJO}} - u_{n,i} \end{aligned} \right\} = \sqrt{2\gamma_n u_{n,i}(1-u_{n,i})(1+o(1))},$$

$$\left. \begin{aligned} u_{n,i} - a_{n,s,\nu,\alpha,i} \\ b_{n,s,\nu,\alpha,i} - u_{n,i} \end{aligned} \right\} = \sqrt{2\gamma_{n,\nu,\alpha}(u_{n,i})u_{n,i}(1-u_{n,i})(1+o(1))},$$

uniformly in $i \in \{0, 1, \dots, n\} \cap [n^\delta, n - n^\delta]$, where $\gamma_n = n^{-1} \log \log n$ and $\gamma_{n,\nu,\alpha}(u) := n^{-1}(C_\nu(u) + \kappa_{\nu,\alpha})$.

Note that $(C_\nu(u) + \kappa_{\nu,\alpha})u(1-u) \rightarrow 0$ as $u \rightarrow \{0, 1\}$. Thus, one can deduce from Theorems 3.5 and 3.7 that

$$\max_{i=0,1,\dots,n} (b_{n,i}^{\text{BJO}} - u_{n,i}) = \max_{i=0,1,\dots,n} (u_{n,i} - a_{n,i}^{\text{BJO}}) = \sqrt{\gamma_n/2}(1+o(1)),$$

$$\max_{i=0,1,\dots,n} (b_{n,i} - u_{n,i}) = \max_{i=0,1,\dots,n} (u_{n,i} - a_{n,i}) = O(n^{-1/2}).$$

REMARK 3.8 (Comparison with Stepanova–Pavlenko [32]). The confidence band $(A_{n,\alpha}^{\text{SP}}, B_{n,\alpha}^{\text{SP}})$ of [32] with the test statistic $T_n^{\text{SP}}(\cdot)$ in (1.9) can be represented as follows: for $0 \leq i \leq n$ and $X_{n:i} \leq x < X_{n:i+1}$,

$$[A_{n,\alpha}^{\text{SP}}(x), B_{n,\alpha}^{\text{SP}}(x)] = [a_{n,\alpha,i}^{\text{SP}}, b_{n,\alpha,i}^{\text{SP}}],$$

where $a_{n,\alpha,0}^{\text{SP}} = 0$, $b_{n,\alpha,0}^{\text{SP}} = b_{n,\alpha,1}^{\text{SP}}$, $a_{n,\alpha,n}^{\text{SP}} = a_{n,\alpha,n-1}^{\text{SP}}$, $b_{n,\alpha,n}^{\text{SP}} = 1$, and for $1 \leq i < n$,

$$[a_{n,\alpha,i}^{\text{SP}}, b_{n,\alpha,i}^{\text{SP}}] = [u_{n,i} \pm n^{-1/2} \kappa_{n,\alpha}^{\text{SP}} \sqrt{u_{n,i}(1-u_{n,i})h(u_{n,i})}] \cap [0, 1].$$

Recall that $h(t) = \log \log(1/[t(1-t)])$. Here, $\kappa_{n,\alpha}^{\text{SP}}$ is the $(1-\alpha)$ -quantile of $T_{n,\alpha}^{\text{SP}}(F_0)$ in case of $F \equiv F_0$, and it converges to the $(1-\alpha)$ -quantile $\kappa_\alpha^{\text{SP}}$ of

$$\sup_{t \in (0,1)} \frac{|\mathbb{U}(t)|}{\sqrt{t(1-t)h(t)}}.$$

Consequently, for fixed $s \in (0, 2]$, $\nu > 3/4$ and $\delta \in (0, 1)$,

$$\frac{b_{n,\alpha,i}^{\text{SP}} - u_{n,i}}{b_{n,s,\nu,\alpha,i} - u_{n,i}}, \frac{u_{n,i} - a_{n,\alpha,i}^{\text{SP}}}{u_{n,i} - a_{n,s,\nu,\alpha,i}} = \frac{\kappa_\alpha^{\text{SP}} \sqrt{h(u_{n,i})}}{\sqrt{2(C_\nu(u_{n,i}) + \kappa_{\nu,\alpha})}}(1+o(1))$$

uniformly in $i \in \{0, 1, \dots, n\} \cap [n^\delta, n - n^\delta]$. But

$$\lim_{u \rightarrow \{0,1\}} \frac{\kappa_\alpha^{\text{SP}} \sqrt{h(u)}}{\sqrt{2(C_\nu(u) + \kappa_{\nu,\alpha})}} = \frac{\kappa_\alpha^{\text{SP}}}{\sqrt{2}} \begin{cases} \geq 1, \\ \rightarrow \infty \text{ as } \alpha \searrow 0, \end{cases}$$

because $h(t)/\log \log(1/t)$ and $C_\nu(t)/\log \log(1/t)$ converge to 1 as $t \searrow 0$. Thus, the confidence band $(A_{n,\alpha}^{\text{SP}}, B_{n,\alpha}^{\text{SP}})$ is asymptotically wider than $(A_{n,s,\nu,\alpha}, B_{n,s,\nu,\alpha})$ in the tail regions for sufficiently small α .

Note that these considerations apply to any choice of the continuous function $h : (0, 1) \rightarrow (0, \infty)$ in (1.9) as long as $h(t)/\log \log(1/t) \rightarrow 1$ as $t \searrow 0$.

REMARK 3.9 (Bahadur and Savage [2] revisited). On $(-\infty, X_{n:1}]$, the upper confidence bounds for F are constant $b_{n,s,\alpha,1}^{\text{BJO}}$ or $b_{n,s,\nu,\alpha,1}$, and this is of order $O(n^{-1} \log \log n)$. Likewise, on $(X_{n:n}, \infty)$, the lower confidence bounds for F are constant $1 - b_{n,s,\alpha,1}^{\text{BJO}}$ or $1 - b_{n,s,\nu,\alpha,1}$. Interestingly, for any $(1-\alpha)$ -confidence band for a continuous distribution function F ,

the upper bound has to be greater than c/n with asymptotic probability at least $e^c\alpha$, and the lower bound has to be smaller than $1 - c/n$ with asymptotic probability at least $e^c\alpha$. This follows from a quantitative version of Theorem 2 of [2], stated as Theorem 3.10 below.

It is also instructive to consider Daniels’ lower confidence bound for a continuous distribution function F , namely

$$P_F(\alpha\mathbb{F}_n(x) \leq F(x) \text{ for all } x \in \mathbb{R}) = 1 - \alpha.$$

THEOREM 3.10. *Let \mathcal{F} be a family of continuous distribution functions, which is convex and closed under translations, that is, $F(\cdot - \mu) \in \mathcal{F}$ for all $F \in \mathcal{F}$ and $\mu \in \mathbb{R}$. Let (A_n, B_n) be a $(1 - \alpha)$ -confidence band for $F \in \mathcal{F}$. Then, for any $F \in \mathcal{F}$ and $\epsilon \in (0, 1)$,*

$$P_F\left(\inf_{x \in \mathbb{R}} B_n(x) < \epsilon\right) \leq (1 - \epsilon)^{-n}\alpha \quad \text{and} \quad P_F\left(\sup_{x \in \mathbb{R}} A_n(x) > 1 - \epsilon\right) \leq (1 - \epsilon)^{-n}\alpha.$$

In our context, \mathcal{F} would be the family of all continuous distribution functions. But the precision bounds in Theorem 3.10 apply to much smaller families \mathcal{F} already, for instance, the family of all convex combinations of $F_o(\cdot - \mu)$, $\mu \in \mathbb{R}$, where F_o is an arbitrary continuous distribution function. For the reader’s convenience, a proof of Theorem 3.10 is provided in Section S.5.

EXAMPLE 3.11 ($s = 1$). The left panel in Figure 2 depicts, for $n = 100$, the 95%-confidence band $(A_{n,1,1,\alpha}, B_{n,1,1,\alpha})$ in case of an idealized standard Gaussian sample with order statistics $X_{n:i} = \Phi^{-1}(i/(n + 1))$. In addition, one sees the Kolmogorov–Smirnov 95%-confidence band $(A_{n,\alpha}^{KS}, B_{n,\alpha}^{KS})$. In the right panel, one sees for the same setting the centered upper bounds $B_{n,1,1,\alpha} - \mathbb{F}_n$, $B_{n,1,\alpha}^{BJO} - \mathbb{F}_n$ and $B_{n,\alpha}^{KS} - \mathbb{F}_n$. Note that a plot of the centered lower bounds $A_{n,1,1,\alpha} - \mathbb{F}_n$, $A_{n,1,\alpha}^{BJO} - \mathbb{F}_n$ and $A_{n,\alpha}^{KS} - \mathbb{F}_n$ would be the reflection of the plots for the centered upper bounds with respect to the point $(0, 0)$. The corresponding critical values $\kappa_{n,1,1,\alpha}$, $\kappa_{n,1,\alpha}^{BJ}$ and $\kappa_{n,\alpha}^{KS}$ have been computed numerically; see Section S.7.

Figure 3 shows the same as the right panel in Figure 2, but with sample sizes $n = 500$ and $n = 4000$ in the left and right panel, respectively.

In the online supplement A, these bands $(A_{n,1,1,\alpha}, B_{n,1,1,\alpha})$ are also compared with the confidence bands of [32], confirming the purely asymptotic result in Remark 3.8.

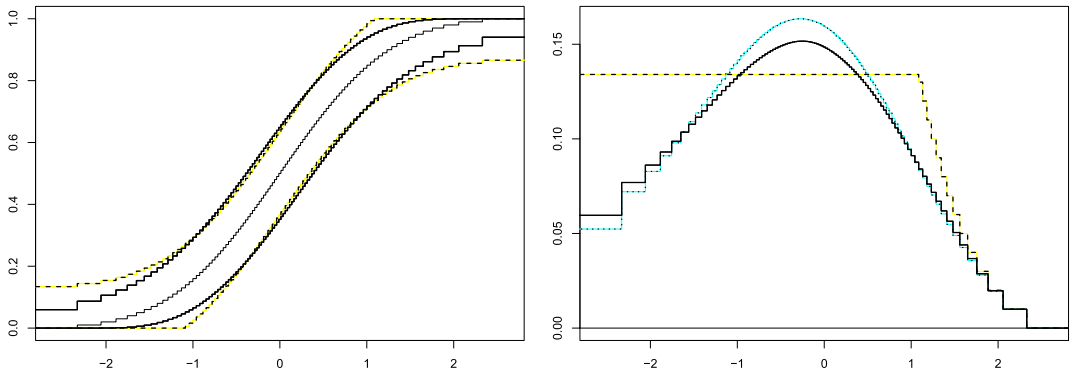


FIG. 2. 95%-confidence bands for $n = 100$. Left panel: $(A_{n,1,1,\alpha}, B_{n,1,1,\alpha})$ (solid) and $(A_{n,\alpha}^{KS}, B_{n,\alpha}^{KS})$ (dashed). Right panel: centered upper bounds $B_{n,1,1,\alpha} - \mathbb{F}_n$ (solid), $B_{n,1,\alpha}^{BJO} - \mathbb{F}_n$ (dotted) and $B_{n,\alpha}^{KS} - \mathbb{F}_n$ (dashed).

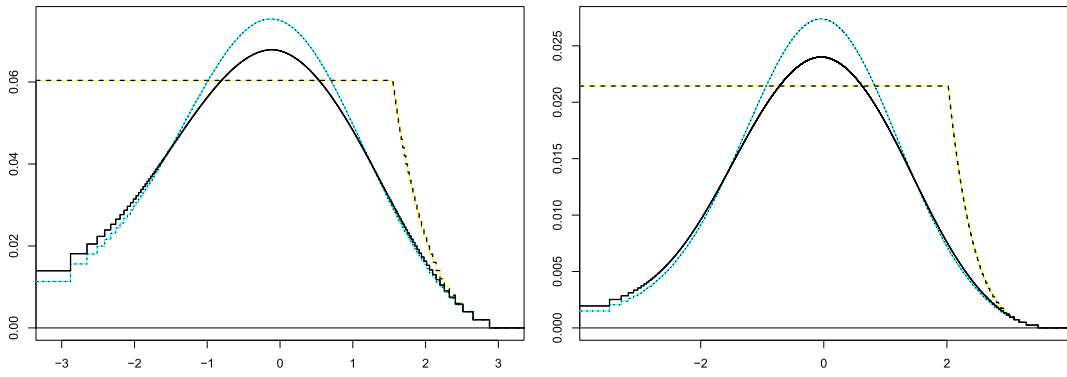


FIG. 3. Centered upper 95%-confidence bounds $B_{n,1,1,\alpha} - \mathbb{F}_n$ (solid), $B_{n,1,\alpha}^{\text{BJO}} - \mathbb{F}_n$ (dotted) and $B_{n,\alpha}^{\text{KS}} - \mathbb{F}_n$ (dashed) for $n = 500$ (left panel) and $n = 4000$ (right panel).

EXAMPLE 3.12 (The impact of s). Figure 4 shows for an idealized Gaussian sample of size $n = 500$, the centered upper 95%-confidence bounds $B_{n,s,1,\alpha} - \mathbb{F}_n$ for $s = 0.6, 1, 1.4$ (left panel) as well as the differences $B_{n,s,1,\alpha} - B_{n,1,1,\alpha}$ for $s = 0.6, 1.4$, right panel. As predicted by Theorem 3.5, the upper bounds $B_{n,s,1}(x)$ are increasing in s for small values of x and decreasing in s for large values of x . The online supplement A contains further plots illustrating the impact of s on our bands. These plots support our claim that choosing s close to 1 is preferable. Other values of s increase the bands' precision somewhere in the tails, but lead to a substantial loss of precision in the central region.

REMARK 3.13 (Discontinuous distribution functions). In the previous considerations, we focused on continuous distribution functions F , and all confidence bands $(A_{n,\alpha}, B_{n,\alpha})$ for F we considered are of the form

$$[A_{n,\alpha}(x), B_{n,\alpha}(x)] = [a_{n,\alpha,i}, b_{n,\alpha,i}] \quad \text{for } x \in [X_{n:i}, X_{n:i+1}) \text{ and } 0 \leq i \leq n$$

with certain numbers $a_{n,\alpha,i}, b_{n,\alpha,i} \in [0, 1]$. Interestingly, such a band has coverage probability at least $1 - \alpha$ for arbitrary, not necessarily continuous distribution functions F ; see Section S.6.

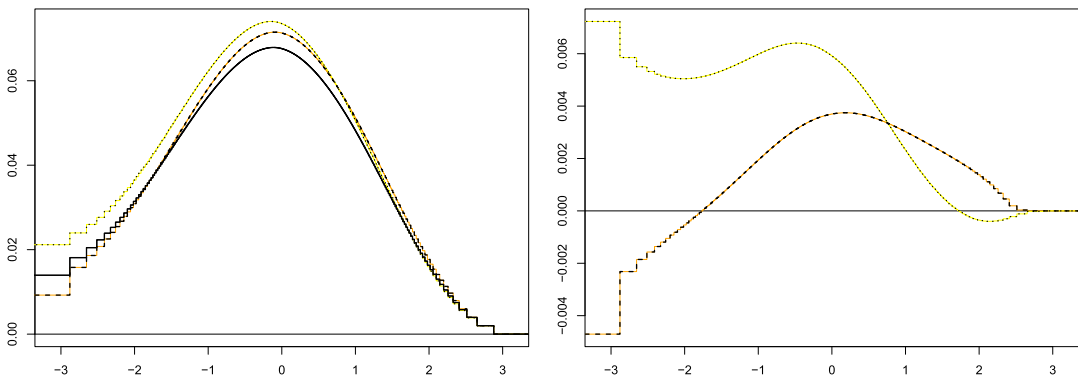


FIG. 4. Upper 95%-confidence bounds for $n = 500$. Left panel: centered bounds $B_{n,s,1,\alpha} - \mathbb{F}_n$ for $s = 0.6$ (dashed), $s = 1.0$ (solid) and $s = 1.4$ (dotted). Right panel: differences $B_{n,s,1,\alpha} - B_{n,1,1,\alpha}$ for $s = 0.6$ (dashed) and $s = 1.4$ (dotted).

4. Proofs for Section 2.

4.1. *Proof of Theorem 2.2.* The following three facts are our essential ingredients.

FACT 4.1 ([5], Theorem 2.2 and Corollary 2.1). There exist on a common probability space a sequence of i.i.d. $U(0, 1)$ random variables $\xi_1, \xi_2, \xi_3, \dots$ and a sequence of Brownian bridge processes $\mathbb{U}^{(1)}, \mathbb{U}^{(2)}, \mathbb{U}^{(3)}, \dots$ such that, for all $0 \leq \delta < 1/4$,

$$\sup_{t \in [1/n, 1-1/n]} \frac{n^\delta |\mathbb{U}_n(t) - \mathbb{U}^{(n)}(t)|}{(t(1-t))^{1/2-\delta}} = O_p(1).$$

FACT 4.2 ([5], Theorem 4.4.1).

$$\sup_{t \in (0,1)} \frac{\mathbb{U}_n(t)^2}{2t(1-t) \log \log n} \rightarrow_p 1.$$

FACT 4.3 ([5], Lemma 4.4.4). For any $1 \leq d_n \leq n$ such that $d_n/n \rightarrow 0$ and $d_n \rightarrow \infty$,

$$\sup_{t \in (0, d_n/n]} \frac{\mathbb{U}_n(t)^2}{2t(1-t) \log \log d_n} \rightarrow_p 1.$$

The same holds with the supremum over $[1 - d_n/n, 1)$.

The asymptotic distribution of $\tilde{T}_{n,v}$ will be derived from the subsequent Lemmas 4.4, 4.5 and 4.6.

LEMMA 4.4. For any sequence of constants $1 \leq d_n \leq n$ such that $d_n/n \rightarrow 0$ and $d_n \rightarrow \infty$ and any choice of $0 < \delta < 1/4$,

$$\sup_{t \in [d_n/n, 1-d_n/n]} \frac{|\mathbb{U}_n(t)^2 - \mathbb{U}^{(n)}(t)^2|}{t(1-t)} = O_p(d_n^{-\delta} (\log \log n)^{1/2}).$$

PROOF. By Fact 4.1, for $0 < \delta < 1/4$,

$$\sup_{t \in [d_n/n, 1-d_n/n]} \frac{|\mathbb{U}_n(t) - \mathbb{U}^{(n)}(t)|}{(t(1-t))^{1/2}} \leq O(d_n^{-\delta}) \sup_{t \in [1/n, 1-1/n]} \frac{n^\delta |\mathbb{U}_n(t) - \mathbb{U}^{(n)}(t)|}{(t(1-t))^{1/2-\delta}} = O_p(d_n^{-\delta}).$$

Together with Fact 4.2 and (1.4), this implies that

$$\begin{aligned} & \sup_{t \in [d_n/n, 1-d_n/n]} \frac{|\mathbb{U}_n(t)^2 - \mathbb{U}^{(n)}(t)^2|}{t(1-t)} \\ & \leq \sup_{t \in [d_n/n, 1-d_n/n]} \frac{|\mathbb{U}_n(t) - \mathbb{U}^{(n)}(t)|}{(t(1-t))^{1/2}} \cdot \left(\frac{|\mathbb{U}_n(t)|}{(t(1-t))^{1/2}} + \frac{|\mathbb{U}^{(n)}(t)|}{(t(1-t))^{1/2}} \right) \\ & = O_p(d_n^{-\delta} (\log \log n)^{1/2}). \end{aligned} \quad \square$$

LEMMA 4.5. For all $v \geq 0$,

$$\sup_{t \in (0, n^{-1} \log n]} \left(\frac{\mathbb{U}_n(t)^2}{2t(1-t)} - C_v(t) \right) \rightarrow_p -\infty.$$

The same holds with the supremum over $(0, n^{-1} \log n]$ replaced by $[1 - n^{-1} \log n, 1)$.

PROOF. Note that with $d_n = \log n$,

$$(4.1) \quad \sup_{t \in (0, d_n/n]} \left(\frac{\mathbb{U}_n(t)^2}{2t(1-t)} - C_v(t) \right) \leq \sup_{t \in (0, d_n/n]} \left(\frac{\mathbb{U}_n(t)^2}{2t(1-t)} - C(d_n/n) \right)$$

since $C_v \geq C$ and C is nonincreasing. By Fact 4.3,

$$\sup_{t \in (0, d_n/n]} \frac{|\mathbb{U}_n(t)^2|}{2t(1-t) \log \log \log n} \rightarrow_p 1,$$

while

$$\frac{C(d_n/n)}{\log \log \log n} = \frac{(1 + o(1)) \log \log n}{\log \log \log n} \rightarrow \infty.$$

Thus, the right-hand side of (4.1) can be written as

$$\begin{aligned} & \sup_{t \in (0, d_n/n]} \left(\frac{\mathbb{U}_n(t)^2}{2t(1-t) \log \log \log n} \cdot \log \log \log n - C(d_n/n) \right) \\ &= \sup_{t \in (0, d_n/n]} \left(\frac{\mathbb{U}_n(t)^2}{2t(1-t) \log \log \log n} - \frac{C(d_n/n)}{\log \log \log n} \right) \log \log \log n \\ &\rightarrow_p (1 - \infty) \cdot \infty = -\infty. \end{aligned} \quad \square$$

LEMMA 4.6. For any fixed $v > 3/4$,

$$\sup_{t \in (0, \rho] \cup [1-\rho, 1)} \left(\frac{\mathbb{U}(t)^2}{2t(1-t)} - C_v(t) \right) \rightarrow -\infty \quad \text{almost surely as } \rho \searrow 0.$$

PROOF. Recall that

$$T_v = \sup_{t \in (0, 1)} \left(\frac{\mathbb{U}(t)^2}{2t(1-t)} - C(t) - vD(t) \right)$$

is finite almost surely for any $v > 3/4$. If we choose $v' \in (3/4, v)$ and write $vD(t) = v'D(t) + (v - v')D(t)$, then we see that for any $\rho \in (0, 1/2]$,

$$\begin{aligned} \sup_{t \in (0, \rho] \cup [1-\rho, 1)} \left(\frac{\mathbb{U}(t)^2}{2t(1-t)} - C(t) - vD(t) \right) &\leq \sup_{t \in (0, \rho] \cup [1-\rho, 1)} (T_{v'} - (v - v')D(t)) \\ &= T_{v'} - (v - v')D(\rho), \end{aligned}$$

because $D(\cdot)$ is symmetric around $1/2$ and monotone decreasing on $(0, 1/2]$. Now the claim follows from $T_{v'} < \infty$ almost surely and $D(\rho) \rightarrow \infty$ as $\rho \searrow 0$. \square

Now we can complete the proof of Theorem 2.2. According to Lemmas 4.5 and 4.6, with $d_n := \log n$,

$$\left. \begin{matrix} \tilde{T}_{n,v} \\ T_v \end{matrix} \right\} = \sup_{t \in [d_n/n, 1-d_n/n]} \left(\frac{1}{2t(1-t)} \left\{ \begin{matrix} \mathbb{U}_n(t)^2 \\ \mathbb{U}(t)^2 \end{matrix} \right\} - C_v(t) \right)$$

with asymptotic probability one. If we replace the Brownian bridge \mathbb{U} with the Brownian bridge $\mathbb{U}^{(n)}$, then Lemma 4.4 implies that the latter two suprema over $[d_n/n, 1-d_n/n]$ differ only by $o_p(1)$. Consequently, $\tilde{T}_{n,v}$ converges in distribution to T_v .

4.2. *Proof of Theorem 2.1.* Note first that in case of $s > 0$,

$$\sup_{t \in (0, \xi_{n:1})} (nK_s(\mathbb{G}_n(t), t) - C_v(\mathbb{G}_n(t), t)) = nK_s(0, \xi_{n:1}) - C_v(\min(\xi_{n:1}, 1/2)) \rightarrow_p -\infty,$$

because $K_s(0, t) = t/s + o(t)$ as $t \searrow 0$ and $E(\xi_{n:1}) = 1/(n+1)$. Since $K_s(1, t) = K_s(0, 1-t)$, $C_v(t) = C_v(1-t)$ and $\xi_{n:1} \stackrel{d}{=} 1 - \xi_{n:n}$,

$$\sup_{t \in [\xi_{n:n}, 1)} (nK_s(\mathbb{G}_n(t), t) - C_v(\mathbb{G}_n(t), t)) = nK_s(1, \xi_{n:n}) - C_v(\max(\xi_{n:n}, 1/2)) \rightarrow_p -\infty.$$

Consequently, it suffices to verify Theorem 2.1 with the modified test statistic

$$T_{n,s,v} := \sup_{t \in [\xi_{n:1}, \xi_{n:n})} (nK_s(\mathbb{G}_n(t), t) - C_v(\mathbb{G}_n(t), t)),$$

provided that we can show that the latter converges in distribution.

In what follows, we show that replacing s with 2 and $C_v(\mathbb{G}_n(t), t)$ with $C_v(t)$ has no effect asymptotically. For these tasks, the following two facts are useful.

FACT 4.7 (Linear bounds for \mathbb{G}_n).

A. By inequality 1, [30, 31], page 415,

$$\sup_{\xi_{n:1} \leq t \leq 1} \frac{t}{\mathbb{G}_n(t)} = O_p(1) \quad \text{and} \quad \sup_{0 \leq t < \xi_{n:n}} \frac{1-t}{1-\mathbb{G}_n(t)} = O_p(1).$$

B. From Daniels' theorem (Theorem 2, [30, 31], p. 341),

$$\sup_{0 < t \leq 1} \frac{\mathbb{G}_n(t)}{t} = O_p(1) \quad \text{and} \quad \sup_{0 \leq t < 1} \frac{1-\mathbb{G}_n(t)}{1-t} = O_p(1).$$

FACT 4.8. For any sequence of constants d_n with $1 \leq d_n \leq n$ such that $d_n/n \rightarrow 0$ and $d_n \rightarrow \infty$,

$$\sup_{d_n/n \leq t \leq 1} \frac{|\mathbb{G}_n(t) - t|}{t} = O_p(d_n^{-1/2})$$

and

$$\sup_{0 \leq t \leq 1-d_n/n} \frac{|\mathbb{G}_n(t) - t|}{1-t} = O_p(d_n^{-1/2})$$

([34], Lemma 3 and Theorem 1S; [30, 31], Chapter 10, Section 5, p. 424). In fact,

$$d_n^{1/2} \sup_{d_n/n \leq t \leq 1} \frac{|\mathbb{G}_n(t) - t|}{t} \rightarrow_d \sup_{0 \leq t \leq 1} |\mathbb{W}(t)|,$$

where \mathbb{W} is a standard Brownian motion; see [26].

A particular consequence of Fact 4.7 is that

$$(4.2) \quad M_{n,1} := \sup_{t \in [\xi_{n:1}, \xi_{n:n})} |\text{logit}(\mathbb{G}_n(t)) - \text{logit}(t)| = O_p(1),$$

where $\text{logit}(t) := \text{log}(t/(1-t))$, and Fact 4.8 implies that

$$(4.3) \quad M_{n,2} := \sup_{t \in [n^{-1} \log n, 1-n^{-1} \log n]} |\text{logit}(\mathbb{G}_n(t)) - \text{logit}(t)| = O_p((\log n)^{-1/2}),$$

with the conventions that $\text{logit}(0) := -\infty$ and $\text{logit}(1) := \infty$. This leads to the following useful bounds.

LEMMA 4.9. For any fixed $s \in \mathbb{R}$,

$$\sup_{t \in [\xi_{n:1}, \xi_{n:n}]} \frac{K_s(\mathbb{G}_n(t), t)}{K_2(\mathbb{G}_n(t), t)} = O_p(1) \quad \text{and} \quad \sup_{t \in [\xi_{n:1}, \xi_{n:n}]} (C_v(t) - C_v(\mathbb{G}_n(t), t)) = O_p(1),$$

where $K_s(t, t)/K_2(t, t) := 1$. Moreover,

$$\sup_{t \in [n^{-1} \log n, 1 - n^{-1} \log n]} \left| \frac{K_s(\mathbb{G}_n(t), t)}{K_2(\mathbb{G}_n(t), t)} - 1 \right| = O_p((\log n)^{-1/2}) \quad \text{and}$$

$$\sup_{t \in [n^{-1} \log n, 1 - n^{-1} \log n]} (C_v(t) - C_v(\mathbb{G}_n(t), t)) = O_p((\log n)^{-1/2}),$$

where $K_s(0, t) = K_s(1, t) := \infty$ in case of $s < 1$.

PROOF. With the auxiliary quantities $M_{n,1}$ in (4.2) and $M_{n,2}$ in (4.3), it follows from the inequalities (S.14) and Lemma S.10 that for $\xi_{n:1} \leq t < \xi_{n:n}$,

$$\begin{aligned} \frac{K_s(\mathbb{G}_n(t), t)}{K_2(\mathbb{G}_n(t), t)} &\leq \exp(|s - 2|M_{n,1}) = O_p(1) \quad \text{and} \\ 0 &\leq C_v(t) - C_v(\mathbb{G}_n(t), t) \leq (1 + \nu)M_{n,1} = O_p(1). \end{aligned}$$

Moreover, for $n^{-1} \log n \leq t \leq 1 - n^{-1} \log n$,

$$\begin{aligned} \left| \frac{K_s(\mathbb{G}_n(t), t)}{K_2(\mathbb{G}_n(t), t)} - 1 \right| &\leq \exp(|s - 2|M_{n,2}) - 1 = O_p((\log n)^{-1/2}) \quad \text{and} \\ 0 &\leq C_v(t) - C_v(\mathbb{G}_n(t), t) \leq (1 + \nu)M_{n,2} = O_p((\log n)^{-1/2}). \end{aligned}$$

(Note that $M_{n,2} = \infty$ if $t < \xi_{n:1}$ or $t \geq \xi_{n:n}$.) \square

Now the statement about the (modified) test statistic $T_{n,s,\nu}$ is an immediate consequence of Theorem 2.2 and the following lemma.

LEMMA 4.10. For $\nu > 3/4$ and any $s \in \mathbb{R}$,

$$T_{n,s,\nu} = \tilde{T}_{n,\nu} + o_p(1).$$

PROOF. With $d_n := \log n$, we know that $\xi_{n:n} > 1 - d_n/n$ with asymptotic probability one, and thus it follows from Fact 4.3 and Lemma 4.9 that

$$\begin{aligned} &\sup_{t \in [\xi_{n:1}, d_n/n]} nK_s(\mathbb{G}_n(t), t) \\ &\leq \sup_{t \in [\xi_{n:1}, 1 - d_n/n]} \frac{K_s(\mathbb{G}_n(t), t)}{K_2(\mathbb{G}_n(t), t)} \sup_{t \in (0, d_n/n]} nK_2(\mathbb{G}_n(t), t) = O_p(\log \log \log n). \end{aligned}$$

On the other hand,

$$\min_{t \in [\xi_{n:1}, d_n/n]} C_v(\mathbb{G}_n(t), t) \geq C(d_n/n) + O_p(1) = (1 + o(1)) \log \log n.$$

Hence,

$$\sup_{t \in [\xi_{n:1}, d_n/n]} (nK_s(\mathbb{G}_n(t), t) - C_v(\mathbb{G}_n(t), t)) \rightarrow_p -\infty,$$

and for symmetry reasons,

$$\sup_{t \in [1 - d_n/n, \xi_{n:n}]} (nK_s(\mathbb{G}_n(t), t) - C_v(\mathbb{G}_n(t), t)) \rightarrow_p -\infty.$$

Since $\tilde{T}_{n,v}$ is equal to

$$\tilde{T}_{n,v}^{\text{restr}} = \sup_{t \in [d_n/n, 1-d_n/n]} (nK_2(\mathbb{G}_n(t), t) - C_v(t))$$

with asymptotic probability one, it suffices to show that

$$T_{n,s,v}^{\text{restr}} := \sup_{t \in [d_n/n, 1-d_n/n]} (nK_s(\mathbb{G}_n(t), t) - C_v(\mathbb{G}_n(t), t)) = \tilde{T}_{n,v}^{\text{restr}} + o_p(1).$$

To this end, note that $\tilde{T}_{n,v}^{\text{restr}} \rightarrow_d T_v$ implies that

$$\sup_{t \in [d_n/n, 1-d_n/n]} nK_2(\mathbb{G}_n(t), t) \leq C_v(d_n/n) + O_p(1) = (1 + o_p(1)) \log \log n.$$

Consequently,

$$\begin{aligned} & |T_{n,s,v}^{\text{restr}} - \tilde{T}_{n,v}^{\text{restr}}| \\ & \leq \sup_{t \in [d_n/n, 1-d_n/n]} |nK_s(\mathbb{G}_n(t), t) - nK_2(\mathbb{G}_n(t), t)| + O_p((\log n)^{-1/2}) \\ & \leq \sup_{t \in [d_n/n, 1-d_n/n]} \left| \frac{K_s(\mathbb{G}_n(t), t)}{K_2(\mathbb{G}_n(t), t)} - 1 \right| \sup_{t \in [d_n/n, 1-d_n/n]} nK_2(\mathbb{G}_n(t), t) + O_p((\log n)^{-1/2}) \\ & = O_p((\log n)^{-1/2})(1 + o_p(1)) \log \log n = o_p(1). \end{aligned} \quad \square$$

It remains to prove the claim that $\kappa_{n,s,v,\alpha} \rightarrow \kappa_{v,\alpha} > 0$. But this follows immediately from the following lemma.

LEMMA 4.11. *Let $G(r) := P(T_v \leq r)$. Then $G(0) = 0$, and G is continuous and strictly increasing on $[0, \infty)$.*

To prove this lemma and other results, we make use of the following well-known result.

FACT 4.12 ([4], Corollary 2.1; [13], Lemma 1.1). The distribution Q of \mathbb{U} is a log-concave measure on $\mathcal{C}[0, 1]$. That means, for Borel sets $\mathcal{B}_0, \mathcal{B}_1 \subset \mathcal{C}[0, 1]$ and $\lambda \in (0, 1)$,

$$\log Q_*((1 - \lambda)\mathcal{B}_0 + \lambda\mathcal{B}_1) \geq (1 - \lambda)Q(\mathcal{B}_0) + \lambda Q(\mathcal{B}_1),$$

where Q_* stands for the inner measure induced by Q , and $(1 - \lambda)\mathcal{B}_0 + \lambda\mathcal{B}_1 := \{(1 - \lambda)g_0 + \lambda g_1 : g_0 \in \mathcal{B}_0, g_1 \in \mathcal{B}_1\}$.

From this fact, one can deduce the following properties of \mathbb{U} .

PROPOSITION 4.13. *For arbitrary functions $h : [0, 1] \rightarrow [0, \infty)$ and $h_o : [0, 1] \rightarrow \mathbb{R}$,*

$$G_1(x) := P(|xh_o + \mathbb{U}| \leq h)$$

is an even, log-concave function of $x \in \mathbb{R}$. Furthermore, if $h_o \geq 0$, then

$$G_2(x) := P(|\mathbb{U}| \leq \sqrt{h + xh_o})$$

is a nondecreasing and log-concave function of $x \geq 0$.

Let \mathbb{W} be a standard Brownian motion process on $[0, 1]$. Then it is well known that $\mathbb{U}(t) := \mathbb{W}(t) - t\mathbb{W}(1)$ defines a Brownian bridge process on $[0, 1]$. The following self-similarity property of the Brownian bridge process \mathbb{U} seems to be less well known.

PROPOSITION 4.14. For fixed numbers $0 \leq a < b \leq 1$, define a stochastic process $\mathbb{Z}_{a,b}$ on $[0, 1]$ as follows:

$$\mathbb{Z}_{a,b}(v) := \mathbb{U}((1 - v)a + vb) - (1 - v)\mathbb{U}(a) - v\mathbb{U}(b),$$

that is, $\mathbb{Z}_{a,b}$ describes the interpolation error when replacing \mathbb{U} on $[a, b]$ with its linear interpolation there. Then the two processes $(\mathbb{U}(t))_{t \in [0,1] \setminus (a,b)}$ and $\mathbb{Z}_{a,b}$ are stochastically independent, and

$$\mathbb{Z}_{a,b} \stackrel{d}{=} \sqrt{b - a}\mathbb{U}.$$

Proofs of Propositions 4.13 and 4.14 are provided in Section S.4.

PROOF OF LEMMA 4.11. Note first that the distribution function $r \mapsto G(r)$ coincides with the function G_2 in Proposition 4.13, where $h(t) := 2t(1 - t)C_v(t)$ and $h_o(t) := 2t(1 - t)$. In particular, $G(r) \leq P(|\mathbb{U}(1/2)| \leq \sqrt{r/2})$, and the latter bound equals 0 for $r = 0$ and is strictly smaller than 1 for any $r \geq 0$.

By Proposition 4.13, $G : [0, \infty) \rightarrow [0, 1]$ is log-concave, and since $G(r) < 1 = \lim_{s \rightarrow \infty} G(s)$ for all $r \geq 0$, this implies that G is continuous and strictly increasing on (r_o, ∞) , where $r_o := \inf\{r > 0 : G(r) > 0\}$. If we can show that $r_o = 0$, then we know that G is, in fact, continuous and strictly increasing on $[0, \infty)$.

To show that $G(r) > 0$ for any $r > 0$, we pick a number $\rho \in (0, 1/2)$ and write T_v as the maximum of the three random variables:

$$\begin{aligned} T_v^{(\rho,1)} &:= \max_{t \in [\rho, 1-\rho]} (\mathbb{U}(t)^2/[2t(1 - t)] - C_v(t)), \\ T_v^{(\rho,2,L)} &:= \max_{t \in (0,\rho]} (\mathbb{U}(t)^2/[2t(1 - t)] - C_v(t)), \\ T_v^{(\rho,2,R)} &:= \max_{t \in [1-\rho,1)} (\mathbb{U}(t)^2/[2t(1 - t)] - C_v(t)). \end{aligned}$$

Then we can write

$$\begin{aligned} G(r) &= P(T_v^{(\rho,1)} \leq r, T_v^{(\rho,2,L)} \leq r, T_v^{(\rho,2,R)} \leq r) \\ &\geq P\left(\max_{t \in [\rho, 1-\rho]} |\mathbb{U}(t)| \leq \delta, T_v^{(\rho,2,L)} \leq 0, T_v^{(\rho,2,R)} \leq 0\right) \end{aligned}$$

with $\delta := \sqrt{2\rho(1 - \rho)r} > 0$.

According to Lemma 4.6, we may choose ρ such that $P(T_v^{(\rho,2,L)} \leq 0) = P(T_v^{(\rho,2,R)} \leq 0) \geq 1/2$. Now we apply Proposition 4.14 twice, first with $[a, b] = [0, \rho]$, and then with $[a, b] = [1 - \rho, 1]$. This shows that \mathbb{U} may be rewritten on $[0, \rho]$ and on $[1 - \rho, 1]$ as follows: for $v \in [0, 1]$,

$$\begin{aligned} \mathbb{U}(\rho v) &= v\mathbb{U}(\rho) + \sqrt{\rho}\mathbb{U}^{(L)}(v), \\ \mathbb{U}(1 - \rho v) &= v\mathbb{U}(1 - \rho) + \sqrt{\rho}\mathbb{U}^{(R)}(v), \end{aligned}$$

where $\mathbb{U}, \mathbb{U}^{(L)}, \mathbb{U}^{(R)}$ are independent Brownian bridge processes. In particular,

$$\begin{aligned} &P(T_v^{(\rho,2,L)} \leq 0 | (\mathbb{U}(t))_{t \in [\rho, 1-\rho]}) \\ &= P(|v\mathbb{U}(\rho) + \sqrt{\rho}\mathbb{U}^{(L)}(v)| \leq \sqrt{2\rho v(1 - \rho v)C_v(\rho v)} \text{ for all } v \in [0, 1] | (\mathbb{U}(t))_{t \in [\rho, 1-\rho]}) \\ &= P(|\mathbb{U}(\rho)v/\sqrt{\rho} + \mathbb{U}^{(L)}(v)| \leq \sqrt{2v(1 - \rho v)C_v(\rho v)} \text{ for all } v \in [0, 1] | (\mathbb{U}(t))_{t \in [\rho, 1-\rho]}) \\ &= G_1(\mathbb{U}(\rho)), \end{aligned}$$

where $G_1(x) := P(|xh_o + \mathbb{U}| \leq h)$ with $h_o(v) := v/\sqrt{\rho}$ and $h(v) := \sqrt{2v(1-\rho v)C_v(\rho v)}$ for $v \in [0, 1]$. Analogously,

$$P(T_v^{(\rho,2,R)} \leq 0 | (\mathbb{U}(t))_{t \in [\rho, 1-\rho]}) = G_1(\mathbb{U}(1-\rho)).$$

According to Proposition 4.13, G_1 is an even, log-concave function on \mathbb{R} . Since $1/2 \leq P(T_v^{(\rho,2,L)} \leq 0) = E[G_1(\mathbb{U}(\rho))]$, there exists a $\delta_o > 0$ such that $G_1(x) \geq 1/2$ for all $x \in [-\delta_o, \delta_o]$. Consequently,

$$\begin{aligned} G(r) &\geq E(1_{\{\|\mathbb{U}\| \leq \delta \text{ on } [\rho, 1-\rho]\}} G_1(\mathbb{U}(\rho)) G_1(\mathbb{U}(1-\rho))) \\ &\geq 4^{-1} P(\|\mathbb{U}\|_\infty \leq \min(\delta, \delta_o)) > 0. \end{aligned}$$

That $P(\|\mathbb{U}\|_\infty \leq \lambda) > 0$ for any $\lambda > 0$ follows, for instance, from the expansion

$$P(\|\mathbb{U}\|_\infty \leq \lambda) = \frac{\sqrt{2\pi}}{8\lambda^2} \exp\left(-\frac{\pi^2}{8\lambda^2}\right) (1 + o(1)) \quad \text{as } \lambda \searrow 0;$$

see [23] or [31], pages 526–527. Alternatively, one could use Proposition 4.13 and separability of $\mathcal{C}[0, 1]$. \square

5. Proofs for Section 3.

5.1. *Proofs for Section 3.1.* PROOF OF THEOREM 3.1. Let $(x_n)_n$ be a sequence in \mathbb{R} such that $\Delta_n(x_n) \rightarrow \infty$. Then, for any fixed $\kappa > 0$,

$$\begin{aligned} P_{F_n}[T_{n,s,v}(F_0) \leq \kappa] &\leq P_{F_n}[x_n \notin [X_{n:1}, X_{n:n}]] \\ (5.1) \qquad \qquad \qquad &+ P_{F_n}[nK_s(\mathbb{F}_n(x_n), F_0(x_n)) \\ &\leq C_v(\mathbb{F}_n(x_n), F_0(x_n)) + \kappa], \end{aligned}$$

where $K_s(u, \cdot) := \infty$ if $s \leq 0$ and $u \in \{0, 1\}$.

To ensure that the first summand on the right-hand side of (5.1) converges to 0, we show that x_n may be chosen such that $d_n/n \leq F_n(x_n) \leq 1 - d_n/n$, where $d_n := \log \log n$. To this end, we have to analyze the auxiliary function H_n in more detail. Elementary calculus reveals that for $t \in [0, 1]$, $(1 + C(t))t(1 - t)$ is an increasing and $1 + C(t)$ is a decreasing function of $t(1 - t) \in [0, 1/4]$. Moreover,

$$1 + C(d_n/n) = (1 + o(1))d_n \quad \text{and} \quad (d_n/n)(1 - d_n/n) = (1 + o(1))d_n/n,$$

whence

$$\min_{t \in [0, 1]} H_n(t) \geq (1 + o(1))n^{-1/2}d_n \quad \text{and} \quad H_n(d_n/n) = (2 + o(1))n^{-1/2}d_n.$$

In particular,

$$|F_n - F_0|(x_n) \geq \Delta_n(x_n)(1 + o(1))d_n/n.$$

Now suppose that $F_n(x_n) < d_n/n$. With $\tilde{x}_n := F_n^{-1}(d_n/n)$, we may conclude that

$$F_n(\tilde{x}_n) \geq F_n(x_n) > |F_n - F_0|(x_n) - d_n/n \geq \Delta_n(x_n)(1 + o(1))d_n/n.$$

In particular, $\max\{d_n/n, F_n(x_n)\}$ is of order $o(F_n(\tilde{x}_n))$, so

$$\Delta_n(\tilde{x}_n) \geq \frac{\sqrt{n}|F_n - F_0|(\tilde{x}_n)}{H_n(F_n)} \geq \frac{(1 + o(1))\sqrt{n}F_n(\tilde{x}_n)}{(2 + o(1))n^{-1/2}d_n} \geq (1/2 + o(1))\Delta_n(x_n) \rightarrow \infty.$$

Analogously, one can show that in case of $F_n(x_n) > 1 - d_n/n$, we may replace x_n with $\tilde{x}_n := F_n^{-1}(1 - d_n/n)$ at the cost of reducing $\Delta_n(x_n)$ by a factor of at most $1/2 + o(1)$.

It remains to show that

$$(5.2) \quad P_{F_n}[nK_s(\mathbb{F}_n(x_n), F_0(x_n)) \leq C_v(\mathbb{F}_n(x_n), F_0(x_n)) + \kappa] \rightarrow 0.$$

By means of the second part of Lemma S.12, the inequality for $K_s(\mathbb{F}_n(x_n), F_0(x_n))$ implies that

$$\begin{aligned} \sqrt{n}|\mathbb{F}_n - F_0|(x_n) &\leq \sqrt{2(C_v(\mathbb{F}_n, F_0) + \kappa) \min\{\mathbb{F}_n(1 - \mathbb{F}_n), F_0(1 - F_0)\}(x_n)} \\ &\quad + 2(C_v(\mathbb{F}_n, F_0) + \kappa)(x_n)/\sqrt{n} \\ &\leq 2 \max(1 + \nu, \kappa) \min\{H_n(\mathbb{F}_n), H_n(F_0)\}(x_n), \end{aligned}$$

because $C_v(\mathbb{F}_n, F_0) \leq \min\{C_v(\mathbb{F}_n), C_v(F_0)\}$, and for the univariate function C_v , it follows from $D \leq C$ that $C_v + \kappa \leq \max(1 + \nu, \kappa)(1 + C)$. Moreover, the assumption that $d_n/n \leq F_n(x_n) \leq 1 - d_n/n$ implies that

$$\frac{h(\mathbb{F}_n)}{h(F_n)}(x_n) \rightarrow_p 1 \quad \text{for } h(t) = t, 1 + C(t), t(1 - t).$$

Consequently, (5.2) would be a consequence of

$$(5.3) \quad P_{F_n}[\sqrt{n}|\mathbb{F}_n - F_0|(x_n) \leq O_p(1) \min\{H_n(F_n), H_n(F_0)\}(x_n)] \rightarrow 0.$$

To bound the left-hand side of (5.3), we consider the quantity

$$M_n := \max\left\{\frac{F_0(1 - F_0)}{F_n(1 - F_n)}(x_n), \frac{F_n(1 - F_n)}{F_0(1 - F_0)}(x_n)\right\} \geq 1$$

and distinguish two cases. Suppose first that $M_n \leq \Delta_n(x_n)$. Since

$$\begin{aligned} \frac{1 + C(F_n)}{1 + C(F_0)}(x_n) \leq 1 &\leq \frac{F_n(1 - F_n)}{F_0(1 - F_0)}(x_n) \leq M_n \quad \text{or} \\ \frac{F_n(1 - F_n)}{F_0(1 - F_0)}(x_n) \leq 1 &\leq \frac{1 + C(F_n)}{1 + C(F_0)}(x_n) \leq 1 + \log M_n, \end{aligned}$$

the definition of H_n implies that

$$\frac{H_n(F_n)}{H_n(F_0)}(x_n) \leq \Delta_n(x_n)^{1/2}.$$

Then it follows from $\sqrt{n}(\mathbb{F}_n - F_n)(x_n) = O_p(\sqrt{F_n(1 - F_n)}(x_n)) = O_p(H_n(F_n(x_n)))$ that

$$\begin{aligned} P_{F_n}[\sqrt{n}|\mathbb{F}_n - F_0|(x_n) \leq O_p(1) \min\{H_n(F_n), H_n(F_0)\}(x_n)] \\ \leq P_{F_n}[\sqrt{n}|F_n - F_0|(x_n) \leq O_p(1) \min\{H_n(F_n), H_n(F_0)\}(x_n) + O_p(H_n(F_n(x_n)))] \\ \leq P_{F_n}[\sqrt{n}|F_n - F_0|(x_n) \leq O_p(\Delta_n(x_n)^{1/2}) \min\{H_n(F_n), H_n(F_0)\}(x_n)] \\ = P_{F_n}[\Delta_n(x_n) \leq O_p(\Delta_n(x_n)^{1/2})] \rightarrow 0. \end{aligned}$$

Now suppose that $M_n \geq \Delta_n(x_n)^{1/2}$. Then

$$\begin{aligned} \frac{|\mathbb{F}_n - F_0|}{|F_n - F_0|}(x_n) &\geq 1 - \frac{|\mathbb{F}_n - F_n|}{|F_n - F_0|}(x_n) \\ &\geq 1 - \frac{|\mathbb{F}_n - F_n|}{|F_n(1 - F_n) - F_0(1 - F_0)|}(x_n) = 1 + O_p(\rho_n) \end{aligned}$$

with

$$\begin{aligned} \rho_n &:= \frac{\sqrt{F_n(1-F_n)}}{\sqrt{n}|F_n(1-F_n) - F_0(1-F_0)|}(x_n) \\ &= \frac{F_n(1-F_n)}{\sqrt{nF_n(1-F_n)}|F_n(1-F_n) - F_0(1-F_0)|}(x_n) \leq \frac{M_n}{(1+o(1))\sqrt{d_n}(M_n-1)} \rightarrow 0. \end{aligned}$$

Consequently,

$$\begin{aligned} P_{F_n}[\sqrt{n}|\mathbb{F}_n - F_0|(x_n) \leq O_p(1) \min\{H_n(F_n), H_n(F_0)\}(x_n)] \\ \leq P_{F_n}[\sqrt{n}|F_n - F_0|(x_n)(1+o_p(1)) \leq O_p(1) \min\{H_n(F_n), H_n(F_0)\}(x_n)] \\ \leq P_{F_n}[\Delta_n(x_n) \leq O_p(1)] \rightarrow 0. \end{aligned} \quad \square$$

PROOF OF COROLLARY 3.2. Since $\|F_n - F_0\|_\infty \leq \varepsilon_n \rightarrow 0$, it suffices to show that (3.2) is satisfied. In what follows, we use frequently the elementary inequalities

$$(5.4) \quad \frac{\phi(x)}{x+1} \leq \Phi(-x) \leq \frac{\phi(x)}{x} \quad \text{for } x > 0,$$

where $\phi(x) := \Phi'(x) = \exp(-x^2/2)/\sqrt{2\pi}$. In particular, as $x \rightarrow \infty$,

$$\begin{aligned} \Phi(-x) &= \exp(-x^2/2 + O(\log x)) \quad \text{and} \\ C(\Phi(x)) &= \log(O(1) + \log(1/\Phi(-x))) = 2\log(x) - \log(2) + o(1). \end{aligned}$$

Now consider two sequences $(x_n)_n$ and $(\mu_n)_n$ tending to ∞ , and let $F_0 = \Phi$, $F_n = (1 - \varepsilon_n)\Phi + \varepsilon_n\Phi(\cdot - \mu_n)$. Then the inequalities (5.4) imply that

$$\begin{aligned} [1 + C(F_0(x_n))]F_0(x_n)(1 - F_0(x_n)) &= [2\log(x_n) + O(1)]\Phi(-x_n)(1 + o(1)) \\ &= \exp[-x_n^2/2 + O(\log(x_n))]. \end{aligned}$$

Moreover,

$$F_0(x_n) - F_n(x_n) = \varepsilon_n(\Phi(\mu_n - x_n) - \Phi(-x_n)) = \varepsilon_n\Phi(\mu_n - x_n)(1 + o(1)),$$

because $\Phi(-x_n) \leq \phi(x_n)/x_n$ while

$$\Phi(\mu_n - x_n) \geq \begin{cases} 1/2 & \text{if } \mu_n \geq x_n, \\ \frac{\phi(x_n - \mu_n)}{x_n - \mu_n + 1} \geq \frac{\phi(x_n) \exp(\mu_n^2/2)}{x_n + 1} & \text{if } \mu_n < x_n. \end{cases}$$

Consequently, $\Delta_n(x_n) \rightarrow \infty$ if

$$(5.5) \quad \frac{n\varepsilon_n\Phi(\mu_n - x_n)}{n^{1/2} \exp[-x_n^2/4 + O(\log(x_n))] + O(\log(x_n))} \rightarrow \infty.$$

In part (a) with $\varepsilon_n = n^{-\beta+o(1)}$ and $\beta \in (1/2, 1)$, we imitate the arguments of [6] and consider

$$\mu_n = \sqrt{2r \log(n)} \quad \text{and} \quad x_n = \sqrt{2q \log(n)}$$

with $0 < r < q \leq 1$. Then by (5.4),

$$\begin{aligned} n\varepsilon_n\Phi(\mu_n - x_n) &= n^{1-\beta-(\sqrt{q}-\sqrt{r})^2+o(1)}, \\ n^{1/2} \exp[-x_n^2/4 + O(\log(x_n))] &= n^{1/2-q/2+o(1)}, \\ O(\log(x_n)) &= n^{o(1)}, \end{aligned}$$

so the left-hand side of (5.5) equals

$$\frac{n^{1-\beta-(\sqrt{q}-\sqrt{r})^2+o(1)}}{n^{1/2-q/2+o(1)} + n^{o(1)}} = \frac{n^{1/2-\beta+q/2-(\sqrt{q}-\sqrt{r})^2+o(1)}}{1 + n^{(q-1)/2+o(1)}} = \frac{n^{1/2-\beta+2\sqrt{r}\sqrt{q}-\sqrt{q}^2/2-r+o(1)}}{1 + n^{(q-1)/2+o(1)}}.$$

The exponent in the numerator is maximal in $q \in (r, 1]$ if $\sqrt{q} = \min\{2\sqrt{r}, 1\}$, that is, $q = \min\{4r, 1\}$, and this leads to

$$\begin{cases} 1/2 - \beta + r & \text{if } r \leq 1/4, \\ 1 - \beta - (1 - \sqrt{r})^2 & \text{if } r \geq 1/4. \end{cases}$$

Thus, when $\beta \in (1/2, 3/4)$ we should choose $\beta - 1/2 < r < 1/4$ and $q = 4r$. When $\beta \in [3/4, 1)$, we should choose $(1 - \sqrt{1 - \beta})^2 < r < 1$ and $q = 1$.

As to part (b), we consider the more general setting that $\varepsilon_n = n^{-\beta+o(1)}$ for some $\beta \in [1/2, 3/4)$, where $\pi_n = \sqrt{n}\varepsilon_n \rightarrow 0$. Note that this scenario covers also a part of part (a), so we establish a connection between the two parts. The constraint that $\pi_n \rightarrow 0$ is trivial when $\beta > 1/2$ but relevant when $\beta = 1/2$. Now we consider

$$\mu_n := \sqrt{2\lambda \log(1/\pi_n)} \quad \text{and} \quad x_n := \sqrt{2q \log(1/\pi_n)}$$

with arbitrary constants $0 < \lambda < q$. Now

$$\begin{aligned} n\varepsilon_n \Phi(\mu_n - x_n) &= n^{1/2}\pi_n \Phi(\mu_n - x_n) \\ &= n^{1/2}\pi_n^{1+(\sqrt{q}-\sqrt{\lambda})^2+o(1)}, \\ n^{1/2} \exp(-x_n^2/4 + O(\log(x_n))) &= n^{1/2}\pi_n^{q/2+o(1)}, \\ O(\log(x_n)) &= \pi_n^{o(1)}, \end{aligned}$$

so the left-hand side of (5.5) equals

$$\frac{n^{1/2}\pi_n^{1+(\sqrt{q}-\sqrt{\lambda})^2+o(1)}}{n^{1/2}\pi_n^{q/2+o(1)} + \pi_n^{o(1)}} = \frac{\pi_n^{1+q/2-2\sqrt{q}\sqrt{\lambda}+\lambda+o(1)}}{1 + n^{-1/2}\pi_n^{-q/2+o(1)}} = \frac{\pi_n^{1+q/2-2\sqrt{q}\sqrt{\lambda}+\lambda+o(1)}}{1 + n^{-1/2+(\beta-1/2)q/2+o(1)}}.$$

The exponent of π_n becomes minimal in $q \in (\lambda, \infty)$ if $q = 4\lambda$. Then we obtain

$$\frac{\pi_n^{1-\lambda+o(1)}}{1 + n^{-1/2+(2\beta-1)\lambda+o(1)}} = \frac{\pi_n^{1-\lambda+o(1)}}{1 + \sqrt{n}^{-(4\beta-2)\lambda-1+o(1)}},$$

and this converges to ∞ if the limiting exponents of π_n and \sqrt{n} are negative. This is the case if $1 < \lambda < 1/(4\beta - 2)$. (Note that $4\beta - 2 < 1$ because $\beta < 3/4$.) \square

PROOF OF LEMMA 3.3. Standard LAN theory implies that $P_{F_n}(S_n) \rightarrow 0$ for arbitrary events $S_n \in \sigma(X_1, \dots, X_n)$ such that $P_{F_0}(S_n) \rightarrow 0$. Thus, for any fixed $0 < \rho < 1/2$, $\varphi_n(X_1, \dots, X_n) \neq \varphi_{n,\rho}(X_1, \dots, X_n)$ with asymptotic probability zero, both under the null and under the alternative hypothesis. Hence, it suffices to show that

$$\limsup_{\rho \rightarrow 0} \limsup_{n \rightarrow \infty} E_{F_n} \varphi_{n,\rho}(X_1, \dots, X_n) \leq \alpha.$$

But $E_{F_n} \varphi_{n,\rho}(X_1, \dots, X_n)$ does not change if we replace f_n with the modified density

$$f_{n,\rho}(x) := \begin{cases} f_n(x) & \text{if } x \notin [x_\rho, y_\rho], \\ c_{n,\rho} f_0(x) & \text{if } x \in [x_\rho, y_\rho], \end{cases}$$

with

$$c_{n,\rho} := \frac{F_n(y_\rho) - F_n(x_\rho)}{1 - 2\rho}.$$

This follows from the fact that the distribution function $F_{n,\rho}$ of $f_{n,\rho}$ satisfies $F_{n,\rho}(x) = F_n(x)$ for $x \notin [x_\rho, y_\rho]$, so the distribution of $\{\mathbb{F}_n(x) : x \notin [x_\rho, y_\rho]\}$ under the alternative hypothesis remains unchanged if we replace f_n with $f_{n,\rho}$. But

$$\sqrt{n}(c_{n,\rho} - 1) \rightarrow \delta_\rho := \frac{A(y_\rho) - A(x_\rho)}{1 - 2\rho},$$

so

$$\sqrt{n}(f_{n,\rho}^{1/2} - f_0^{1/2}) \rightarrow \frac{1}{2}a_\rho f_0^{1/2} \quad \text{in } L_2(\lambda)$$

with

$$a_\rho(x) = \begin{cases} a(x) & \text{if } x \notin [x_\rho, y_\rho], \\ \delta_\rho & \text{if } x \in [x_\rho, y_\rho]. \end{cases}$$

Hence, the asymptotic power of the test $\varphi_{n,\rho}$ under the alternative is bounded by the asymptotic power of the optimal test of F_0 versus $F_{n,\rho}$ at level α , so

$$\limsup_{n \rightarrow \infty} E_{F_n} \varphi_{n,\rho}(X_1, \dots, X_n) \leq \Phi(\Phi^{-1}(\alpha) + \|a_\rho\|_{L_2(F_0)}).$$

But

$$\begin{aligned} \|a_\rho\|_{L_2(F_0)}^2 &= \int_{(-\infty, x_\rho) \cup (y_\rho, \infty)} a^2 dF_0 + (1 - 2\rho)\delta_\rho^2 \\ &= \int_{(-\infty, x_\rho) \cup (y_\rho, \infty)} a^2 dF_0 + \frac{(A(y_\rho) - A(x_\rho))^2}{(1 - 2\rho)} \end{aligned}$$

converges to 0 as $\rho \searrow 0$, so $\Phi(\Phi^{-1}(\alpha) + \|a_\rho\|_{L_2(F_0)}) \rightarrow \alpha$ as $\rho \searrow 0$. \square

PROOF OF THEOREM 3.4. Let $\rho \in (0, 1/2)$ be fixed. The test statistic $T_{n,s,v}$ for the uniform empirical process may be written as the maximum of $T_{n,s,v}^{(\rho,1)}$ and $T_{n,s,v}^{(\rho,2)}$, where

$$\begin{aligned} T_{n,s,v}^{(\rho,1)} &:= \sup_{t \in \mathcal{T}_{n,s} \cap [\rho, 1-\rho]} (nK_s(\mathbb{G}_n(t), t) - C_v(\mathbb{G}_n(t), t)), \\ T_{n,s,v}^{(\rho,2)} &:= \sup_{t \in \mathcal{T}_{n,s} \setminus [\rho, 1-\rho]} (nK_s(\mathbb{G}_n(t), t) - C_v(\mathbb{G}_n(t), t)). \end{aligned}$$

Here, $\mathcal{T}_{n,s} := (0, 1)$ if $s > 0$ and $\mathcal{T}_n := \{\xi_{n:1}, \xi_{n:n}\}$ if $s \leq 0$. A supremum over the empty set is defined to be $-\infty$. The proofs of Theorems 2.2 and 2.1 can be easily adapted to show that

$$T_{n,s,v}^{(\rho,1)} \rightarrow_d T_v^{(\rho,1)} \quad \text{and} \quad T_{n,s,v}^{(\rho,2)} \rightarrow_d T_v^{(\rho,2)} := \max\{T_v^{(\rho,2,L)}, T_v^{(\rho,2,R)}\},$$

where $T_v^{(\rho,1)}$, $T_v^{(\rho,2,L)}$ and $T_v^{(\rho,2,R)}$ are defined as in the proof of Lemma 4.11. In particular, since $C_v(1/2) = 0$ and $\mathbb{U}(1/2) \neq 0$ almost surely,

$$\begin{aligned} \liminf_{n \rightarrow \infty} P(T_{n,s,v}^{(\rho,1)} > 0) &= 1, \\ \limsup_{n \rightarrow \infty} P(T_{n,s,v}^{(\rho,2)} \geq 0) &\leq \pi_0(\rho) := P(T_v^{(\rho,2)} \geq 0). \end{aligned}$$

Note that $\pi_0(\rho) \rightarrow 0$ as $\rho \rightarrow 0$ by virtue of Lemma 4.6.

Now we consider the goodness-of-fit test statistic $T_{n,s,v}(F_0)$. It is the maximum of $T_{n,s,v}^{(\rho,1)}(F_0)$ and $T_{n,s,v}^{(\rho,2)}(F_0)$. Here, $T_{n,s,v}^{(\rho,j)}(F_0)$ is defined as $T_{n,s,v}^{(\rho,j)}$, where $t \in \mathcal{T}_{n,s}$ is replaced with $x \in \mathbb{R}$ if $s > 0$ and $x \in [X_{n:1}, X_{n:n}]$ if $s \leq 0$, $[\rho, 1 - \rho]$ is replaced with $[x_\rho, y_\rho] = [F_0^{-1}(\rho), F_0^{-1}(1 - \rho)]$, and $(\mathbb{G}_n(t), t)$ is replaced with $(\mathbb{F}_n(x), F_0(x))$. Under the null hypothesis, $T_{n,s,v}^{(\rho,j)}(F_0)$ has the same distribution as $T_{n,s,v}^{(\rho,j)}$ for $j = 1, 2$. This convergence and standard LAN theory imply that under the alternative hypothesis,

$$\liminf_{n \rightarrow \infty} P_{F_n}(T_{n,s,v}^{(\rho,1)}(F_0) > 0) = 1,$$

$$\limsup_{n \rightarrow \infty} P_{F_n}(T_{n,s,v}^{(\rho,2)}(F_0) \geq 0) \leq \pi_A(\rho) := \Phi(\Phi^{-1}(\pi_0(\rho)) + \|a\|_{L_2(F_0)}).$$

With standard empirical process theory one can show that under the alternative hypothesis,

$$\sqrt{n}(\mathbb{F}_n - F_0) \rightarrow_d \mathbb{U} \circ F_0 + A$$

in the space $\ell^\infty(\mathbb{R})$ of bounded functions on \mathbb{R} , equipped with the supremum norm $\|\cdot\|_\infty$. Moreover, for arbitrary bounded functions h, h_n on \mathbb{R} such that $\|h_n - h\|_\infty \rightarrow 0$,

$$nK_s(F_0 + n^{-1/2}h_n, F_0) - C_v(F_0 + n^{-1/2}h_n, F_0) \rightarrow h^2/[2F_0(1 - F_0)] - C_v(F_0)$$

uniformly on $[x_\rho, y_\rho]$. By virtue of an extended continuous mapping theorem, for example, [33], Theorem 1.11.1, page 67, one can conclude that

$$T_{n,s,v}^{(\rho,1)}(F_0) \rightarrow_d T_v^{(\rho,1)}(A),$$

where $T_v^{(\rho,j)}(A)$ is defined as $T_v^{(\rho,j)}$ with $\mathbb{U} + A \circ F_0^{-1}$ in place of \mathbb{U} . Finally, note that the distribution Q_A of $\mathbb{U} + A \circ F_0^{-1}$ is absolutely continuous with respect to the distribution Q_0 of \mathbb{U} , where $\log(dQ_A/dQ_0)$ has distribution $N(-\|a\|_{L_2(F_0)}^2/2, \|a\|_{L_2(F_0)}^2)$ under Q_0 . This follows from [31] (Section 4.1 and especially Theorem 4.1.5, p. 157), or [33] (Section 3.10). Consequently,

$$P(T_v^{(\rho,2)}(A) \geq 0) \leq \pi_A(\rho).$$

All in all, we may conclude that

$$P_{F_n}(T_{n,s,v}(F_0) \leq 0) \leq P_{F_n}(T_{n,s,v}^{(\rho,1)}(F_0) \leq 0) \rightarrow 0,$$

and for fixed $r > 0$,

$$\begin{aligned} \limsup_{n \rightarrow \infty} P_{F_n}(T_{n,s,v}(F_0) \leq r) &\leq \limsup_{n \rightarrow \infty} P_{F_n}(T_{n,s,v}^{(\rho,1)}(F_0) \leq r) \\ &\leq P(T_v^{(\rho,1)}(A) \leq r) \\ &\leq P(T_v(A) \leq r) + P(T_v^{(\rho,2)}(A) > r) \\ &\leq P(T_v(A) \leq r) + \pi_A(\rho), \\ \limsup_{n \rightarrow \infty} P_{F_n}(T_{n,s,v}(F_0) \geq r) &\leq \limsup_{n \rightarrow \infty} P_{F_n}(T_{n,s,v}^{(\rho,1)}(F_0) < r) \\ &\quad + \limsup_{n \rightarrow \infty} P_{F_n}(T_{n,s,v}^{(\rho,2)}(F_0) \geq r) \\ &\leq P(T_v^{(\rho,1)}(A) \geq r) + \pi_A(\rho) \\ &\leq P(T_v(A) \geq r) + \pi_A(\rho). \end{aligned}$$

Since $\pi_A(\rho) \rightarrow 0$ as $\rho \searrow 0$, this proves that $T_{n,s,v}(F_0)$ converges in distribution to $T_v(A)$ under the alternative hypothesis.

The convergence claimed in the second part of the theorem follows from the first part together with convergence of the critical values $\kappa_{n,s,v,\alpha}$ to $\kappa_{v,\alpha}$. The inequality claimed in the second part is a consequence of Anderson’s [1] inequality or Proposition 4.13 with $h_o := A \circ F_0^{-1}$ and $h(t) := \sqrt{2t(1-t)}(C_v(t) + \kappa_{v,\alpha})$.

The third part of the theorem follows from the fact that for any $t \in (0, 1)$,

$$\begin{aligned} P(T_v(A) > \kappa_{v,\alpha}) &\geq P\left(\frac{(\mathbb{U} + A \circ F_0^{-1})^2(t)}{2t(1-t)} > C_v(t) + \kappa_{v,\alpha}\right) \\ &\geq \Phi\left(\frac{|A(F_0^{-1}(t))|}{\sqrt{t(1-t)}} - \sqrt{2C_v(t) + 2\kappa_{v,\alpha}}\right) \\ &= \Phi\left(\frac{|A(F_0^{-1}(t))|}{\sqrt{t(1-t)}} - \sqrt{2C(t)} - b_{v,\alpha}(t)\right), \end{aligned}$$

where $b_{v,\alpha} := (2vD + 2\kappa_{v,\alpha})/(\sqrt{2C} + 2vD + 2\kappa_{v,\alpha} + \sqrt{2C})$ is bounded on $(0, 1)$. \square

5.2. *Proofs for Section 3.2.* For notational convenience, we suppress the dependence of the confidence bounds on s, v and α and just write $a_{n,i}^{BJO}, a_{n,i}, b_{n,i}^{BJO}$ and $b_{n,i}$.

PROOF OF THEOREM 3.5. Note first that $H_s(u, t) = \gamma H_s(u/\gamma, t/\gamma)$ for arbitrary $u \geq 0, t > 0$ and $\gamma > 0$.

Now we prove the claim for the upper bounds $b_{n,i}^{BJO} = 1 - a_{n,n-i}^{BJO}$ and $b_{n,i} = 1 - a_{n,n-i}$. For any integer $i \in [0, n^\delta]$, let

$$x_{n,i} := u_{n,i}/\gamma_n = i/\log \log n.$$

For fixed $\lambda > 0$, let

$$\tilde{b}_{n,i} := u_{n,i} + \lambda\gamma_n(B_s(x_{n,i}) - x_{n,i}) = \gamma_n(x_{n,i} + \lambda(B_s(x_{n,i}) - x_{n,i})) > u_{n,i}.$$

It follows from $x + s \leq B_s(x) \leq x + 1 + \sqrt{2x + 1}$ that

$$\lambda s \gamma_n \leq \tilde{b}_{n,i} \leq \lambda \gamma_n B_s(n^\delta / \log \log n) = (\lambda + o(1))n^{\delta-1}.$$

On the one hand, if $\lambda > 1$, then it follows from the first inequality in (S.15) that

$$nK_s(u_{n,i}, \tilde{b}_{n,i}) \geq nH_s(u_{n,i}, \tilde{b}_{n,i}) = n\gamma_n H_s(x_{n,i}, x_{n,i} + \lambda(B_s(x_{n,i}) - x_{n,i})) \geq n\gamma_n \lambda,$$

because $H_s(x_{n,i}, x_{n,i} + t(B_s(x_{n,i}) - x_{n,i}))$ is convex in t with values 0 for $t = 0$ and 1 for $t = 1$. And if $\lambda < 1$, the second inequality in (S.15) implies that

$$\begin{aligned} nK_s(u_{n,i}, \tilde{b}_{n,i}) &\leq nH_s(u_{n,i}, \tilde{b}_{n,i})/(1 - \tilde{b}_{i,n})^+ \\ &= n\gamma_n H_s(x_{n,i}, x_{n,i} + \lambda(B_s(x_{n,i}) - x_{n,i}))/ (1 - \tilde{b}_{n,i}) \\ &\leq n\gamma_n \lambda / (1 - (\lambda + o(1))n^{\delta-1}) = n\gamma_n (\lambda + o(1)). \end{aligned}$$

On the other hand, $\kappa_{n,s,\alpha}^{BJ} = (1 + o(1))n\gamma_n$ and

$$\begin{aligned} C_v(u_{i,n}, \tilde{b}_{i,n}) + \kappa_{n,s,v,\alpha} &= C_v(\tilde{b}_{i,n}) + \kappa_{n,s,v,\alpha} \\ &\begin{cases} \leq C_v(\lambda s \gamma_n) + \kappa_{n,s,v,\alpha} = (1 + o(1))n\gamma_n, \\ \geq C_v((\lambda + o(1))n^{\delta-1}) + \kappa_{n,s,v,\alpha} = (1 + o(1))n\gamma_n. \end{cases} \end{aligned}$$

Consequently, for any fixed $\lambda > 1$ and sufficiently large n ,

$$nK_s(u_{n,i}, \tilde{b}_{n,i}) > \max\{C_v(u_{n,i}, \tilde{b}_{n,i}) + \kappa_{n,s,v,\alpha}, \kappa_{n,s,\alpha}^{BJ}\},$$

and thus

$$\max\{b_{n,i}^{\text{BJO}} - u_{n,i}, b_{n,i} - u_{n,i}\} \leq \lambda\gamma_n(B_s(x_{n,i}) - x_{n,i})$$

for all integers $i \in [0, n^\delta]$. Likewise, for any fixed $\lambda \in (0, 1)$ and sufficiently large n ,

$$nK_s(u_{n,i}, \tilde{b}_{n,i}) < \min\{C_v(u_{n,i}, \tilde{b}_{n,i}) + \kappa_{n,s,v,\alpha} + \kappa_{n,s,\alpha}^{\text{BJ}}\},$$

and thus

$$\min\{b_{n,i}^{\text{BJO}} - u_{n,i}, b_{ni} - u_{n,i}\} \geq \lambda\gamma_n(B_s(x_{n,i}) - x_{n,i})$$

for all integers $i \in [0, n^\delta]$.

The differences $u_{n,i} - a_{n,i}^{\text{BJO}} = b_{n,n-i}^{\text{BJO}} - u_{n,n-i}$ and $u_{n,i} - a_{n,i} = b_{n,n-i} - u_{n,n-i}$ can be treated analogously. For each integer $i \in [1, n^\delta]$ and fixed $\lambda > 0$, let $x_{n,i} = u_{n,i}/\gamma_n = i/\log n$ as before and

$$\tilde{a}_{n,i} := u_{n,i} + \lambda\gamma_n(A_s(x_{n,i}) - x_{n,i}) = \gamma_n(x_{n,i} + \lambda(A_s(x_{n,i}) - x_{n,i})) < u_{n,i}.$$

On the one hand, if $\lambda > 1$ and $\tilde{a}_{n,i} > 0$, then $A_s(x_{i,n}) > 0$ and

$$nK_s(u_{n,i}, \tilde{a}_{n,i}) \geq nH_s(u_{n,i}, \tilde{a}_{n,i}) = n\gamma_n H_s(x_{n,i}, x_{n,i} + \lambda(A_s(x_{n,i}) - x_{n,i})) \geq n\gamma_n\lambda,$$

because $H_s(x_{n,i}, x_{n,i} + t(A_s(x_{n,i}) - x_{n,i}))$ is convex in $t \in [0, \lambda]$ with values 0 for $t = 0$ and 1 for $t = 1$. And if $\lambda < 1$, then

$$\begin{aligned} nK_s(u_{n,i}, \tilde{a}_{n,i}) &\leq nH_s(u_{n,i}, \tilde{a}_{n,i})/(1 - u_{in}) \\ &= n\gamma_n H_s(x_{n,i}, x_{n,i} + \lambda(A_s(x_{n,i}) - x_{n,i}))/ (1 - u_{n,i}) \\ &\leq n\gamma_n\lambda/(1 - n^{\delta-1}). \end{aligned}$$

On the other hand, $\kappa_{n,s,\alpha}^{\text{BJ}} = (1 + o(1))n\gamma_n$ and

$$\begin{aligned} C_v(u_{i,n}, \tilde{a}_{i,n}) + \kappa_{n,s,v,\alpha} &= C_v(u_{i,n}) + \kappa_{n,s,v,\alpha} \\ &\begin{cases} \leq C_v(n^{-1}) + \kappa_{n,s,v,\alpha} = (1 + o(1))n\gamma_n, \\ \geq C_v(\min\{n^{\delta-1}, 1/2\}) + \kappa_{n,s,v,\alpha} = (1 + o(1))n\gamma_n. \end{cases} \end{aligned}$$

Consequently, for any fixed $\lambda > 1$ and sufficiently large n ,

$$\max\{u_{n,i} - a_{n,i}^{\text{BJO}}, u_{n,i} - a_{n,i}\} \leq \lambda\gamma_n(x_{n,i} - A_s(x_{n,i}))$$

for all integers $i \in [1, n^\delta]$. Likewise, for any fixed $\lambda \in (0, 1)$ and sufficiently large n ,

$$\min\{u_{n,i} - a_{n,i}^{\text{BJO}}, u_{ni} - a_{n,i}\} \geq \lambda\gamma_n(x_{n,i} - A_s(x_{n,i}))$$

for all integers $i \in [1, n^\delta]$. \square

PROOF OF THEOREM 3.7. We only prove the bounds for $a_{n,i}$ and $b_{n,i}$. The bounds for $a_{n,i}^{\text{BJO}}$ and $b_{n,i}^{\text{BJO}}$ can be derived analogously with obvious modifications. Moreover, since $u_{n,i} - a_{n,i} = b_{n,n-i} - u_{n,n-i}$, it suffices to prove the bounds for $b_{n,i}$ only. For a fixed factor $\lambda > 0$ and any integer $i \in [n^\delta, n - n^\delta]$, let

$$\tilde{b}_{n,i} := u_{n,i} + \lambda\sqrt{2\gamma_n(u_{n,i})u_{n,i}(1 - u_{n,i})}.$$

Note that

$$\begin{aligned} 0 &\leq \frac{\tilde{b}_{n,i} - u_{n,i}}{u_{n,i}(1 - u_{n,i})} \\ &\leq \lambda \sqrt{2n^{-1}(C_v(n^{\delta-1}) + \kappa_{v,\alpha})n^{1-\delta}(1 - n^{\delta-1})^{-1}} \\ &= O(n^{-\delta/2}(\log \log n)^{1/2}), \end{aligned}$$

whence

$$c_n := \max_{n^\delta \leq i \leq n - n^\delta} |\text{logit}(\tilde{b}_{n,i}) - \text{logit}(u_{n,i})| = o(1).$$

On the one hand, the inequalities (S.14) imply that uniformly in $n^\delta \leq i \leq n - n^\delta$,

$$\begin{aligned} nK_s(u_{n,i}, \tilde{b}_{n,i}) &= nK_{1-s}(\tilde{b}_{n,i}, u_{n,i}) \\ &= (1 + o(1))nK_2(\tilde{b}_{n,i}, u_{n,i}) \\ &= (1 + o(1))\lambda^2(C_v(u_{n,i}) + \kappa_{v,\alpha}). \end{aligned}$$

On the other hand, Lemma S.10 and Theorem 2.1 imply that uniformly in $n^\delta \leq i \leq n - n^\delta$,

$$|C_v(u_{n,i}, \tilde{b}_{n,i}) + \kappa_{n,s,v,\alpha} - C_v(u_{n,i}) - \kappa_{v,\alpha}| \leq (1 + v)c_n + |\kappa_{n,s,v,\alpha} - \kappa_{v,\alpha}| = o(1).$$

Consequently, for fixed $\lambda > 1$ and sufficiently large n ,

$$nK_s(u_{n,i}, \tilde{b}_{n,i}) > C_v(u_{n,i}, \tilde{b}_{n,i}) + \kappa_{n,s,v,\alpha},$$

and thus

$$b_{n,i} - u_{n,i} \leq \lambda \sqrt{2\gamma_n(u_{n,i})u_{n,i}(1 - u_{n,i})}$$

for all integers $i \in [n^\delta, n - n^\delta]$. Likewise, for fixed $\lambda \in (0, 1)$ and sufficiently large n ,

$$nK_s(u_{n,i}, \tilde{b}_{n,i}) < C_v(u_{n,i}, \tilde{b}_{n,i}) + \kappa_{n,s,v,\alpha},$$

and thus

$$b_{n,i} - u_{n,i} \geq \lambda \sqrt{2\gamma_n(u_{n,i})u_{n,i}(1 - u_{n,i})}$$

for all integers $i \in [n^\delta, n - n^\delta]$. \square

Acknowledgments. The authors owe thanks to David Mason for pointing out the relevance of the tools of Csörgő et al. [5] for some of the results presented here. We are also grateful to Günther Walther for stimulating conversations about likelihood ratio tests in non-parametric settings and to Rudy Beran for pointing out the interesting results of Bahadur and Savage [2]. Constructive comments of two referees and an associate editor are gratefully acknowledged.

Funding. The first author was supported in part by the Swiss National Science Foundation. The second author was supported in part by NSF Grant DMS-1104832 and NI-AID Grant 2R01 AI291968-04.

SUPPLEMENTARY MATERIAL

Supplement A to: “A new approach to tests and confidence bands for distribution functions” (DOI: [10.1214/22-AOS2249SUPPA](https://doi.org/10.1214/22-AOS2249SUPPA); .pdf). Further technical details and proofs; additional numerical examples [9].

Supplement B to: “A new approach to tests and confidence bands for distribution functions” (DOI: [10.1214/22-AOS2249SUPPB](https://doi.org/10.1214/22-AOS2249SUPPB); .zip). Details and implementation of the new bands: zipped arxiv with R files [10].

REFERENCES

- [1] ANDERSON, T. W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.* **6** 170–176. MR0069229 <https://doi.org/10.2307/2032333>
- [2] BAHADUR, R. R. and SAVAGE, L. J. (1956). The nonexistence of certain statistical procedures in non-parametric problems. *Ann. Math. Stat.* **27** 1115–1122. MR0084241 <https://doi.org/10.1214/aoms/1177728077>
- [3] BERK, R. H. and JONES, D. H. (1979). Goodness-of-fit test statistics that dominate the Kolmogorov statistics. *Z. Wahrsch. Verw. Gebiete* **47** 47–59. MR0521531 <https://doi.org/10.1007/BF00533250>
- [4] BORELL, C. (1974). Convex measures on locally convex spaces. *Ark. Mat.* **12** 239–252. MR0388475 <https://doi.org/10.1007/BF02384761>
- [5] CSÖRGŐ, M., CSÖRGŐ, S., HORVÁTH, L. and MASON, D. M. (1986). Weighted empirical and quantile processes. *Ann. Probab.* **14** 31–85. MR0815960
- [6] DONOHO, D. and JIN, J. (2004). Higher criticism for detecting sparse heterogeneous mixtures. *Ann. Statist.* **32** 962–994. MR2065195 <https://doi.org/10.1214/009053604000000265>
- [7] DÜMBGEN, L. and SPOKOINY, V. G. (2001). Multiscale testing of qualitative hypotheses. *Ann. Statist.* **29** 124–152. MR1833961 <https://doi.org/10.1214/aos/996986504>
- [8] DÜMBGEN, L. and WALTHER, G. (2008). Multiscale inference about a density. *Ann. Statist.* **36** 1758–1785. MR2435455 <https://doi.org/10.1214/07-AOS521>
- [9] DÜMBGEN, L. and WELLNER, J. (2023). Supplement A to “A New Approach to Tests and Confidence Bands for Distribution Functions.” <https://doi.org/10.1214/22-AOS2249SUPPA>
- [10] DÜMBGEN, L. and WELLNER, J. (2023). Supplement B to “A New Approach to Tests and Confidence Bands for Distribution Functions.” <https://doi.org/10.1214/22-AOS2249SUPPB>
- [11] EICKER, F. (1979). The asymptotic distribution of the suprema of the standardized empirical processes. *Ann. Statist.* **7** 116–138. MR0515688
- [12] ERDŐS, P. (1942). On the law of the iterated logarithm. *Ann. of Math. (2)* **43** 419–436. MR0006630 <https://doi.org/10.2307/1968801>
- [13] GAENSSLER, P., MOLNÁR, P. and ROST, D. (2007). On continuity and strict increase of the CDF for the sup-functional of a Gaussian process with applications to statistics. *Results Math.* **51** 51–60. MR2375638 <https://doi.org/10.1007/s00025-007-0257-1>
- [14] GONTSCHARUK, V., LANDWEHR, S. and FINNER, H. (2016). Goodness of fit tests in terms of local levels with special emphasis on higher criticism tests. *Bernoulli* **22** 1331–1363. MR3474818 <https://doi.org/10.3150/14-BEJ694>
- [15] INGSTER, Y. I. (1997). Some problems of hypothesis testing leading to infinitely divisible distributions. *Math. Methods Statist.* **6** 47–69. MR1456646
- [16] ITÔ, K. and MCKEAN, H. P. JR. (1974). *Diffusion Processes and Their Sample Paths. Die Grundlehren der Mathematischen Wissenschaften, Band 125*. Springer, Berlin. Second printing, corrected. MR0345224
- [17] JAESCHKE, D. (1979). The asymptotic distribution of the supremum of the standardized empirical distribution function on subintervals. *Ann. Statist.* **7** 108–115. MR0515687
- [18] JAGER, L. and WELLNER, J. A. (2007). Goodness-of-fit tests via phi-divergences. *Ann. Statist.* **35** 2018–2053. MR2363962 <https://doi.org/10.1214/0009053607000000244>
- [19] JANSSEN, A. (1995). Principal component decomposition of non-parametric tests. *Probab. Theory Related Fields* **101** 193–209. MR1318192 <https://doi.org/10.1007/BF01375824>
- [20] LEHMANN, E. L. and ROMANO, J. P. (2005). *Testing Statistical Hypotheses*, 3rd ed. *Springer Texts in Statistics*. Springer, New York. MR2135927
- [21] MASON, D. M. and SCHUENEMEYER, J. H. (1983). A modified Kolmogorov–Smirnov test sensitive to tail alternatives. *Ann. Statist.* **11** 933–946. MR0707943 <https://doi.org/10.1214/aos/1176346259>
- [22] MILBRODT, H. and STRASSER, H. (1990). On the asymptotic power of the two-sided Kolmogorov–Smirnov test. *J. Statist. Plann. Inference* **26** 1–23. MR1073112 [https://doi.org/10.1016/0378-3758\(90\)90091-8](https://doi.org/10.1016/0378-3758(90)90091-8)
- [23] MOGUL’SKIĬ, A. A. (1979). The law of the iterated logarithm in Chung’s form for function spaces. *Teor. Veroyatn. Primen.* **24** 399–407. MR0532454
- [24] O’REILLY, N. E. (1974). On the weak convergence of empirical processes in sup-norm metrics. *Ann. Probab.* **2** 642–651. MR0383486 <https://doi.org/10.1214/aop/1176996610>
- [25] OWEN, A. B. (1995). Nonparametric likelihood confidence bands for a distribution function. *J. Amer. Statist. Assoc.* **90** 516–521. MR1340504
- [26] RÉNYI, A. (1969). On some problems in the theory of order statistics. *Bull. Inst. Int. Stat.* **42** 165–176. MR0258212

- [27] RÉVÉSZ, P. (1982/83). A joint study of the Kolmogorov–Smirnov and the Eicker–Jaeschke statistics. *Statist. Decisions* **1** 57–65. MR0685588
- [28] ROHDE, A. and DÜMBGEN, L. (2013). Statistical inference for the optimal approximating model. *Probab. Theory Related Fields* **155** 839–865. MR3034794 <https://doi.org/10.1007/s00440-012-0414-7>
- [29] SCHMIDT-HIEBER, J., MUNK, A. and DÜMBGEN, L. (2013). Multiscale methods for shape constraints in deconvolution: Confidence statements for qualitative features. *Ann. Statist.* **41** 1299–1328. MR3113812 <https://doi.org/10.1214/13-AOS1089>
- [30] SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. Wiley, New York. MR0838963
- [31] SHORACK, G. R. and WELLNER, J. A. (2009). *Empirical Processes with Applications to Statistics*. Classics in Applied Mathematics **59**. SIAM, Philadelphia, PA. Reprint of the 1986 original [MR0838963]. MR3396731 <https://doi.org/10.1137/1.9780898719017.ch1>
- [32] STEPANOVA, N. and PAVLENKO, T. (2018). Goodness-of-fit tests based on sup-functionals of weighted empirical processes. *Teor. Veroyatn. Primen.* **63** 358–388. MR3796493 <https://doi.org/10.4213/tvp5160>
- [33] VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer Series in Statistics. Springer, New York. With applications to statistics. MR1385671 <https://doi.org/10.1007/978-1-4757-2545-2>
- [34] WELLNER, J. A. (1978). Limit theorems for the ratio of the empirical distribution function to the true distribution function. *Z. Wahrsch. Verw. Gebiete* **45** 73–88. MR0651392 <https://doi.org/10.1007/BF00635964>