Publ. Math. Debrecen
76/3 (2010), 287–302

The Rosenblatt coefficient of dependence and the ASCLT for some classes of weakly dependent random sequences

By RITA GIULIANO (Pisa)

Dedicated to the 100th anniversary of the birthday of Béla Gyires

Abstract. Using some Berry–Esseen type results, we prove a new bound for the Rosenblatt coefficient of the normalized partial sums of a sequence of random variables, weakly dependent in some sense; this bound is used to prove an Almost Sure Central Limit Theorem for the same sequence.

1. Introduction

The Almost Sure Central Limit Theorem (ASCLT from now on) is a classical result in the asymptotic theory of random sequences. It has been originally proved for a sequence of independent identically distributed random variables ($\mathbf{E}[X_1] = 0$, $\mathbf{E}[X_1^2] = 1$) $(X_n)_{n\geq 1}$ with partial sums $(S_n)_{n\geq 1}$ (see [3], [12] and [19]) and it can be stated as follows: P-almost surely,

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} 1_A \left(\frac{S_k}{\sqrt{k}} \right) = \mu(A),$$

for all Borel sets $A \subseteq \mathbb{R}$ such that $\lambda(\partial A) = 0$ (1_A is the indicator function of the set A). Here and in the sequel λ denotes the Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$,

Mathematics Subject Classification: Primary: 60F05; Secondary 60G10.
Key words and phrases: Rosenblatt coefficient, measure of dependence, strong mixing coefficients, uniform dependence coefficients, Almost Sure Central Limit Theorem.
This work is partially supported by the M.I.U.R. Italy.

while μ stands for the standard Gaussian measure on $\mathbb{R}$, i.e.

$$\mu(A) = \int_A \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, \lambda(dx), \quad A \in \mathcal{B}(\mathbb{R}).$$

The ASCLT has been recently extended to more general sequences: the typical framework is that of a stationary sequence, weakly dependent in some sense. For instance the paper [15] deals with associated, α- mixing and ρ-mixing sequences; in [5] sequences dependent in the sense of Doukhan and Louhichi are considered. See also the papers [1], [7], [8], [11], [13], [14].

The point of view of the present paper is somewhat different. More precisely, the aim of the present paper is to prove an ASCLT for sequences such that their partial sums S_n (properly normalized) have a "good" speed of convergence to the Gaussian law. The term "good" will be specified later, see the statement of Theorem 2.4. There are many examples of such sequences: for instance the case of m-dependent sequences (already considered in the paper [10]).

Our point of view allows to study also some non-stationary sequences: we shall consider a strongly mixing sequence and a sequence satisfying another condition of dependence, introduced by Rio in [18].

For such sequences we prove a general result (Theorem 2.4 of this paper), which is, in some sense, a generalization of the ASCLT to some kind of Borel sets A such that ∂A is not necessarily of Lebesgue measure 0. We deduce the ASCLT as a corollary of Theorem 2.4 (Corollary 2.5).

Our method of proof is based on a new bound for the Rosenblatt coefficient of normalized partial sums (Theorem 2.2), which is of its own interest and generalizes an analogous result previously obtained in [9] for i.i.d. sequences $(X_n)_{n\geq 1}$.

The paper is organized as follows: Section 2 contains the statements of the main results (Theorems 2.2 and 2.4 and Corollary 2.5); Section 3 contains some examples (in particular Proposition 3.9 shows that Rio's notion of weak dependence has some natural property of stability); in Section 4 we give the proof of Theorem 2.2 and in Section 5 the proof of Theorem 2.4 and Corollary 2.5.

Throughout the whole paper, the symbol C denotes a constant, which may not have the same value in all cases.

2. The main results

Some preliminary definitions and notations are needed.

Let X, Y be two random variables. The *Rosenblatt coefficient* of dependence of X and Y is defined as

$$\sup_{A,x} \big| P(X \in A,\ Y \leq x) - P(X \in A)P(Y \leq x) \big|,$$

where the sup is taken for $A \in \mathcal{B}(\mathbb{R})$ and $x \in \mathbb{R}$.

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $(X_i)_{i\in\mathbb{Z}}$ a sequence of real random variables defined on $(\Omega, \mathcal{F}, P)$. We shall put $S_n = X_1 + X_2 + \cdots + X_n$, $v_n = \mathrm{Var}(S_n)$ and

$$U_n = \frac{S_n}{\sqrt{v_n}}.$$

Let $p \geq 1$ be a fixed integer ; let $(Y_i)_{i\in\mathbb{Z}}$ an independent copy of $(X_i)_{i\in\mathbb{Z}}$, and put

$$V_n = \begin{cases} \dfrac{Y_1 + \ldots Y_n}{\sqrt{v_n}} & \text{for } n \leq p \\ \dfrac{Y_1 + \ldots Y_p + X_{p+1} + \ldots X_n}{\sqrt{v_n}} & \text{for } n > p. \end{cases}$$

For every integer n put

$$\Pi_n = \sup_{x\in\mathbb{R}} |P(U_n \leq x) - \Phi(x)|; \qquad \Pi'_n = \sup_{x\in\mathbb{R}} |P(V_n \leq x) - \Phi(x)|,$$

where Φ is the distribution function of the standard normal law.

Remark 2.1. Notice that Π'_n needs not be small for large n, since the denominator $\sqrt{v_n}$ in the definition of V_n is not the right normalization in general; anyway, in the particular cases that we shall study, it does happen that it goes to 0 as $n \to \infty$ (see Section 3).

The first result proved in this paper concerns the Rosenblatt coefficient of dependence of U_p and U_q, for $p < q$:

Theorem 2.2. *Let $(X_i)_{i\in\mathbb{Z}}$ be a sequence of centered random variables. Then there exists a constant C, depending on the sequence $(X_i)_{i\in\mathbb{Z}}$ only, such that, for every pair of integers p, q with $p < q$, the following bound holds*

$$\sup_{A,x} \big| P(U_p \in A,\ U_q \leq x) - P(U_p \in A)P(U_q \leq x) \big| \leq C \left(\sqrt[4]{\frac{v_p}{v_q}} + \Pi_q + \Pi'_q \right)$$

(where the sup *is taken for $A \in \mathcal{B}(\mathbb{R})$ and $x \in \mathbb{R}$)*

Theorem 2.2 can be used to prove the second main result of this paper (Theorem 2.4 here below).

For a fixed Borel set $A \subseteq \mathbb{R}$, consider the two sequences (T_n) and (W_n) defined respectively as

$$T_n = \frac{\sum_{i=1}^n 1_A(U_{2^i})}{n}; \quad W_n = \frac{\sum_{i=1}^n \frac{1}{i} 1_A(U_i)}{\log n}, \quad n \geq 1.$$

For every positive constant α put

$$\phi_\alpha(n) = \frac{v_n}{n^\alpha}. \tag{2.3}$$

Theorem 2.4. *In addition to the hypotheses of Theorem 2.2, assume that there exist two constants $\gamma > 0$ and $\delta \geq 0$ such that*

$$\max\{\Pi_n, \Pi'_n\} \leq C \frac{\log^\delta n}{n^\gamma},$$

where C depends on the sequence $(X_i)_{i \in \mathbb{Z}}$ only. Assume moreover that there exists $\alpha > 0$ such that the function ϕ_α is either non-decreasing or ultimately bounded (i.e. there exist two constants M and N such that $0 < M \leq \phi_\alpha(n) < N$, for sufficiently large n). Let $A \subseteq \mathbb{R}$ be a finite union of intervals. Then, P-a.s. the two sequences $(T_n)_{n \geq 1}$ and $(W_n)_{n \geq 1}$ have the same limit points as $n \to \infty$.

Recall that we denote by λ the Lebesgue measure on $\mathbb{R}$ and by μ the standard Gaussian measure on $\mathbb{R}$, i.e.

$$\mu(A) = \int_A \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, \lambda(dx), \qquad A \in \mathcal{B}(\mathbb{R}).$$

Theorem 2.4 has the following consequence:

Corollary 2.5 (ASCLT)**.** *With the same assumtions as in Theorem 2.4, there exists a P-null set Γ such that, for every $\omega \in \Gamma^c$, we have*

$$\lim_{n \to \infty} \frac{\sum_{i=1}^n \frac{1}{i} 1_A(U_i)}{\log n} = \mu(A)$$

for every Borel set $A \subseteq \mathbb{R}$ such that $\lambda(\partial A) = 0$.

3. Applications to some non-stationary random sequences

Example 3.1. Given the sequence of random variables $(X_i)_{i\in\mathbb{Z}}$, consider the strong mixing coefficients of $(X_i)_{i\in\mathbb{Z}}$, defined as

$$\alpha(m) \doteq \sup_{k\geq 1}\big\{\sup\{|P(A\cap B)-P(A)P(B)|,\ A\in\mathcal{F}_k,\ B\in\mathcal{G}_{k+m}\}\big\},$$

where $\mathcal{F}_\ell$ is the σ-algebra generated by $(X_i)_{i\leq\ell}$, $\mathcal{G}_\ell$ is the σ-algebra generated by $(X_i)_{i\geq\ell}$. Assume that for some $\delta>1$

$$C_1 \doteq \sup_n \mathbf{E}[X_n^{2+\delta}] < \infty \tag{3.2}$$

and that

$$C_2 \doteq \sum_m m\big(\alpha(m)\big)^{\frac{\delta-1}{2+\delta}} < +\infty. \tag{3.3}$$

Under these assumptions, Theorem 2 of [20] states that, for every $h\in BL(\mathbb{R})$

$$\mathbf{E}[h(U_n)-h(N)] \leq C\|h\|_{BL}\frac{n}{v_n^{3/2}},$$

where N is a random variable with standard normal law and C is a constant depending on the sequence $(X_i)_{i\in\mathbb{Z}}$ only.

More precisely, the bound of Theorem 2 of [20] is, apart from a multiplying absolute constant,

$$\|h\|_{BL}\left(\frac{\sum_{i=1}^n \mathbf{E}[|X_i|^3]}{v_n^{3/2}} + \frac{n}{v_n^{3/2}}C_1^{3/(2+\delta)}C_2\right) \leq \|h\|_{BL}\big(C_3 + C_1^{3/(2+\delta)}C_2\big)\frac{n}{v_n^{3/2}},$$

where $C_3 \doteq \sup_n \mathbf{E}[|X_n|^3] < +\infty$.

Notice that the strong mixing coefficients of the sequence $\ldots, Y_0, Y_1, Y_2, \ldots, Y_p, X_{p+1}, \ldots$ are clearly not greater than those of $(X_i)_{i\in\mathbb{Z}}$; arguing as in the proof of Theorem 2 of [20], it is possible to prove that also

$$\mathbf{E}[h(V_n)-h(N)] \leq C\|h\|_{BL}\frac{n}{v_n^{3/2}},$$

for every $h\in BL(\mathbb{R})$ and where C is again a constant depending on the sequence $(X_i)_{i\in\mathbb{Z}}$ only. (The difference is in the denominators: more precisely, Theorem 2 of [20] can be directly applied to the (well normalized) sequence

$$W_n = \begin{cases} \dfrac{Y_1+\ldots Y_n}{\sqrt{v_n}} & \text{for } n\leq p \\[2ex] \dfrac{Y_1+\ldots Y_p + X_{p+1}+\ldots X_n}{\sqrt{v_p+v_{n-p}}} & \text{for } n>p; \end{cases}$$

nevertheless, the proof of Theorem 2 carries over with no changes also for V_n, since in each estimation needed in such proof, the denominator can be taken out by linearity as a common factor, hence has no influence on the final result). If in addition ϕ (defined in (2.3)) verifies one of the assumptions of Theorem 2.4 with $\alpha > 2/3$, then $\frac{n}{v_n^{3/2}} \leq C\frac{1}{n^{\frac{3\alpha-2}{2}}}$. Hence

$$\max\big\{\mathbf{E}[h(V_n)-h(N)], \mathbf{E}[h(U_n)-h(N)]\big\} \leq C\|h\|_{BL}\frac{1}{n^{\frac{3\alpha-2}{2}}},$$

and from this relation, using a similar argument as in the proof of Lemma 4.9 (see Section 4 of the present paper), it is not difficult to get the inequality

$$\max\{\Pi_n, \Pi_n'\} \leq \frac{C}{n^{\frac{3\alpha-2}{2}}}.$$

Hence, applying Corollary 2.5, we have the following result:

Theorem 3.4. *Let $X_{i\in\mathbb{Z}}$ be a sequence of random variables verifying* (3.2) *and* (3.3)*. If ϕ (defined in* (2.3)*) satisfies one of the assumptions of Theorem 2.4, then the ASCLT (i.e. Corollary 2.5) holds for $(X_i)_{i\in\mathbb{Z}}$.*

Example 3.5. Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{A}$ a sub-σ-algebra of $\mathcal{F}$ and X a random variable defined on $(\Omega, \mathcal{F}, P)$ and with values in a metric space $(\mathcal{G}, \delta)$. In [18] the following measure of dependence between X and $\mathcal{A}$ is introduced:

Definition 3.6. Put

$$\varphi(\mathcal{A}, X) = \sup_{f\in\mathcal{L}_1(\mathcal{G},\delta)} \big|\big|\mathbf{E}[f(X)|\mathcal{A}] - \mathbf{E}[f(X)]\big|\big|_\infty,$$

where $\mathcal{L}_1(\mathcal{G}, \delta)$ is the set of 1-lipshitzian functions defined on $(\mathcal{G}, \delta)$ and taking values in $[0, 1]$.

We call $\varphi(\mathcal{A}, X)$ the *uniform Rio mixing coefficient* between X and $\mathcal{A}$.

In the same paper [18] the *uniform dependence coefficients* of a sequence $(X_i)_{i\in\mathbb{Z}}$ of real-valued random variables defined on $(\Omega, \mathcal{F}, P)$ are defined as follows:

Definition 3.7. Let $\mathcal{F}_k$ be the σ-algebra generated by $(X_i)_{i\leq k}$. Put $\varphi_0 = 1$ and, for every integer $r \geq 1$,

$$\varphi_r = \sup_{\substack{k\in\mathbb{Z}\\ r\leq r_1<r_2<r_3}} \varphi\big(\mathcal{F}_k, (X_{k+r_1}, X_{k+r_2}, X_{k+r_3})\big).$$

Then $(\varphi_r)_{r\geq 0}$ is the sequence of the *uniform dependence coefficients* of $(X_i)_{i\in\mathbb{Z}}$.

Remark 3.8. The random vector $(X_{k+r_1}, X_{k+r_2}, X_{k+r_3})$ takes values in the metric space $(\mathbb{R}^3, d)$, where d is the euclidean distance. In what follows we shall write $\mathcal{L}_1(\mathbb{R}^3)$ in place of $\mathcal{L}_1(\mathbb{R}^3, d)$.

For a fixed integer p, define the sequence $(Z_i^{(p)})_{i\in\mathbb{Z}}$ as follows

$$Z_i^{(p)} = \begin{cases} Y_i & \text{for } i \le p \\ X_i & \text{for } i \ge p+1 \end{cases}$$

Denote by $(\tilde{\varphi}_r^{(p)})_{r\ge 0}$ the uniform dependence coefficients of the sequence $(Z_i^{(p)})$. The following (rather natural) result holds good:

Proposition 3.9. *For every integer $r \ge 0$ we have*

$$\sup_{p\in\mathbb{N}} \tilde{\varphi}_r^{(p)} \le \varphi_r.$$

The proof of Proposition 3.9, though rather technical, presents no difficulty, hence is omitted for the sake of brevity. The interested reader can find it on the author's home page, at the address

`http://www.dm.unipi.it/~giuliano/`

Using Theorem 2 of [18] and Proposition 3.9 the following result can be proved. Notice that again no stationariness assumption is needed.

Theorem 3.10. *Let $(X_i)_{i\in\mathbb{Z}}$ be a sequence of real centered random variables, bounded by a constant M.*
Denote $w_n = v_n - v_{n-1}$, and assume that there exists an integer n_0 such that, for $n \ge n_0$, we have

$$w_n \ge \frac{1}{2}. \tag{3.11}$$

Assume that the sequence $(\varphi_r)_{r\ge 0}$ of the uniform dependence coefficients satisfies the condition

$$\sum_{r\ge 0} r\varphi_r < \infty. \tag{3.12}$$

Then, for every integer n,

$$\max\{\Pi_n, \Pi'_n\} \le \frac{C}{\sqrt{n}},$$

where C is a constant depending on the sequence $(X_i)_{i\in\mathbb{Z}}$ only.

Hence, if the function ϕ defined in (2.3) verifies one of the assumptions of (2.4), the ASCLT is in force.

4. The proof of Theorem 2.2

We start by collecting some preliminary results and lemmas.

Definition 4.1. The *concentration function* of a r.v. S is defined as

$$Q(\epsilon) = \sup_{x\in\mathbb{R}} P(x \le S \le x+\epsilon), \quad \epsilon \in \mathbb{R}^+.$$

In the sequel we denote by Q_n the concentration function of U_n.

The following result gives an estimate of Q_n. Its form is similar to the one given in [16] for a sequence of i.i.d. random variables.

Lemma 4.2. *For every* $\epsilon \in \mathbb{R}^+$,

$$Q_n(\epsilon) \le C\big(\epsilon + \Pi_n\big),$$

where C *is an absolute constant.*

Proof. Denote by F_n the distribution function of U_n and put $F_n(x^-) = \lim_{t\uparrow x^-} F_n(t)$. Then

$$\sup_{x\in\mathbb{R}} \big|F_n(x^-) - \Phi(x)\big| = \sup_{x\in\mathbb{R}} \lim_{t\uparrow x^-} \big|F_n(t) - \Phi(t)\big| \le \Pi_n,$$

so that

$$\max\big\{|F_n(x+\epsilon) - \Phi(x+\epsilon)|, |F_n(x^-) - \Phi(x)|\big\} \le \Pi_n.$$

Hence

$$\begin{aligned} P(x \le U_n \le x+\epsilon) &= F_n(x+\epsilon) - F_n(x^-) \le |F_n(x+\epsilon) - \Phi(x+\epsilon)| \\ &+ |F_n(x^-) - \Phi(x)| + \Phi(x+\epsilon) - \Phi(x) \le 2\Pi_n + \frac{1}{\sqrt{2\pi}}\,\epsilon \le C\,(\epsilon + \Pi_n). \quad \square \end{aligned}$$

The following lemma is stated in [2] without proof:

Lemma 4.3. *If* S *and* T *are random variables, then, for every pair of real numbers* a, b *with* $b \ge 0$ *we have*

$$P(S+T \le a-b) - P(|T| > b) \le P(S \le a) \le P(S+T \le a+b) + P(|T| > b).$$

Proof. The first inequality follows from the inclusion

$$\{S+T \le a-b\} \subseteq \{S \le a\} \cup \{|T| > b\}.$$

The second inequality follows from the first one applied to the pair of random variables $S+T, -T$ and to the pair of numbers $a+b, b$. $\square$

We now begin the proof of Theorem 2.2. Let p, q be two integers with $p \leq q$; let $(Y_i)_{i\in\mathbb{Z}}$ an independent copy of $(X_i)_{i\in\mathbb{Z}}$, and put

$$V_q = \frac{Y_1 + \ldots Y_p + X_{p+1} + \ldots X_q}{\sqrt{v_q}}.$$

Put moreover

$$Z = V_q - U_q = \frac{(Y_1 - X_1) + \cdots + (Y_p - X_p)}{\sqrt{v_q}} = \frac{R_p}{\sqrt{v_q}}.$$

If we set

$$H = \{U_p \in A\}, \quad K = \{U_q \leq x\},$$

our aim is to give a bound for $|P(H \cap K) - P(H)P(K)|$.

Let $\epsilon > 0$ be any positive real number and put

$$K_1 = \{V_q \leq x - \epsilon\}, \quad K_2 = \{V_q \leq x + \epsilon\}, \quad F = \{|Z| > \epsilon\}.$$

By Lemma 4.3 (applied to $S = U_q$, $T = Z$, $a = x$, $b = \epsilon$) we can write

$$P(K_1) - P(F) \leq P(K) \leq P(K_2) + P(F).$$

Hence

$$\begin{aligned}
&|P(H \cap K) - P(H)P(K)| \leq \max\{|P(H \cap K) - P(K_1)P(H) + P(F)P(H)|,\\
&|P(H \cap K) - P(K_2)P(H) - P(F)P(H)|\}\\
&\quad \leq \max\{|P(H \cap K) - P(K_1)P(H)|, |P(H \cap K) - P(K_2)P(H)|\} + P(F). \qquad (4.4)
\end{aligned}$$

In what follows we shall estimate the three quantities in the last member, i.e. $|P(H \cap K) - P(K_1)P(H)|$, $|P(H \cap K) - P(K_2)P(H)|$ and $P(F)$. We start with $P(F)$. We have

$$P(F) = P(|R_p| > \epsilon\sqrt{v_q}) \leq \frac{\mathbf{E}\big[|R_p|\big]}{\epsilon\sqrt{v_q}} \leq \frac{\mathrm{Var}^{1/2}(R_p)}{\epsilon\sqrt{v_q}}. \qquad (4.5)$$

Now, since $(X_i)_{i\in\mathbb{Z}}$ and $(Y_i)_{i\in\mathbb{Z}}$ are independent and have the same law,

$$\begin{aligned}
\mathrm{Var}(R_p) &= \sum_{i=1}^{p} \mathrm{Var}(Y_i - X_i) + 2 \sum_{1\leq i<j\leq p} \mathrm{Cov}((Y_i - X_i), (Y_j - X_j))\\
&= 2\sum_{i=1}^{p} \mathrm{Var}(X_i) + 4 \sum_{1\leq i<j\leq p} \mathrm{Cov}(X_i, X_j) = 2\mathrm{Var}(S_p) = 2v_p. \qquad (4.6)
\end{aligned}$$

From (4.5) and the (4.6) we conclude that

$$P(F) \le \frac{\sqrt{2}}{\epsilon} \sqrt{\frac{v_p}{v_q}}. \tag{4.7}$$

We now pass to the terms $|P(H \cap K) - P(K_1)P(H)|$ and $|P(H \cap K) - P(K_2)P(H)|$. We give the details only for $|P(H \cap K) - P(K_2)P(H)|$, since the proof is almost identical for the other quantity (see Remark 4.15).

We need some more lemmas.

Lemma 4.8. *Let g be a Lipschitzian function defined on $\mathbb{R}$, with Lipschitz constant β. Then*

$$\big|\mathbf{E}[g(U_q)] - \mathbf{E}[g(V_q)]\big| \le \sqrt{2}\,\beta \sqrt{\frac{v_p}{v_q}}.$$

Proof. Arguing as for relation (4.5) and using (4.6) we get

$$\begin{aligned}|\mathbf{E}[g(U_q)] - \mathbf{E}[g(V_q)]| &\le \mathbf{E}[|g(U_q) - g(V_q)|] \le \beta\,\mathbf{E}[|U_q - V_q|] \\ &= \beta\,\frac{\mathbf{E}[|R_p|]}{\sqrt{v_q}} \le \frac{\beta\,\mathrm{Var}^{1/2}(R_p)}{\sqrt{v_q}} \le \sqrt{2}\,\beta \sqrt{\frac{v_p}{v_q}}.\end{aligned}$$

□

In the sequel we denote by $\tilde{Q}_q$ the concentration function of V_q.

Lemma 4.9. *Let $z \in \mathbb{R}$ and $g = 1_{(-\infty,z]}$. Then, for every $\delta > 0$ we have*

$$\big|\mathbf{E}[g(U_q)] - \mathbf{E}[g(V_q)]\big| \le \frac{C}{\delta}\sqrt{\frac{v_p}{v_q}} + Q_q(\delta) + \tilde{Q}_q(\delta).$$

Proof. Put

$$h(t) = \left(1 + \frac{z-t}{\delta}\right) 1_{(z,z+\delta]}(t), \qquad \tilde{g}(t) = g(t) + h(t).$$

Then $\tilde{g}$ is Lipschitzian with Lipschitz constant $1/\delta$, so that, by Lemma 4.8,

$$\big|\mathbf{E}[\tilde{g}(U_q)] - \mathbf{E}[\tilde{g}(V_q)]\big| \le \frac{C}{\delta}\sqrt{\frac{v_p}{v_q}}. \tag{4.10}$$

On the other hand, h has support contained in $(z, z+\delta]$ and is bounded by 1, hence we have trivially

$$\big|\mathbf{E}[h(U_q) - h(V_q)]\big| \le Q_q(\delta) + \tilde{Q}_q(\delta). \tag{4.11}$$

Now, recalling that $g = \tilde{g} - h$, we can write

$$\begin{aligned}\big|\mathbf{E}[g(U_q)] - \mathbf{E}[g(V_q)]\big| &= \big|\mathbf{E}[(\tilde{g} - h)(U_q)] - \mathbf{E}[(\tilde{g} - h)(V_q)]\big| \\ &\le \big|\mathbf{E}[\tilde{g}(U_q)] - \mathbf{E}[\tilde{g}(V_q)]\big| + \big|\mathbf{E}[h(U_q) - h(V_q)]\big|,\end{aligned}$$

and the conclusion follows from relations (4.10) and (4.11). □

The following lemma concerns the concentration function $\tilde{Q}_q$ of V_q. It can be proved exactly as Lemma 4.2.

Lemma 4.12. *There is an absolute constant C such that for every $\epsilon \in \mathbb{R}^+$*

$$\tilde{Q}_n(\epsilon) \le C\big(\epsilon + \Pi_n'\big).$$

We go back to the proof of the main result (2.2). Since H and K_2 are independent we can write

$$\begin{aligned}|P(H\cap K) - P(K_2)P(H)| &= P(H)\big|P(K|H) - P(K_2|H)\big| \\ &= P(H)\big|\mathbf{E}_H[f(U_q)] - \mathbf{E}_H[g(V_q)]\big|,\end{aligned}$$

where $f = 1_{(-\infty,x]}$ and $g = 1_{(-\infty,x+\epsilon]}$. We denote by $\mathbf{E}_H$ the expectation with respect to the probability law $P(\cdot|H)$. By summing and subtracting $\mathbf{E}_H[g(U_q)]$, we see that the above quantity is not greater than

$$\begin{aligned}P(H)\big|\mathbf{E}_H[g(U_q)] - \mathbf{E}_H[g(V_q)]\big| + P(H)\mathbf{E}_H[|f-g|(U_q)] \\ = |\mathbf{E}[g(U_q)] - \mathbf{E}[g(V_q)]\big| + \mathbf{E}[|f-g|(U_q)] \le \frac{C}{\epsilon}\sqrt{\frac{v_p}{v_q}} + 2\,Q_q(\epsilon) + \tilde{Q}_q(\epsilon),\end{aligned} \tag{4.13}$$

using Lemma 4.9 and observing that the function $f - g$ is bounded by 1 and has the interval $(x, x+\epsilon]$ as its support.

Remark 4.14. The above proof, and in particular relation (4.13) explains why Π_q and Π_q' appear in the formulas (despite the fact that replacing the first p of the X_i's with the corresponding Y_i's reduces the amount of dependence of the sequence, as Proposition 3.9 shows).

Remark 4.15. The proof for for $|P(H\cap K) - P(K_1)P(H)|$ is identical to the proof for $|P(H\cap K) - P(K_2)P(H)|$, the only difference is in the fact one needs to take $g = 1_{(-\infty,x-\epsilon]}$ and add and subtract $\mathbf{E}_H[f(V_q)]$. This gives the inequality

$$|P(H\cap K) - P(K_1)P(H)| \le \frac{C}{\epsilon}\sqrt{\frac{v_p}{v_q}} + Q_q(\epsilon) + 2\,\tilde{Q}_q(\epsilon), \tag{4.16}$$

instead of (4.13).

We now insert relations (4.7), (4.13) and (4.16) into (4.4), and obtain

$$|P(H\cap K) - P(H)P(K)| \le \frac{C}{\epsilon}\sqrt{\frac{v_p}{v_q}} + 2Q_q(\epsilon) + 2\tilde{Q}_q(\epsilon) \le C\left(\frac{1}{\epsilon}\sqrt{\frac{v_p}{v_q}} + \epsilon + \Pi_q + \Pi_q'\right),$$

by Lemmas 4.2 and 4.12. The above relation holds for every $\epsilon > 0$; by passing to the infimum in ϵ, we get

$$|P(H\cap K) - P(H)P(K)| \le C\left(\sqrt[4]{\frac{v_p}{v_q}} + \Pi_q + \Pi_q'\right). \qquad \square$$

5. The proof of Theorem 2.4 and of the ASCLT

Let's start with the proof of Theorem 2.4. It is sufficient to consider the case in which A is of the form $A = (-\infty, x]$. The proof is split in two steps: (i) here below and (ii) on p. 284.

Put

$$a_n = \log_2\left(1 + \frac{1}{n}\right). \tag{5.1}$$

(i) Here we prove that (T_n) and (H_n) have the same limit points, where

$$H_n = \frac{\sum_{i=1}^{2^n} a_i 1_A(U_i)}{n}.$$

This is equivalent to proving that the sequence

$$T_n - H_n + \frac{a_{2^n} 1_A(U_{2^n})}{n} = \frac{\sum_{i=1}^{n} 1_A(U_{2^i}) - \sum_{i=1}^{2^n-1} a_i 1_A(U_i)}{n}$$

tends to 0 as $n \to \infty$, P-a.s. The numerator of the fraction in the second member above can be written as

$$\begin{aligned}\sum_{i=1}^{n} 1_A(U_{2^i}) - \sum_{i=1}^{n} \sum_{j=2^{i-1}}^{2^i-1} a_j 1_A(U_j) &= \sum_{i=1}^{n} \left(1_A(U_{2^i}) - \sum_{j=2^{i-1}}^{2^i-1} a_j 1_A(U_j)\right) \\ &= \sum_{i=1}^{n} \sum_{j=2^{i-1}}^{2^i-1} a_j \big(1_A(U_{2^i}) - 1_A(U_j)\big),\end{aligned}$$

(note that $\sum_{j=2^{i-1}}^{2^i-1} a_j = \log_2(2^i) - \log_2(2^{i-1}) = 1$). Put now

$$R_i = \sum_{j=2^{i-1}}^{2^i-1} a_j \big(1_A(U_{2^i}) - 1_A(U_j)\big). \tag{5.2}$$

Then we must prove that, P-a.s.

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} R_i}{n} = 0.$$

We write

$$\frac{\sum_{i=1}^{n} R_i}{n} = \frac{\sum_{i=1}^{n} \big(R_i - \mathbf{E}[R_i]\big)}{n} + \frac{\sum_{i=1}^{n} \mathbf{E}[R_i]}{n} = \frac{\sum_{i=1}^{n} \tilde{R}_i}{n} + \frac{\sum_{i=1}^{n} \mathbf{E}[R_i]}{n},$$

and we shall consider separately the two summands above.

For the first one we shall apply the Gaal–Koksma Law (see [17], p. 134) to the sequence $(\tilde{R}_n)_{n\geq 1}$:

Theorem 5.3 (Gaal–Koksma Strong Law of Large Numbers). *Let $(X_n)_{n\geq 1}$ be a sequence of centered random variables with finite variance. Suppose that there exists a constant $\alpha > 0$ such that, for all integers $m \geq 0$, $n \geq 0$*

$$\mathbf{E}\Big[\Big(\sum_{i=m+1}^{m+n} X_i\Big)^2\Big] \leq C\big((m+n)^\alpha - m^\alpha\big), \tag{5.4}$$

for a suitable constant C independent on m and n. Then, for each $\delta > 0$,

$$\sum_{i=1}^{n} X_i = O\big(n^{\alpha/2}(\log n)^{2+\delta}\big), \quad P\text{-a.s.}$$

Remark 5.5. The condition $m \geq 0$ can be replaced by $m \geq m_0$ (for a suitable m_0).

We need a bound for $\mathbf{Cov}(\tilde{R}_i, \tilde{R}_j)$. It is easily seen that, for $i \leq j$,

$$\mathrm{Cov}(\tilde{R}_i, \tilde{R}_j) = \sum_{h=2^{i-1}}^{2^i-1} \sum_{k=2^{j-1}}^{2^j-1} a_h a_k \big(\mathcal{C}(2^i, 2^j) - \mathcal{C}(h, 2^j) - \mathcal{C}(2^i, k) + \mathcal{C}(h, k)\big),$$

where

$$\mathcal{C}(p,q) = \mathrm{Cov}(1_A(U_p), 1_A(U_q)) = P(U_p \in A,\ U_q \in A) - P(U_p \in A)P(U_q \in A).$$

By Theorem (2.1) there exists a constant C such that, for sufficiently large j and for every p, q with $2^{i-1} \leq p \leq 2^i$ and $2^{j-1} \leq q \leq 2^j$, we have

$$C(p,q) \leq C\left(\sqrt[4]{\frac{v_p}{v_q}} + \frac{\log^\delta q}{q^\gamma}\right) \leq C\left(\frac{p}{q}\right)^\eta \leq C\, 2^{-\eta|i-j|},$$

where $\eta = (\gamma/2) \wedge (\alpha/4)$. We obtain, for large j,

$$\mathrm{Cov}(\tilde{R}_i, \tilde{R}_j) \leq C 2^{-\eta|i-j|} \sum_{h=2^{i-1}}^{2^i-1} a_h \sum_{k=2^{j-1}}^{2^j-1} a_k = C 2^{-\eta|i-j|}.$$

In particular $\mathbf{E}[\tilde{R}_i^2] \leq C$. In order to use the Gaal–Koksma Law, we evaluate, for large m,

$$\begin{aligned}
\mathbf{E}\Big[\Big(\sum_{i=m+1}^{m+n} \tilde{R}_i\Big)^2\Big] &= \mathbf{E}\Big[\sum_{i=m+1}^{m+n} \tilde{R}_i^2 + 2 \sum_{m+1 \leq i < j \leq m+n} \tilde{R}_i \tilde{R}_j\Big] \\
&\leq Cn + 2C \sum_{m+1 \leq i < j \leq m+n} 2^{-\eta|i-j|} = Cn + 2C \sum_{r=1}^{n-1} (n-r) 2^{-r\eta} \\
&\leq Cn + 2Cn \sum_{r=0}^{n-1} 2^{-r\eta} \leq Cn = C\big[(m+n) - m\big].
\end{aligned}$$

Hence the condition in the Gaal–Koksma Law holds with $\alpha = 1$ and we obtain

$$\sum_{i=1}^{n} \tilde{R}_i = O\big(\sqrt{n}(\log n)^{2+\delta}\big), \quad P\text{-a.s.},$$

which implies

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} \tilde{R}_i}{n} = 0, \quad P\text{-a.s.}$$

We now prove that

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} \mathbf{E}[R_i]}{n} = 0.$$

By Cesaro's Theorem, it will be sufficient to prove that

$$\lim_{n\to\infty} \mathbf{E}[R_n] = \lim_{n\to\infty} \sum_{j=2^{n-1}}^{2^n-1} a_j\big(P(U_{2^n} \in A) - P(U_j \in A)\big) = 0$$

(recall formula (5.2)). This is immediate by the relation $\sum_{j=2^{i-1}}^{2^i-1} a_j = \log_2(2^i) - \log_2(2^{i-1}) = 1$ and by the assumption on Π_n, which implies

$$\lim_{n\to\infty} P(U_n \in A) = \mu(A).$$

(ii) We now prove that (H_n) and (W_n) have the same limit points. First, observe that

$$W_n = \frac{\sum_{i=1}^{n} \frac{1}{i \log 2} 1_A(U_i)}{\log_2 n}.$$

Since the sequences (a_n) (see definition (5.1)) and (b_n), where $b_n = \frac{1}{n \log 2}$, are equivalent as $n \to \infty$, this amounts to showing that (H_n) has the same limit points as

$$V_n = \frac{\sum_{i=1}^{n} a_i 1_A(U_i)}{\log_2 n}.$$

This is easy since, for $2^r \le n < 2^{r+1}$ we can write

$$\frac{\sum_{i=1}^{2^r} a_i 1_A(U_i)}{r+1} \le V_n \le \frac{\sum_{i=1}^{2^{r+1}} a_i 1_A(U_i)}{r}. \qquad \square$$

We pass to the proof of the ASCLT (Corollary 2.5). Consider first a Borel set A of the form $A = (-\infty, x]$. The Gaal–Koksma–Law applied to the sequence

$$1_A(U_{2^i}) - P(U_{2^i} \in A)$$

gives, P-a.s,

$$\lim_{n\to\infty}\left(T_n - \frac{\sum_{i=1}^n P(U_{2^i}\in A)}{n}\right) = \lim_{n\to\infty}\frac{\sum_{i=1}^n \big(1_A(U_{2^i}) - P(U_{2^i}\in A)\big)}{n} = 0.$$

by an argument similar to that used above for the sequence $(\tilde{R}_n)_{n\geq 1}$ (see p. 13 for the definition of $(\tilde{R}_n)_{n\geq 1}$). On the other hand, again by Cesaro's Theorem and the assumption on Π_n,

$$\lim_{n\to\infty}\frac{\sum_{i=1}^n P(U_{2^i}\in A)}{n} = \lim_{n\to\infty} P(U_{2^n}\in A) = \mu(A).$$

Hence we get

$$\lim_{n\to\infty} T_n = \mu(A), \quad P\text{-a.s.} \tag{5.6}$$

Now, classical techniques (similar to those used in the Glivenko–Cantelli Theorem, see for instance [6], p. 59) yield that the P-null set Γ such that (5.6) holds for $\omega\in\Gamma^c$ is independent on A, and it is henceforth immediate that, on Γ^c, (5.6) holds also for Borel sets A that are finite unions of disjoint intervals.

For a general set A with $\lambda(\delta A)=\mu(\delta A)=0$, fix $\epsilon>0$ and let A_ϵ and B_ϵ be finite unions of disjoint intervals such that

$$A_\epsilon\subseteq A\subseteq B_\epsilon \quad\text{and}\quad \mu(B_\epsilon\setminus A_\epsilon)<\epsilon.$$

Then

$$\frac{\sum_{i=1}^n 1_{A_\epsilon}(U_{2^i})}{n}\leq T_n\leq\frac{\sum_{i=1}^n 1_{B_\epsilon}(U_{2^i})}{n};$$

hence, by passing to the limit as $n\to\infty$, we get, for $\omega\in\Gamma^c$,

$$\mu(A_\epsilon)\leq\liminf_{n\to\infty} T_n(\omega)\leq\limsup_{n\to\infty} T_n(\omega)\leq\mu(B_\epsilon); \tag{5.7}$$

since

$$\mu(A_\epsilon)\leq\mu(A)\leq\mu(B_\epsilon)\leq\mu(A_\epsilon)+\epsilon \tag{5.8}$$

by passing to the limit as $\epsilon\to 0$ in (5.8) and after in (5.7) we deduce that $\lim_{n\to\infty} T_n(\omega)$ exists for $\omega\in\Gamma^c$ and moreover

$$\lim_{n\to\infty} T_n(\omega)=\mu(A), \qquad \omega\in\Gamma^c. \qquad \square$$

Acknowledgements. The author wishes to thank two anonymous referees for helpful comments and improvements.

References

[1] I. Berkes and E. Csaki, A universal result in almost sure central limit theory, *Stoch. Proc. Appl.* **94**(1), 2001, 105–134.

[2] J. R. Blum, D. L. Hanson and J. I. Rosenblatt, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, Vol. 1, 1963, 389–393.

[3] G. A. Brosamler, An almost everywhere central limit theorem, *Math. Proc. Cambridge Philos. Soc.* **104**, no. 3, 1988, 561–574.

[4] J.-R. Chazottes and S.Gouzel, On almost-sure versions of classical limit theorems for dynamical systems, *Probab. Theory Related Fields* **138**, no. 1–2 (2007), 195–234.

[5] M. Dudzinski, A note on the almost sure central limit theorem for some dependent random variables, *Statist. Probab. Lett.* **61**, 2003, 31–40.

[6] R. Durrett, Probability: Theory and Examples, *Thomson*, 2005.

[7] I. Fazekas and Z. Rychlik, Almost sure functional limit theorems, *Ann. Univ. Mariae Curie-Sklodowska Sect. A* **56**, 2002, 1–18.

[8] I. Fazekas and Z. Rychlik, Almost sure central limit theorems for random fields, *Math. Nachr.* **259** , 1, 2003, 12–18.

[9] R. Giuliano Antonini, On the Rosenblatt coefficient for normalized sums of real random variables, *Rend. Accad. Naz. Sci. XL Mem. Mat. Appl.* (5) **24**, 2000, 111–120.

[10] R. Giuliano Antonini, The Rosenblatt coefficient of dependence for m-dependent random sequences with applications to the ASCLT, *Theory of Stochastic Processes* **14**, no. 1 (2008), 30–38.

[11] K. Joag-Dev and F. Proschan, Negative association of random variables, with applications, *Ann. Statist.* **11**, no. 1, 1983, 286–295.

[12] M. Lacey and W. Philipp, A note on the almost sure central limit theorem, *Statist. Probab. Lett.* **9**, 1990, 201–205.

[13] P. Major, Almost sure functional limit theorems, Part I. The general case, *Studia Scient. Math. Hung.* **34**, 1998, 273–304.

[14] P. Matula, On the almost sure central limit theorem for associated random variables, *Probab. Math. Statist.* **18**, no. 2, 1998, 411–416.

[15] M. Peligrad and Q. Shao, A note on the almost sure central limit theorem for weakly dependent random variables, *Statist. Probab. Lett.* **22**, 1995, 131–136.

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

[17] W. Philipp and W. Stout, Almost sure invariance principles for partial sums of weakly dependent random variables, *Mem. Amer. Math. Soc.* **161**, 1975, 389–393.

[18] E. Rio, Sur le théorème de Berry–Esseen pour les suites faiblement dépendantes, *Probab. Theory Relat. Fields* **104**, 1996, 255–282.

[19] P. Schatte, On strong versions of the almost sure central limit theorem, *Math. Nachr.* **137**, 1988, 249–256.

[20] J. Sunklodas, On normal approximation for strongly mixing random variables, *Acta Appl. Math.* **97**, 2007, 251–260.

RITA GIULIANO
DIPARTIMENTO DI MATEMATICA "L. TONELLI"
LARGO B. PONTECORVO, 5
56100 PISA
ITALY

E-mail: giuliano@dm.unipi.it
URL: http://www.dm.unipi.it/~giuliano/

(Received June 22, 2009; revised March 25, 2010)