An L^p-view of the Bahadur–Kiefer theorem

Miklós Csörgő *

Department of Mathematics and Statistics, Carleton University, Ottawa, Ont K1S 5B6, Canada. E-mail: `mcsorgo@math.carleton.ca`

Zhan Shi **

Laboratoire de Probabilités, Université Paris VI, 4 Place Jussieu, F-75252 Paris Cedex 05, France. E-mail: `shi@ccr.jussieu.fr`

Summary. Let α_n and β_n be respectively the uniform empirical and quantile processes, and define $R_n = \alpha_n + \beta_n$, which usually is referred to as the Bahadur–Kiefer process. The well-known Bahadur–Kiefer theorem confirms the following remarkable equivalence: $\|R_n\|/\sqrt{\|\alpha_n\|} \sim n^{-1/4}(\log n)^{1/2}$ almost surely, as n goes to infinity, where $\|f\| = \sup_{0\le t\le 1} |f(t)|$ is the L^∞-norm. We prove that $\|R_n\|_2/\sqrt{\|\alpha_n\|_1} \sim n^{-1/4}$ almost surely, where $\|\cdot\|_p$ is the L^p-norm. It is interesting to note that there is no longer any logarithmic term in the normalizing function. More generally, we show that $n^{1/4}\|R_n\|_p/\sqrt{\|\alpha_n\|_{(p/2)}}$ converges almost surely to a finite positive constant whose value is explicitly known. We also extend our result to a more general Bahadur–Kiefer process as well.

Keywords. Empirical process, quantile process, Bahadur–Kiefer representation, L^p-modulus of continuity for Brownian motion, Brownian bridge, Kiefer process.

1991 Mathematics Subject Classification. Primary 60F25; Secondary 62G30.

---

* Research supported by an NSERC Canada Grant at Carleton University, Ottawa, and by a Paul Erdős Visiting Professorship of the Paul Erdős Summer Research Center of Mathematics, Budapest.

** Research supported by a Fellowship of the Paul Erdős Summer Research Center of Mathematics, Budapest.

1. Introduction

Let $\{U_i\}_{i\geq 1}$ be a sequence of independent and identically distributed random variables, whose common law is the uniform distribution in $(0,1)$. Define the uniform empirical process

$$\alpha_n(t) \stackrel{\text{def}}{=} n^{1/2}\big(F_n(t) - t\big), \qquad 0 \leq t \leq 1,$$

where $F_n(\cdot)$ is the empirical distribution function based on the first n observations, i.e.,

$$F_n(t) \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{U_i \leq t\}}, \qquad 0 \leq t \leq 1.$$

Likewise, we can define the uniform empirical quantile process

$$\beta_n(t) \stackrel{\text{def}}{=} n^{1/2}\big(F_n^{-1}(t) - t\big), \qquad 0 \leq t \leq 1,$$

where $F_n^{-1}(t) \stackrel{\text{def}}{=} \inf\{s > 0 : F_n(s) \geq t\}$ (for $0 < t \leq 1$) and $F_n^{-1}(0) \stackrel{\text{def}}{=} F_n^{-1}(0+)$ is the inverse function (quantile function) of F_n. The process

$$R_n(t) \stackrel{\text{def}}{=} \alpha_n(t) + \beta_n(t), \qquad 0 \leq t \leq 1,$$

which is often referred to as the $(0,1)$–uniform **Bahadur–Kiefer process**, enjoys some remarkable properties. Let us recall the following Bahadur–Kiefer representation theorem.

Theorem A (Kiefer [17], Shorack [24], Deheuvels and Mason [11]). *We have,*

$$\lim_{n\to\infty} n^{1/4} (\log n)^{-1/2} \frac{\|R_n\|}{\sqrt{\|\alpha_n\|}} = 1, \qquad \text{a.s.,} \tag{1.1}$$

where $\|f\| \stackrel{\text{def}}{=} \sup_{0\leq t\leq 1} |f(t)|$ *denotes the uniform sup-norm of* f.

Together with some well-known laws of the iterated logarithm (LIL's) for α_n (cf. Fact 3.2 in Section 3 for the exact statement), (1.1) immediately implies the following:

$$\limsup_{n\to\infty} n^{1/4} (\log n)^{-1/2} (\log_2 n)^{-1/4} \|R_n\| = 2^{-1/4}, \qquad \text{a.s.,} \tag{1.2}$$

$$\liminf_{n\to\infty} n^{1/4} (\log n)^{-1/2} (\log_2 n)^{1/4} \|R_n\| = \frac{\pi^{1/2}}{8^{1/4}}, \qquad \text{a.s.,} \tag{1.3}$$

where $\log_2 n \stackrel{\text{def}}{=} \log(\log n)$.

The study of the Bahadur–Kiefer representation was initiated by Bahadur [2], who proved a "pointwise" version of (1.1). Kiefer [17] pointed out Theorem A, even though he only proved the convergence in probability, and omitted the proof of the theorem due to its extreme length. The upper bound in (1.1) was proved by Shorack [24], and the lower bound by Deheuvels and Mason [11]. We mention that a simplified proof of the lower bound was since discovered by Einmahl [14]. For a detailed discussion of various aspects of the Bahadur–Kiefer theorem, as well as extensions to sequential empirical processes, we refer to Csörgő and Szyszkowicz [10].

Looking at Theorem A, it is remarkable that the ratio between $\|R_n\|$ and $\sqrt{\|\alpha_n\|}$, suitably normalized, should almost surely converge to a constant. A natural question would be whether it remains true if the uniform sup-norm $\|\cdot\|$ is replaced by, say, the L^2-norm $\|\cdot\|_2$. For example, one might wonder if either $\|R_n\|_2/\sqrt{\|\alpha_n\|}$ or $\|R_n\|_2/\sqrt{\|\alpha_n\|_2}$ would still be of order of magnitude which is around $n^{-1/4}(\log n)^{1/2}$, perhaps with an extra term of some power of $\log_2 n$.

Somewhat surprisingly, the answer is no: we should use a **different** normalizing function. Moreover, under the new normalization, the ratio between $\|R_n\|_2$ and the square root of α_n under the L^1-norm, converges again to a constant limit with probability one. More precisely, we have the following L^p version of the Bahadur–Kiefer representation. Throughout the paper, we write $\|f\|_p \stackrel{\text{def}}{=} (\int_0^1 |f(t)|^p \, \mathrm{d}t)^{1/p}$.

Theorem 1.1. *Let $2 \le p < \infty$ and $q \stackrel{\text{def}}{=} p/2$. Then*

$$\lim_{n\to\infty} n^{1/4} \frac{\|R_n\|_p}{\sqrt{\|\alpha_n\|_q}} = c_0(p), \qquad \text{a.s.,} \tag{1.4}$$

where

$$c_0(p) \stackrel{\text{def}}{=} (\mathbb{E}|\mathcal{N}|^p)^{1/p} = \sqrt{2}\left(\frac{\Gamma((p+1)/2)}{\sqrt{\pi}}\right)^{1/p}, \tag{1.5}$$

and $\mathcal{N}$ denotes a Gaussian $\mathcal{N}(0,1)$ variable. In particular,

$$\lim_{n\to\infty} n^{1/4} \frac{\|R_n\|_2}{\sqrt{\|\alpha_n\|_1}} = 1, \qquad \text{a.s.}$$

Remark 1.2. The condition $p \ge 2$ is to ensure that $\|\cdot\|_q$ is a true metric norm. When $0 < p < 2$, it is no longer a metric. Our proof shows that in this case, we still have the following weaker version of Theorem 1.1: for $0 < p < 2$, there exist two finite constants $c_1(p) > 0$ and $c_2(p) > 0$, depending on p, such that

$$c_1(p) \le \liminf_{n\to\infty} n^{1/4} \frac{\|R_n\|_p}{\sqrt{\|\alpha_n\|_q}} \le \limsup_{n\to\infty} n^{1/4} \frac{\|R_n\|_p}{\sqrt{\|\alpha_n\|_q}} \le c_2(p), \qquad \text{a.s.}$$

Remark 1.3. The reason for which the normalizing function in Theorem 1.1 differs from the one in Theorem A will become clear in Section 3.

From (1.4), it is possible to deduce the almost sure asymptotics of R_n under the L^p-norm.

Corollary 1.4. *For* $2 \le p < \infty$,

$$\limsup_{n\to\infty}\, n^{1/4}(\log_2 n)^{-1/4}\, \|R_n\|_p = 2^{1/4}\, c_0(p)\, \sqrt{c_3(q)}\,, \qquad \text{a.s.},$$
$$\liminf_{n\to\infty}\, n^{1/4}(\log_2 n)^{1/4}\, \|R_n\|_p = c_0(p)\, \sqrt{c_4(q)}\,, \qquad \text{a.s.},$$

where $c_0(p)$ *is as in* (1.5), $q = p/2$, *and* $c_3(q) \in (0,\infty)$ *and* $c_4(q) \in (0,\infty)$ *are defined by*

$$c_3(q) \stackrel{\text{def}}{=} \frac{2^{-(q-1)/q} q^{-1/2}(q+2)^{(q-2)/(2q)}}{\int_0^1 (1-x^q)^{-1/2}\,\mathrm{d}x}$$
$$= \frac{2^{-(q-1)/q} q^{1/2}(q+2)^{(q-2)/(2q)}}{B(1/2,1/q)}\,, \tag{1.6}$$
$$c_4(q) \stackrel{\text{def}}{=} \inf_{f\in\mathcal{C}} \Big(\int_{-\infty}^{\infty} |x|^q\, f(x)\,\mathrm{d}x \Big)^{1/q}. \tag{1.7}$$

Here, $B(\,\cdot\,,\,\cdot\,)$ *is the usual beta function, and* $\mathcal{C}$ *is the set of probability densities* f *such that*

$$\frac{1}{8}\int_{-\infty}^{\infty} \frac{(f'(y))^2}{f(y)}\,\mathrm{d}y \le 1.$$

Remark 1.5. Comparing this corollary with (1.2)–(1.3), it is immediately noted that R_n has rather different asymptotics under L^p- and L^∞- norms. This is in complete constrast to the situation for the empirical process α_n. Indeed, $\|\alpha_n\|$ and $\|\alpha_n\|_p$ satisfy almost the same LIL's (from both limsup and liminf points of view), except for the constants, cf. Lemmas 3.3 and 3.8 and Fact 3.2 in Section 3 (they are stated for the Kiefer process, but in view of the KMT strong invariance in Fact 3.1, one can immediately deduce the corresponding LIL's for α_n).

Remark 1.6. The value of the constant $c_4(q)$ in (1.7) is in general implicit, except for $q = 1$ or 2. Indeed, $c_4(2) = 1/\sqrt{8}$, and it follows from the proof of Lemma 3.3 (cf. Section 3) and Takács [30, Theorem 1] that

$$c_4(1) = \sqrt{\frac{2\,|a_1'|^3}{27}}\,,$$

where $a_1' < 0$ denotes the largest real root of $\mathrm{Ai}'(\,\cdot\,)$, the derivative of the Airy function $\mathrm{Ai}(\,\cdot\,)$.

As a consequence of (1.1), the continuity of $\|\cdot\|$ on $C[0,1]$ and the fact that we have (cf. Doob [13], or Theorem 1.5.1 in [8])

$$\mathbb{P}(\|B\| \le x) = 1 + 2\sum_{k=1}^{\infty}(-1)^k \exp(-2k^2x^2), \quad x > 0, \tag{1.8}$$

where $\{B(t);\ 0 \le t \le 1\}$ is a Brownian bridge, we also have the following corollary.

Corollary 1.7 (Kiefer [17]). *For $x > 0$,*

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}\big(n^{1/4}(\log n)^{-1/2}\|R_n\| \le x\big) &= \mathbb{P}\{\sqrt{\|B\|} \le x\} \\ &= 1 + 2\sum_{k=1}^{\infty}(-1)^k \exp(-2k^2x^4).\end{aligned} \tag{1.9}$$

In a similar vein, from Theorem 1.1 we conclude the following L^p analogue of (1.9).

Corollary 1.8. *With $2 \le p < \infty$ and $q = p/2$ we have*

$$\lim_{n\to\infty} \mathbb{P}\big(n^{1/4}\|R_n\|_p \le x\big) = \mathbb{P}\Big(c_0(p)\sqrt{\|B\|_q} \le x\Big), \quad x > 0, \tag{1.10}$$

where $\{B(t);\ 0 \le t \le 1\}$ is a Brownian bridge.

For a Brownian bridge $\{B(t);\ 0 \le t \le 1\}$, Smirnov [26], Anderson and Darling [1] established the following result (cf., e.g., [8, Theorem 1.5.2]):

$$\mathbb{P}\Big((\|B\|_2)^2 \le x\Big) = 1 - \frac{2}{\pi}\sum_{k=1}^{\infty}(-1)^{k+1}\int_{(2k-1)\pi}^{2k\pi} \frac{\exp(-t^2x/2)}{\sqrt{-t\sin t}}\,\mathrm{d}t, \qquad x > 0. \tag{1.11}$$

Consequently, with $p = 4$ and hence $q = 2$, (1.10) and (1.11) yield

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}\big(n^{1/4}\|R_n\|_4 \le x\big) &= \mathbb{P}\big(c_0(4)\sqrt{\|B\|_2} \le x\big) \\ &= 1 - \frac{2}{\pi}\sum_{k=1}^{\infty}(-1)^{k+1}\int_{(2k-1)\pi}^{2k\pi} \frac{\exp(-t^2x^4/6)}{\sqrt{-t\sin t}}\,\mathrm{d}t, \qquad x > 0,\end{aligned} \tag{1.12}$$

on account of $c_0(4) = 3^{1/4}$.

The convergence in distribution of the appropriately normed functionals $\|R_n\|$ and $\|R_n\|_p$, respectively as in (1.9) and (1.10), is of special interest from the practical point of view of constructing classes of goodness-of-fit statistics for a large family of distributions (cf. Theorems B and 4.1 via Remark 4.2 in Section 4 (Appendix)).

The respective statements of Corollaries 1.7 and 1.8 combined also imply that the $(0,1)$–uniform Bahadur–Kiefer process $\{R_n(t);\ 0 \le t \le 1\}$ cannot be so normalized that it would converge weakly to a nondegenerate random element Y of $D[0,1]$ (endowed with the Skorohod J_1 topology). Indeed, if $a_n R_n \xrightarrow{\mathcal{D}} Y$ in $D[0,1]$ were to be true with any sequence $\{a_n\}$ of positive real numbers, then the latter would have to yield both (1.9) and (1.10) simultaneously, without any further renormalization, and this of course is impossible. Consequently, Corollaries 1.7 and 1.8 cannot result from any standard finite dimensional distributions and tightness type arguments on $D[0,1]$. This fact has already been established in 1972, via a different route, by Vervaat [31], [32] (for further comments along these lines cf. also Zitikis [33, Section 1 and Remark 6.1]), which we now summarize by restating it here as a corollary to Theorems A and 1.1 via the combined statements of Corollaries 1.7 and 1.8.

Corollary 1.9 (Vervaat [31], [32]). *The weak convergence*

$$a_n R_n \xrightarrow{\mathcal{D}} Y, \quad n \to \infty, \tag{1.13}$$

for the $(0,1)$*–uniform Bahadur–Kiefer process* $\{R_n(t);\ 0 \le t \le 1\}$ *cannot hold true in the space* $D[0,1]$, *endowed with the Skorohod* J_1 *topology, for any sequence* $\{a_n\}$ *of positive real numbers and any nondegenerate random element* Y *of* $D[0,1]$.

The rest of the paper is organized as follows. Section 2 is devoted to some preliminaries for the modulus of continuity of the Brownian motion and the Brownian bridge under the L^p-norm. The latter results are of interest on their own. Theorem 1.1 and Corollary 1.4 are proved in Section 3. We extend our results to more general Bahadur–Kiefer processes in Section 4 (Appendix).

Notation. Throughout the paper, $c_5, c_6, \cdots, c_{18}$ stand for some finite positive constants. We write $a_n \sim b_n$ $(n \to \infty)$ to denote $\lim_{n\to\infty} a_n/b_n = 1$.

Technical Remark. (i) When dealing with a Brownian bridge $\{B(t);\ 0 \le t \le 1\}$, some of our statements involve t in the left neighbourhood of 0 or in the right neighbourhood of 1. This can be rigorously justified, for example, by bringing in some independent Brownian bridges for $t \in [-1, 0]$ and for $t \in [1, 2]$. However, since these pieces do not influence any of the results, we shall not give any further discussion about this, in order not to make the proof tedious. The remark also applies to Brownian motion $W(t)$ when t is in the left neighbourhood of 0.

(ii) Unless stated otherwise, we shall be dealing with index n which ultimately goes to infinity; as a consequence, even without further mention, our statements should be understood for the situation when n is sufficiently large.

(iii) Our use of "almost surely" is not systematic.

2. Modulus of continuity

Let $\{W(t);\, t \geq 0\}$ be a standard one-dimensional Brownian motion. Throughout the section, we fix $2 \leq p < \infty$ and write $q \stackrel{\text{def}}{=} p/2$.

The main result of this section is the following probability estimate for the modulus of continuity of W under the L^p-norm. Observe that it is very different from Lévy's usual modulus of continuity theorem.

Proposition 2.1. *Let*

$$\Lambda_1(h) \stackrel{\text{def}}{=} \int_0^1 |W(s+h) - W(s)|^p \, \mathrm{d}s.$$

For any $\varepsilon > 0$, there exists $c_5 = c_5(\varepsilon, p)$ such that for all $0 < h \leq 1/2$,

$$\mathbb{P}\Big(\Big| \Lambda_1(h) - h^q \, \mathbb{E}(|\mathcal{N}|^p) \Big| > \varepsilon \, h^q \Big) \leq c_5 \, h^4, \tag{2.1}$$

where $\mathcal{N}$ is as before a Gaussian $\mathcal{N}(0,1)$ variable.

Remark 2.2. By means of a standard argument (cf. for example Csörgő and Révész [8, pp. 26–27]) and (2.1), one easily obtains:

$$\lim_{h \to 0} \frac{1}{\sqrt{h}} \Big(\int_0^1 |W(s+h) - W(s)|^p \, \mathrm{d}s \Big)^{1/p} = c_0(p), \qquad \text{a.s.},$$

where $c_0(p)$ is defined in (1.5). This should be compared with Lévy's well-known modulus of continuity theorem:

$$\lim_{h \to 0} \frac{1}{\sqrt{2h \log(1/h)}} \sup_{0 \leq s \leq 1} \sup_{0 \leq u \leq h} |W(s+u) - W(s)| = 1, \qquad \text{a.s.}$$

The proof of the proposition relies on the following moment inequality for partial sums, which is a particular case of Theorem 2.10 of Petrov [23, p. 62].

Fact 2.3. *Let $\{X_i\}_{i\geq 1}$ be a sequence of iid variables with $\mathbb{E}(X_1)=0$, such that $\mathbb{E}(X_1^8)<\infty$. Then*

$$\mathbb{E}\Big[\Big(\sum_{i=1}^{n} X_i\Big)^8\Big] \leq c_6\, n^4\, \mathbb{E}(X_1^8),$$

where c_6 is an absolute constant.

Proof of Proposition 2.1. It suffices to treat the situation when h is in the (positive) neighbourhood of 0. Let $M=M(h)\stackrel{\text{def}}{=}[1/(2h)]+1$. We have,

$$\begin{aligned}
\Lambda_1(h) &\leq \sum_{m=1}^{2M}\int_{(m-1)h}^{mh} |W(s+h)-W(s)|^p\,\mathrm{d}s\\
&= \sum_{j=1}^{M}\int_{(2j-2)h}^{(2j-1)h} |W(s+h)-W(s)|^p\,\mathrm{d}s\\
&\quad +\sum_{j=1}^{M}\int_{(2j-1)h}^{2jh} |W(s+h)-W(s)|^p\,\mathrm{d}s\\
&\stackrel{\text{def}}{=} \Lambda_2(h)+\Lambda_3(h), \qquad (2.2)
\end{aligned}$$

with obvious notation. Clearly, $(\int_{(2j-2)h}^{(2j-1)h}|W(s+h)-W(s)|^p\,\mathrm{d}s)_{1\leq j\leq M}$ are iid variables, each distributed as $h^{q+1}\,\Xi$, where

$$\Xi \stackrel{\text{def}}{=} \int_0^1 |W(s+1)-W(s)|^p\,\mathrm{d}s.$$

Therefore,

$$\Lambda_2(h) \;\stackrel{\text{law}}{=}\; h^{q+1}\sum_{j=1}^{M} Y_j,$$

where (Y_j) are iid variables, each having the law of Ξ, and "$\stackrel{\text{law}}{=}$" stands for identity in law. Since $\Xi\leq(2\sup_{0\leq t\leq 2}|W(t)|)^p$, we immediately deduce that Ξ admits finite moments of any order. Therefore, by Chebyshev's inequality and Fact 2.3, for any $\varepsilon>0$,

$$\begin{aligned}
\mathbb{P}\Big(\,|\Lambda_2(h)-h^{q+1}M\,\mathbb{E}(\Xi)|>\varepsilon h^{q+1}M\Big)
&\leq (\varepsilon h^{q+1}M)^{-8}\,\mathbb{E}\Big[\big(\Lambda_2(h)-h^{q+1}M\,\mathbb{E}(\Xi)\big)^8\Big]\\
&\leq (\varepsilon M)^{-8}c_6\,M^4\,\mathbb{E}\big[(\Xi-\mathbb{E}\Xi)^8\big]\\
&\leq c_7\,h^4.
\end{aligned}$$

Since $\Lambda_3(h)$ has the same distribution as $\Lambda_2(h)$, we obtain, in view of (2.2),

$$\mathbb{P}\Big(\Lambda_1(h) - 2h^{q+1}M\mathbb{E}(\Xi) > 2\varepsilon h^{q+1}M\Big) \le 2c_7\, h^4. \tag{2.3}$$

On the other hand, instead of (2.2), if we use the relation

$$\Lambda_1(h) \ge \sum_{j=1}^{M-1}\Big(\int_{(2j-2)h}^{(2j-1)h} + \int_{(2j-1)h}^{2jh}\Big)|W(s+h)-W(s)|^p\,\mathrm{d}s,$$

the same argument yields that

$$\mathbb{P}\Big(\Lambda_1(h) - 2h^{q+1}(M-1)\,\mathbb{E}(\Xi) < -2\varepsilon h^{q+1}(M-1)\Big) \le c_8\, h^4. \tag{2.4}$$

Combining (2.3) and (2.4) yields that, for all $\varepsilon > 0$,

$$\mathbb{P}\Big(\Big|\,\Lambda_1(h) - h^q\,\mathbb{E}(\Xi)\Big| > \varepsilon\, h^q\Big) \le c_9\, h^4.$$

Since $\mathbb{E}(\Xi) = \mathbb{E}(|\mathcal{N}|^p)$, this completes the proof of (2.1). $\square$

Looking at the proof of Proposition 2.1, we realize that the positivity of h has played no role at all, i.e. the argument works out also for negative h (when $|h|$ is small). Therefore, we can state the following "two-sided" version of the proposition: for any $\varepsilon > 0$ and $0 < |h| \le 1/2$,

$$\mathbb{P}\Big(\Big|\,\Lambda_1(h) - |h|^q\,\mathbb{E}(|\mathcal{N}|^p)\Big| > \varepsilon\,|h|^q\Big) \le c_{10}\, h^4,$$

or, more conveniently,

$$\mathbb{P}\Big(\Big|\,(\Lambda_1(h))^{1/p} - c_0(p)\,\sqrt{|h|}\,\Big| > \varepsilon\,\sqrt{|h|}\Big) \le c_{11}\, h^4, \tag{2.5}$$

where $c_0(p)$ is the constant in (1.5).

Consider now a standard one-dimensional Brownian bridge process $\{B(t);\, 0 \le t \le 1\}$. It is well-known that B can be realized as

$$B(t) = W(t) - tW(1), \qquad 0 \le t \le 1.$$

Using this representation and the Minkowski inequality, we have

$$\Big|\,(\Lambda_1(h))^{1/p} - \Big(\int_0^1 |B(s+h)-B(s)|^p\,\mathrm{d}s\Big)^{1/p}\,\Big| \le |h\,W(1)|. \tag{2.6}$$

On the other hand, by the usual estimate for Gaussian tails, for any $c > 0$,

$$\mathbb{P}\big(|W(1)| > c\,|h|^{-1/2}\big) \le 2\exp\big(-\frac{c^2}{2|h|}\big). \tag{2.7}$$

Combining (2.5)–(2.7) (and replacing ε by $\varepsilon/2$ in (2.5)) yields that

$$\begin{aligned}
\mathbb{P}\Big[\, \Big| \Big(\int_0^1 |B(s+h)-B(s)|^p \,\mathrm{d}s\Big)^{1/p} - c_0(p)\sqrt{|h|}\,\Big| > \varepsilon\sqrt{|h|}\,\Big] \\
&\le \mathbb{P}\Big[\, \Big| (\Lambda_1(h))^{1/p} - \Big(\int_0^1 |B(s+h)-B(s)|^p \,\mathrm{d}s\Big)^{1/p} \Big| > \frac{\varepsilon}{2}\sqrt{|h|}\,\Big] \\
&\qquad + \mathbb{P}\Big[\, \Big| (\Lambda_1(h))^{1/p} - c_0(p)\sqrt{|h|}\,\Big| > \frac{\varepsilon}{2}\sqrt{|h|}\,\Big] \\
&\le \mathbb{P}\Big(|W(1)| > \frac{\varepsilon}{2\sqrt{|h|}}\Big) + c_{12}\, h^4 \\
&\le 2\exp\big(-\frac{\varepsilon^2}{8|h|}\big) + c_{12}\, h^4.
\end{aligned}$$

So we have proved the following result which will be useful in Section 3 in the study of the Bahadur–Kiefer representation.

Proposition 2.4. *Let $\{B(t);\, 0 \le t \le 1\}$ be a Brownian bridge, and fix $\varepsilon > 0$. There exists $c_{13} = c_{13}(\varepsilon, p) > 0$ such that whenever $0 < |h| \le 1/2$,*

$$\mathbb{P}\Big[\, \Big| \Big(\int_0^1 |B(s+h)-B(s)|^p \,\mathrm{d}s\Big)^{1/p} - c_0(p)\sqrt{|h|}\,\Big| > \varepsilon\sqrt{|h|}\,\Big] \le c_{13}\, h^4. \tag{2.8}$$

3. Proof of Theorem 1.1 and Corollary 1.4

Let α_n be the uniform empirical process defined in Section 1. We first recall the well-known Komlós–Major–Tusnády (**KMT**) strong approximation theorem.

Fact 3.1 (Komlós, Major and Tusnády [18]). (*Possibly in an enlarged probability space*), *there exists a coupling for the empirical process α_n and an iid sequence of standard Brownian bridges $\{B_i\}_{i\ge 1}$, such that*

$$\Big\| \alpha_n - \frac{1}{\sqrt{n}}\sum_{i=1}^n B_i \Big\| \;=\; \mathcal{O}\Big(\frac{(\log n)^2}{\sqrt{n}}\Big), \qquad \text{a.s.}, \tag{3.1}$$

where $\|\cdot\|$ *denotes as before the uniform sup-norm.*

We shall be working on the independent Brownian bridges $(B_i)_{i\ge 1}$ introduced in (3.1). For notational convenience, we write

$$K_n(t) \stackrel{\text{def}}{=} \sum_{i=1}^{n} B_i(t), \qquad 0 \le t \le 1. \tag{3.2}$$

In the literature, $K_n(t)$, as a process indexed by (t, n), is referred to as the **Kiefer process**. We now recall two important versions of the LIL for the empirical process. In view of (3.1), it is equivalent to state them for K_n.

Fact 3.2 (Chung [4], Smirnov [27], Mogulskii [20]). *We have,*

$$\limsup_{n\to\infty} \frac{\|K_n\|}{\sqrt{2n\log_2 n}} = \frac{1}{2}, \qquad \text{a.s.} \tag{3.3}$$

$$\liminf_{n\to\infty} \frac{\sqrt{\log_2 n}}{\sqrt{n}}\, \|K_n\| = \frac{\pi}{\sqrt{8}}, \qquad \text{a.s.} \tag{3.4}$$

The following result will be useful in the proof of Theorem 1.1.

Lemma 3.3. *For* $q \ge 1$,

$$\liminf_{n\to\infty} \frac{\sqrt{\log_2 n}}{\sqrt{n}}\, \|K_n\|_q = c_4(q), \qquad \text{a.s.,} \tag{3.5}$$

where $c_4(q) \in (0, \infty)$ *is the constant defined in* (1.7).

Proof. Though the general statement of Lemma 3.3 seems to be new, it was implicitly proved by Donsker and Varadhan [12]. Indeed, according to Borovkov and Mogulskii [3], there exists $c_{14} = c_{14}(q) \in (0, \infty)$ such that

$$\lim_{x\to 0} x^2 \log \mathbb{P}\big(\|W\|_q < x \big) = -c_{14}, \tag{3.6}$$

where W is a Brownian motion. From this, a standard argument (cf. for example Shorack and Wellner [25, pp. 527–529]) yields

$$\liminf_{T\to\infty} \frac{\sqrt{\log_2 T}}{\sqrt{T}}\, \|W(\,\cdot\,T)\|_q = \sqrt{c_{14}}\,, \qquad \text{a.s.} \tag{3.7}$$

On the other hand, Donsker and Varadhan [12] proved that the "lim inf" expression in (3.7) equals $c_4(q)$, where the constant $c_4(q)$ is defined in (1.7). Therefore $c_{14} = (c_4(q))^2$.

To prove the lemma, note that (3.6) holds also for the Brownian bridge B in lieu of W, i.e.

$$\lim_{x\to 0} x^2 \log \mathbb{P}(\,\|B\|_q < x) = -(c_4(q))^2.$$

Applying the usual Borel–Cantelli argument readily completes the proof of Lemma 3.3. □

Now we recall some results concerning the oscillations of K_n. Write

$$\omega_n(h) \stackrel{\text{def}}{=} \sup_{0\le s\le t\le 1,\, t-s\le h} |K_n(t) - K_n(s)|, \qquad 0 < h < 1, \tag{3.8}$$

throughout the section.

Fact 3.4 (Stute [29]). *For any non-increasing sequence of positive numbers $(a_n)_{n\ge 1}$ such that $n \mapsto na_n$ is non-decreasing and that $\log(1/a_n)/\log_2 n \to \infty$,*

$$\lim_{n\to\infty} \frac{\omega_n(a_n)}{\sqrt{2na_n \log(1/a_n)}} = 1, \qquad \text{a.s.} \tag{3.9}$$

Fact 3.5 (Mason et al. [19]). *If a_n is non-increasing and na_n is non-decreasing such that $\log(1/a_n)/\log_2 n \to \varrho \in [0,\infty)$,*

$$\lim_{n\to\infty} \frac{\omega_n(a_n)}{\sqrt{2na_n \log_2 n}} = \sqrt{\varrho + 1}\,, \qquad \text{a.s.} \tag{3.10}$$

The next is a simple observation. A discrete-time version of this was (somewhat implicitly) stated in Einmahl [14, p. 530]. Let $\{B(t);\, 0 \le t \le 1\}$ as before be a standard Brownian bridge.

Fact 3.6. *Fix $0 < u < v < 1$. The process*

$$\Big\{ \frac{B(u+(v-u)t) - tB(v) - (1-t)B(u)}{\sqrt{v-u}};\ 0 \le t \le 1 \Big\}$$

is again a Brownian bridge. Furthermore, it is independent of $\sigma\{B(s);\, 0 \le s \le u\} \vee \sigma\{B(s);\, v \le s \le 1\}$, where $\sigma\{\,\cdot\,\}$ stands for the σ-algebra induced by the process or variables between the braces.

Let us start the proof of Theorem 1.1. As before, we fix $2 \le p < \infty$, and write $q \stackrel{\text{def}}{=} p/2$. The first step in the proof is the following preliminary estimate, which will later lead to a law of large numbers.

Lemma 3.7. *Let $\varepsilon > 0$ and $n > N^6 \ge n_0$. Define, for each $0 \le i \le N-1$,*

$$\begin{aligned} b_{i,n} &\stackrel{\text{def}}{=} n^{-1/2} B(\frac{i}{N}), \\ \Lambda_4(i,n) &\stackrel{\text{def}}{=} \Big(\int_{i/N}^{(i+1)/N} |B(t - b_{i,n}) - B(t)|^p \,\mathrm{d}t \Big)^{1/p}, \\ \varphi_\varepsilon(x) &\stackrel{\text{def}}{=} \varepsilon N^{-1/p} |x|^{1/2} + 2N^{1-1/p} n^{1/6} |x|, \qquad x \in \mathbb{R}. \end{aligned}$$

When n_0 is sufficiently large,

$$\begin{aligned} &\mathbb{P}\Big(\Big| \Lambda_4(i,n) - c_0(p)\, N^{-1/p} |b_{i,n}|^{1/2} \Big| > \varphi_\varepsilon(b_{i,n}) \Big) \\ &\qquad \le c_{15}\, n^{-4/3} N^4 + 2\exp(-2n^{1/3}), \end{aligned} \tag{3.11}$$

where $c_0(p)$ is as in (1.5).

Proof. Define the σ–algebra

$$\mathcal{F}_{n,N} \stackrel{\text{def}}{=} \sigma\Big\{ b_{j,n};\, 0 \le j \le N \Big\}.$$

For each $0 \le i \le N-1$, let

$$\xi_i(t) \stackrel{\text{def}}{=} \sqrt{N} \Big(B(\frac{i+t}{N}) - tB(\frac{i+1}{N}) - (1-t)B(\frac{i}{N}) \Big), \qquad 0 \le t \le 1,$$

which, according to Fact 3.6, is a Brownian bridge independent of $\mathcal{F}_{n,N}$. (Actually $\{\xi_i\}_{0 \le i \le N-1}$ are independent Brownian bridges, though we will not use this). Define

$$\begin{aligned} \Lambda_5(y,i,n) &\stackrel{\text{def}}{=} \Big(\int_{i/N}^{(i+1)/N} |B(s+y) - B(s)|^p \,\mathrm{d}s \Big)^{1/p}, \qquad y \in \mathbb{R}, \\ E_n &\stackrel{\text{def}}{=} \Big\{ \max_{0 \le j \le N} |b_{j,n}| < n^{-1/3} \Big\}. \end{aligned}$$

Observe that

$$\Lambda_5(y,i,n) = N^{-1/2-1/p} \Big(\int_0^1 |g(t)|^p \,\mathrm{d}t \Big)^{1/p},$$

where

$$\begin{aligned} g(t) &\stackrel{\text{def}}{=} \sqrt{N} \Big(B(\frac{i+t+yN}{N}) - B(\frac{i+t}{N}) \Big) \\ &= \xi_i(t+yN) - \xi_i(t) + yN^{3/2} (B(\frac{i+1}{N}) - B(\frac{i}{N})). \end{aligned}$$

By the Hölder inequality, on the event E_n, we have

$$\Big| \Lambda_5(y,i,n) - N^{-1/2-1/p}\Big(\int_0^1 |\xi_i(t+yN)-\xi_i(t)|^p \,\mathrm{d}t\Big)^{1/p}\Big|$$
$$\le |y|\, N^{1-1/p}\, |B(\frac{i+1}{N}) - B(\frac{i}{N})|$$
$$\le 2N^{1-1/p} n^{1/6}\, |y|.$$

Write the conditional probability $\mathbb{P}^{\mathcal{F}_{n,N}}(\,\cdot\,) \stackrel{\mathrm{def}}{=} \mathbb{P}(\,\cdot\,|\mathcal{F}_{n,N})$. Note that E_n is an $\mathcal{F}_{n,N}$–measurable event. Therefore, applying (2.8) to the Brownian bridge ξ_i yields that, for any $\mathcal{F}_{n,N}$–measurable random variable Y with $|YN| < n^{-1/6}$,

$$\mathbb{1}_{E_n}\, \mathbb{P}^{\mathcal{F}_{n,N}}\Big(\Big| \Lambda_5(Y,i,n) - c_0(p)\, N^{-1/p}\sqrt{|Y|}\,\Big| > \varphi_\varepsilon(Y)\Big) \le c_{13}\,(YN)^4. \tag{3.12}$$

Choose $Y \stackrel{\mathrm{def}}{=} -\,b_{i,n}$, which effectively is $\mathcal{F}_{n,N}$–measurable and satisfies $|YN| < n^{-1/6}$ on E_n. Note that $\Lambda_5(-b_{i,n},i,n) = \Lambda_4(i,n)$. Take the expectation on both sides of (3.12) to see that

$$\mathbb{P}\Big(\Big| \Lambda_4(i,n) - c_0(p)\, N^{-1/p} |b_{i,n}|^{1/2}\Big| > \varphi_\varepsilon(-b_{i,n});\ E_n\Big) \le c_{13}\, n^{-4/3} N^4.$$

On the other hand,

$$\mathbb{P}(E_n^c) \le \mathbb{P}\Big(\|B\| \ge n^{1/6}\Big) \le 2\mathbb{P}\Big(\sup_{0\le t\le 1} B(t) \ge n^{1/6}\Big) = 2\exp(-2n^{1/3}).$$

(For the exact distribution of $\sup_{0\le t\le 1} B(t)$, cf. for example Csörgő and Révész [8, p. 43]). This completes the proof of Lemma 3.7. □

Proof of Theorem 1.1. Let α_n be an empirical process, whose associated iid KMT Brownian bridges $(B_i)_{i\ge 1}$ are defined via (3.1). Let K_n be as in (3.2). Recall the following strong approximation theorem due to Csörgő and Szyszkowicz [10, Theorem 4.1]: as n goes to infinity,

$$\Big\| R_n - \frac{K_n - K_n(\,\cdot\, - n^{-1}K_n)}{\sqrt{n}}\Big\| = \mathcal{O}\big(n^{-3/8}(\log n)^{3/4}(\log_2 n)^{1/8}\big), \qquad \text{a.s.} \tag{3.13}$$

In view of Fact 3.1 and Lemma 3.3, the proof of Theorem 1.1 is equivalent to showing the following: for $2 \le p < \infty$ and $q = p/2$,

$$\lim_{n\to\infty} \frac{\|K_n - K_n(\,\cdot\, - n^{-1}K_n)\|_p}{\sqrt{\|K_n\|_q}} = c_0(p), \qquad \text{a.s.} \tag{3.14}$$

(Since we do not really need a result as strong as (3.13), we point out that the KMT theorem (i.e. Fact 3.1) – together with Fact 3.4, Lemma 3.3 and some elementary computations – also suffice to imply the equivalence between (3.14) and Theorem 1.1. Observations in this direction have already been made by several authors, cf. for example Deheuvels and Mason [11], Csörgő and Szyszkowicz [10]).

To prove (3.14), let us write

$$N = N(n) \stackrel{\text{def}}{=} [\,(\log n)^{4p}\,],$$

the integer part of $(\log n)^{4p}$. Applying (3.10) to the sequence $a_n = 1/N$ yields that, almost surely for all sufficiently large n,

$$\max_{0\le i\le N-1} \sup_{i/N\le t\le (i+1)/N} |K_n(t) - K_n(\frac{i}{N})| \le 2\sqrt{4p+1}\,\frac{n^{1/2}(\log_2 n)^{1/2}}{N^{1/2}}. \tag{3.15}$$

Note that, for any fixed $r \ge 1$, there exists a finite constant $c_{16} = c_{16}(r)$, such that

$$|x^r - y^r| \le c_{16}\,\big(\,|x-y|^r + y^{r-1}|x-y|\big), \qquad x\ge 0,\ y\ge 0. \tag{3.16}$$

We can use this inequality for $r = q$, $x = |K_n(i/N)|$ and $y = |K_n(t)|$, to see that, almost surely as n goes to infinity, uniformly for $0 \le i \le N-1$ and $t \in [i/N, (i+1)/N]$,

$$\begin{aligned}
&|K_n(t)|^q - |K_n(\frac{i}{N})|^q \\
&\quad = \mathcal{O}\Big(\frac{n^{q/2}(\log_2 n)^{q/2}}{N^{q/2}}\Big) + \mathcal{O}\Big(\frac{n^{1/2}(\log_2 n)^{1/2}}{N^{1/2}}\,\|K_n\|^{q-1}\Big)\\
&\quad = \mathcal{O}\Big(\frac{n^{q/2}(\log_2 n)^{q/2}}{N^{q/2}}\Big) + \mathcal{O}\Big(\frac{n^{q/2}(\log_2 n)^{q/2}}{N^{1/2}}\Big),
\end{aligned}$$

the last identity following from the Chung–Smirnov LIL (cf. (3.3)). Integrating over $t \in [i/N, (i+1)/N]$ and then summing over i, we obtain,

$$\|K_n\|_q^q - \frac{1}{N}\sum_{i=0}^{N-1} |K_n(\frac{i}{N})|^q = \mathcal{O}\Big(\frac{n^{q/2}(\log_2 n)^{q/2}}{N^{1/2}}\Big), \qquad \text{a.s.},$$

which, according to Lemma 3.3, is $o(\|K_n\|_q^q)$, almost surely. As a consequence,

$$\|K_n\|_q^q \sim \frac{1}{N}\sum_{i=0}^{N-1} |K_n(\frac{i}{N})|^q, \qquad \text{a.s.} \tag{3.17}$$

Now fix $0 < \varepsilon < 1$. Write, for each $0 \le i \le N-1$,

$$
\begin{aligned}
k_{i,n} &\stackrel{\text{def}}{=} K_n\big(\frac{i}{N}\big),\\
\Lambda_6(i,n) &\stackrel{\text{def}}{=} \Big(\int_{i/N}^{(i+1)/N} |K_n(t-\frac{k_{i,n}}{n}) - K_n(t)|^p \, \mathrm{d}t\Big)^{1/p},\\
\psi_\varepsilon(x) &\stackrel{\text{def}}{=} \varepsilon N^{-1/p}|x|^{1/2} + 2N^{1-1/p}n^{-1/3}\,|x|, \qquad x \in \mathbb{R}.
\end{aligned}
$$

Since for each n, $n^{-1/2}K_n$ is a Brownian bridge, applying Lemma 3.7 to $n^{-1/2}K_n$ (instead of to B) yields that, for $n \ge 1$,

$$
\begin{aligned}
&\mathbb{P}\Big(\Big|\,\Lambda_6(i,n) - c_0(p)\,N^{-1/p}|k_{i,n}|^{1/2}\,\Big| > \psi_\varepsilon(k_{i,n}),\ \text{for some } 0 \le i < N\Big)\\
&\qquad \le \sum_{i=0}^{N-1} \mathbb{P}\Big(\Big|\,\Lambda_6(i,n) - c_0(p)\,N^{-1/p}|k_{i,n}|^{1/2}\,\Big| > \psi_\varepsilon(k_{i,n})\Big)\\
&\qquad \le c_{15}\, n^{-4/3}N^5 + 2N\exp\big(-2n^{1/3}\big).
\end{aligned}
$$

The expression on the right hand side being summable for n, we can use the Borel–Cantelli lemma to see that, almost surely for all large n and all $0 \le i \le N-1$,

$$
\Big|\,\Lambda_6(i,n) - c_0(p)\,N^{-1/p}|k_{i,n}|^{1/2}\,\Big| \le \psi_\varepsilon(k_{i,n}).
$$

Let us apply (3.16) to $r = p$, $x = \Lambda_6(i,n)$ and $y = c_0(p)\,N^{-1/p}|k_{i,n}|^{1/2}$. For this choice of (r,x,y), we have $y \le c_0(p)\,\psi_\varepsilon(k_{i,n})/\varepsilon$, which implies that, almost surely for all large n and all $0 \le i \le N-1$,

$$
\begin{aligned}
&\Big|\int_{i/N}^{(i+1)/N} |K_n(t-\frac{k_{i,n}}{n}) - K_n(t)|^p \, \mathrm{d}t - N^{-1}|k_{i,n}|^q\,\mathbb{E}(|\mathcal{N}|^p)\Big|\\
&\qquad \le c_{16}\,\Big(\big(\psi_\varepsilon(k_{i,n})\big)^p + \big(\frac{c_0(p)}{\varepsilon}\big)^{p-1}\big(\psi_\varepsilon(k_{i,n})\big)^p\Big)\\
&\qquad \le \frac{c_{17}}{\varepsilon^{p-1}}\,\big(\psi_\varepsilon(k_{i,n})\big)^p\\
&\qquad \le c_{18}\,\varepsilon\, N^{-1}|k_{i,n}|^q + \frac{c_{18}}{\varepsilon^{p-1}}\,N^{p-1}n^{-p/3}\,|k_{i,n}|^p,
\end{aligned}
$$

where $c_{17} = c_{17}(p)$ and $c_{18} = c_{18}(p)$ depend only on p. Summing over i gives that, for any $\varepsilon > 0$, when n is sufficiently large,

$$
\begin{aligned}
&\Big|\sum_{i=0}^{N-1}\int_{i/N}^{(i+1)/N} |K_n(t-\frac{k_{i,n}}{n}) - K_n(t)|^p \, \mathrm{d}t - \frac{\mathbb{E}(|\mathcal{N}|^p)}{N}\sum_{i=0}^{N-1}|K_n(\frac{i}{N})|^q\Big|\\
&\qquad \le \frac{c_{18}\,\varepsilon}{N}\sum_{i=0}^{N-1}|K_n(\frac{i}{N})|^q + \frac{c_{18}}{\varepsilon^{p-1}}\,N^{p-1}\,n^{-p/3}\sum_{i=0}^{N-1}|K_n(\frac{i}{N})|^p. \qquad (3.18)
\end{aligned}
$$

According to (3.17), for large n,

$$\frac{1}{N}\sum_{i=0}^{N-1}|K_n(\frac{i}{N})|^p \sim \|K_n\|_p^p \le \|K_n\|^p \le (n\log_2 n)^q,$$

the last inequality following from the Chung–Smirnov LIL, cf. (3.3). Hence, as n goes to infinity,

$$N^{p-1}\, n^{-p/3}\sum_{i=0}^{N-1}|K_n(\frac{i}{N})|^p = \mathcal{O}\Big(n^{p/6}N^p(\log_2 n)^q\Big), \qquad \text{a.s.},$$

which, in view of Lemma 3.3 and (3.17), gives that

$$N^{p-1}\, n^{-p/3}\sum_{i=0}^{N-1}|K_n(\frac{i}{N})|^p = o(\|K_n\|_q^q) = o\Big(\frac{1}{N}\sum_{i=0}^{N-1}|K_n(\frac{i}{N})|^q\Big).$$

Going back to (3.18), we obtain: for any $0<\varepsilon<1$ and all large n,

$$\Big|\sum_{i=0}^{N-1}\int_{i/N}^{(i+1)/N}|K_n(t-\frac{k_{i,n}}{n})-K_n(t)|^p\,\mathrm{d}t - \frac{\mathbb{E}(|\mathcal{N}|^p)}{N}\sum_{i=0}^{N-1}|K_n(\frac{i}{N})|^q\Big|$$
$$\le \frac{c_{18}\,\varepsilon}{N}\sum_{i=0}^{N-1}|K_n(\frac{i}{N})|^q + o\Big(\frac{1}{N}\sum_{i=0}^{N-1}|K_n(\frac{i}{N})|^q\Big).$$

Since c_{18} does not depend on ε, and since $\varepsilon>0$ can be as small as possible, we conclude that almost surely,

$$\sum_{i=0}^{N-1}\int_{i/N}^{(i+1)/N}|K_n(t-\frac{k_{i,n}}{n})-K_n(t)|^p\,\mathrm{d}t \sim \frac{\mathbb{E}(|\mathcal{N}|^p)}{N}\sum_{i=0}^{N-1}|K_n(\frac{i}{N})|^q$$
$$\sim \mathbb{E}(|\mathcal{N}|^p)\,\|K_n\|_q^q, \tag{3.19}$$

the last line following from (3.17).

We are now ready to complete the proof of Theorem 1.1. Indeed, by (3.15) and applying Fact 3.4 to $a_n = 2\sqrt{4p+1}\,n^{-1/2}N^{-1/2}(\log_2 n)^{1/2}$, we have

$$\max_{0\le i\le N-1}\ \sup_{i/N\le t\le (i+1)/N}|K_n(t-\frac{k_{i,n}}{n})-K_n(t-\frac{K_n(t)}{n})|$$
$$= \mathcal{O}\Big(n^{1/4}(\log n)^{1/2-p}(\log_2 n)^{1/4}\Big), \qquad \text{a.s.},$$

which, in view of (3.16), implies that uniformly for $0\le i\le N-1$ and $t\in[i/N,(i+1)/N]$,

$$|K_n(t-\frac{k_{i,n}}{n})-K_n(t)|^p - |K_n(t-\frac{K_n(t)}{n})-K_n(t)|^p$$
$$= \mathcal{O}\Big(n^{p/4}(\log n)^{p/2-p^2}(\log_2 n)^{p/4}\Big)$$
$$+\,\mathcal{O}\Big(n^{1/4}(\log n)^{1/2-p}(\log_2 n)^{1/4}\,(\Lambda_7(n))^{p-1}\Big), \qquad \text{a.s.}, \tag{3.20}$$

where

$$\Lambda_7(n) \stackrel{\text{def}}{=} \sup_{0\le t\le 1} |K_n(t-\frac{K_n(t)}{n}) - K_n(t)|.$$

Recall $\omega_n(\,\cdot\,)$ from (3.8). Since $\|K_n\| \le \sqrt{n\log_2 n}$ almost surely for all large n (cf. (3.3)), by Fact 3.4, we have, for large n,

$$\begin{aligned}\Lambda_7(n) &\le \omega_n\big(n^{-1/2}(\log_2 n)^{1/2}\big)\\ &= \mathcal{O}\Big(n^{1/4}(\log n)^{1/2}(\log_2 n)^{1/4}\Big), \qquad \text{a.s.}\end{aligned}$$

(Actually, it can be deduced from Theorem A that $\Lambda_7(n) \sim (\log n)^{1/2}\|K_n\|^{1/2}$, which in turn gives us the exact asymptotics of $\Lambda_7(n)$. For more details, cf. (A.1.11) of Csörgő and Horváth [6, p. 417]). In view of (3.20), and integrating with respect to $t \in [i/N, (i+1)/N]$ and then summing over i, we obtain: almost surely when n goes to infinity,

$$\begin{aligned}&\sum_{i=0}^{N-1}\int_{i/N}^{(i+1)/N} |K_n(t-\frac{k_{i,n}}{n}) - K_n(t)|^p\,\mathrm{d}t\\ &\quad= \int_0^1 |K_n(t-\frac{K_n(t)}{n}) - K_n(t)|^p\,\mathrm{d}t + \mathcal{O}\Big(n^{p/4}(\log n)^{-p/2}(\log_2 n)^{p/4}\Big).\end{aligned}$$

Together with (3.19) and (3.5), this gives

$$\int_0^1 |K_n(t-\frac{K_n(t)}{n}) - K_n(t)|^p\,\mathrm{d}t \;\sim\; \mathbb{E}(|\mathcal{N}|^p)\,\|K_n\|_q^q, \qquad \text{a.s.}$$

We have therefore proved (3.14), hence completed the proof of Theorem 1.1. □

To check Corollary 1.4, we need the following estimate.

Lemma 3.8. *For any* $q \ge 1$,

$$\limsup_{n\to\infty} \frac{\|K_n\|_q}{\sqrt{2n\log_2 n}} = c_3(q), \qquad \text{a.s.,} \tag{3.21}$$

where $c_3(q)$ *is defined in* (1.6).

Proof. Lemma 3.8 actually is known, cf. Gajek et al. [16] for a direct proof. However, it turns out that it can also be deduced, by means of a simple argument, from some classical results for empirical processes via the KMT strong invariance. So we outline the argument here, which might be of some interest.

That the "limsup" expression in (3.21) should be equal to a constant of particular form, is a straightforward consequence of Finkelstein's functional LIL for the empirical process. In fact, according to Finkelstein [15],

$$\limsup_{n\to\infty} \frac{\|K_n\|_q}{\sqrt{2n\log_2 n}} = \sup_{f\in\mathcal{F}} \|f\|_q\,, \qquad \text{a.s.},$$

where $\mathcal{F} \stackrel{\text{def}}{=} \{f\,:\, f(t)=\int_0^t \dot f(s)\,\mathrm{d}s,\;\; f(1)=0,\;\; \int_0^1 (\dot f(s))^2\,\mathrm{d}s \le 1\}$ is the so-called Finkelstein's set. Fortunately, to get the exact value of $\sup_{f\in\mathcal{F}}\|f\|_q$, we do not have to do any technical computation. Indeed, Strassen [28] solved a variational problem and calculated the value of $\sup_{f\in\mathcal{S}}\|f\|_q$, where $\mathcal{S} \stackrel{\text{def}}{=} \{f\,:\, f(t)=\int_0^t \dot f(s)\,\mathrm{d}s,\;\; \int_0^1 (\dot f(s))^2\,\mathrm{d}s \le 1\}$ is Strassen's set. From this, a simple argument using symmetry and scaling readily yields the value of $\sup_{f\in\mathcal{F}}\|f\|_q$, cf. [9] for more details. □

Proof of Corollary 1.4. Follows from Theorem 1.1, Fact 3.1, Lemmas 3.8 and 3.3. □

4. Appendix

Let $\{X_i\}_{i\ge1}$ be a sequence of real valued independent and identically distributed random variables with a right continuously defined distribution function F, and let $\widetilde F_n(\cdot)$ be the empirical distribution function based on the first n observations, i.e.,

$$\widetilde F_n(x) \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{X_i\le x\}}, \qquad x\in\mathbb{R}.$$

Define the general empirical process

$$\widetilde\alpha_n(x) \stackrel{\text{def}}{=} n^{1/2}\big(\tilde F_n(x)-F(x)\big), \qquad x\in\mathbb{R}, \tag{4.1}$$

and its "corresponding" empirical quantile process

$$\widetilde\beta_n(t) \stackrel{\text{def}}{=} n^{1/2}\big(\widetilde F_n^{-1}(t)-F^{-1}(t)\big), \qquad 0\le t\le 1, \tag{4.2}$$

where

$$\widetilde F_n^{-1}(t) \stackrel{\text{def}}{=} \inf\{x:\, \widetilde F_n(x)\ge t\},\; 0<t\le 1,\; \widetilde F_n^{-1}(0) \stackrel{\text{def}}{=} \widetilde F_n^{-1}(0+),$$

and similarly $F^{-1}(\cdot)$ in terms of $F(\cdot)$, are the empirical and theoretical quantile functions, i.e., the inverse functions of $\widetilde F_n(\cdot)$ and $F(\cdot)$ respectively.

Clearly, with any continuous distribution function $F(\cdot)$ the random variables $\{U_i \stackrel{\text{def}}{=} F(X_i)\}_{i\geq 1}$ are independent $(0,1)$–uniform random variables and, in terms of these uniformly distributed random variables, we have

$$\begin{aligned}\widetilde{\alpha}_n\big(F^{-1}(t)\big) &= n^{1/2}\big(\widetilde{F}_n\big(F^{-1}(t)\big) - t\big) \\ &= n^{1/2}\big(F_n(t) - t\big) = \alpha_n(t), \quad 0 \leq t \leq 1,\end{aligned} \tag{4.3}$$

as well as

$$n^{1/2}\Big(F\Big(\widetilde{F}_n^{-1}(t)\Big) - t\Big) = n^{1/2}\big(F_n^{-1}(t) - t\big) = \beta_n(t), \qquad 0 \leq t \leq 1. \tag{4.4}$$

Moreover, by the mean value theorem, we can write

$$\begin{aligned}\widetilde{\beta}_n(t) &\stackrel{\text{def}}{=} n^{1/2}\big(\widetilde{F}_n^{-1}(t) - F^{-1}(t)\big) \\ &= n^{1/2}\big(F^{-1}\big(F\big(\widetilde{F}_n^{-1}(t)\big)\big) - F^{-1}(t)\big) \\ &= n^{1/2}\big(F^{-1}\big(F_n^{-1}(t)\big) - F^{-1}(t)\big) \\ &= \beta_n(t)/f\big(F^{-1}(\theta_n(t))\big), \qquad 0 < t < 1,\end{aligned} \tag{4.5}$$

where $F_n^{-1}(t) \wedge t < \theta_n(t) < F_n^{-1}(t) \vee t$, $0 < t < 1$, *provided of course that we have* $\mathrm{d}F^{-1}(t)/\mathrm{d}t = 1/f(F^{-1}(t)) < \infty$ for $t \in (0,1)$, i.e., provided that F is an absolutely continuous distribution function (with respect to Lebesgue measure) with a strictly positive density function $f = F'$ on the real line. The function $f(F^{-1}(t))$ is called the **density quantile function**, and $1/f(F^{-1}(t))$ the **quantile density function** in Parzen [21], [22].

In view of (4.4) and (4.5) it is clear that if one were to study the empirical quantile process $\widetilde{\beta}_n$ on $[0,1]$ via its $(0,1)$–uniform version β_n, then one should *renormalize* the former by multiplying it by its density quantile function $f(F^{-1})$. Hence we assume that $f = F'$ exists on the real line and, as in Csörgő and Révész [7], we define the **general empirical quantile process** δ_n by

$$\delta_n(t) \stackrel{\text{def}}{=} f(F^{-1}(t))\widetilde{\beta}_n(t), \qquad 0 \leq t \leq 1, \tag{4.6}$$

with $\widetilde{\beta}_n$ as in (4.2), and the **general Bahadur–Kiefer process** $\widetilde{R}_n$ by

$$\widetilde{R}_n(t) \stackrel{\text{def}}{=} \alpha_n(t) + \delta_n(t), \qquad 0 \leq t \leq 1, \tag{4.7}$$

with $\alpha_n(t) = \widetilde{\alpha}_n(F^{-1}(t))$ (cf. (4.1) and (4.3)).

Csörgő and Révész [7] initiated the study of the almost sure asymptotic behaviour of the processes δ_n and $\widetilde{R}_n$ under the following assumptions on their underlying distribution function F:

(4.8) (i) F is twice differentiable on (a, b), where $a \stackrel{\text{def}}{=} \sup\{x : F(x) = 0\}$, $b \stackrel{\text{def}}{=} \inf\{x : F(x) = 1\}$, $-\infty \le a < b \le \infty$,

(ii) $F'(x) = f(x) > 0, \quad x \in (a, b)$,

(iii) for some constant $\gamma > 0$ we have

$$\sup_{0<t<1} \frac{t(1-t)|f'(F^{-1}(t))|}{f^2(F^{-1}(t))} \le \gamma,$$

(iv) the two limits $A \stackrel{\text{def}}{=} \lim_{x \downarrow a} f(x)$ and $B \stackrel{\text{def}}{=} \lim_{x \uparrow b} f(x)$ are finite, and

(iv$a$) both limits A and B are positive, or

(iv$b$) if the limit $A = 0$ (resp., $B = 0$), then assume also that the density function f is nondecreasing on a right neighbourhood of a (resp., nonincreasing on a left neighbourhood of b).

Under these conditions Csörgő and Révész [7] show that the general empirical quantile process δ_n of (4.6) is almost surely $o(n^{-1/2+\varepsilon})$ near, with any small $\varepsilon > 0$, to the uniform empirical quantile process β_n of (4.4) in the uniform sup-norm $\|\cdot\|$. This, in turn, on account of (cf. R_n and $\widetilde{R}_n$ via (4.3) and (4.4))

$$\begin{aligned} \{\widetilde{R}_n(t) - R_n(t);\ 0 \le t \le 1,\ n = 1, 2, \dots\} \\ = \{\delta_n(t) - \beta_n(t);\ 0 \le t \le 1,\ n = 1, 2, \dots\}, \end{aligned} \tag{4.9}$$

leads to an extension of the Bahadur–Kiefer theory of quantiles for $\widetilde{R}_n$ under appropriate subsets of the just listed conditions in (4.8). For related details of these extensions we refer to Csörgő and Révész ([7], [8, Sections 5.2 and 5.3]), Csörgő ([5, Chapter 6]), Csörgő and Horváth ([6, Section 6.5]), and Csörgő and Szyszkowicz ([10, Section 5]). In particular, as an extension of Theorem A and that of its consequences, we have (cf. Csörgő and Szyszkowicz [10, Theorem 5.1 and Corollaries 5.A and 5.B]) the following conclusions. For brevity, we write from now on:

$$\widetilde{S}_n(t) \stackrel{\text{def}}{=} \widetilde{R}_n(t)\, \mathbb{1}_{\{1/(n+1) \le t \le n/(n+1)\}}, \qquad 0 \le t \le 1.$$

Theorem B. *Given the conditions* (4.8) (i), (ii) *and* (iii) *on* F, *we have*

$$\lim_{n\to\infty} n^{1/4} (\log n)^{-1/2} \frac{\|\widetilde{S}_n\|}{\sqrt{\|\alpha_n\|}} = 1, \qquad \text{a.s.,} \tag{4.10}$$

$$\limsup_{n\to\infty} n^{1/4}(\log n)^{-1/2}(\log_2 n)^{-1/4}\|\widetilde{S}_n\| = 2^{-1/4}, \qquad \text{a.s.,} \tag{4.11}$$

$$\liminf_{n\to\infty} n^{1/4}(\log n)^{-1/2}(\log_2 n)^{-1/4}\|\widetilde{S}_n\| = \frac{\pi^{1/2}}{8^{1/4}}, \qquad \text{a.s.,} \tag{4.12}$$

and, as $n \to \infty$,

$$n^{-1/4}(\log n)^{-1/2}\|\widetilde{S}_n\| \xrightarrow{\mathcal{D}} \sqrt{\|B\|}, \tag{4.13}$$

where $\{B(t);\ 0 \le t \le 1\}$ *is a Brownian Bridge.*

Moreover, if in addition to the conditions (4.8)(i),(ii) *and* (iii), *we have also* (4.8) (iv) *and* (iv$a$), *or* (4.8) (iv) *and* (iv$b$), *then* (4.10)–(4.13) *remain true for* $\widetilde{R}_n$ *in the place of* $\widetilde{S}_n$.

Along similar lines now, as an extension of Theorem 1.1 and that of Corollary 1.4, we have also the following theorem.

Theorem 4.1. *Let* $2 \le p < \infty$ *and* $q = p/2$. *Then, assuming the conditions* (4.8) (i), (ii) *and* (iii) *on* F, *we have*

$$\lim_{n\to\infty} n^{1/4}\frac{\|\widetilde{S}_n\|_p}{\sqrt{\|\alpha_n\|_q}} = c_0(p), \qquad \text{a.s.,} \tag{4.14}$$

$$\limsup_{n\to\infty} n^{-1/4}(\log_2 n)^{-1/4}\|\widetilde{S}_n\|_p = 2^{1/4}c_0(p)\sqrt{c_3(q)}, \qquad \text{a.s.,} \tag{4.15}$$

$$\liminf_{n\to\infty} n^{1/4}(\log_2 n)^{-1/4}\|\widetilde{S}_n\|_p = c_0(p)\sqrt{c_4(q)}, \qquad \text{a.s.,} \tag{4.16}$$

and, as $n \to \infty$,

$$n^{1/4}\|\widetilde{S}_n\|_p \xrightarrow{\mathcal{D}} c_0(p)\sqrt{\|B\|_q}, \tag{4.17}$$

where $\{B(t);\ 0 \le t \le 1\}$ *is a Brownian bridge, and* $c_0(p)$, $c_3(q)$ *and* $c_4(q)$ *are as in* (1.5), (1.6) *and* (1.7) *respectively.*

Moreover, if in addition to the conditions (4.8) (i), (ii) *and* (iii), *we have also* (4.8) (iv) *and* (iv$a$), *or* (4.8) (iv) *and* (iv$b$), *then* (4.14)–(4.17) *remain true for* $\widetilde{R}_n$ *in the place of* $\widetilde{S}_n$.

Remark 4.2. For the sake of applications of Theorems B and 4.1 to goodness-of-fit problems, we call attention to (4.13) and (4.17), as well as to their respective versions without the indicator function $t \mapsto \mathbb{1}_{\{1/(n+1)\le t\le n/(n+1)\}}$. In particular, it is of interest to note that we have (1.9) in terms of (4.13), as well as (1.12) in terms of (4.17) with $p = 4$

and $q = 2$, which both accommodate the whole random sample $X_1, X_2, \ldots, X_n$ of size $n \geq 1$ on F under the conditions (4.8) (i), (ii) and (iii). Moreover, from the goodness-of-fit point of view in general, (1.10) in terms of (4.17) is a brand new alternative to (1.9) in terms of (4.13).

Proof of Theorem 4.1. Let K_n be as in (3.2). We have (cf. Csörgő and Szyszkowicz [10, Theorems 5.2 and 5.3 respsectively]): as $n \to \infty$, under the conditions (4.8) (i), (ii) and (iii) on F,

$$\Big\| \widetilde{S}_n - \frac{K_n - K_n(\cdot - n^{-1}K_n)}{\sqrt{n}} \Big\| = O\big(n^{-3/8}(\log n)^{3/4}(\log_2 n)^{1/8}\big), \qquad \text{a.s.,} \tag{4.18}$$

and, on assuming also (4.8) (iv) and (iv$a$), or (4.8) (iv) and (iv$b$),

$$\Big\| \widetilde{R}_n - \frac{K_n - K_n(\cdot - n^{-1}K_n)}{\sqrt{n}} \Big\| = O\big(n^{-3/8}(\log n)^{3/4}(\log_2 n)^{1/8}\big), \qquad \text{a.s.} \tag{4.19}$$

Consequently, the rest of the proof of (4.14) of Theorem 4.1 is identical to that of Theorem 1.1, while the respective proofs of (4.15) and (4.16) are now similarly contained in that of Corollary 1.4. □

References

[1] Anderson, T.W. and Darling, D.A.: Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Statist.* 23 (1952) 193–212.

[2] Bahadur, R.R.: A note on quantiles in large samples. *Ann. Math. Statist.* 37 (1966) 577–580.

[3] Borovkov, A.A. and Mogulskii, A.A.: On probabilities of small deviations for stochastic processes. *Siberian Adv. Math.* 1 (1991) 39–63.

[4] Chung, K.L.: An estimate concerning the Kolmogorov limit distribution. *Trans. Amer. Math. Soc.* 67 (1949) 36–50.

[5] Csörgő, M.: *Quantile Processes with Statistical Applications.* SIAM, Philadelphia, PA, 1983.

[6] Csörgő, M. and Horváth, L.: *Weighted Approximations in Probability and Statistics.* Wiley, Chichester, 1993.

[7] Csörgő, M. and Révész, P.: Strong approximations of the quantile process. *Ann. Statist.* 6 (1978) 882–894.

[8] Csörgő, M. and Révész, P.: *Strong Approximations in Probability and Statistics.* Academic Press, New York, 1981.

[9] Csörgő, M., Shi, Z. and Yor, M.: Some asymptotic properties of the local time of the uniform empirical process. *Bernoulli.* (To appear).

[10] Csörgő, M. and Szyszkowicz, B.: Sequential quantile and Bahadur–Kiefer processes. In: *Handbook of Statistics* 16, *Order Statistics: Theory and Methods* (N. Balakrishnan and C.R. Rao, eds.) Ch. 21, pp. 631-688, 1988 Elsevier Science B.V.

[11] Deheuvels, P. and Mason, D.M.: Bahadur–Kiefer-type processes. *Ann. Probab.* 18 (1990) 669–697.

[12] Donsker, M.D. and Varadhan, S.R.S.: On laws of the iterated logarithm for local times. *Comm. Pure Appl. Math.* 30 (1977) 707–753.

[13] Doob, J.L.: Heuristic approach to the Kolmogorov–Smirnov theorems. *Ann. Math. Statist.* 20 (1949) 393–403.

[14] Einmahl, J.H.J.: A short and elementary proof of the main Bahadur–Kiefer theorem. *Ann. Probab.* 24 (1996) 526–531.

[15] Finkelstein, H.: The law of the iterated logarithm for empirical distributions. *Ann. Math. Statist.* 42 (1971) 607–615.

[16] Gajek, L., Kałuszka, M. and Lenic, A.: The law of the iterated logarithm for L_p-norms of empirical processes. *Statist. Probab. Lett.* 28 (1996) 107–110.

[17] Kiefer, J.: Deviations between the sample quantile process and the sample df. In: *Nonparametric Techniques in Statistical Inference*, pp. 299–319. Cambridge University Press, Cambridge, 1970.

[18] Komlós, J., Major, P. and Tusnády, G.: An approximation of partial sums of independent R.V.'s and the sample DF. I. *Z. Wahrsch. verw. Gebiete* 32 (1975) 111–131.

[19] Mason, D.M., Shorack, G.R. and Wellner, J.A.: Strong limit theorems for oscillation moduli of the uniform empirical process. *Z. Wahrsch. verw. Gebiete* 65 (1983) 83–97.

[20] Mogulskii, A.A.: On the law of the iterated logarithm in Chung's form in functional space. *Th. Probab. Appl.* 24 (1979) 405–413.

[21] Parzen, E.: Nonparametric statistical data modeling. With comments by John W. Tukey, Roy E. Welsch, William F. Eddy, D.V. Lindley, Michael E. Tarter and Edwin L. Crow, and a rejoinder by the author. *J. Amer. Statist. Assoc.* 74 (1979) 105–131.

[22] Parzen, E.: A density quantile function perspective on robust estimation. In: *Proceedings of a Seminar on Robust Estimation.* (Eds. R. Launer and G. Wilkinson.) Academic Press, New York, 1979.

[23] Petrov, V.V.: *Limit Theorems of Probability Theory.* Oxford Science Publications, Oxford, 1995.

[24] Shorack, G.R.: Kiefer's theorem via the Hungarian construction. *Z. Wahrsch. verw. Gebiete* 61 (1982) 369–373.

[25] Shorack, G.R. and Wellner, J.A.: *Empirical Processes with Applications to Statistics.* Wiley, New York, 1986.

[26] Smirnov, N.V.: On the distribution of the von Mises ω^2–criterion (in Russian). *Matem. Sbornik.* 5 (1937) 973–993.

[27] Smirnov, N.V.: Approximate laws of distribution of random variables from empirical data (in Russian). *Uspehi Mat. Nauk.* 10 (1944) 179–206.

[28] Strassen, V.: An invariance principle for the law of the iterated logarithm. *Z. Wahrsch. verw. Gebiete* 3 (1964) 211–226.

[29] Stute, W.: The oscillation behavior of empirical processes. *Ann. Probab.* 10 (1982) 86–107.

[30] Takács, L.: On the distribution of the integral of the absolute value of the Brownian motion. *Ann. Appl. Probab.* 3 (1993) 186–197.

[31] Vervaat, W.: Functional central limit theorems for processes with positive drift and their inverses. *Z. Wahrsch. Verw. Gebiete* 23 (1972) 245–253.

[32] Vervaat, W.: *Success Epochs in Bernoulli Trials: with Applications to Number Theory.* Mathematical Centre Tracts 42 (1972) (Second edition in 1977). Matematisch Centrum, Amsterdam.

[33] Zitikis, R.: The Vervaat processes. In: *Asymptotic Methods in Probability and Statistics – a Volume in Honour of Miklós Csörgő* (B. Szyszkowicz, ed.) 1998 Elsevier Science B.V. (To appear).