Monatsh Math 153, 183–204 (2008)
DOI 10.1007/s00605-007-0519-8
Printed in The Netherlands

Monatshefte für
Mathematik

Entropy conditions for subsequences of random variables with applications to empirical processes

By

István Berkes[1,*], **Walter Philipp**[2], and **Robert Tichy**[1,†]

[1] Technical University Graz, Austria
[2] University of Illinois, Champaign, IL, USA

Communicated by J. Schoißengeier

Received April 13, 2007; accepted in final form November 20, 2007
Published online February 1, 2008 © Springer-Verlag 2008

Abstract. We introduce new entropy concepts measuring the size of a given class of increasing sequences of positive integers. Under the assumption that the entropy function of $\mathscr{A}$ is not too large, many strong limit theorems will continue to hold uniformly over all sequences in $\mathscr{A}$. We demonstrate this fact by extending the Chung-Smirnov law of the iterated logarithm on empirical distribution functions for independent identically distributed random variables as well as for stationary strongly mixing sequences to hold uniformly over all sequences in $\mathscr{A}$. We prove a similar result for sequences $(n_k\omega)$ mod 1 where the sequence (n_k) of real numbers satisfies a Hadamard gap condition.

2000 Mathematics Subject Classification: 11K38, 60F15
Key words: Discrepancy, entropy, random numbers

1. Introduction

Let $X_1, X_2, \dots$ be a sequence of random variables and $\mathscr{A}$ a class of increasing sequences of positive integers. The purpose of our paper is to investigate under what conditions the sequence (X_n) satisfies the strong law of large numbers uniformly over $\mathscr{A}$ in the sense that

$$\lim_{N\to\infty} \frac{1}{N} \sup_{(p_k)\,\in\,\mathscr{A}} \left| \sum_{p_k \leqslant N} X_{p_k} \right| = 0 \quad \text{a.s.} \tag{1.1}$$

When relation (1.1) is valid, we will also be interested in the speed of convergence; in particular we will investigate when the uniform law of the iterated logarithm

$$\limsup_{N\to\infty} \frac{1}{\sqrt{N \log\log N}} \sup_{(p_k)\,\in\,\mathscr{A}} \left| \sum_{p_k \leqslant N} X_{p_k} \right| < \infty \quad \text{a.s.} \tag{1.2}$$

holds.

[*] Research supported by FWF grant S9603-N13 and OTKA grants K 61052 and K 67961.
[†] Research supported by FWF grants S9603-N13 and S9611-N13.

With great sadness, we inform the reader that Walter Philipp passed away on July 19, 2006, at the age of 69, near Graz, Austria. – I. Berkes and R.F. Tichy.

Our work was motivated by recent results of Mauduit and Sárközy on pseudorandom behavior. Given independent random variables $X_1, X_2, \ldots$ taking the values ± 1 with probability $1/2 - 1/2$, it is not hard to prove that

$$\lim_{N\to\infty} \frac{1}{N} \max_{a \geqslant 1, b \geqslant 0} \left| \sum_{\substack{k \geqslant 1 \\ ak+b \leqslant N}} X_{ak+b} \right| = 0 \quad \text{a.s.}$$

Thus, if (x_n) is a 'truly' random ± 1 sequence, then the quantity

$$K_N := \max_{a \geqslant 1, b \geqslant 0} \left| \sum_{\substack{k \geqslant 1 \\ ak+b \leqslant N}} x_{ak+b} \right| \tag{1.3}$$

is $o(N)$. Hence for a computer generated ± 1 sequence $(x_1, \ldots, x_N)$ the quantity (1.3), introduced by Mauduit and Sárközy in [14] (see Knuth [10], p. 148 for a related 'serial test'), can be used as a measure of pseudorandomness. More generally, for a sequence $(x_1, \ldots, x_N)$ in $[0, 1)$, Mauduit and Sárközy introduced the quantity

$$W_N(x_1, \ldots, x_N) := \max_{a \geqslant 1, b \geqslant 0} \left| \sum_{\substack{k \geqslant 1 \\ ak+b \leqslant N}} (1(x_{ak+b} \leqslant 1/2) - 1/2) \right| \tag{1.4}$$

which they called the *well-distribution measure* of the sequence. In a long series of papers (see e.g. [15], [16], [4], [13] and the references therein), they investigated the quantities K_N and W_N for several interesting sequences (x_n) defined by number-theoretic algorithms. A good estimate for K_N and W_N means a high degree of pseudorandomness, but to assess how this is related to "true" random behavior, one needs a probabilistic analysis of the well-distribution measure, i.e. to study the behavior of K_N and W_N for typical classes of random sequences. The purpose of our paper is to provide such an analysis.

In addition to the well-distribution measure W_N, several related quantities (e.g. the corresponding correlation measures) are of considerable interest in pseudorandomness studies (see e.g. [14], [13], [1], [4]), but in the present paper we will not deal with such quantities.

Note that in (1.1) we use the norming N^{-1} for all sums $\sum_{p_k \leqslant N} X_{p_k}$, although many of them consist of less than N elements. This choice is the natural one for our arithmetic applications, and changing (1.1) to

$$\lim_{N\to\infty} \sup_{(p_k) \in \mathscr{A}} \frac{1}{\#\{k : p_k \leqslant N\}} \left| \sum_{p_k \leqslant N} X_{p_k} \right| = 0 \quad \text{a.s.} \tag{1.5}$$

would actually result in a useless concept. If, for example, $\mathscr{A}$ consists of all increasing arithmetic progressions in $\mathbb{N}$, then there are many sets in $\mathscr{A}$ having exactly one element under N and thus (1.5) is false even if X_n are i.i.d. binomial variables.

Given a sequence (x_n) and a class $\mathscr{A}$ of subsequences of $\mathbb{N}$, we define a generalization of the well-distribution measure of Mauduit and Sárközy by

$$W_N^{(\mathscr{A})}(x_1, \ldots, x_N) := \sup_{(p_k) \in \mathscr{A}} \sup_{0 \leqslant t \leqslant 1} \left| \sum_{p_k \leqslant N} (1(x_{p_k} \leqslant t) - t) \right|. \tag{1.6}$$

When $(x_1,\ldots,x_N)$ is understood, we shall simply write $W_N(\mathscr{A})$ instead of $W_N^{(\mathscr{A})}(x_1,\ldots,x_N)$. Note that, in contrast to (1.4), definition (1.6) involves a sup in t as well. This is quite natural, since the value $1/2$ in (1.4) was arbitrary. Note that in the case when the class $\mathscr{A}$ consists of the single sequence $\mathbb{N}$, the quantity in (1.6) reduces to $ND_N(x_1,x_2,\ldots,x_N)$, where

$$D_N = D_N(x_1,\ldots,x_N) := \sup_{0\leqslant t\leqslant 1}\left|\frac{1}{N}\operatorname{card}(k\leqslant N: x_k\leqslant t)-t\right|$$

is the discrepancy of $(x_1,\ldots,x_N)$. Clearly, the sup in t makes the estimation of $W_N(\mathscr{A})$ more difficult, but our results will be easy to compare with known discrepancy estimates in the literature.

Let $\eta_1,\eta_2,\ldots$ be a sequence of random variables in $[0,1)$. Analogously to (1.1), we will say that (η_n) satisfies the Glivenko-Cantelli theorem uniformly over $\mathscr{A}$ if

$$\lim_{N\to\infty}\frac{1}{N}\sup_{(p_k)\in\mathscr{A}}\sup_{0\leqslant t\leqslant 1}\left|\sum_{p_k\leqslant N}(1(\eta_{p_k}\leqslant t)-t)\right| = 0 \quad \text{a.s.,} \tag{1.7}$$

where $1(A)$ denotes the indicator function of the set A. Relation (1.7) expresses a certain uniformity in the behavior of empirical distribution functions of subsequences of (η_k), a requirement tailored for the specific needs of pseudorandom behavior. It seems this kind of subsequential uniformity has not been studied in the probabilistic literature so far. On the other hand, uniform convergence of empirical processes with respect to sets in Euclidean spaces has a wide literature going back to the 1970's. Let (η_n) be a sequence of i.i.d. random variables, uniformly distributed over the unit cube $\mathbb{K}^d$ of $\mathbb{R}^d$, and let $\mathscr{C}$ be a class of Borel sets $\subseteq \mathbb{K}^d$. Put

$$Z_N(C) = \sum_{k\leqslant N}(1(\eta_k\in C)-\mu(C)), \quad C\in\mathscr{C}$$

where μ is the Lebesgue measure. It is known that the validity of the uniform strong law and LIL, i.e.

$$\lim_{N\to\infty}\sup_{C\in\mathscr{C}}\frac{1}{N}|Z_N(C)| = 0 \quad \text{a.s.} \tag{1.8}$$

and

$$\limsup_{N\to\infty}\frac{\sup_{C\in\mathscr{C}}|Z_N(C)|}{\sqrt{N\log\log N}} < \infty \quad \text{a.s.} \tag{1.9}$$

are closely connected with the geometry of the class $\mathscr{C}$, namely how many sufficiently separated elements of the class $\mathscr{C}$ exist, or how closely the elements of $\mathscr{C}$ can be approximated by "special" sets. For example, let $N_I(\delta,\mathscr{C})$ denote the smallest number r of measurable sets $A_1,\ldots,A_r$ in $\mathbb{K}^d$ such that for every $C\in\mathscr{C}$ there exist A_i,A_j, $1\leqslant i<j\leqslant r$ such that $A_i\subset C\subset A_j$ and $\mu(A_j\backslash A_i)<\delta$ ("metric entropy with inclusion"). Then the validity of the uniform LIL and CLT is closely related to the finiteness of the entropy integral

$$\int_0^1(\log N_I(x^2,\mathscr{C}))^{1/2}dx.$$

(See e.g. Dudley [6], [7], Dudley and Philipp [8].) Another important geometric property relevant for the uniform strong law (1.8), discovered by Vapnik and Červonenkis, is how finite sets $\{x_1, \dots, x_N\}$ in $\mathbb{R}^d$ can be "shattered" by elements of $\mathscr{C}$, i.e. how many different sets of the form $\{x_1, \dots, x_N\} \cap C$, $C \in \mathscr{C}$ exist. In fact, a necessary and sufficient condition for (1.8) can be given in terms of this quantity; see e.g. Pollard [21], p. 22.

In this paper, we will develop entropy concepts for classes of subsequences of $\mathbb{N}$ and use them to study uniform subsequential limit theorems of the type (1.1), (1.2) and (1.7). Quite naturally, the behavior of the quantities in (1.1), (1.2), (1.7) is connected with the size of the class $\mathscr{A}$ and we will see that various "sequential" analogues of entropy measures in $\mathbb{R}^d$ will lead to substantial information on empirical processes. Beside the simplest case of i.i.d. variables η_n, we will study some classes of dependent sequences as well, in particular mixing and lacunary sequences.

Throughout our paper, we will assume that the class $\mathscr{A}$ contains the sequence $\mathbb{N}$. This assumption implies that W_N is bounded below by N times the discrepancy of the same sequence and this will permit us to compare our results with classical discrepancy bounds in the literature. Apart from the lower bounds in the LIL in Theorems 1, 2, 6, all our results remain valid without this assumption.

Given a class $\mathscr{A}$ of subsequences of $\mathbb{N}$, for each $N \geqslant 1$ let $\mathscr{A}_N$ denote the collection of the restrictions of these subsequences to the segment $[1, 2, \dots, N]$ of the first N positive integers, i.e.

$$\mathscr{A}_N := \{A \cap [1, 2, \dots, N] : A \in \mathscr{A}\}.$$

Clearly

$$\mathscr{A}_N = \bigcup_{r \geqslant 1} \mathscr{A}_N(r)$$

where $\mathscr{A}_N(r)$ denotes the class of sets $A \in \mathscr{A}_N$ for which $N2^{-r} <$ card $A \leqslant N2^{-(r-1)}$. We call

$$\psi(\mathscr{A}; N, r) := \operatorname{card} \mathscr{A}_N(r) \tag{1.10}$$

the entropy function of the class $\mathscr{A}$.

Next, let (η_k) be a sequence of random variables in $[0, 1)$. In the simplest case of Theorem 1, the η_k will be independent, with each η_k having uniform distribution over $[0, 1)$, i.e.

$$P(\eta_k \leqslant t) = t, \quad 0 \leqslant t \leqslant 1, \ k \geqslant 1. \tag{1.11}$$

In Theorems 2 and 4 we permit η_k to have asymptotically uniform distribution over $[0, 1)$.

Theorem 1. *Let* (η_k) *be a sequence of independent random variables with uniform distribution* (1.11) *over* [0,1). *Let* $\mathscr{A}$ *be a class of subsequences of* $\mathbb{N}$ *with entropy function* ψ *satisfying*

$$\psi(\mathscr{A}; N, r) \leqslant \exp(B \cdot 2^{r/2} \log \log N), \quad r \geqslant 0, \ N \geqslant 10 \tag{1.12}$$

for some constant $B > 0$. *Then with probability* 1,

$$1/4 \leqslant \limsup_{N\to\infty} (N \log\log N)^{-1/2} W_N(\mathscr{A}) \leqslant C$$

for some constant C, *depending only on the constant* B *in* (1.12).

Theorem 1 gives the precise growth speed of $W_N(\mathscr{A})$ in the almost everywhere sense under the entropy condition (1.12). It is worth mentioning that in the case when $\mathscr{A}$ is the class of arithmetic progressions, the recent paper of Alon *et al.* [1] (for a preliminary study see Cassaigne, Mauduit and Sárközy [4]) shows that the precise order of magnitude of $W_N(\mathscr{A})$ in probability ("typical value") is $O(\sqrt{N})$. They also determine the typical value of the corresponding correlations.

As Theorem 6 below will show, Theorem 1 remains valid, under a more stringent entropy condition, for a large class of mixing sequences (η_k) of random variables. Applications include e.g. continued fraction digits and digits in other classical expansions. Before, however, stating this general result, we will consider a particularly simple and arithmetically interesting dependent sequence, namely the sequence $\eta_k = \eta_k(\omega) = n_k\omega$ mod 1 for rapidly increasing sequences (n_k) of integers. This is not covered by the mixing theory, but it will exemplify the methods applied in this field.

It is easy to see that the class $\mathscr{A}_0$ of arithmetic progressions satisfies $\psi(\mathscr{A}_0; N, r) \leqslant C2^{2r}$ and thus Theorem 1 applies for this class. $\mathscr{A}_0$ is, however, a fairly small class and we will show now that condition (1.12) permits much larger classes $\mathscr{A}$ than $\mathscr{A}_0$. To see this we first construct, for each integer $r \geqslant 1$, a class $\mathscr{A}^{(r)}$ of sequences of positive integers such that each sequence in $\mathscr{A}^{(r)}$, intersected with $[2^n, 2^{n+1})$, has 2^{n-r} elements, the so obtained finite sequences are all different and for $n \geqslant n_0(r)$ their number is $2^{2^{[r/2]}}$. To this end, choose $2^{2^{[r/2]}}$ subsets $H_1^{(n)}, \dots, H_{2^{2^{[r/2]}}}^{(n)}$ of $\{2^n, \dots, 2^{n+1}-1\}$ with cardinality 2^{n-r}. Since $\binom{2^n}{2^{n-r}} \to \infty$, this is possible for $n \geqslant n_0(r)$; for $1 \leqslant n < n_0(r)$ we choose a single subset $H_0^{(n)}$ of $\{2^n, \dots, 2^{n+1}-1\}$ with 2^{n-r} elements. Let $\mathscr{A}^{(r)}$ consist of the sequences $\cup_{n<n_0(r)} H_0^{(n)} \cup_{n \geqslant n_0(r)} H_j^{(n)}$, $j = 1, 2, \dots, 2^{2^{[r/2]}}$; this class obviously has the above properties. Let now $\mathscr{A} = \cup_{r=1}^{\infty} \mathscr{A}^{(r)}$. Clearly, all sequences in $\mathscr{A}^{(r)}$ have between $N2^{-r-1}$ and $N2^{-r+1}$ elements in $[1, \dots, N]$ for any $N \geqslant 1$ and thus if a sequence $(n_k) \in \mathscr{A}$ has between $N2^{-r}$ and $N2^{-(r-1)}$ elements in $[1, N]$ then (n_k) belongs to $\mathscr{A}_r$ or $\mathscr{A}_{r-1}$. Hence

$$\psi(\mathscr{A}; N, r) \leqslant 2^{2^{(r-1)/2}} + 2^{2^{r/2}} \leqslant 2^{2^{r/2}+1} \leqslant \exp(2 \cdot 2^{r/2}).$$

In the construction of the sequences in $\mathscr{A}^{(r)}$ above, we chose their finite segments in $[2^n, 2^{n+1})$ separately and the number of choices for the finite segment in $[2^n, 2^{n+1})$ was $2^{2^{[r/2]}}$ (actually we can have a little more by (1.12)), which is much more than the number $O(2^{2r})$ of arithmetic progressions in $[2^n, 2^{n+1})$ having between 2^{n-r} and 2^{n-r+1} elements. This shows that Theorem 1 permits much larger classes $\mathscr{A}$ than arithmetic progressions. Note that once the $2^{2^{[r/2]}}$ sequences in each interval $[2^n, 2^{n+1})$ were chosen, we combined them to infinite sequences by "patching together" the first, second, …, $2^{2^{[r/2]}}$-th sequence in the intervals. If, alternatively, we combine each finite sequence in an individual interval with all other finite sequences in other intervals, we get a class $\mathscr{A}^{(r)}$ satisfying $\psi(\mathscr{A}^{(r)}; 2^{n+1}, r) \geqslant 2^{2^{[r/2]}n}$ and thus (1.12) fails by a large margin. Similarly, the class $\mathscr{A}$ of all sequences (n_k) which are linear in each interval $[2^n, 2^{n+1})$ fails (1.12) by a large margin.

By a result of Finkelstein [9], the sequence

$$f_N(t) = (N \log \log N)^{-1/2} \sum_{k \leqslant N} (1(\eta_k \leqslant t) - t) \tag{1.13}$$

of normalized empirical distribution functions is, with probability 1, relatively compact with respect to the uniform topology and the class of its limit points is the class $\mathscr{C}_0$ of absolute continuous functions $x(t), t \in [0,1]$ satisfying

$$x(0) = x(1) = 0, \quad \int_0^1 x'(t)^2 dt \leqslant 1.$$

It is easily seen that all functions in $\mathscr{C}_0$ belong to the Lip $(1/2)$ class, and thus for any $\varepsilon > 0$ there exists with probability 1 a random index $N_0 = N_0(\varepsilon)$ such that

$$|f_N(t) - f_N(s)| \leqslant C|t-s|^{1/2} + \varepsilon \tag{1.14}$$

for all $0 \leqslant s, t \leqslant 1$, and all $N \geqslant N_0(\varepsilon)$, where C is an absolute constant. The last relation is a substantial sharpening of the ordinary LIL

$$\limsup_{N \to \infty} \sup_{0 \leqslant t < 1} |f_N(t)| < \infty \quad \text{a.s.}$$

As we will prove, a similar sharpening of Theorem 1 holds. Define, for a fixed sequence $\mathbf{p} = (p_k) \in \mathscr{A}$ and $0 \leqslant t \leqslant 1$,

$$f_{N,\mathbf{p}}(t) := (N \log \log N)^{-1/2} \sum_{p_k \leqslant N} (1(\eta_{p_k} \leqslant t) - t). \tag{1.15}$$

Then under the hypotheses of Theorem 1 the following result holds. For each $0 < \alpha < 1/2$ and $\varepsilon > 0$ there is with probability 1 a random index $N_0 = N_0(\varepsilon)$ such that

$$|f_{N,\mathbf{p}}(t) - f_{N,\mathbf{p}}(s)| \leqslant C|t-s|^{\alpha} + \varepsilon \tag{1.16}$$

for all $0 \leqslant s, t \leqslant 1$, all $(p_k) \in \mathscr{A}$ and all $N \geqslant N_0$, where the constant C depends only on B in (1.12).

Our next theorem concerns lacunary sequences $\{n_k \omega\}$, where $\{\cdot\}$ denotes fractional part. By a classical result of Weyl (see e.g. [11], pp. 32–33), for any increasing sequence (n_k) of integers, $\{n_k \omega\}$ is uniformly distributed for almost every ω in the sense that its discrepancy D_N tends to 0 as $N \to \infty$. This fact and the simplicity of its definition make $\{n_k \omega\}$ a natural object for a pseudorandomness study, and in fact a number of results in our paper will deal with this sequence. Very few sharp results on the discrepancy of $\{n_k \omega\}$ exist in the literature; precise asymptotics are known only for $n_k = k$ and rapidly increasing (n_k). Philipp [19] proved that if (n_k) satisfies the Hadamard gap condition

$$n_{k+1}/n_k \geqslant q > 1, \quad k = 1, 2, \ldots. \tag{1.17}$$

then the discrepancy of $\{n_k \omega\}$ satisfies the law of the iterated logarithm, i.e. for almost all $\omega \in [0,1)$ we have

$$\frac{1}{4} \leqslant \limsup_{N \to \infty} \frac{N D_N(\{n_k \omega\})}{\sqrt{N \log \log N}} \leqslant C(q), \tag{1.18}$$

where $C(q) \ll 1/(q-1)$. Our next theorem proves an LIL for the well distribution measure $W_N(\mathscr{A})$ of this sequence, extending substantially Philipp's result.

Theorem 2. *Let (n_k) be a sequence of real numbers satisfying a Hadamard gap condition* (1.7) *and let $\eta_k = \eta_k(\omega) = \{n_k\omega\}$. Let the class $\mathscr{A}$ be a class of subsequences of $\mathbb{N}$ with entropy function satisfying*

$$\psi(\mathscr{A}; N, r) \leqslant B \cdot 2^{r^\beta} \tag{1.19}$$

for some constants $B > 0$ and $\beta > 0$. Then with probability 1

$$1/4 \leqslant \limsup_{N \to \infty} (N \log\log N)^{-1/2} W_N(\mathscr{A}) \leqslant C$$

for some constant $C < \infty$, depending only on B, β and q.

As we noted earlier, the class $\mathscr{A}$ of arithmetic progressions satisfies (1.19) with $\beta = 2$. A construction similar to that discussed after Theorem 1 shows that for large β, (1.19) permits considerably larger classes than the class of arithmetic progressions.

Again we shall prove an estimate of the modulus of continuity of the empirical process. Define $f_{N,\mathbf{p}}$ as in (1.15). In analogy with (1.16) we shall obtain under the hypotheses of Theorem 2 that for each $\varepsilon > 0$ there is with probability 1 a random index $N_0(\varepsilon)$ such that

$$|f_{N,\mathbf{p}}(t) - f_{N,\mathbf{p}}(s)| \leqslant C|t-s|^{1/32} + \varepsilon \tag{1.20}$$

for all $0 \leqslant s, t \leqslant 1$ all $(p_k) \in \mathscr{A}$ and all $N \geqslant N_0(\varepsilon)$.

The second entropy concept is based on the Hamming distance of sequences of integers. For $N \geqslant 1$ we define the (normalized) distance of two sequences A and B of integers by

$$d(A, B; N) = \frac{1}{N} \sum_{n \leqslant N} |1(n \in A) - 1(n \in B)|.$$

Given a class $\mathscr{A}$ of increasing sequences of positive integers, we define the entropy function κ by

$$\kappa(\mathscr{A}; \delta, N) := \sup\{m : \text{there exist } A_1, \ldots, A_m \in \mathscr{A}_N \text{ such that } d(A_i, A_j; N) > \delta \text{ for all } i \neq j\}. \tag{1.21}$$

Clearly κ is a non-increasing function of $\delta \geqslant 0$.

Theorem 3. *Let (η_k) be a sequence of independent random variables with the uniform distribution* (1.11) *over* $[0, 1)$. *Let $\mathscr{A}$ be a class of increasing sequences of positive integers with entropy function $\kappa(\mathscr{A}; \delta, N)$ growing not faster than a polynomial in $1/\delta$* (*depending only on $\mathscr{A}$*). *Then with probability* 1

$$W_N(\mathscr{A}) \ll \sqrt{N}(\log \kappa(\mathscr{A}; N^{-\alpha}, N) + (\log\log N)^{1/2}) \quad \textit{for any } \alpha > 1/2.$$

The same result holds if $\eta_k = \{n_k\omega\}$, where (n_k) is a sequence of real numbers satisfying the Hadamard gap condition (1.7).

As an example, consider a Vapnik-Červonenkis (VC) class $\mathscr{A}$ in the set $\mathbb{N}$ of positive integers. For any finite set $F \subset \mathbb{N}$, let $\Delta^{\mathscr{A}}(F)$ be the number of different subsets $F \cap A, A \in \mathscr{A}$. For $n = 1, 2, \ldots$ let

$$m^{\mathscr{A}}(n) := \max(\Delta^{\mathscr{A}}(F) : \text{ card } F = n).$$

Clearly $m^{\mathscr{A}}(n) \leqslant 2^n$. Let

$$v = V(\mathscr{A}) := \begin{cases} \inf\{n : m^{\mathscr{A}}(n) < 2^n\} \\ +\infty \text{ if } m^{\mathscr{A}}(n) = 2^n \text{ for all } n. \end{cases}$$

If $V(\mathscr{A}) < +\infty$, then $\mathscr{A}$ is called a VC class in $\mathbb{N}$. We recall a result of Dudley [6, Lemma 7.13] or Dudley [7, p. 105] measuring the size of VC classes. Let Γ be the set of all laws on $\mathbb{N}$ of the form

$$n^{-1} \sum_{j=1}^{N} \delta_{x(j)}$$

for unit point masses $\delta_{x(j)}$ on $x(j) \in \mathbb{N}$, $j = 1, 2, \ldots, n;\ n = 1, 2, \ldots$ where the $x(j)$ need not be distinct. For $\delta > 0$ and $\gamma \in \Gamma$ let

$$\kappa^*(\mathscr{A};\ \gamma, \delta) := \sup\{m : \text{there exist } A_1, \ldots, A_m \in \mathscr{A} \\ \text{such that } \gamma(A_i \Delta A_j) > \delta \text{ for } i \neq j\}$$

and put

$$\kappa^*(\mathscr{A};\ \delta) := \sup\{\kappa^*(\mathscr{A};\ \gamma, \delta) : \gamma \in \Gamma\}.$$

Lemma 1. [6], [7]. *If $\mathscr{A}$ is a VC class in $\mathbb{N}$ with $V(\mathscr{A}) = v$, then there is a constant K depending only on v such that*

$$\kappa^*(\mathscr{A};\ \delta) \leqslant K\delta^{-v} |\log \delta|^v \quad \textit{for all } \delta > 0.$$

Hence, if $\mathscr{A}$ is a VC class in $\mathbb{N}$, the entropy function κ defined in (1.21) does not grow faster than a polynomial in $1/\delta$.

Corollary 1. *Let (η_k) be a sequence of independent random variables with uniform distribution* (1.11) *over* $[0, 1)$ *or $\eta_k = \{n_k \omega\}$ with a Hadamard lacunary (n_k). Then if $\mathscr{A}$ is a VC class in $\mathbb{N}$, with probability* 1 *we have*

$$W_N(\mathscr{A}) \ll \sqrt{N} \log N.$$

Theorem 4. *Let (n_k) be an increasing sequence of integers and let $\eta_k = \eta_k(\omega) = \{n_k \omega\}$. Let $\mathscr{A}$ be a class of subsequences of $\mathbb{N}$ with entropy function satisfying* (1.19) *for some positive constants β and B. Then with probability* 1,

$$\begin{aligned} W_N(\mathscr{A}) &\ll N^{\frac{\beta}{1+\beta}} (\log N)^{\frac{3}{1+\beta}+\varepsilon} && \textit{if } \beta > 1, \\ &\ll N^{\frac{1}{2}} (\log N)^{2+\varepsilon} && \textit{if } \beta = 1, \\ &\ll N^{\frac{1}{2}} (\log N)^{\frac{3}{2}+\varepsilon} && \textit{if } \beta < 1. \end{aligned}$$

Note that we do not make here any growth or arithmetic condition on the (n_k). In the case when $\mathscr{A} = \mathscr{L}$ is the class of arithmetic progressions in $\mathbb{N}$, Mauduit and Sárközy [15], [16] proved that for almost every ω

$$W_N(\{k\omega\}) \ll N^{1/2}(\log N)^{1+\varepsilon} \quad \text{a.e.}$$

$$W_N(\{k^2\omega\}) \ll N^{3/5}(\log N)^{2/5+\varepsilon} \quad \text{a.e.}$$

and for $k = 3, 4, \ldots$

$$W_N(\{k^r\omega\}) \ll N^{1-\alpha_r} \quad \text{a.e.}$$

with some (explicitly given) constant $\alpha_r > 0$. They also proved that the above relations, with a slightly smaller exponent of the log, hold for any irrational ω whose partial quotients in the continued fraction expansion remain bounded. For the case $\mathscr{A} = \mathscr{L}$, Philipp and Tichy [20] proved that for any increasing sequence (n_k) of integers we have

$$W_N(\{n_k\omega\}) \ll N^{2/3}(\log N)^{1+\varepsilon} \quad \text{a.e.} \tag{1.22}$$

Note that

$$\psi(\mathscr{L}, N, r) \leqslant C \cdot 2^{2r} \tag{1.23}$$

and thus the case $\beta = 2$ in Theorem 4 and (1.23) yield the result of Philipp and Tichy.

Theorem 5. *Let (n_k) be an increasing sequence of integers and let $\eta_k = \eta_k(\omega) = \{n_k\omega\}$. Let $\mathscr{A}$ be a class of subsequences of $\mathbb{N}$ with entropy function $\kappa(\mathscr{A}; \delta, N) \leqslant C\delta^{-v}$ for some $v \geqslant 0$, where C depends only on $\mathscr{A}$. Then with probability* 1,

$$W_N(\mathscr{A}) \ll N^{\frac{v+1}{v+2}}(\log N)^{\frac{3}{v+2}+\varepsilon}, \quad \varepsilon > 0.$$

Finally, we formulate a theorem which extends our results for mixing sequences of random variables. Let (ξ_n) be a strictly stationary sequence of random variables satisfying a strong mixing condition

$$|P(AB) - P(A)P(B)| \leqslant \alpha(n) \downarrow 0 \tag{1.24}$$

for all $A \in \mathscr{F}_1^k$ and $B \in \mathscr{F}_{k+n}^\infty$. Here $\mathscr{F}_a^b$ denotes the σ-field generated by $\{\xi_n, a \leqslant n \leqslant b\}$. Let f be a measurable mapping from the space of infinite sequences $(\alpha_1, \alpha_2, \ldots)$ of real numbers into the real line. Define

$$\eta_n = f(\xi_n, \xi_{n+1}, \ldots), \quad n \geqslant 1 \tag{1.25}$$

and

$$\eta_{mn} = E(\eta_n | \mathscr{F}_n^{n+m}), \quad m, n \geqslant 1.$$

We assume that η_n can be closely approximated by η_{mn} in the form

$$E|\eta_n - \eta_{mn}| \leqslant \phi(m) \downarrow 0 \tag{1.26}$$

for all $m, n \geqslant 1$. This means that the functions $f(\xi_n, \xi_{n+1}, \ldots)$ can be closely approximated by functions of finitely many variables.

Sequences of the above type appear in many arithmetic applications. For example, the partial quotients in the continued fraction expansion of a number ξ chosen at random in $(0,1)$ according to the Gaussian measure $P(C) = (\log 2)^{-1} \int_C (1+x)^{-1} dx$ are stationary and satisfy the strong mixing condition (1.24) with an exponentially decreasing $\alpha(n)$. (See e.g. [12], Chapter 9.) Similar results hold for the digits in several other expansions. Condition (1.24) also holds, with exponentially decreasing $\alpha(n)$, for a large class of Markov processes; for example, for ξ_n defined by a stochastic recurrence relation $\xi_n = g(\xi_{n-1}, \varepsilon_n)$, where ε_n is an i.i.d. sequence.

Theorem 6. *Let* (ξ_n) *be a strictly stationary sequence of random variables satisfying the strong mixing condition* (1.24) *with*

$$\alpha(n) \ll n^{-p}, \quad p \geqslant 8. \tag{1.27}$$

Suppose that the random variables η_n *defined by* (1.25) *are uniformly distributed over* $[0,1)$ *and that they satisfy* (1.26) *with*

$$\phi(m) \ll m^{-q}, \quad q \geqslant 12. \tag{1.28}$$

Let $\mathscr{A}$ *be a class of increasing subsequences of* $\mathbb{N}$ *with entropy function satisfying*

$$\psi(\mathscr{A}; N, r) \leqslant B \cdot 2^{r\beta} \tag{1.29}$$

for some constants $B > 0$ *and*

$$0 \leqslant \beta \leqslant \min(p/5 - 1,\ q/5 - 3). \tag{1.30}$$

Then with probability 1

$$1/4 \leqslant \limsup_{N\to\infty} (N \log\, \log N)^{-1/2} W_N(\mathscr{A}) \leqslant C$$

for some constant $C < +\infty$.

Again we have a stronger result, expressing the Lipschitz property of the normalized empirical distribution functions $f_{N,\mathbf{p}}$. Specifically, for each $\varepsilon > 0$ there is with probability 1 a random index $N_0 = N_0(\varepsilon)$ such that

$$|f_{N,\mathbf{p}}(t) - f_{N,\mathbf{p}}(s)| \leqslant C_1 |t-s|^{1/100} + \varepsilon \tag{1.31}$$

for all $0 \leqslant s < 1$, all $(p_k) \in \mathscr{A}$ and all $N \geqslant N_0$. The constant C_1 only depends on the constants implied by $\ll$ in (1.27), (1.28) and the constant B in (1.29).

Our paper is organized as follows. Theorem 2 is proved in complete detail in Section 2. The proofs of Theorem 1 and Theorem 6 are considerably simpler and are given in Section 4. Theorem 3 is proved in Section 3 in the lacunary case; since the i.i.d. can be proved in the same way, we will omit it. Finally, Theorems 4 and 5 are proven in Section 5.

2. Proof of Theorem 2

Clearly, sequences $(p_k) \in \mathscr{A}$ having at most $\sqrt{N}$ elements in $[1,N]$ contribute to the supremum in (1.7) by at most $\sqrt{N}$ and thus in the proof of Theorem 2 (and in fact all proofs in our paper) we can restrict the definition of W_N to sequences $(p_k) \in \mathscr{A}$ having more than $\sqrt{N}$ elements in $[1,N]$.

Let (n_k) be a sequence of real numbers satisfying the Hadamard gap condition (1.17) and let $\eta_k = \eta_k(\omega) = \{n_k\omega\}$. For $0 \leqslant s < t \leqslant 1$ we set

$$x_n(s,t) := 1(s \leqslant \eta_n < t) - (t-s). \tag{2.1}$$

Finally, let β be the constant in the entropy condition (1.19). The key of the proof of Theorem 2 is the following exponential inequality, which is a sharpening of [19, Proposition 4.2.1].

Proposition 1. *Let $N \geqslant 1$ be an integer and let $R \geqslant 1$. Suppose that $\ell := t-s \geqslant N^{-3/2}$. Then for some constant $A \geqslant 1$ we have as $N \to \infty$*

$$P\left(\max_{Q \leqslant N}\left|\sum_{k=1}^{Q} x_k(s,t)\right| \geqslant AR\ell^{1/32}(N\log\log N)^{1/2}\right)$$
$$\ll \exp(-16R\ell^{-1/32}\log\log N) + R^{-8\beta}N^{-2\beta}$$

where A and the constant implied by $\ll$ depend only on q and β.

Proof. We follow the proof of [19, Proposition 4.2.1]. First, we note that by the argument in [19, p. 338] we can assume without loss of generality that $q \geqslant 16$. Next, for each $k = 1,2,\dots,$ define r_k to be the largest integer r such that

$$2^r \leqslant n_k 4^{k^{1/4}}. \tag{2.2}$$

Let $\mathscr{F}_k$ denote the σ-field generated by the dyadic intervals

$$U_{\nu k} = [\nu 2^{-r_k}, (\nu+1)2^{-r_k}) \quad \nu = 0,1,2,\dots,2^{r_k}-1.$$

Then, as in the proof of [19, Lemma 4.2.2] we have for $k \geqslant 0$ and $j \geqslant 1$

$$E(x_{j+k}|\mathscr{F}_j) \ll \ell \cdot 4^{j^{1/4}} 16^{-k} \quad \text{a.s.} \tag{2.3}$$

where the constant implied by $\ll$ is absolute.

As in [19], we define blocks $H_1, I_1, H_2, I_2, \dots$ of consecutive integers inductively: both H_j and I_j consist of $2[j^{1/2}]$ consecutive integers and there are no gaps between the blocks. Thus $H_1 = \{1,2\}$, $I_1 = \{3,4\},\dots,H_4 = \{13,14,15,16\}$, $I_4 = \{17,18,19,20\},\dots$ Let h_j be the largest number of H_j. For $N \geqslant 1$ let $M = M_N$ be defined by

$$h_{M-1} < N \leqslant h_M, \tag{2.4}$$

then

$$h_M - h_{M-1} = 4[M^{1/2}] \ll N^{1/3}. \tag{2.5}$$

As in [19, (4.2.5)] we discretize the $x_\nu, \nu \in H_j$ by setting

$$\xi_\nu := E(x_\nu|\mathscr{F}_{h_j}) \quad \nu \in H_j. \tag{2.6}$$

We introduce the block sums

$$w_j = \sum_{\nu \in H_j} x_\nu, \quad y_j = \sum_{\nu \in H_j} \xi_\nu = E(w_j|\mathscr{F}_{h_j}). \tag{2.7}$$

Then as in [19, Lemma 4.2.3],

$$||x_k - \xi_k||_2 \ll 2^{-k^{1/4}}.$$

The proof of [19, Lemma 4.2.4] with (8β)-th instead of sixth moments yields □

Lemma 2. *As* $N \to \infty$,

$$P\left(\sum_{j \leqslant M} |y_j - w_j| \geqslant R\ell^{1/32} N^{1/2} \right) \ll R^{-8\beta} N^{-2\beta}.$$

Relation [19, (4.2.8)] continues to hold with $2^{j^{1/8}}$ on the right side and thus the proof of [19, Lemma 4.2.5] yields

$$E(w_j^2 | \mathscr{F}_{h_{j-1}}) \ll \ell j^{1/2} \quad \text{a.s.} \tag{2.8}$$

where the constant implied by $\ll$ is absolute. Also Lemmas 4.2.6, 4.2.7, and 4.2.8 in [19] remain valid as they stand, yielding

$$\sum_{n=N+1}^{h_M} x_n \ll \ell^{1/8} N^{1/2} \tag{2.9}$$

and

$$y_j = Y_j + v_j \tag{2.10}$$

where $(Y_j, \mathscr{L}_j)$ is a martingale difference sequence with $\mathscr{L}_j = \sigma(y_1, \ldots, y_j)$, satisfying

$$v_j = E(y_j | \mathscr{L}_{j-1}) \ll \ell \cdot 16^{-j^{1/4}} \quad \text{a.s.} \tag{2.11}$$

and

$$\sum_{j \leqslant M} E(Y_j^2 | \mathscr{L}_{j-1}) \leqslant D\ell N \quad \text{a.s.} \tag{2.12}$$

where the constant D and the constants implied by $\ll$ are absolute.

Finally we replace [19, Lemma 4.2.9] by the following lemma.

Lemma 3. *As* $N \to \infty$

$$P\left(\max_{k \leqslant M} \left| \sum_{j \leqslant k} Y_j \right| > 8RD\ell^{1/32} (N \log \log N)^{\frac{1}{2}} \right) \ll \exp(-16R\ell^{-1/32} \log \log N).$$

For the proof we choose in the proof of [19, Lemma 4.2.9] the parameters c, λ and K as follows:

$$c = 2M^{\frac{1}{2}}, \quad \lambda = 2\ell^{-1/16}(\log \log M)^{\frac{1}{2}} M^{-3/4}, \quad K = 4RD\ell^{3/8} M^{3/2}.$$

□

Treating the block sums

$$z_j := \sum_{\nu \in I_j} \xi_\nu$$

in the same way, and taking (2.9), (2.11) and Lemma 2 into account we finally obtain the estimate as claimed in Proposition 1.

Proposition 2. *Let $(p_k) \in \mathscr{A}$ and let $N \geqslant 1$, $R \geqslant 1$. Let $\phi(N)$ denote the largest k such that $p_k \leqslant N$ and assume that $\phi(N) \geqslant N^{1/2}$. Finally, suppose that $t - s \geqslant N^{-3/4}$. Then*

$$P\left(\max_{Q \leqslant N}\left|\sum_{p_k \leqslant Q} x_{p_k}(s,t)\right| \geqslant AR(t-s)^{1/32}(\phi(N)\log\log N)^{1/2}\right)$$
$$\ll \exp(-14R(t-s)^{-1/32}\log\log N) + R^{-8\beta}\phi(N)^{-2\beta}$$

where both $A \geqslant 1$ and the constant implied by $\ll$ depend only on q and β.

Proof. Since there are $\phi(N)$ terms with $p_k \leqslant N$, Proposition 1 implies

$$P\left(\max_{Q \leqslant N}\left|\sum_{p_k \leqslant Q} x_{p_k}(s,t)\right| \geqslant AR(t-s)^{1/32}(\phi(N)\log\log\phi(N))^{1/2}\right)$$
$$\ll \exp(-16R(t-s)^{-1/32}\log\log\phi(N)) + R^{-8\beta}\phi(N)^{-2\beta}.$$

Here we used the fact that

$$t - s \geqslant N^{-3/4} \geqslant \phi(N)^{-3/2}$$

by the assumptions of Proposition 2. Next observe that by $N^{1/2} \leqslant \phi(N) \leqslant N$, $\log\log\phi(N)$ differs from $\log\log N$ by not more than 1 and thus their ratio is between $14/16$ and 1 for $N \geqslant N_0$. Hence the probability in question does not exceed

$$\exp(-14R(t-s)^{-1/32}\log\log N) + R^{-8\beta}\phi(N)^{-2\beta}.$$

□

Proposition 3. *Let $N \geqslant 1$, $R \geqslant 1$ and suppose that $t - s \geqslant N^{-3/4}$. Then*

$$P\left(\max_{Q \leqslant N}\max_{(p_k) \in \mathscr{A}_N}\left|\sum_{p_k \leqslant Q} x_{p_k}(s,t)\right| \geqslant AR(t-s)^{1/32}(N\log\log N)^{1/2}\right)$$
$$\ll \exp(-12R(t-s)^{-1/32}\log\log N) + R^{-8\beta}N^{-2\beta},$$

where both $A \geqslant 1$ and the constant implied by $\ll$ depend only on q and β.

Proof. We partition $\mathscr{A}_N$ into

$$\mathscr{A}_N = \bigcup_{r \geqslant 0} \mathscr{A}_N(r). \tag{2.13}$$

As we noted at the beginning of this section, it suffices to consider those r's such that $2^r \leqslant \sqrt{N}$ and thus using the entropy condition (1.19) and applying Proposition 2 with R replaced by $R2^{r/2}$ and $\phi(N) = N2^{-r}$ we get

$$P\left(\max_{Q \leqslant N}\max_{(p_k) \in \mathscr{A}_N(r)}\left|\sum_{p_k \leqslant Q} x_{p_k}(s,t)\right| \geqslant AR2^{r/2}(t-s)^{1/32}(N2^{-r}\log\log N)^{1/2}\right)$$
$$\ll 2^{\beta r}(\exp(-14R2^{r/2}(t-s)^{-1/32}\log\log N) + (R2^{r/2})^{-8\beta}(N2^{-r})^{-2\beta}). \tag{2.14}$$

Summing (2.14) over all considered r in (2.13) we obtain the result.

We now can finish the proof of Theorem 2 using the familiar chaining argument. For $N \geqslant 10$ let m and M be defined by

$$m = m(N) = [(\log \log N)^{1/2}],\ M = M(N) = \left[\frac{\log N}{2\log 2}\right] + 4. \tag{2.15}$$

□

We write s and t in binary form and obtain

$$s = a2^{-m} + \sum_{i=m+1}^{M} \sigma_i 2^{-i} + \theta_1 2^{-M}$$

and

$$t = b2^{-m} + \sum_{i=m+1}^{M} \tau_i 2^{-i} + \theta_2 2^{-M},$$

where $\sigma_i = 0, 1$ and $\tau_i = 0, 1$ and a and b are integers with $0 \leqslant a, b \leqslant 2^m$ and $0 \leqslant \theta_1, \theta_2 \leqslant 1$. Given a sequence $\mathbf{p} = (p_k)$ of positive integers, we also write

$$Z(s,t) := Z^{(\mathbf{p})}(Q; s,t) := \left| \sum_{p_k \leqslant Q} x_{p_k}(s,t) \right|.$$

We observe that for $s < r < t$

$$Z(s,t) \leqslant Z(s,r) + Z(r,t), \tag{2.16}$$

$$Z(r,t) \leqslant Z(s,t) + Z(s,r). \tag{2.17}$$

Thus

$$\begin{aligned} Z(s,t) \leqslant{} & Z(a2^{-m}, b2^{-m}) + \sum_{i=m+1}^{M} Z(a_i 2^{-i}, (a_i+1)2^{-i}) \\ & + \sum_{i=m+1}^{M} Z(b_i 2^{-i}, (b_i+1)2^{-i}) + Z(a_{M+1}2^{-M}, (a_{M+1}+1)2^{-M}) \\ & + Z(b_{M+1}2^{-M}, (b_{M+1}+1)2^{-M}) + 2Q2^{-M}, \end{aligned} \tag{2.18}$$

where a_i, b_i $(m < i \leqslant M+1)$ are integers with $0 \leqslant a_i, b_i < 2^i$. The last term is explained by the fact that for $0 \leqslant h < 2^M$ and $0 \leqslant \theta \leqslant 1$,

$$Z(h2^{-M}, (h+\theta)2^{-M}) \leqslant Z(h2^{-M}, (h+1)2^{-M}) + 2^{-M}$$

by an application of (2.16), (2.17). We define the following events:

$$\begin{aligned} E_N(a,b) = \Bigg\{ & \max_{\substack{Q \leqslant N \\ (p_k) \in \mathscr{A}_N}} Z^{(\mathbf{p})}(Q; a2^{-m}, b2^{-m}) \\ & \geqslant A((b-a)2^{-m})^{1/32} (N \log \log N)^{1/2} \Bigg\} \end{aligned} \tag{2.19}$$

$$E_N = \bigcup_{0 \leqslant a,b \leqslant 2^m} E_N(a,b)$$

$$F_N(i,a) = \left\{ \max_{\substack{Q \leqslant N \\ (p_k) \in \mathscr{A}_N}} Z^{(\mathbf{p})}(Q; a2^{-i}, (a+1)2^{-i}) \geqslant A2^{-i/32}(N \log\log N)^{1/2} \right\} \quad (2.20)$$

and

$$F_N = \bigcup_{m < i \leqslant M} \bigcup_{0 \leqslant a < 2^i} F_N(i,a).$$

Here, A is the constant from Proposition 3. Using Proposition 3 with $R = 1$ we obtain

$$P(E_N(a,b)) \ll \exp(-12 \log\log N)$$

and so

$$P(E_N) \ll 2^{2m} \exp(-12 \log\log N) \ll \exp(-10 \log\log N) = (\log N)^{-10}.$$

Similarly, with $R = 1$ and $t - s = 2^{-i}$,

$$P(F_N(i,a)) \ll \exp(-12 \cdot 2^{i/32} \log\log N) + N^{-2\beta}$$

and so

$$\begin{aligned} P(F_N) &\ll \sum_{m < i \leqslant M} 2^i \exp(-12 \cdot 2^{i/32} \log\log N) + 2^M N^{-2\beta} \\ &\ll \exp(-10 \log\log N) = (\log N)^{-10}. \end{aligned}$$

(Note that in the applications of Proposition 3 the condition $t - s \geqslant N^{-3/4}$ is satisfied.) Consequently,

$$\sum_{p=1}^{\infty} P(E_{2^p} \cup F_{2^p}) < \infty.$$

Hence the Borel-Cantelli lemma implies that with probability 1 only finitely many of the events E_{2^p} of F_{2^p} occur. Let N be sufficiently large and let p be such that $2^{p-1} < N \leqslant 2^p$. Then by (2.18) we have with probability 1 for all $0 \leqslant s < t \leqslant 1$, $N \geqslant N_0(\varepsilon)$

$$\begin{aligned} &\max_{Q \leqslant N} \max_{(p_k) \in \mathscr{A}} Z^{(\mathbf{p})}(Q; s, t) \\ &\quad \leqslant A((b-a)2^{-m(2^p)})^{1/32}(2^p \log\log 2^p)^{1/2} \\ &\qquad + 4A \sum_{m(2^p) < i \leqslant M(2^p)} 2^{-i/32}(2^p \log\log 2^p)^{1/2} + 2^{p+1}2^{-M(2^p)} \\ &\quad \leqslant 4A[(t-s)^{1/32} + \varepsilon](N \log\log N)^{1/2} + 4N^{1/2} \\ &\quad \leqslant 8A[(t-s)^{1/32} + \varepsilon](N \log\log N)^{1/2}. \end{aligned}$$

This proves (1.20) and thus Theorem 2.

3. Proof of Theorem 3

We prove the theorem first in the lacunary case, i.e. for the sequence $\eta_k = \{n_k\omega\}$. Fix $1/2 < \alpha < 1$ and $0 \leqslant s < t \leqslant 1$. By the hypotheses of the theorem, we can choose $\beta > 0$ such that

$$\kappa(\mathscr{A}; \delta, N) \ll \delta^{-\beta/2} \tag{3.1}$$

where the constant implied by $\ll$ depends only on $\mathscr{A}$. For simplicity we set

$$\kappa(\delta) := \kappa(\mathscr{A}; \delta^{\alpha}, [1/\delta]). \tag{3.2}$$

By Proposition 1 we have for any sequence (p_k) of positive integers and any $R \geqslant 1$, $0 < \varepsilon \leqslant 1/32$ and $t - s \geqslant 2^{-3r/2}$ as $r \to \infty$

$$P\left(\max_{Q \leqslant 2^r} \left| \sum_{p_k \leqslant Q} x_{p_k}(s,t) \right| \geqslant AR2^{\frac{1}{2}}r(t-s)^{\varepsilon}(\log\kappa(2^{-r}) + (\log r)^{\frac{1}{2}}) \right)$$
$$\ll \begin{cases} \exp(-16R(t-s)^{-\varepsilon}\log\kappa(2^{-r})\log^{\frac{1}{2}}r) + R^{-8\beta}2^{-2r\beta} & \text{if } \log\kappa(2^{-r}) > \log^{\frac{1}{2}}r \\ \exp(-16R(t-s)^{-\varepsilon}\log r) + R^{-8\beta}2^{-2r\beta} & \text{if } \log\kappa(2^{-r}) \leqslant \log^{\frac{1}{2}}r \end{cases} \tag{3.3}$$

for some constant $A \geqslant 1$. (In the case of the first line of (3.3) we apply Proposition 1 with R replaced by $R\log\kappa(2^{-r})(\log r)^{-1/2}$). Let

$$\delta := AR(t-s)^{\varepsilon}2^{-r/2} \tag{3.4}$$

and $\mathscr{B} = \{(p_k^{(1)}), \ldots, (p_k^{(M)})\}$ a maximal set of sequences in $\mathscr{A}$ with pairwise distance $>\delta$ with respect to the normalized Hamming distance $d(\cdot,\cdot,2^r)$. Then

$$M = \kappa(\mathscr{A}, \delta, 2^r) \leqslant \kappa(\mathscr{A}; 2^{-\alpha r}, 2^r) = \kappa(2^{-r})$$

since

$$\delta \geqslant (t-s)^{\varepsilon}2^{-r/2} \geqslant 2^{-r(3\varepsilon/2+1/2)} \geqslant 2^{-\alpha r},$$

provided we choose $\varepsilon > 0$ so small that $3\varepsilon/2 + 1/2 < \alpha$. Clearly, for any $(q_k) \in \mathscr{A}$ there is a $(p_k) \in \mathscr{B}$ with $d((p_k),(q_k),2^r) \leqslant \delta$, which implies that for any $Q \leqslant 2^r$ the sums $\sum_{p_k \leqslant Q} x_{p_k}(s,t)$ and $\sum_{q_k \leqslant Q} x_{q_k}(s,t)$ differ at most by $\delta 2^r = AR(t-s)^{\varepsilon}2^{r/2}$. Hence using (3.3) we get

$$P\left(\max_{(q_k) \in \mathscr{A}} \max_{Q \leqslant 2^r} \left| \sum_{q_k \leqslant Q} x_{q_k}(s,t) \right| \geqslant 2AR2^{\frac{1}{2}r}(t-s)^{\varepsilon}(\log\kappa(2^{-r}) + \log^{\frac{1}{2}}r) \right)$$
$$\ll \exp(-8R(t-s)^{-\varepsilon}(\log\kappa(2^{-r}) + \log^{\frac{1}{2}}r)\log^{\frac{1}{2}}r + \log\kappa(2^{-r}))$$
$$+ R^{-8\beta}\kappa(2^{-r})2^{-2r\beta}$$
$$\ll \exp(-4R(t-s)^{-\varepsilon}\log r) + R^{-8\beta}2^{-3r\beta/2} \tag{3.5}$$

by distinguishing the cases $\log\kappa(2^{-r}) > \log^{\frac{1}{2}}r$ and $\log\kappa(2^{-r}) \leqslant \log^{\frac{1}{2}}r$ and by using (3.1) in the last step.

Relation (3.5) is analogous to Proposition 3 and the proof of Theorem 3 in the lacunary case can now be completed by the same chaining argument that was used the proof of Theorem 2. The proof for i.i.d. uniform random variables η_k is the same, except that instead of Proposition 1 we use the analogous exponential bound given by Lemma 4 below.

4. Proof of Theorems 1 and 6

The proof of Theorem 1 follows the pattern of the proof of Theorem 2. Let (η_n) be an i.i.d. sequence with the uniform distribution (1.11) and define $x_n(s,t)$ by (2.1). We replace Proposition 1 by the following lemma which is an immediate consequence of Bernstein's inequality and Skorokhod's maximal inequality.

Lemma 4. *Let $N \geqslant 1$, $R \geqslant 1$, $0 < \alpha < 1/2$. There exist $\gamma = \gamma(\alpha) > 1$, $\rho = \rho(\alpha) > 0$ such that for $\ell = t - s \geqslant N^{-\gamma}$ we have*

$$P\left(\max_{Q \leqslant N}\left|\sum_{k=1}^{Q} x_k(s,t)\right| \geqslant 6R\ell^{\alpha}(N \log\, \log N)^{1/2}\right)$$
$$\ll \exp(-2R\ell^{-(1-2\alpha)} \log\, \log N) + \exp(-RN^{\rho})$$

where the constant implied by $\ll$ depends on α.

Proof. Clearly, the $x_k(s,t)$ are independent random variables with mean 0 and variance $\ell(1-\ell) \leqslant \ell$. Hence Bernstein's inequality (see e.g. Petrov [18], pp. 57–58) implies that the probability

$$P\left(\left|\sum_{k=1}^{Q} x_k(s,t)\right| \geqslant x\right)$$

can be bounded by $2\exp(-x^2/4Q\ell)$ if $0 \leqslant x \leqslant Q\ell(1-\ell)$ and by $2\exp(-x/4)$ if $x > Q\ell(1-\ell)$. Thus for any $x \geqslant 0$ we have

$$P\left(\left|\sum_{k=1}^{Q} x_k(s,t)\right| \geqslant x\right) \leqslant 2\exp(-x^2/4Q\ell) + 2\exp(-x/4).$$

Choose γ so that $1 < \gamma < 1/(2\alpha)$. Then for any $1 \leqslant Q \leqslant N$, $\ell \geqslant N^{-\gamma}$ we have

$$P\left(\left|\sum_{k=1}^{Q} x_k(s,t)\right| \geqslant 3R\ell^{\alpha}(N \log\, \log N)^{1/2}\right)$$
$$\leqslant 2\exp(-9R^2\ell^{2\alpha}N \log\, \log N/4Q\ell) + 2\exp(-3R\ell^{\alpha}(N \log\, \log N)^{1/2}/4)$$
$$\leqslant 2\exp(-2R\ell^{-(1-2\alpha)} \log\, \log N) + 2\exp(-RN^{1/2-\gamma\alpha}).$$

□

Using Skorokhod's inequality (see e.g. Breiman [3], p. 45) completes the proof.

Lemma 5. *Let $(p_k) \in \mathscr{A}$ and let $N \geqslant 1, R \geqslant 1$, $0 < \alpha < 1/2$. Let $\phi(N)$ denote the largest k such that $p_k \leqslant N$ and assume that $\phi(N) \geqslant N^{1/2}$. Finally, suppose that $\ell = t - s \geqslant N^{-\gamma/2}$, where $\gamma = \gamma(\alpha) > 1$ is the constant in Lemma 4. Then*

$$P\left(\max_{Q \leqslant N}\left|\sum_{p_k \leqslant Q} x_{p_k}(s,t)\right| \geqslant 6R\ell^{\alpha}(\phi(N) \log\, \log N)^{1/2}\right)$$
$$\ll \exp(-R \cdot \ell^{-(1-2\alpha)} \log\, \log N) + \exp(-RN^{\rho/2})$$

where $\rho = \rho(\alpha) > 0$ is the constant in Lemma 4.

The proof is an easy modification of the proof of Proposition 2.

Lemma 6. *Let $N \geqslant 1, R \geqslant 1, 0 < \alpha < 1/2$ and suppose that $\ell = t - s \geqslant N^{-\gamma/2}$, where $\gamma = \gamma(\alpha) > 1$ is the constant in Lemma* 4. *Let B denote the constant in* (1.12). *Then we have*

$$P\left(\max_{Q \leqslant N}\max_{(p_k) \in \mathscr{A}_N}\left|\sum_{p_k \leqslant Q} x_{p_k}(s,t)\right| \geqslant 12BR\ell^{\alpha}(N\log\log N)^{1/2}\right)$$
$$\ll \exp(-R \cdot \ell^{-(1-2\alpha)}\log\log N) + \exp(-RN^{\rho/2})$$

where $\rho = \rho(\alpha) > 0$ is the constant in Lemma 4.

Proof. We follow the proof of Proposition 3. We partition $\mathscr{A}_N$ as in (2.13), and, similarly as in the proof of Proposition 3, it suffices to consider those r's for which $N2^{-r} \geqslant \sqrt{N}$. Applying Lemma 5 with R replaced by $2BR2^{r/2}$ and $\phi(N) = N2^{-r}$ and using (1.12) it follows that the probability in the statement of Lemma 6 does not exceed

$$\sum_{r \geqslant 0, 2^r \leqslant \sqrt{N}} P\left(\max_{Q \leqslant N}\max_{(p_k) \in \mathscr{A}_N(r)}\left|\sum_{p_k \leqslant Q} x_{p_k}(s,t)\right| \geqslant 12BR2^{r/2} \cdot \ell^{\alpha}(N2^{-r}\log\log N)^{1/2}\right)$$
$$\ll \sum_{r \geqslant 0}\exp(-2BR2^{r/2}\ell^{-(1-2\alpha)}\log\log N + B2^{r/2} \cdot \log\log N)$$
$$+\sum_{r \geqslant 0}\exp(-2BR2^{r/2}N^{\rho/2} + B2^{r/2} \cdot \log\log N)$$
$$\ll \exp(-R\ell^{-(1-2\alpha)}\log\log N) + \exp(-RN^{\rho/2}).$$ □

The remainder of the proof of Theorem 1 can now be completed as in Section 2.

The proof of Theorem 6 also follows the pattern of the proof of Theorem 2. We will need the following exponential bound.

Proposition 4. *Assume the conditions of Theorem* 6 *and let $x_n(s,t)$ be defined by* (2.1). *Let $N \geqslant 1$ and $R \geqslant 1$ and suppose that $\ell := t - s \geqslant N^{-2}$. Then for some constant $A \geqslant 1$ depending only on p and q we have as $N \to \infty$*

$$P\left(\max_{Q \leqslant N}\left|\sum_{k=1}^{Q} x_k(s,t)\right| \geqslant AR\ell^{1/120}(N\log\log N)^{1/2}\right)$$
$$\ll \exp(-6R\ell^{-1/120}\log\log N) + R^{-p/4}N^{-p/10} + R^{-3}N^{2-q/5}.$$

Proposition 4 is similar to [19, Proposition 3.3.1], but the term $R^{-2}N^{-1.03}$ there is replaced by a term depending on p, q, which improves if p and q are increasing. The proof follows the proof of [19, Proposition 3.3.1] with minor changes. Since the changes are routine, we will leave the details to the reader.

Proposition 5. *Let $N \geqslant 1$ and suppose that $\ell \geqslant N^{-1}$. Then as $N \to \infty$ we have for some $\delta = \delta(p,q) > 1/2$*

$$P\left(\max_{Q \leqslant N}\max_{(p_k) \in \mathscr{A}_N}\left|\sum_{p_k \leqslant Q} x_{p_k}(s,t)\right| \geqslant AR\ell^{1/120}(N\log\log N)^{1/2}\right)$$
$$\ll \exp(-4R\ell^{-1/120}\log\log N) + R^{-2}N^{-\delta}.$$

Proof. We fix first $(p_k) \in \mathscr{A}$ and let $\phi(N)$ denote the largest k with $p_k \leqslant N$. Since the sequence (x_{p_k}) is mixing with an even better mixing rate, under the assumptions $\phi(N) \geqslant \sqrt{N}$, $\ell = t - s \geqslant N^{-1}$ Proposition 4 implies

$$P\left(\max_{Q \leqslant N} \left| \sum_{p_k \leqslant Q} x_{p_k}(s,t) \right| \geqslant AR\ell^{1/120}(\phi(N) \log \log N)^{1/2}\right)$$
$$\ll \exp(-5R\ell^{-1/120} \log \log N) + R^{-p/4}\phi(N)^{-p/10} + R^{-3}\phi(N)^{2-q/5}. \qquad (4.1)$$

We now partition $\mathscr{A}_N$ as in (2.13) and apply (4.1) with R replaced by $R2^{r/2}$ and $\phi(N) = N2^{-r}$. As in our earlier proofs, it suffices to consider the case $N2^{-r} \geqslant \sqrt{N}$. Letting β denote the constant in the entropy condition (1.29), an upper bound for the probability in Proposition 5 is obtained by multiplying the bound in (4.1) by $2^{r\beta}$ and sum over the indicated r's. The sum of the first terms is

$$\ll \sum_r 2^{r\beta} \exp(-5R \cdot 2^{r/2}\ell^{-1/120} \log \log N) \ll \exp(-4R\ell^{-1/120} \log \log N).$$

The sum of the second terms is

$$\begin{aligned}
&\ll (R2^{r/2})^{-p/4} \sum_{2^r \leqslant \sqrt{N}} (N2^{-r})^{-p/10} \cdot 2^{r\beta} \\
&\leqslant R^{-p/4}N^{-p/10} \sum_{2^r \leqslant \sqrt{N}} 2^{r\beta} \\
&\ll R^{-p/4}N^{-p/10+\beta/2} \\
&\ll R^{-p/4}N^{-\delta}
\end{aligned}$$

for some $\delta > 1/2$. We used here the fact that $\beta < p/5 - 1$ by (1.30). Finally, the sum of the third terms is

$$\begin{aligned}
&\ll R^{-3} \sum_{2^r \leqslant \sqrt{N}} (N2^{-r})^{2-q/5}2^{r\beta} = R^{-3}N^{2-q/5} \sum_{2^r \leqslant \sqrt{N}} 2^{r(\beta+q/5-2)} \\
&\ll R^{-3}N^{2-q/5}\sqrt{N}^{\beta+q/5-2} = R^{-3}N^{1-q/10+\beta/2} \ll R^{-3}N^{-\delta}
\end{aligned}$$

for some $\delta > 1/2$, using the fact that $\beta < q/5 - 3$ by (1.30). This completes the proof of Proposition 4. □

The proof of Theorem 6 can now be completed by using the chaining argument in Theorem 2.

5. Proof of Theorems 4 and 5

Assume the conditions of Theorem 4. Fix $N \geqslant 1$, $r \geqslant 1$ and let (p_k) be a fixed sequence in $[1,N]$ such that $(p_k) \in \mathscr{A}_N(r)$. By the Erdős-Turán inequality (see e.g. [5], p. 15 or [11], p. 114) we have for any $1 \leqslant Q \leqslant N$

$$\sup_{0 \leqslant t \leqslant 1} \left| \sum_{p_k \leqslant Q} (1(\eta_{p_k}(\omega) \leqslant t) - t) \right| \leqslant \frac{6R}{H} + 6 \sum_{1 \leqslant h \leqslant H} \frac{1}{h} \left| \sum_{p_k \leqslant Q} e(hn_{p_k}\omega) \right|.$$

Here $R = \#\{k : p_k \leqslant Q\}$, $e(x) = \exp(2\pi i x)$ and $H \geqslant 1$ is arbitrary. Clearly $R \leqslant N$ and thus

$$\max_{Q \leqslant N} \sup_{0 \leqslant t \leqslant 1} \left| \sum_{p_k \leqslant Q} (1(\eta_{p_k}(\omega) \leqslant t) - t) \right|^2$$
$$\leqslant \frac{72N^2}{H^2} + 72 \left(\sum_{1 \leqslant h \leqslant H} \frac{1}{h} \max_{Q \leqslant N} \left| \sum_{p_k \leqslant Q} e(hn_{p_k}\omega) \right| \right)^2 .$$

By Hunt's inequality (see e.g. [17]) we have

$$E \left(\max_{Q \leqslant N} \left| \sum_{p_k \leqslant Q} e(hn_{p_k}\omega) \right|^2 \right) \leqslant C \sum_{p_k \leqslant N} 1 \leqslant CN2^{-(r-1)}$$

and thus choosing $H = N$ and using Minkowski's inequality we get

$$E \left(\max_{Q \leqslant N} \sup_{0 \leqslant t \leqslant 1} \left| \sum_{p_k \leqslant Q} (1(\eta_{p_k} \leqslant t) - t) \right|^2 \right)$$
$$\ll N2^{-r} \log^2 N + 1 \ll N2^{-r} \log^2 N. \tag{5.1}$$

(To justify the last step, we note that without loss of generality we can assume that $N2^{-(r-1)} \geqslant 1$, since otherwise $\mathscr{A}_N(r)$ is empty.) Since the number of sequences $(p_k) \in \mathscr{A}_N(r)$ is at most $B \cdot 2^{r\beta}$ by the assumptions of Theorem 4, we have for any $\alpha > 0$, $\tau > 0$ (to be chosen suitably later),

$$P \left(\max_{(p_k) \in \mathscr{A}_N(r)} \max_{Q \leqslant N} \sup_{0 \leqslant t \leqslant 1} \left| \sum_{p_k \leqslant Q} (1(\eta_{p_k} \leqslant t) - t) \right| \geqslant 2N^{\alpha} (\log N)^{\tau} \right)$$
$$\ll N^{1-2\alpha} (\log N)^{2-2\tau} \cdot 2^{r(\beta-1)}. \tag{5.2}$$

Without loss of generality we can assume that $N2^{-(r-1)} \geqslant N^{\alpha}(\log N)^{\tau}$, i.e.

$$2^r \leqslant 2N^{1-\alpha} (\log N)^{-\tau} \tag{5.3}$$

since otherwise the absolute value of the sum in (5.2) would be less than $N^{\alpha}(\log N)^{\tau}$. Summing the probability bounds in (5.2) over all r subject to (5.3) and choosing α and τ according to the following table

β	α	τ
>1	$\beta/(1+\beta)$	$(3+\varepsilon)/(1+\beta)$
$=1$	$\frac{1}{2}$	$2+\varepsilon$
<1	$\frac{1}{2}$	$\frac{3}{2}+\varepsilon$

we obtain letting $N = 2^{m+1}$, $m = 1, 2, \ldots$

$$P \left(\max_{2^m < Q \leqslant 2^{m+1}} \max_{(p_k) \in \mathscr{A}} \sup_{0 \leqslant t \leqslant 1} \left| \sum_{p_k \leqslant Q} (1(\eta_{p_k} \leqslant t) - t) \right| \geqslant C^* 2^{m\alpha} m^{\tau} \right)$$
$$\ll (\log 2^m)^{-(1+\varepsilon')} \ll m^{-(1+\varepsilon')}$$

for some $C^* > 0$, $\varepsilon' > 0$. We apply the convergence part of the Borel-Cantelli lemma and obtain the conclusion of Theorem 4.

Turning to the proof of Theorem 5, let $(p_k) \in \mathscr{A}$. From (5.1) it follows that

$$E\left(\max_{Q \leqslant N} \sup_{0 \leqslant t \leqslant 1} \left| \sum_{p_k \leqslant Q} (1(\eta_{p_k} \leqslant t) - t) \right|^2 \right) \ll N \log^2 N. \tag{5.4}$$

Let

$$\delta := N^{-\frac{1}{v+2}} (\log N)^{\frac{3+\varepsilon}{v+2}}$$

and let $\mathscr{B} = \{(p_k^{(1)}), \ldots, (p_k^{(M)})\}$ be a maximal set of sequences in $\mathscr{A}$ with pairwise distance $>\delta$ with respect to the normalized Hamming distance $d(\cdot,\cdot,N)$. By the assumptions of Theorem 5 we have

$$M = \kappa(\mathscr{A}; \delta, N) \leqslant C\delta^{-v} = CN^{\frac{v}{v+2}} (\log N)^{-v\frac{3+\varepsilon}{v+2}}.$$

Clearly, for any $(q_k) \in \mathscr{A}$ there is a $(p_k) \in \mathscr{B}$ with $d((p_k),(q_k),N) \leqslant \delta$, which implies that for any $Q \leqslant N$ the sums $\sum_{p_k \leqslant Q} x_{p_k}(s,t)$ and $\sum_{q_k \leqslant Q} x_{q_k}(s,t)$ differ at most by

$$\delta N = N^{\frac{v+1}{v+2}} (\log N)^{\frac{3+\varepsilon}{v+2}}.$$

Hence using (5.4) we get

$$\begin{aligned} &P\left(\max_{Q \leqslant N} \max_{(p_k) \in \mathscr{A}} \sup_{0 \leqslant t \leqslant 1} \left| \sum_{p_k \leqslant Q} (1(\eta_{p_k} \leqslant t) - t) \right| \geqslant 2N^{\frac{v+1}{v+2}} (\log N)^{\frac{3+\varepsilon}{v+2}} \right) \\ &\qquad \leqslant P\left(\max_{Q \leqslant N} \max_{(p_k) \in \mathscr{B}} \sup_{0 \leqslant t \leqslant 1} \left| \sum_{p_k \leqslant Q} (1(\eta_{p_k} \leqslant t) - t) \right| \geqslant N^{\frac{v+1}{v+2}} (\log N)^{\frac{3+\varepsilon}{v+2}} \right) \\ &\qquad \leqslant M \cdot N \log^2 N \cdot N^{-2\frac{v+1}{v+2}} (\log N)^{-2\frac{3+\varepsilon}{v+2}} = (\log N)^{-(1+\varepsilon)}. \end{aligned}$$

We let $N = 2^m, m = 1, 2, \ldots,$ apply the Borel-Cantelli lemma and obtain the conclusion of Theorem 5.

References

[1] Alon N, Kohayakawa Y, Mauduit C, Moreira CG, Rödl V (2007) Measures of pseudorandomness for finite sequences: typical values. Proc London Math Soc **95**: 778–812

[2] Berkes I, Philipp W, Tichy RF (2007) Pseudorandom numbers and entropy conditions. J Complexity Theory **23**: 516–527

[3] Breiman L (1968) Probability. New York: Addison-Wesley

[4] Cassaigne J, Mauduit C, Sárközy A (2002) On finite pseudorandom binary sequences VII: The measures of pseudorandomness. Acta Arith **103**: 97–118

[5] Drmota M, Tichy RF (1997) Sequences, Discrepancies and Applications. Lect Notes Math **1651**. Berlin Heidelberg New York: Springer

[6] Dudley RM (1978) Central limit theorems for empirical measures. Ann Prob **6**: 899–929; Corr ibid **7**: 909–911

[7] Dudley RM (1984) A course in empirical processes. Lect Notes Math **1097**: 1–142. Berlin Heidelberg New York: Springer

[8] Dudley RM, Philipp W (1984) Invariance principles of Banach space valued random elements and empirical processes. Z Wahrscheinlichkeitstheorie verw Geb **62**: 509–552

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

[10] Knuth DE (1981) The Art of Computer Programming, Vol 2, 2nd edn. New York: Addison-Wesley

[11] Kuipers L, Niederreiter H (1974) Uniform Distribution of Sequences. New York: Wiley

[12] Lévy P (1937) Théorie de l'Addition des Variables Aléatoires. Paris: Gauthier-Villars

[13] Mauduit C, Niederreiter H, Sárközy A (2007) On pseudorandom $[0, 1)$ and binary sequences. Publ Math Debrecen, to appear

[14] Mauduit C, Sárközy A (1997) On finite pseudorandom binary sequences I: Measure of pseudorandomness, the Legendre symbol. Acta Arith **82**: 365–377

[15] Mauduit C, Sárközy A (2000) On finite pseudorandom binary sequences, V. On $\{n\alpha\}$ and $\{n^2\alpha\}$ sequences. Monatsh Math **129**: 197–216

[16] Mauduit C, Sárközy A (2000) On finite pseudorandom binary sequences, VI. On $\{n^k\alpha\}$ sequences. Monatsh Math **130**: 281–298

[17] Mozzochi CJ (1971) On the pointwise convergence of Fourier series. Lect Notes Math **199**. Berlin Heidelberg New York: Springer

[18] Petrov VV (1995) Limit Theorems of Probability Theory. Cambridge: Univ. Press

[19] Philipp W (1977) A functional law of the iterated logarithm for empirical disribution functions of weakly dependent random variables. Ann Probab **5**: 319–350

[20] Philipp W, Tichy RF (2002) Metric theorems for distribution measures of pseudorandom sequences. Monatsh Math **135**: 321–326

[21] Pollard D (1984) Convergence of Stochastic Processes. Berlin Heidelberg New York: Springer

Authors' addresses: István Berkes, Department of Statistics, Technical University Graz, Steyrergasse 17/IV, A-8010 Graz, Austria, e-mail: berkes@tugraz.at; Walter Philipp, Department of Statistics, University of Illinois, 725 S. Wright Street, Champaign, IL 61820, USA; Robert F. Tichy, Department of Analysis and Computational Number Theory, Technical University Graz, Steyrergasse 30, A-8010 Graz, Austria, e-mail: tichy@tugraz.at