On the strong approximation of the non-overlapping \( k \)-spacings process with application to the moment convergence rates

Sergio Alvarez-Andrade  
Laboratoire de Mathématiques Appliquées, Université de Technologie de Compiègne, B.P. 529, 60205 Compiègne cedex, France  
email: sergio.alvarez@utc.fr

Salim Bouzebda  
Laboratoire de Mathématiques Appliquées, Université de Technologie de Compiègne, B.P. 529, 60205 Compiègne cedex, France  
email: salim.bouzebda@utc.fr

Nabil Nessigha  
Laboratoire de Mathématiques Appliquées, Université de Technologie de Compiègne, B.P. 529, 60205 Compiègne cedex, France  
email: nabil.nessigha@gmail.fr

Abstract. In the present work, we establish the strong approximations of the empirical \( k \)-spacings process \( \{\alpha_n(x) : 0 \leq x < \infty\} \) (cf. (3)). We state the moment convergence rates results for this process.

1 Introduction

Let \( U_1, U_2, \ldots \), be independent and identically distributed (i.i.d.) uniform on \([0, 1]\) random variables (r.v.’s) defined on the same probability space \( (\Omega, \mathcal{A}, \mathbb{P}) \). Denote by \( 0 =: U_{0,n} \leq U_{1,n} \leq \cdots \leq U_{n-1,n} \leq U_{n,n} := 1 \), the order statistics of

2010 Mathematics Subject Classification: 62E20; 60F15  
Key words and phrases: stochastic processes; strong approximations; Gaussian processes; moment convergence rates, \( k \)-spacings process
Non-overlapping k-spacings process

\[ U_1, U_2, \ldots, U_{n-1}, \text{ on } [0, 1]. \]

The corresponding non-overlapping k-spacings are then defined by

\[
\begin{align*}
D_{i,n}^k & := U_{ik,n} - U_{(i-1)k,n}, \quad \text{for } 1 \leq i \leq N - 1, \\
D_{N,n}^k & := 1 - U_{(N-1)k,n},
\end{align*}
\]  \(1\)

where \(N = \lfloor n/k \rfloor\), with \(\lfloor u \rfloor \leq u < \lfloor u \rfloor + 1\) denoting the integer part of \(u\).

When \(k = 1\), i.e., \(N = n\), the k-spacings reduce to the usual 1-spacings (or simple spacings) see for instance [18], [19], [20], [45] and [46]. Some useful bibliographical references related to k-spacings can be seen in [2], namely the applications for the goodness-of-fit tests, Kolmogorov-Smirnov tests (see, for instance, [22] and [23]).

We will use in the sequel the normalized non-overlapping k-spacings \(\{D_{i,n}^k : 1 \leq i \leq N\} \) for a fixed \(k \geq 1\), as \(n \to \infty\), the distribution function of \(kD_{i,n}^k\) (which is independent of the index \(i\) with \(1 \leq i \leq N-1\)) converges to the distribution function \(F_k(\cdot)\), of a standard gamma random variable with expectation \(k\), given by

\[
F_k(t) := \frac{1}{(k-1)!} \int_0^t x^{k-1} e^{-x} dx. \tag{2}
\]

For each choice of \(k \geq 1\), the empirical k-spacings process is defined by

\[
\alpha_n(x) := N^{1/2} \left( \hat{F}_n(x) - F_k(x) \right), \quad \text{for } x > 0, \tag{3}
\]

where \(\hat{F}_n(\cdot)\) is the empirical distribution function of \(\{kD_{i,n}^k : 1 \leq i \leq N\}\), defined for \(n \geq m\), by

\[
\hat{F}_n(x) := \frac{1}{N} \sum_{i=1}^N \mathbb{1}_{\{kD_{i,n}^k \leq x\}}, \quad \text{for } x \in \mathbb{R}, \tag{4}
\]

with \(\mathbb{1}_{\{A\}}\) denoting the indicator function of the event \(A\). This paper aims to obtain a refinement of the strong approximation results for \(\{\alpha_n(x) : 0 \leq x < \infty, n \geq 1\}\) obtained by [7]. To prove the invariance principle, we use the same method developed in [7], which is based on the following representation of simple spacings given by [39]. In the sequel of this section, we use a notation similar to that of [7]. Let \(E_1, E_2, \ldots\) denote an i.i.d. sequence of exponential r.v.'s with mean 1, then, for each \(n > 1\), we have the following representation of the non-overlapping k-spacings

\[
\left\{ D_{i,n}^k : 1 \leq i \leq N - 1, D_{N,n}^k \right\}
\]
$$d = \left\{ \frac{\sum_{\ell=i}^{i+k-1} E_{\ell}}{S_n} : i = 1, k + 1, \ldots, \left( \left\lfloor \frac{n}{k} \right\rfloor - 1 \right) k + 1, \frac{\sum_{\ell=\left\lceil \frac{n}{k} \right\rceil}^{n} E_{\ell}}{S_n} \right\}, \quad (5)$$

where

$$S_n := \sum_{i=1}^{n} E_i.$$ 

In particular, if $n = Nk$ is an integer multiple of $k$, then

$$\left\{ D_{i,n}^k, 1 \leq i \leq N \right\} \overset{d}{=} \left\{ \frac{Y_i}{T_N}, 1 \leq i \leq N \right\}, \quad (6)$$

where

$$Y_i := \sum_{\ell=\left(\left( i-1 \right) k + 1 \right)}^{ik} E_{\ell}, \text{ for } i = 1, 2, \ldots, N, \quad (7)$$

is a sequence of i.i.d. r.v.'s with distribution function $F_k(\cdot)$ and $T_N = \sum_{i=1}^{N} Y_i$. Now, we denote by $G_N(\cdot)$ the empirical distribution function of the sequence $Y_1, \ldots, Y_N$, defined by

$$G_N(x) := \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}_{(Y_i \leq x)}, \text{ for } x \in \mathbb{R}^+.$$

(8)

Let $\{\beta_N(x) : 0 \leq x < \infty, N \geq 1\}$ be the corresponding empirical processes, defined by

$$\beta_N(x) := \sqrt{N} \left( G_N(x) - F_k(x) \right), \text{ for } x \in \mathbb{R}^+. \quad (9)$$

By (6), we have the following representation

$$\{\alpha_{Nk}(x), 0 \leq x < \infty\} \overset{d}{=} \left\{ \alpha_N^1(x) = \beta_N \left( x \frac{T_N}{Nk} \right) + \mathcal{R}_N(x), 0 \leq x < \infty \right\}, \quad (10)$$

where

$$\mathcal{R}_N(x) = N^{1/2} \left( F_k \left( x \frac{T_N}{Nk} \right) - F_k(x) \right).$$

In [18], more than 60 references are given on this subject, with statistical applications such as testing uniformity or goodness-of-fit tests. Further, [23] obtained tables for the limiting distribution of the Kolmogorov-Smirnov (K-S) statistic based on the spacing process. Here, we mention that [39] was the
first to suggest the use of the K-S and Cramér-von Mises functionals of the spacing process. [22] consider the k-spacings and characterized the limiting distribution of the statistics

\[
W_n(g, k) = N^{-1/2} \sum_{i=1}^{N} \left( g\left(NkD_{i,N}^k\right) - \alpha \right),
\]

(11)

where \( g(\cdot) \) is a smooth function, \( k \) is fixed and \( \alpha = \mathbb{E}[g(Y)] \), \( Y \) is a rv with a density function \( f_k(y) \). These statistics \( W_n(g, k) \) can be used for testing goodness-of-fit to a uniform distribution. [42] discovered that the Greenwood test (one corresponding to \( g(x) = x^2, \ x \geq 0 \) is asymptotically optimal among tests based on simple spacings (i.e., \( k = 1 \)). [22] demonstrated that the Greenwood-type test based on disjoint higher-order spacings is asymptotically superior to the conventional Greenwood-type test based on simple spacings (\( k = 1 \)). [35] demonstrated that, for any fixed spacing size \( k \), the Greenwood type test is optimal among type statistics-based tests (11). [35] discovered that the Greenwood type test based on overlapping k-spacings is superior to the corresponding test based on disjoint m-spacings for any fixed spacing size \( k \).

A known limitation of tests based on symmetric sum functions of spacings is their inability to detect alternatives converging to the null distribution at a rate faster than \( n^{-1/4} \), for more details, refer to [44]. For applications of the spacing in statistical tests and others we may refer to [25, 24], [45], [46], [8].

In the present work, we establish strong approximations of the process \( \{\alpha_n(x) : 0 \leq x < \infty\} \) defined respectively in (3) in Section 2. Motivated by [15]'s results for the uniform empirical process \( \{\eta_n(t) : 0 \leq t \leq 1\} \), we state a result of moment convergence rates for the process \( \{\alpha_n(x) : 0 \leq x < \infty\} \). Mathematical developments are given in Section 4.

## 2 Strong approximations

### 2.1 Preliminaries

Let us recall some useful (in this work) gaussian processes. Let \( W = \{W(s) : s \geq 0\} \) and \( B = \{B(u) : u \in [0,1]\} \) be the standard Wiener process and Brownian bridge, that is, the centered Gaussian processes with continuous sample paths and covariance functions

\[
\mathbb{E}(W(s)W(t)) = s \land t, \quad \text{for} \quad s, t \geq 0
\]

and

\[
\mathbb{E}(B(u)B(v)) = u \land v - uv, \quad \text{for} \quad u, v \in [0,1].
\]
In the sequel, the underlying probability space \((\Omega, \mathcal{A}, P)\) is assumed to be rich enough, in the sense that an independent sequence of Gaussian processes, which are independent of the originally given i.i.d. a sequence of random vectors can be constructed on this probability space, see for instance [17], Lemma A1 in [12] and [34]. Before stating the main result of this section, let us recall some useful results.

**Theorem 1 ([7])** Given the process \(\{\alpha_n(x) : 0 \leq x < \infty\}\) constructed from a sequence \(U_1, U_2, \ldots\) of i.i.d random variables of uniform law on \([0, 1]\) and defined on a space of probability eventually enlarged version of \((\Omega, \mathcal{A}, P)\), there exists a sequence of Brownian bridges \(B_1, B_2, \ldots\), defined on \((\Omega, \mathcal{A}, P)\) such that, for all \(0 \leq x < \infty\), if,

\[
V_n(x) = B_N(F_k(x)) - \frac{1}{k} x f_k(x) \int_0^\infty B_N(F_k(u)) du,
\]

then with probability 1,

\[
\sup_{0 \leq x < \infty} |\alpha_n(x) - V_n(x)| = O \left( n^{-1/4}(\log n)^{3/4} \right), \quad \text{as} \quad n \to \infty.
\]

Notice that the approximating Gaussian processes, for \(k = 1\), and \(x \in [0, 1]\) is given by

\[
V_n(x) = B_n(x) + (1 - x) \log(1 - x) \int_0^1 \frac{B_n(u)}{1 - u} du.
\]

**2.2 Main result**

Our next theorem describes strong approximations.

**Theorem 2** There exists a sequence of Gaussian processes \(\{V_n(x) : 0 \leq x \leq \infty, n \geq 1\}\), such that the following properties hold. We have

\[
\mathbb{E}V_n(x) = 0,
\]

and

\[
\mathbb{E}V_n(x)V_n(y) = \min (F_k(x), F_k(y)) - F_k(x)F_k(y) - \frac{1}{k} xy F_k(x) F_k(y).
\]

Moreover, for all \(x > 0\) and \(n\) great enough, we have

\[
\mathbb{P} \left( \sup_{0 \leq t \leq \infty} |\alpha_n(t) - V_n(t)| \geq N^{-1/4}(A_1(\log N)^{3/4} + x) \right) \leq B_1 \exp(-C_1x),
\]

where \(A_1, B_1\) and \(C_1\) are positive constants.
The proof of Theorem 2 will be done in Section 4. We will consider the cases $n = [Nk]$ and $k(N - 1) < n < Nk$.

**Remark 1** The proof of this Theorem is based on the proof of Theorem 3.1 of [7]. More precisely, for reader convenience, the relations that we need are given in pages 10 and 11 of this paper. For instance, the relations (32), (33), (35)-(38) were stated in the proof of [7]'s Theorem.

**Remark 2** We highlight that Theorem 2 is more refined than the previous works [2] and [7]. More precisely, we obtain the analog theorem of KMT result, permitting the following result

\[
\sup_{0 \leq t \leq \infty} |\alpha_n(t) - V_n(t)| = O(N^{-1/4}((\log N)^{3/4})),
\]

in an almost sure sense that is not possible with previous results.

Let $\mathcal{D}(A)$ be the space of right-continuous real-valued functions defined on $A$ which have left-hand limits, equipped with the Skorohod topology; refer to [13] for further details on this problem.

**Corollary 1** Let $\Phi(\cdot)$ be a functional defined on the space $\mathcal{D}(\mathbb{R})$, satisfying a Lipschitz condition

\[
|\Phi(v) - \Phi(w)| \leq L \sup_{t \in \mathbb{R}} |v(t) - w(t)|.
\]

Assume further that the distribution of the r.v. $\Phi(V_n(\cdot))$ has a bounded density. Then, as $n \to \infty$,

\[
\sup_{x \in \mathbb{R}} |\mathbb{P}\{\Phi(\alpha_n(\cdot)) \leq x\} - \mathbb{P}\{\Phi(\Phi(V_n(\cdot))) \leq x\}| = O(N^{-1/4}((\log N)^{3/4})).
\]

For more comments on this kind of result, we may refer to [4, 5].

### 3 Moment convergence rates

Zhang in [49] investigated the uniform empirical process and obtained the precise asymptotics in the Baum-Katz and Davis law of large numbers given by [26] and [27] for a sequence of i.i.d. random variables. For further details we refer to [47], [30], [36], [48], [31, 32], [3, 1] and the references therein. The
legendary paper by [33] introducing the concept of “complete convergence” is to be cited here. The last mentioned reference generated a series of papers, in particular Baum and Katz (1965)’s seminal work (cf. [9]), which provided necessary and sufficient conditions for the convergence of the series
\[
\sum_{n=1}^{\infty} n^{r/p - 2} \mathbb{P}\left( \left| \sum_{i=1}^{n} X_i \right| \geq \epsilon \right)
\]
for suitable values of \( r \) and \( p \). One result, among others, of [49] reads as follows: for \( 1 \leq p < 2, r > p \), we have, for \( g_n(\epsilon) = \epsilon n^{1/p} \),
\[
\lim_{\epsilon \to 0} \epsilon^2 (r-p)/(2-p) \sum_{n=1}^{\infty} n^{r/p - 2} \mathbb{E}\{\left\| \alpha_n \right\|_\infty - g_n(\epsilon)\}^+
= \frac{p}{r-p} \mathbb{E}\left[ \left\| B \right\|_\infty^{2(r-p)/(2-p)} \right].
\] (17)

Motivated by the moment convergence rates for the uniform empirical process \( \eta_n(u) : 0 \leq u \leq 1, n \geq 1 \) stated by [15], see Theorem 3 below, we state analog results for the empirical \( k \)-spacing process \( \alpha_n(x) : 0 \leq x < \infty \) defined respectively in (3), with \( \{x\}_+ = \max\{x, 0\} \).

**Theorem 3** Let \( a > -1 \), then
\[
\lim_{\epsilon \to 0} \left( 1 - \frac{a + 1}{4\epsilon^2} \right)^{1/2} \sum_{n=1}^{+\infty} n^a \mathbb{E}\{\left\| \eta_n \right\| - \epsilon \sqrt{2\log n}\}^+ = \frac{\sqrt{\pi/2}}{a + 1},
\]
and
\[
\lim_{\epsilon \to 0} \left( 1 - \frac{a + 1}{4\epsilon^2} \right)^{1/2} \sum_{n=1}^{+\infty} \frac{(\log n)^a}{n} \mathbb{E}\{\left\| \eta_n \right\| - \epsilon \sqrt{2\log n}\}^+ = \frac{\sqrt{\pi/2}}{a + 1}.
\]

Let us recall Proposition 2.1 of [15], stating the following result for the Brownian bridge \( \{B(t) : 0 \leq t \leq 1\} \).

**Theorem 4** Let \( a > -1, a_n = o(1/\log n) \), then
\[
\lim_{\epsilon \to 0} \left( 1 - \frac{a + 1}{4\epsilon^2} \right)^{1/2} \sum_{n=1}^{+\infty} n^a \mathbb{E}\{\left\| B \right\| - (\epsilon + a_n) \sqrt{2\log n}\}^+ = \frac{\sqrt{\pi/2}}{a + 1}.
\]

In the following theorem, we state our result for the empirical process \( \alpha_n(x) \) defined respectively in (3).
Theorem 5 Let $a > -1$, then
\[
\lim_{\epsilon \searrow \frac{\sqrt{a+1}}{2}} \left( 1 - \frac{a+1}{4\epsilon^2} \right)^{1/2} \sum_{n=1}^{+\infty} n^a \mathbb{E}[\|\alpha_n\| - \epsilon \sqrt{2 \log n}]_+ = \frac{\sqrt{\pi/2}}{a+1},
\]
and
\[
\lim_{\epsilon \searrow \frac{\sqrt{a+1}}{2}} \left( 1 - \frac{a+1}{4\epsilon^2} \right)^{1/2} \sum_{n=1}^{+\infty} \frac{(\log n)^a}{n} \mathbb{E}[\|\alpha_n\| - \epsilon \sqrt{2 \log n}]_+ = \frac{\sqrt{\pi/2}}{a+1}.
\]

4 Proof

This section is devoted to the proof of our results. The previously defined notation continues to be used in the following. In this section, we introduce some technical lemma’s (with proofs) useful for the proof of Theorem 2.

Let us recall the following theorem.

Theorem 6 [34] On a suitable probability space, we may define the uniform empirical process $\{\beta_N(x) : 0 \leq x < \infty, n \geq 1\}$, in combination with a sequence of Brownian bridges $\{B_N^{(1)}(t) : 0 \leq t \leq 1\}$, such that, for any $n \in \mathbb{N}^*$ and any positive number $x$,
\[
P\left\{ \sup_{0 \leq x < \infty} |\beta_N(u) - B_n(F_k(x))| \geq \frac{1}{\sqrt{n}} \left( A_2 \log n + x \right) \right\} \leq B_2 \exp(-C_2x),
\]
(18)
where $A_2$, $B_2$ and $C_2$ are suitable absolute constants.

The following lemma is in the spirit of Lemma 3.1 of [7], but we give an exponential rate. Our proof follows the lines of the just mentioned lemma, namely decompositions given on page 20 of this paper and relation (22) are stated by [7].

Lemma 1 We have, for each $\epsilon > 0$, and all $n \geq m$ sufficiently large
\[
P \left( \left| N^{1/2} \left( \frac{T_N}{Nk} - 1 \right) - \frac{1}{k} \int_0^\infty t \mathbb{E}_N^{(1)}(F_k(t)) \right| > \frac{A_3(\log N)^2 + x}{\sqrt{N}} \right) \leq B_3 \exp(-C_3x),
\]
(19)
where $A_3$, $B_3$ and $C_3$ are suitable absolute constants and $x \leq \frac{Nk}{2}$. 

Proof of Lemma 1.

It is readily checked that,

\[
\frac{T_N}{Nk} = \frac{1}{Nk} \sum_{i=1}^{N} Y_i = \frac{1}{k} \int_0^\infty t dG_N(t) \quad \text{and} \quad \int_0^\infty t dF_k(t) = k. \tag{20}
\]

From which, we obtain readily that

\[
N^{1/2} \left( \frac{T_N}{N} - k \right) = \int_0^\infty t d\beta_N(t) = -\int_0^\infty \beta_N(t) dt. \tag{21}
\]

Let \(\lambda_N\) be a sequence of positive numbers. Making use of the triangle inequality, we infer that

\[
\left| \int_0^\infty \beta_N(t) dt - \int_0^\infty B_N^{(1)}(F_k(t)) dt \right| \\
\leq \int_0^{\lambda_N} \left| \beta_N(t) - B_N^{(1)}(F_k(t)) \right| dt \\
+ \left| \int_{\lambda_N}^\infty B_N^{(1)}(F_k(t)) dt \right| \\
+ \left| \int_{\lambda_N}^\infty \beta_N(t) dt \right|.
\]

We have the inequality

\[
\int_0^{\lambda_N} \left| \beta_N(t) - B_N^{(1)}(F_k(t)) \right| dt \\
\leq \sup_{0 \leq x < \infty} \left| \beta_N(t) - B_N^{(1)}(F_k(t)) \right| \int_0^{\lambda_N} dt \\
= \lambda_N \sup_{0 \leq x < \infty} \left| \beta_N(t) - B_N^{(1)}(F_k(t)) \right|.
\]

Making use of Theorem 6, we obtain

\[
\mathbb{P} \left( \int_0^{\lambda_N} \left| \beta_N(t) - B_N^{(1)}(F_k(t)) \right| dt > \lambda_N N^{-1/2} (A \log N + x) \right) \\
\leq \mathbb{P} \left( \lambda_N \sup_{0 \leq x < \infty} \left| \beta_N(t) - B_N^{(1)}(F_k(t)) \right| > \lambda_N N^{-1/2} (A \log N + x) \right)
\]
\[\mathbb{P}\left( \sup_{0 \leq x < \infty} \left| \beta_N(t) - B_N^{(1)}(F_k(t)) \right| > N^{-1/2}(A \log N + x) \right) \]
\[\leq B e^{-C x}.\]

Let us now study the behavior of \(\int_{\Lambda_N}^{\infty} |B_N^{(1)}(F_k(t))| \, dt\). We have that the variance can be written as (see [21])

\[\sigma_1^2 = \mathbb{E} \left[ \int_{\Lambda_N}^{\infty} B_N^{(1)}(F_k(t)) \, dt \int_{\Lambda_N}^{\infty} B_N^{(1)}(F_k(s)) \, ds \right] \]
\[= \int_{\Lambda_N}^{\infty} \int_{\Lambda_N}^{\infty} (F_k(t) \wedge F_k(s) - F_k(t)F_k(s)) \, ds \, dt \]
\[= \int_{\Lambda_N}^{\infty} \left\{ (1 - F_k(s)) \int_{\Lambda_N}^{s} F_k(t) \, dt + F_k(s) \left( \int_{s}^{\infty} (1 - F_k(t)) \, dt \right) \right\} \, ds.\]

By [7], there exists \(t_0 > 0\) such that

\[1 - F_k(t) \leq 2 \exp \left( -\frac{t}{2} \right), \text{ if } t \geq t_0. \tag{22}\]

From (22), we have

\[\sigma_1^2 \leq \sigma_{1,N}^2 = C e^{-\Lambda_N},\]

for some positive constant \(C\).

Finally, taking in account that \(\int_{\Lambda_N}^{\infty} B_N^{(1)}(F_k(t)) \, dt\) is a Gaussian-centered variable, with variance \(\sigma_1^2\), we have, for all \(x > 0\),

\[\mathbb{P}\left( \left| \int_{\Lambda_N}^{\infty} B_N^{(1)}(F_k(t)) \, dt \right| \geq \frac{x}{N^{1/2}} \right) \]
\[\leq \mathbb{P}\left( \frac{1}{\sigma_1} \left| \int_{\Lambda_N}^{\infty} B_N^{(1)}(F_k(t)) \, dt \right| \geq \frac{x}{\sqrt{N}} \right) \]
\[\leq 2 \frac{\sigma_{1,N} N^{1/2}}{x} e^{-\frac{x^2}{2 \sigma_{1,N}^2 N}} \leq \frac{C}{x} e^{-x^2},\]

where the inequality on the right-hand-side is obtained by using (22), jointly with \(\lambda_N = C_1 \log N\).

Finally, we must study the behavior of \(\left| \int_{\Lambda_N}^{\infty} \beta_N(t) \, dt \right|\). Tacking in account that \(T_N\) is a \(\gamma(Nk, 1)\) r.v., we have

\[\mathbb{P}\left( T_N \geq \sqrt{2Nkx} + x \right) + \mathbb{P}\left( -T_N \geq \sqrt{2Nkx} + x \right) \leq 2 e^{-x}.\]
The last inequality was obtained from page 29 of [14]. Then,
\[ \mathbb{P}(T_N - Nk \geq x) \leq \mathbb{P}\left(T_N \geq \sqrt{2Nk}x + x\right) \leq e^{-x} \]
for \(0 \leq x \leq \frac{Nk}{2}\). Moreover, from (22), we have
\[ \mathbb{P}(T_N \leq Nk - x) \leq C e^{-Nk+x} \leq e^{-x} \]
for \(Nk \geq t_0 + x\) and \(x \leq \frac{Nk}{2}\).

The following Lemma is in the spirit of Lemma 3.2 of [7], but we give an exponential rate. Our proof follows the lines of the just mentioned Lemma, namely relations (25)-(27) given on page 23 of this paper are stated by [7].

**Lemma 2** For each \(\varepsilon > 0\) and \(n \geq k\), we have
\[
\mathbb{P}\left(\sup_{0 \leq x \leq \infty} \left| B_N^{(1)}(\frac{F_k(x)}{Nk}) - B_N^{(1)}(F_k(x))\right| > N^{-1/4}(A_4 \log^{3/4} N + x)\right) \leq B_4 \exp(-C_4 x),
\]
where \(A_4, B_4\) and \(C_4\) are positive constants.

**Proof of lemma 2.** The random variable \(\int_0^\infty B_N^{(1)}(F_k(t))\,dt\) has a normal distribution, with expectation 0 and finite variance, given by
\[
\sigma_2^2 = \mathbb{E}\left[\left(\int_0^\infty B_N^{(1)}(F_k(t))\,dt\right)^2\right] = \int_0^\infty \left\{(1 - F_k(s)) \int_s^\infty F_k(t)\,dt + F_k(s) \left(\int_s^\infty (1 - F_k(t))\,dt\right)\right\} ds.
\]
Hence, we have the inequality
\[
\mathbb{P}\left(\frac{1}{\sigma_2} \left|\int_0^\infty B_N^{(1)}(F_k(t))\,dt\right| > \frac{x}{\sigma_2}\right) \leq \frac{2\sigma_2}{\sigma_2} e^{-\frac{x^2}{2\sigma_2^2}}. (23)
\]
We have
\[
\mathbb{P}\left(\left|\frac{T_N}{Nk} - 1\right| > \frac{A_3(\log N)^{1/2}}{N^{1/2}} + x\right)
\]
\[
\begin{align*}
&\leq \mathbb{P}\left( \left| N^{1/2} \left( \frac{H}{Nk} - 1 \right) - \frac{1}{k} \int_0^\infty B_N^{(1)}(F_k(t)) \, dt \right| > \frac{A_3 (\log N)^{1/2}}{N^{1/2}} + x \right) \\
&\quad + \mathbb{P}\left( \left| \frac{1}{k} \int_0^\infty B_N^{(1)}(F_k(t)) \, dt \right| > x \right) \\
&\leq B_3 e^{-C_3 x} + 2 \frac{\sigma^2}{xk} e^{-C_4 x^2} \text{ for } x > 0,
\end{align*}
\]

where we have used Lemma 1 jointly with (23).

By Taylor expansions, we readily obtain
\[
|F_k \left( \frac{x H}{Nk} \right) - F_k(x)| = x f_k(x_N) \left| \frac{H}{Nk} - 1 \right|, \tag{24}
\]

where
\[
|x_N - x| \leq x \left| \frac{H}{Nk} - 1 \right|.
\]

Let \(0 < \delta < 1\) and define \(A_N(\delta)\) by
\[
A_N(\delta) = \left\{ \omega : \left| \frac{H}{Nk} - 1 \right| \leq \delta \right\}. \tag{25}
\]

Now, by choosing \(N\) sufficiently large so that
\[
\frac{A_3}{N^{1/2}} (\log N)^{1/2} \leq \delta,
\]

and using (24), we get that
\[
\mathbb{P}(A_N^c(\delta)) \leq e^{-x}.
\]

In addition, we have for each \(x_N \in A_N(\delta),\)
\[
x f_k(x_N) \leq \frac{(1 + \delta)^{m-1}}{\Gamma(k)} x^k e^{-(1-\delta)x}, \tag{26}
\]

which is bounded on \([0, \infty)\). Now, we let
\[
A_{12} = A_{11} \sup_{0 \leq x < \infty} \frac{(1 + \delta)^{k-1}}{\Gamma(k)} x^k e^{-(1-\delta)x}. \tag{27}
\]

Recall the following elementary fact
\[
\mathbb{P}(A) \leq \mathbb{P}(B^c) + \mathbb{P}(A \cap B),
\]
then, for large enough $N$, we infer that we have

$$
\mathbb{P}\left( \sup_{0 \leq x < \infty} \left| F_k\left(x \frac{T_N}{Nk}\right) - F_k(x) \right| > A_{11}N^{-1/2}\left((\log aN)^{1/2} + x\right) \right) \\
\leq \mathbb{P}\left( A_k^{0}(\delta) \right) + \mathbb{P}\left( A_N(\delta) \text{ and } \sup_{0 \leq x < \infty} \left| F_k\left(x \frac{T_N}{Nk}\right) - F_k(x) \right| > A_{11}N^{-1/2}\left((\log aN)^{1/2} + x\right) \right) \\
> A_{11}N^{-1/2}\left((\log aN)^{1/2} + x\right)
$$

(28)

(29)

This completes the proof of Lemma 2.

**Proof of Theorem 2.**

We are going to give the main steps of the proof. Assume first that $n = Nk$. Keep in mind the representation (10) for the empirical process of $k$-spacings.
We are aimed to prove the following, for $N \geq 1$, $x > 0$,

$$
\mathbb{P}\left( \sup_{0 \leq x \leq \infty} |\alpha_N^1(x) - V_N^*(x)| \geq N^{-1/2}(A_3(\log N)^{3/4} + x) \right) \leq B_3 \exp(-C_3x),
$$

(30)

where

$$
V_N^*(x) = B_N^{(1)}(F_k(x)) - \frac{1}{k}xf_k(x) \int_0^\infty B_N^{(1)}(F_k(y)) dy.
$$

(31)

By second-order Taylor expansion in the second term of (10), we get

$$
\sup_{0 \leq x < \infty} |\alpha_N^1(x) - V_N^*(x)| \leq I_N^{(1)} + I_N^{(2)} + I_N^{(3)} + I_N^{(4)},
$$

(32)

with

$$
I_N^{(1)} = \sup_{0 \leq x < \infty} |\beta_N \left( x \frac{T_N}{Nk} \right) - B_N^{(1)} \left( F_k \left( x \frac{T_N}{Nk} \right) \right) |,
$$

$$
I_N^{(2)} = \sup_{0 \leq x < \infty} |B_N^{(1)} \left( F_k \left( x \frac{T_N}{Nk} \right) \right) - B_N^{(1)}(F_k(x))|,
$$

$$
I_N^{(3)} = N^{1/2} \left( \frac{T_N}{Nk} - 1 \right)^2 \sup_{0 \leq x < \infty} |x^2f_k'(x_N)|,
$$

$$
I_N^{(4)} = \sup_{0 \leq x < \infty} \left| \frac{xf_k(x_N)}{k} \right| \left( N^{1/2} \left( \frac{T_N}{Nk} - 1 \right) - \int_0^\infty t dB_N^{(1)}(F_k(t)) \right),
$$

where

$$
|x_N - x| \leq x \left| \frac{T_N}{Nk} - 1 \right|.
$$

The behavior of $I_N^{(1)}$, is obtained from Theorem 6. The behavior of $I_N^{(2)}$, is obtained from Lemma 2, the behavior of $I_N^{(3)}$ is obtained from Lemma 1. For the behavior of $I_N^{(4)}$, we have

$$
\mathbb{P}\left( \left| \frac{T_N}{Nk} - 1 \right| \geq \frac{A_{33}}{\sqrt{Nk}} \right) \leq \mathbb{P}\left( T_N \geq A_{33}\sqrt{Nk} + Nk \right)
$$

$$
+ \mathbb{P}\left( -T_N \geq A_{33}\sqrt{Nk} + Nk \right)
$$

$$
\leq 2e^{-Nk}.
$$

The last inequality was obtained from page 29 of [14]. Let us define

$$
B_{N}^\delta = \left\{ \omega : \left| \frac{T_N}{Nk} - 1 \right| \leq \delta \right\}
$$
in the same spirit of the $A^\delta_N$ of [7], on $B^\delta_N$, we have

$$\sup_{0 \leq x < \infty} |x^2 f'_k(x) N| = M^{(1)} < \infty,$$

and consider also $N$ large enough to make $\frac{A_{33}}{Nk} \leq \delta$, then

$$p \left( \left( \frac{T_N}{Nk} - 1 \right)^2 \sup_{0 \leq x < \infty} |x^2 f'_k(x) N| \geq \frac{1}{Nk} \right)
\leq \mathbb{P} (B^\delta_N) + \mathbb{P} \left( B^\delta_N \text{ and } \left( \frac{T_N}{Nk} - 1 \right)^2 \geq \frac{A_{34}}{Nk} \right)
\leq 2e^{-Nk} \leq 2e^{-x},$$

(33)

for large enough $N$.

Hence together with Lemma 4.4.4 of [16], we can define a sequence of Gaussian processes $\{V_{Nk}(x) : 0 \leq x < \infty\}, N = 1, 2, \ldots$, such that for each $N$, we have

$$\{\alpha_{Nk}(x), V_{Nk}(y) : 0 \leq x, y < \infty\}^d \{\alpha^1_N(x), V^*_N(y) : 0 \leq x, y < \infty\}. \quad (34)$$

This completes the proof Theorem 2 for the case where $n = Nk$.

Now, we prove the general case where $k(N-1) < n < Nk$. It follows from (5) that

$$\{\alpha_{n}(x), 0 \leq x < \infty\}
\overset{d}{=} \left\{ N^{1/2} \left( G_{N,k} \left( x \frac{S_n}{Nk} \right) - F_k(x) \right), 0 \leq x < \infty \right\}, \quad (35)$$

where

$$G_{N,k}(x) = \frac{1}{N} \sum_{i=1}^{N-1} \mathbb{I}_{\{Y_i < x\}} + \frac{1}{N} \mathbb{I}_{\{\sum_{\ell=(N-1)k+1}^{n} \xi_{\ell} < x\}}. \quad (36)$$

Notice that we have the following fact

$$\sup_{0 \leq x < \infty} \left| G_{N,k} \left( x \frac{S_n}{Nk} \right) - G_{N-1} \left( x \frac{S_n}{Nk} \right) \right| \leq \frac{1}{N} + \frac{1}{N(N-1)}, \quad (37)$$

and
for \( N \) large enough and \( x \geq 0 \), where we have used (33). Set

\[
\mathcal{P} = \mathbb{P} \left( \sup_{0 \leq x < \infty} N^{1/2} \left| G_{N,k} \left( x \frac{S_n}{Nk} \right) - F_k(x) \right| - V^*_{N-1}(x) \right.
\]

\[
> A_{19} N^{-1/2} \left( \log aN \right)^{1/2} \bigg) \bigg).
\]

By the use of (30) in connection with (37), we infer that

\[
\mathcal{P} \leq \mathbb{P} \left( \sup_{0 \leq x < \infty} N^{1/2} \left| F_k \left( x \frac{S_n}{Nk} \right) - F_k(x) \right| > A_{20} N^{-1/2} \left( \log aN \right)^{1/2} \bigg)
\]

\[
+ \mathbb{P} \left( \sup_{0 \leq x < \infty} N^{1/2} \left| V^*_{N-1} \left( x \frac{S_n}{Nk} \right) - V^*_{N-1}(x) \right| > A_{21} N^{-1/2} \left( \log aN \right)^{1/2} \bigg)
\]

\[
+ B_3 N^{-\epsilon}.
\]

Once more, by a first order of the Taylor expansion, we have

\[
N^{1/2} \left| F_k \left( x \frac{S_n}{Nk} \right) - F_k(x) \right| = x f_k(x), N^{1/2} \left| \frac{S_n - T_{N-1}}{T_{N-1}} \right|,
\]

where

\[
|x_N - x| \leq x \left| \frac{S_n - T_{N-1}}{T_{N-1}} \right|.
\]

By combining Lemma 1 with (38), it follows that

\[
\mathbb{P} \left( \left| \frac{S_n}{T_{N-1}} - 1 \right| > A_{22} N^{-1/2} \right) \leq B_{22} e^{-x}.
\]

By arguing similarly as in the proof (28), we obtain that

\[
\mathbb{P} \left( \sup_{0 \leq x < \infty} N^{1/2} \left| F_k \left( x \frac{S_n}{T_{N-1}} \right) - F_k(x) \right| > A_{20} N^{-1/2} \left( \log aN \right)^{1/2} \bigg) \leq B_{20} e^{-x}.
\]

Now, by definition (31), (43), and through a similar argument as that used at the end of the proof of Lemma 2, we get

\[
\mathbb{P} \left( \sup_{0 \leq x < \infty} N^{1/2} \left| V^*_{N-1} \left( x \frac{S_n}{T_{N-1}} \right) - V^*_{N-1}(x) \right| > A_{21} N^{-1/2} \left( \log aN \right)^{1/2} \bigg) \leq B_{21} e^{-x}.
\]
Then, making use of the equations (40), (43) and (44), we obtain

\[
\mathbb{P}\left( \sup_{0 \leq x < \infty} \left| N^{1/2} \left( G_{N,k} \left( x \frac{S_n}{Nk} \right) - F_k(x) \right) - V_{N-1}^*(x) \right| > A_{23} N^{-1/2} (\log aN)^{1/2} \right) \leq B_{23} e^{-x}. \tag{45}
\]

Again, by Lemma 4.4.4 of [16] and (35), we can get a sequence of Gaussian processes \( \{ V_n(x) : 0 \leq x < \infty \} \) such that for each \( N \), we have

\[
\{ \alpha_n(x), V_n(y) : 0 \leq x, y < \infty \} \quad \overset{d}{=} \quad \left\{ N^{1/2} \left( G_{N,k} \left( x \frac{S_n}{Nk} \right) - F_k(x) \right), V_{N-1}^*(y) : 0 \leq x, y < \infty \right\}.
\]

The last equation completes the proof of Theorem 2. \( \square \)

**Proof of Theorem 5**

This Theorem follows the proof of Theorem 1.1 of [15] and is based on the following two technical lemmas.

Recall that the Gaussian processes \( \{ V_n^*(x) : 0 \leq x \leq \infty, n \geq 1 \} \) is such that

\[
V_n^*(x) = B_N^{(1)}(F_k(x)) - \frac{1}{k} x f_k(x) \int_0^\infty B_N^{(1)}(F_k(y)) dy,
\]

satisfies

\[
\mathbb{E} V_n^*(x) = 0, \tag{46}
\]

and

\[
\mathbb{E} V_n^*(x) V_n^*(y) = \min (F_k(x), F_k(y)) - F_k(x)F_k(y) - \frac{1}{k} xy F_k(x) F_k(y). \tag{47}
\]

**Lemma 3** For all \( x > 0 \)

\[
\mathbb{P}( \| V_n^* \| \geq x ) \sim 2e^{-2x^2}, \text{ as } x \to \infty.
\]

**Proof of Lemma 3** Recall that from (23), we have the inequality

\[
\mathbb{P} \left( \frac{1}{k\sigma^2} \left| \int_0^\infty B_N^{(1)}(F_k(t)) dt \right| > u \right) \leq 2 \frac{\sigma^2}{uk} e^{-\frac{u^2}{2\sigma^2}}. \tag{48}
\]
This last relation, jointly with (26) gives the behavior of
\[
\mathbb{P}\left( \sup_{0 \leq x \leq \infty} \left| \frac{x f_k(x)}{k \sigma_2} \int_0^\infty B_N^{(1)}(F_k(t)) \, dt \right| > u \right).
\]
Moreover, it is well known that for all \( x > 0 \),
\[
\mathbb{P}(\|B(t)\| \geq x) = 2 \sum_{k=1}^\infty (-1)^{k+1} e^{-2k^2x^2}.
\]
In particular,
\[
\mathbb{P}(\|B(t)\| \geq x) \sim 2e^{-2x^2},
\]
see, for instance, Lemma 2.1 of [15]. This result completes the proof of Lemma 3.

Inspired from Proposition 2.1 of [15], we have

**Lemma 4** Let \( a > -1 \), \( a_n = o(1/\log n) \), then
\[
\lim_{\epsilon \downarrow \sqrt{\frac{a+1}{4e^2}}} \left( 1 - \frac{a + 1}{4e^2} \right)^{1/2} \sum_{n=1}^\infty n^a \mathbb{E}\left\{ \|V_N - (\epsilon + a_n)\sqrt{2\log n}\| \right\}^+ = \frac{\sqrt{\pi/2}}{a + 1}.
\]

**Proof of Lemma 4.** The proof follows the lines of the proof of Proposition 2.1 of [15] by tacking \( V_N \) in the place of \( B \).

For \( p < -1/2 \), we have
\[
\mathbb{E}\left\{ \sup_{0 \leq x < \infty} |V_N(x)| - \left( \epsilon \sqrt{2\log n} + (\log n)^p \right) \right\}^+ + - \mathbb{E}\left\{ \sup_{0 \leq x < \infty} |\alpha_n(x) - V_N(x)| - (\log n)^p \right\}^+ \leq \mathbb{E}\left\{ \sup_{0 \leq x < \infty} |\alpha_n(x)| - \epsilon \sqrt{2\log n} \right\}^+ \leq \mathbb{E}\left\{ \sup_{0 \leq x < \infty} |V_N(x)| - \left( \epsilon \sqrt{2\log n} - (\log n)^p \right) \right\}^+ + - \mathbb{E}\left\{ \sup_{0 \leq x < \infty} |\alpha_n(x) - V_N(x)| - (\log n)^p \right\}^+.
\]

The rest of the proof follows the line of the proof of Theorem 1.1 of [15].
Acknowledgements

The authors would like to thank the referee for a carefully and thoroughly compiled list of places where the presentation could be improved. The paper has benefited from those points.

References


*Received: March 21, 2022*