ANNALES DE L'I. H. P., SECTION B

SANA LOUHICHI

Weak convergence for empirical processes of associated sequences

Annales de l'I. H. P., section B, tome 36, nº 5 (2000), p. 547-567
<http://www.numdam.org/item?id=AIHPB_2000__36_5_547_0>

Ann. Inst. Henri Poincaré, Probabilités et Statistiques 36, 5 (2000) 547–567

Weak convergence for empirical processes of associated sequences

by

Sana LOUHICHI [1]
UMR C 8628 CNRS, Université de Paris-Sud, Bât. 425, Mathématiques, F-91405 Orsay Cedex, France

Manuscript received in 29 October 1998, revised 22 March 2000

ABSTRACT. – We establish a weak convergence theorem for empirical processes of stationary and associated random variables having the uniform marginal distribution. To carry out the proof, we develop a tightness criterion for the empirical process constructed from any stationary sequence fulfilling a suitable moment inequality. We apply the result to stationary non mixing moving average sequences with positive coefficients. Based on this class of linear processes, we compare mixing and association. © 2000 Éditions scientifiques et médicales Elsevier SAS

Key words: Weak convergence, Empirical processes, Bracketing number, Association, Mixing, Linear processes

AMS classification: 60E15, 60F05, 60F17, 60G05

RÉSUMÉ. – Nous montrons le TLC empirique pour des suites strictement stationnaires, associées et distribuées suivant la loi uniforme. Nous donnons, en particulier, un critère assurant la tension du processus empirique construit à partir d'une suite stationnaire vérifiant une inégalité de moments convenable. Nous appliquons notre résultat aux processus linéaires non mélangeant à coefficients positifs. En se basant sur cette

[1] E-mail: Sana.Louhichi@math.u-psud.fr.

classe de processus linéaires, nous comparons le mélange et l'association.

1. INTRODUCTION, NOTATIONS AND PREVIOUS RESULTS

Let $(X_n)_{n\in\mathbb{Z}}$ be a stationary sequence of random variables (r.v's) on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let F be the common distribution function of $(X_n)_{n\in\mathbb{Z}}$. The empirical distribution function F_n of $X_1, \dots, X_n$ is defined as:

$$F_n(x) := F_n(x, \omega) = \frac{1}{n} \sum_{1\leqslant i\leqslant n} \mathbf{1}_{X_i(\omega)\leqslant x}, \quad x \in \mathbb{R}.$$

The empirical process G_n based on the observations $X_1, \dots, X_n$ is defined by:

$$G_n(x) := G_n(x, \omega) = \sqrt{n}\big[F_n(x, \omega) - F(x)\big]. \tag{1}$$

Let $D[-\infty, +\infty]$ be the space of *cadlag* functions on $[-\infty, +\infty]$ having finite limits at $\pm\infty$. Suppose that $D[-\infty, +\infty]$ is equipped with the Skorohod topology. The usual Empirical Central Limit Theorem (ECLT) gives conditions under which the empirical process $\{G_n(x),\ x \in \mathbb{R}\}$ converges in distribution, as a random element of $D[-\infty, +\infty]$, to a Gaussian process G with zero mean and covariance

$$\operatorname{Cov}\big(G(x), G(y)\big) = \sum_{k\in\mathbb{Z}} \operatorname{Cov}(\mathbf{1}_{X_0\leqslant x}, \mathbf{1}_{X_k\leqslant y}). \tag{2}$$

The proof of such theorem requires two steps:

Step 1. Establish the convergence of finite-dimensional distributions.

Step 2. Establish the tightness property.

In general, it remains to prove step 2 since step 1 follows from a suitable central limit theorem, usually well known.

For the sake of simplicity, we suppose in the sequel that the marginal distribution function F is *continuous* on $\mathbb{R}$. This restriction allows to suppose that the marginal law is $\mathcal{U}([0, 1])$: the uniform law over $[0, 1]$ (cf. Billingsley [2]).

The purpose of this paper is to study the convergence of the empirical process for associated sequences. Let us recall that $(X_n)_{n\in\mathbb{Z}}$ is a sequence of *associated* r.v's if for every finite subcollection $X_{i_1},\ldots,X_{i_n}$ and every pair of coordinatewise non decreasing functions $h,k:\mathbb{R}^n\to\mathbb{R}$

$$\mathrm{Cov}\big(h(X_{i_1},\ldots,X_{i_n}),k(X_{i_1},\ldots,X_{i_n})\big)\geqslant 0,$$

whenever the covariance is defined. This definition was introduced by Esary et al. [7], mainly for the sake of applications in reliability and statistics.

Let us briefly recall what is known about this problem.

Assume that X_0 has a *bounded density* f. Up to now, the fi-di convergence of G_n was proved only under the summability condition $\sum_{n>0}\mathrm{Cov}^{1/3}(X_1,X_n)<\infty$ (cf. Yu [15]). Hence the condition $\mathrm{Cov}(X_1,X_n)=\mathcal{O}(n^{-3})$ seems to be necessary for the empirical central limit theorem.

Yu [15] was the first to prove an ECLT for stationary associated sequences. He supposed that

$$\mathrm{Cov}\big(F(X_1),F(X_n)\big)=\mathcal{O}\big(n^{-b}\big),\quad \text{for some } b>15/2. \tag{3}$$

Next Shao and Yu [13] weakened condition (3). Their condition ensuring the ECLT was:

$$\mathrm{Cov}\big(F(X_1),F(X_n)\big)=\mathcal{O}\big(n^{-b}\big),\quad \text{for some } b>\big(3+\sqrt{33}\big)/2. \tag{4}$$

Since X_0 has a bounded density f, the convergence of G_n holds under a condition on the covariance of the original r.v's:

$$\mathrm{Cov}(X_1,X_n)=\mathcal{O}\big(n^{-b}\big),\quad \text{for some } b>\big(3+\sqrt{33}\big)/2; \tag{5}$$

as it can be seen by writing

$$\begin{aligned}\mathrm{Cov}\big(F(X_1),F(X_n)\big)&=\iint f(s)f(y)\mathrm{Cov}(\mathbf{1}_{X_1\leqslant x},\mathbf{1}_{X_n\leqslant y})\,\mathrm{d}x\,\mathrm{d}y\\&\leqslant \|f\|_\infty^2\mathrm{Cov}(X_1,X_n).\end{aligned} \tag{6}$$

The convergence of the empirical process, G_n, is also known to hold for sequences $(X_n)_{n\in\mathbb{Z}}$ satisfying some conditions of weak dependence called mixing. Let us recall the mixing coefficients and the mixing assumptions yielding the convergence of the empirical process $\{G_n(x),\ x\in[0,1]\}$.

As a measure of dependence, Volkonskii and Rozanov [14] introduced the β-mixing coefficients, defined for any two σ-algebras $\mathcal{A}$ and $\mathcal{B}$ by

$$\beta(\mathcal{A},\mathcal{B}) = \frac{1}{2}\sup\sum_i\sum_j\left|\mathrm{Cov}(\mathbf{1}_{A_i},\mathbf{1}_{B_j})\right|,$$

where the sup is taken over all pairs of finite partitions A_i and B_j of Ω such that $A_i \in \mathcal{A}$ for each i and $B_j \in \mathcal{B}$ for each j. The β-mixing coefficients of the strictly stationary sequence $(X_n)_{n\in\mathbb{Z}}$ are defined, for each $n \in \mathbb{N}$, by

$$\begin{aligned}&\beta_n = \beta(\mathcal{A}_0,\mathcal{B}_n), \quad \text{where } \mathcal{A}_0 = \sigma(X_i, i \leqslant 0) \text{ and}\\ &\mathcal{B}_n = \sigma(X_i, i \geqslant n).\end{aligned} \tag{7}$$

The sequence $(X_n)_{n\in\mathbb{Z}}$ is said to be *β-mixing* or *absolutely regular* (a.r.) if the mixing coefficients β_n tend to 0, as n tends to ∞.

For β-mixing sequences, the convergence of G_n follows, as a particular case in Doukhan et al. [6], if

$$\beta_n = \mathcal{O}\big(n^{-1}(\mathrm{Log}\, n)^{-a}\big), \quad \text{for some } a > 2. \tag{8}$$

Recall that the fi-di convergence of G_n needs the mixing condition $\sum_{n>0}\beta_n < \infty$. Hence condition (8) is nearly optimal.

It is interesting to compare mixing and association. We note first that an important property of associated random variables is that noncorrelation implies independence (see e.g. Newman [9]); the only alternative frame for this to hold is the Gaussian one. This means that one may hope that dependence will appear in this case only through the covariance structure, and also justifies the study of such processes: indeed a covariance is much easier to compute than a mixing coefficient. Unfortunately, a main inconvenience of mixing is that there are only few mixing models for which the mixing coefficients can be explicitly evaluated. Examples of such models are linear processes, that we intend to focus on in this paper.

Suppose $\varepsilon := (\varepsilon_i,\ i \in \mathbb{Z})$ is a stationary sequence of independent r.v's fulfilling $\mathbb{E}\varepsilon_0 = 0$, $\sigma_0^2 := \mathbb{E}\varepsilon_1^2 < \infty$. Let $a := (a_n)_{n\in\mathbb{Z}}$ be a sequence of real numbers such that $\sum_{r\in\mathbb{Z}} a_r^2 < \infty$. Then the random sequence $(X_k,\ k \in \mathbb{Z})$:

$$X_k := \sum_{j=-\infty}^{\infty} a_j\varepsilon_{k-j}, \tag{9}$$

is well defined, strictly stationary and $\mathbb{E}X_1^2 < \infty$. In the sequel, we denote by $\tilde{\mathcal{L}}_2(a, \varepsilon)$ the class of linear sequences, defined by (9) and by $\mathcal{L}_2(a, \varepsilon)$ the subset of $\tilde{\mathcal{L}}_2(a, \varepsilon)$ for which $a_i = 0$ for $i < 0$.

If the distribution of ε_0 has an absolutely continuous density in $\mathbb{L}^1$, then the linear sequence $(X_k,\ k \in \mathbb{Z})$ in $\mathcal{L}_2(a, \varepsilon)$ is a.r. with

$$\beta_n \leqslant K \sum_{j \geqslant n} \Bigl(\sum_{k \geqslant j} |a_k| \Bigr)^{2/3}, \quad \text{for some constant } K,$$

as soon as $\sum_{j \geqslant 0} (\sum_{k \geqslant j} |a_k|)^{2/3} < \infty$ (cf. Pham and Tran [12]).

Up to our knowledge, there are no conditions yielding the mixing property for non-causal linear sequences of the set $\tilde{\mathcal{L}}_2(a, \varepsilon)$ (cf. Doukhan [5] for a survey of literature about this question).

We note also that association and mixing define two distinct but not disjoint classes of processes; as it is shown by the following examples taken from the class $\mathcal{L}_2(a, \varepsilon)$.

Associated but not mixing sequences. The following example is well known.

Suppose $\varepsilon = (\varepsilon_i,\ i \in \mathbb{Z})$ is i.i.d. with $P(\varepsilon_i = 1) = P(\varepsilon_i = 0) = 1/2$. Let $a = (2^{-k},\ k \geqslant 1)$. Then the linear process $(X_k) \in \mathcal{L}_2(a, \varepsilon)$ so defined is associated (cf. $(\mathcal{P}_2)$ and $(\mathcal{P}_4)$ of Esary et al. [7]). However (X_k) fails to be mixing, since $\beta_n \geqslant 1/4$ (cf. Bradley [3] and the references therein). Moreover, X_1 has the uniform on [0, 1] marginal law, and, for $n \in \mathbb{N}$,

$$r(n) := \mathrm{Cov}(X_1, X_n) = c^2 \sigma_0^2 2^{-n},$$

which decreases exponentially fast to 0. Hence, Yu's [15] result yields the convergence of the empirical process G_n constructed from this sequence (X_k).

Associated and mixing sequences. Suppose that the requirement of Theorem 2.1 in Pham and Tran [12] holds. Suppose moreover that $a_k \geqslant 0$ for all $k \in \mathbb{N}$. Then the linear process $(X_k) \in \mathcal{L}_2(a, \varepsilon)$ so defined, is at the same time associated and β-mixing.

Mixing but not associated sequences. Suppose that the sequence $a = (a_k,\ k \in \mathbb{N})$ satisfies $a_0 a_1 < 0$ and $a_i = 0$ for all $i \geqslant 2$. Then $X_n = a_0 \varepsilon_n + a_1 \varepsilon_{n-1}$ is not associated since $\mathrm{Cov}(X_n, X_{n-1}) = a_0 a_1 \sigma_0^2 < 0$ (which is in contradiction with the definition of association). However the sequence (X_n) is mixing, since it is m-dependent.

Another example of classes that make the difference between mixing and association is the Gaussian one. Indeed Gaussian processes are associated if and only if their covariance function is positive (cf. Pitt [11]).

Let us now come back to the main purpose of this paper, which is the study of the convergence of the empirical process under association.

In the following section, we give the main result. Our condition yielding the ECLT under association requires $b > 4$ which improves on condition (4) (cf. Theorem 1 below). We apply the result to linear sequences and we discuss the results known in this area (cf. Corollary 1 and the remarks below). In particular, we compare the results obtained when the linear processes are viewed as associated sequences or as mixing r.v's. In section three, we prove the results and we emphasise the necessity to have $b > 4$, of course with the method that we propose.

An outline of the proofs is the following. By adapting an approach of Andrews and Pollard [1], we approximate G_n by an empirical process indexed by suitable regular functions that belong to a finite class of functions. Hence to prove the tightness property, it remains to control some moment of the oscillation of the empirical process indexed by those suitable regular functions. This will be done using a Rosenthal moment inequality for regular functions of associated r.v's proved in Shao and Yu [13]. The tightness property then holds if we estimate the variance quantities that come from Rosenthal's inequality. Those variance estimates differ from the estimates given in Shao and Yu [13] (cf. Lemma 2 below). Hence, to carry out the proof of the tightness property, we develop a general criterion yielding the tightness of the empirical process constructed from any stationary sequence fulfilling a suitable moment inequality (cf. Proposition 1).

Finally, the paper is concluded with an appendix dedicated to Proposition 1.

2. MAIN RESULT AND APPLICATION

THEOREM 1. – *Let $(X_n)_{n\in\mathbb{Z}}$ be a stationary associated sequence with continuous marginal distribution F. Assume that, for $n \in \mathbb{N}^*$,*

$$\mathrm{Cov}\big(F(X_1), F(X_n)\big) = \mathcal{O}\big(n^{-b}\big), \quad \textit{for } b > 4. \tag{10}$$

Then

$$G_n(.) \to G(.) \quad \textit{in } D[-\infty, +\infty],$$

where $G_n(.)$ *is defined by* (1) *and* G *is the zero-mean Gaussian process with covariance defined by* (2).

The above theorem applied to the sequences $(X_k)_{k\in\mathbb{Z}}$ in $\tilde{\mathcal{L}}_2(a,\varepsilon)$ yields:

COROLLARY 1. – *Let* $(X_i)_{i\in\mathbb{Z}}$ *be a stationary sequence that belongs to the set* $\tilde{\mathcal{L}}_2(a,\varepsilon)$. *Suppose that* $a_i \geqslant 0$ *for each* i *in* $\mathbb{Z}$. *Suppose moreover that the law of* X_1 *has a bounded density. If, for* $n \in \mathbb{N}^*$,

$$\sum_{k\in\mathbb{Z}} a_k a_{k-n} = \mathcal{O}\big(n^{-b}\big) \quad \textit{for } b > 4, \tag{11}$$

then the conclusion of Theorem 1 *holds.*

Proof of Corollary 1. – It suffices to note that the linear sequence is associated (since $a_i \geqslant 0$ for each $i \in \mathbb{Z}$) and that

$$\mathrm{Cov}\big(F(X_1), F(X_n)\big) \leqslant \|f\|_\infty^2 \mathrm{Cov}(X_1, X_n) = \sigma_0^2 \|f\|_\infty^2 \sum_{k\in\mathbb{Z}} a_k a_{k-n}. \qquad \square$$

Remark. – Let $(X_i)_{i\in\mathbb{Z}} \in \mathcal{L}_2(a,\varepsilon)$. Suppose that $a_k = \mathcal{O}(|k|^{-a})$.

(1) If $(X_i)_{i\in\mathbb{Z}}$ satisfies the requirement of Theorem 2.1 in Pham and Tran [12], then the sequence $(X_i)_{i\in\mathbb{Z}}$ is β-mixing if $a > 5/2$ and the coefficients β_n fulfil $\beta_n \leqslant cn^{5/3-2a/3}$. Hence condition (8) is satisfied only if $a > 4$.

(2) Suppose now that the requirement of Corollary 1 holds. Then $r(n) \leqslant cn^{-a}$ and the convergence of G_n holds if $a > 4$.

Hence association or mixing require the same condition on the decay of (a_k) for the ECLT. However, the association property follows as soon as the coefficients a_k are positive, while the β-mixing property requires some additional conditions on the law of ε_0.

An immediate consequence of Corollary 1 is the following:

COROLLARY 2. – *Let* $(X_i)_{i\in\mathbb{Z}}$ *be a stationary sequence in* $\tilde{\mathcal{L}}_2(a,\varepsilon)$ *fulfilling condition* (11). *Suppose that* $a_j \geqslant 0$ *for each* $j \in \mathbb{Z}$ *and that*

$$\mathrm{Card}\{j \in \mathbb{Z}, a_j > 0\} = \infty. \tag{12}$$

Suppose moreover that, for some $\delta > 0$ *and for any* $u \in \mathbb{R}$

$$\big|\mathbb{E}(\mathrm{e}^{\mathrm{i}u\varepsilon_0})\big| \leqslant C\big(1+|u|\big)^{-\delta}. \tag{13}$$

Then the conclusion of Theorem 1 *holds for the sequence* $(X_i)_{i\in\mathbb{Z}}$.

Proof of Corollary 2. – Note that conditions (12) and (13) guarantee the existence of a bounded density for X_0 (cf. Giraitis and Surgailis [8]). Those conditions, together with (11), yield then Corollary 1, which in turns implies Corollary 2. □

3. PROOF OF THE MAIN RESULT

As we have already noticed in Section 1, the proof of Theorem 1 requires two steps. The proof of the first step follows from the CLT for weakly associated random vectors of Burton et al. [4], we recall that it requires $b > 3$.

In order to prove the stochastic equicontinuity, we adapt Andrews and Pollard's [1] approach for mixing sequences to the context of associated sequences.

In the following subsection, we give a criterion ensuring the stochastic equicontinuity property for empirical processes indexed by a class of functions $\mathcal{F}$. This criterion will be available for any stationary sequence fulfilling a suitable moment inequality. Next (cf. Section 3.2), we apply the stochastic equicontinuity criterion to the empirical process G_n (defined by (1)) when the observations are stationary and associated.

3.1. A general tightness criterion

Let $(X_n)_{n\in\mathbb{Z}}$ be a stationary sequence of r.v's. Let $\mathcal{F}$ be a class of real-valued functions uniformly bounded by 1. Suppose that $\mathcal{F}$ is equipped with the seminorm $\rho(f) = \|f(X_1)\|_2$ ($\|Z\|_p$ denotes the norm $(\mathbb{E}|Z|^p)^{1/p}$ of a random variable Z). The empirical process G_n indexed by the class of functions $\mathcal{F}$ is defined by:

$$G_n(f) = \frac{1}{\sqrt{n}} \sum_{1\leqslant i\leqslant n} \bigl(f(X_i) - \mathbb{E}f(X_i)\bigr). \tag{14}$$

Hence the usual empirical process G_n on $[0, 1]$ defined by (1) is the process G_n indexed by the class:

$$\mathcal{F} = \{f_t : x \to \mathbf{1}_{x\leqslant t},\ t \in [0, 1]\}. \tag{15}$$

Let, for $\delta > 0$, $\mathcal{F}(\delta)$ be the set of real-valued functions uniformly bounded by 1 such that their first derivatives are uniformly bounded by $1/\delta$. We adapt the following definition of the bracketing number in order to make the proofs as easy as possible.

DEFINITION 1. – *The bracketing number* $\mathcal{N}(\delta) = \mathcal{N}(\delta, \mathcal{F})$ *is the smallest value of* N *for which there exist functions* $f_1, \ldots, f_N$ *in* $\mathcal{F}(\delta)$ *such that for each* f *in* $\mathcal{F}$ *there exists* i, j *for which* $f_i \leqslant f \leqslant f_j$ *and* $\mathbb{E}(f_j - f_i)(X_0) \leqslant C\delta$, *for a positive constant* C *that depends only on the sequence* $(X_n)_{n\in\mathbb{Z}}$.

For each $k \in \mathbb{N}$, *let* $\mathcal{F}_k$ *be the finite subclass of* $\mathcal{F}(2^{-k})$ *with cardinality* $\mathcal{N}_k := \mathcal{N}(2^{-k}, \mathcal{F})$ *defined by* $\mathcal{F}_k := \{f_1, \ldots, f_{\mathcal{N}_k}\}$.

We introduce the following assumption.

ASSUMPTION $\mathcal{A}(r)$. – *For any real numbers* $r > 2$, $\mu > 0$, *there exists a positive constant* $C_{r,\mu}$ *for which the following moment inequality holds for any functions* f_k, g_k *in* $\mathcal{F}_k$

$$\begin{aligned}\mathbb{E}\big|S_n(f_k - g_k) - \mathbb{E}S_n(f_k - g_k)\big|^r \\ \leqslant C_{r,\mu}\big\{n^{1+\mu}2^{2k} + n^{r/2}\rho^{r/2}(f_k - g_k)\big\},\end{aligned} \tag{16}$$

where $S_n(f) = f(X_1) + \cdots + f(X_n)$ *and* $\rho(f) = \|f(X_1)\|_2$.

As it will be seen in the sequel, associated sequences fulfill Assumption $\mathcal{A}(r)$.

The following proposition gives a maximal inequality for an arbitrary stationary sequence fulfilling Assumption $\mathcal{A}(r)$. This maximal inequality yields the tightness of the empirical process indexed by a suitable class of functions $\mathcal{F}$. Let us note that the forthcoming maximal inequality requires conditions on the integrability of the bracketing numbers introduced in Definition 1.

PROPOSITION 1. – *Let* $(X_n)_{n\in\mathbb{Z}}$ *be a stationary sequence of r.v's that satisfies Assumption* $\mathcal{A}(r)$, *for* $r > 2$, $\mu > 0$. *Let* $\mathcal{F}$ *be a class of real-valued functions uniformly bounded by* 1, *whose bracketing numbers* (*see Definition* 1) *satisfy*

$$\lim_{n\to+\infty} \frac{n^{2-r/2+\mu}}{a_n^2}\mathcal{N}\left(\frac{a_n}{\sqrt{n}}, \mathcal{F}\right) = 0, \quad \textit{and}$$

$$\int_0^1 x^{-3/4}\mathcal{N}^{1/r}(x, \mathcal{F})\,\mathrm{d}x < +\infty, \tag{17}$$

for the same r *and for a sequence* (a_n) *that decreases to* 0 *at infinity. Then for each* $\varepsilon > 0$ *and* $\delta > 0$ *there exists an integer* m *depending only on* ε, r *and on the bracketing number* $\mathcal{N}(., \mathcal{F})$ *for which*

$$\left\| \sup_{\rho(f-g)\leqslant\delta} \left|G_n(f) - G_n(g)\right| \right\|_r$$
$$\leqslant \varepsilon + \mathcal{N}_m^{2/r} \sup_{\rho(f-g)\leqslant 2\delta} \left\|G_n(f) - G_n(g)\right\|_r, \tag{18}$$

where $\mathcal{N}_m := \mathcal{N}(2^{-m}, \mathcal{F})$.

Hence the stochastic equicontinuity follows from (18), if we choose $\delta > 0$ such that

$$\limsup_{n\to+\infty} \mathcal{N}_m^{2/r} \sup_{\rho(f-g)\leqslant 2\delta} \left\|G_n(f) - G_n(g)\right\|_r \leqslant \varepsilon.$$

Let us note that the first condition on the bracketing number (see conditions (17)) requires $r > 4$ (recall that (a_n) decreases to 0 at infinity).

In the following subsection, we prove Theorem 1 (Proposition 1 is proved in Appendix A).

3.2. Proof of Theorem 1

To prove Theorem 1, we apply Proposition 1 with $\mathcal{F}$ as defined by (15). Let us precise, in such a case, the approximating class $\mathcal{F}_k$ and the bracketing numbers $\mathcal{N}(x, \mathcal{F})$ (cf. Definition 1).

For some $k \in \mathbb{N}$, we denote by T_k the covering set of [0, 1] defined by

$$T_k = \{p2^{-k},\ p = 1, \ldots, 2^k\}.$$

For each t_k in T_k, we define the step function f_{t_k} by:

$$f_{t_k}(x) = \mathbf{1}_{x<t_k-2^{-k}} - (2^k x - 2^k t_k)\mathbf{1}_{x\in[t_k-2^{-k}, t_k]}. \tag{19}$$

Let $\mathcal{F}_k$ be the set of all those functions f_{t_k} whenever t_k runs over the covering set T_k. The class $\mathcal{F}_k$, previously defined, approximates $\mathcal{F}$ (defined in (15)). Certainly, for each f_t in $\mathcal{F}$:

$$f_{t_k}(x) \leqslant f_t(x) \leqslant f_{t_k+2^{1-k}}(x) =: g_{t_k}(x) \quad \text{where } t_k = 2^{-k}[t2^k],$$
$$\mathbb{E}(f_{t_k+2^{1-k}} - f_{t_k})(X_1) \leqslant 3 \times 2^{-k}, \quad \|f_{t_k}\|_\infty \leqslant 1 \quad \text{and} \quad \|f'_{t_k}\|_\infty \leqslant 2^k,$$

square brackets denoting the integer part, as usual. The correspondence between $\mathcal{F}$ and T (respectively $\mathcal{F}_k$ and T_k) implies that $\mathcal{N}(x, \mathcal{F}) = \mathcal{O}(1/x)$ (respectively $\mathcal{F}_k$ and T_k have the same cardinality: 2^k). Consequently, the integral condition on the covering number stated in Proposition 1 is satisfied as soon as $r > 4$, while the first condition needs

$r > 5 + 2\mu$ (see conditions (17)). Therefore, in order to prove the tightness property, we will first check Assumption $\mathcal{A}(r)$ for some $r > 5$. We make use of the following lemma.

LEMMA 1 (Shao and Yu [13]). – *Let $r > 2$. Let f be a real valued function bounded by 1 with bounded first derivative. Suppose that $(X_n)_{n\in\mathbb{Z}}$ is a sequence of stationary and associated r.v's such that, for $n \in \mathbb{N}^*$,*

$$\mathrm{Cov}(X_1, X_n) = \mathcal{O}(n^{-b}), \quad \textit{for some } b > r - 1.$$

Then, for any $\mu > 0$ there exists some positive constant k_μ independent of the function f for which

$$\mathbb{E}|S_n(f) - \mathbb{E}S_n(f)|^r \leqslant k_\mu\left(n^{1+\mu}\|f'\|_\infty^2 + n^{r/2}\left(\sum_{j=1}^{n}|\mathrm{Cov}(f(X_1), f(X_j))|\right)^{r/2}\right). \tag{20}$$

Inequality (20) is known as Rosenthal's inequality. The term $n^{1+\mu}\|f'\|_\infty^2$ in inequality (20) is replaced, for independent sequences, by $n\mathbb{E}|f|^r(X_1)$ (cf. Petrov [10]). This difference will affect the conditions yielding the convergence of the empirical process.

Define for some fixed s, and for a and h in $[0, 1]$, the step function $f_{a,h}$ by:

$$f_{a,h}(x) = \begin{cases} 0 & \text{if } x \leqslant s - a, \\ 1 + \dfrac{x - s}{a} & \text{if } s - a \leqslant x \leqslant s, \\ 1 & \text{if } s \leqslant x \leqslant s + h, \\ 1 + \dfrac{s + h - x}{a} & \text{if } s + h \leqslant x \leqslant s + h + a, \\ 0 & \text{if } x \geqslant s + h + a. \end{cases}$$

Clearly the function $f_{a,h}$ depends on s, we will write it $f_{s,a,h}$ in any confusing situation.

We recall that our purpose is to check Assumption $\mathcal{A}(r)$, that means that the functions f_{t_k} of the approximating subset $\mathcal{F}_k$ satisfy the moment inequality (16).

For this, let t_k and t'_k be two arbitrary real numbers in the covering set T_k. Consider the functions f_{t_k}, $f_{t'_k}$ defined as in (19).

We assume without loss of generality that $t_k < t'_k$, hence $t_k - 2^{-k} < t_k \leqslant t'_k - 2^{-k} < t'_k$. Therefore $f_{t'_k} - f_{t_k}$ is one of the functions $f_{a,h}$ previously defined, with $s := s_1 = t_k$, $a := a_1 = 2^{-k}$ and $h := h_1 = t'_k - t_k - 2^{-k}$. We note that if $h \neq 0$ then necessarily $a \leqslant h$.

Now suppose that condition (10) holds. Let $\nu = b - 4$ and $r = 5 + 3\mu$ for some $\mu \in]0, \nu/3[$. Lemma 1 yields:

$$\mathbb{E}\left|S_n(f_{t_k} - f_{t'_k}) - \mathbb{E}S_n(f_{t_k} - f_{t'_k})\right|^r \leqslant C_r\Bigg\{n^{1+\mu}2^{2k} + n^{r/2}\left(\sum_{j=1}^{+\infty}\left|\operatorname{Cov}(f_{s_1,a_1,h_1}(X_1), f_{s_1,a_1,h_1}(X_j))\right|\right)^{r/2}\Bigg\}, \tag{21}$$

where C_r is a positive constant that does not depend on the class $\mathcal{F}$.

Hence Assumption $\mathcal{A}(r)$ holds if we control the covariance quantity of the last inequality. This is the purpose of the forthcoming lemma.

LEMMA 2. – *Let $(X_n)_{n\in\mathbb{Z}}$ be a stationary associated sequence of uniform-$[0,1]$ random variables. Assume that there exists a positive constant K such that, for $r \in \mathbb{N}^*$,*

$$\operatorname{Cov}(X_1, X_r) \leqslant K r^{-b}, \quad \textit{for some } b > 4. \tag{22}$$

Then, there exists a positive constant C depending only on K and b for which

$$\sum_{j=1}^{+\infty}\left|\operatorname{Cov}(f_{a,h}(X_1), f_{a,h}(X_j))\right| \leqslant C\rho(f_{a,h}),$$

where the functions $f_{a,h}$ are previously defined with $0 \leqslant a \leqslant h$ or $h = 0$ and $a > 0$.

Before proving Lemma 2, we continue the proof of Theorem 1.

Inequality (21), together with Lemma 2, yields:

$$\mathbb{E}\left|S_n(f_{t_k} - f_{t'_k}) - \mathbb{E}S_n(f_{t_k} - f_{t'_k})\right|^r \leqslant C_r\left\{n^{1+\mu}2^{2k} + n^{r/2}\rho^{r/2}(f_{t_k} - f_{t'_k})\right\}. \tag{23}$$

Therefore Assumption $\mathcal{A}(r)$ holds. Proposition 1 yields then:

$$\left\|\sup_{\rho(f_t - f_s)\leqslant\delta}\left|G_n(f_t) - G_n(f_s)\right|\right\|_r \leqslant \varepsilon + \mathcal{N}_m^{2/r}\sup_{\rho(f_t - f_s)\leqslant 2\delta}\left\|G_n(f_t) - G_n(f_s)\right\|_r, \tag{24}$$

where the functions f_t, f_s are in the set $\mathcal{F}$ defined by (15), δ and ε are arbitrary positive real-numbers, m is a fixed real-number depending only on r, ε and on the covering number $\mathcal{N}(., \mathcal{F})$.

We now evaluate the moment quantity $\|G_n(t) - G_n(s)\|_r$. Some elementary estimations (cf. also (5.27) of Shao and Yu [13]) yield:

$$\begin{aligned}&\big\|G_n(t) - G_n(s)\big\|_r\\&\quad\leqslant C\left(n^{\frac{-(r-4-2\mu)}{2(r+2)}} + \left(\sum_{j=1}^{+\infty}\big|\mathrm{Cov}\big(f_{s,0,t-s}(X_1), f_{s,0,t-s}(X_j)\big)\big|\right)^{1/2}\right).\end{aligned}$$

The above inequality, together with (24) and Lemma 2, implies:

$$\begin{aligned}&\Big\|\sup_{\rho(f_t-f_s)\leqslant\delta}\big|G_n(t) - G_n(s)\big|\Big\|_r\\&\quad\leqslant \varepsilon + \mathcal{N}_m^{2/r}\sup_{\rho(f_t-f_s)\leqslant 2\delta}\big\|G_n(t) - G_n(s)\big\|_r\\&\quad\leqslant \varepsilon + C\mathcal{N}_m^{2/r}\big(n^{-(r-4-2\mu)/(2(r+2))} + \delta^{1/2}\big).\end{aligned}$$

Therefore, we obtain, noting that $r > 4 + 2\mu$ and taking the limit in the last inequality,

$$\limsup_{n\to+\infty}\Big\|\sup_{\rho(f_t-f_s)\leqslant\delta}\big|G_n(t) - G_n(s)\big|\Big\|_r \leqslant \varepsilon + C\mathcal{N}_m^{2/r}\delta^{1/2}.$$

m is fixed, we can therefore choose δ in such a way that $C\mathcal{N}_m^{2/r}\delta^{1/2} \leqslant \varepsilon$.

3.2.1. Proof of Lemma 2

The proof of this lemma needs the following preparatory lemma.

LEMMA 3. – *Let (X_1, X_2) be a vector of associated r.v's with the uniform-*[0, 1] *marginal law. Let $f_{a,h}$ be the function previously defined.*
If $h = 0$, $a > 0$, then:

$$\big|\mathrm{Cov}\big(f_{a,0}(X_1), f_{a,0}(X_2)\big)\big| \leqslant 5\rho(f_{a,0})\mathrm{Cov}^{1/4}(X_1, X_2).$$

Suppose now that $h \neq 0$, $0 \leqslant a \leqslant h$. Then the following inequalities hold:

(1) *If $a \geqslant \min(\mathrm{Cov}^{1/4}(X_1, X_2), (\mathrm{Cov}(X_1, X_2)/h)^{1/3})$, then*

$$\big|\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_2)\big)\big| \leqslant \sqrt{6}h^{1/2}\mathrm{Cov}^{1/4}(X_1, X_2). \qquad (25)$$

(2) *If $a < \min(\mathrm{Cov}^{1/4}(X_1, X_2), (\mathrm{Cov}(X_1, X_2)/h)^{1/3})$, then*

$$-\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_2)\big) \leqslant 81h^{2/3}\mathrm{Cov}^{1/3}(X_1, X_2) \qquad (26)$$

and moreover,

- *If* $h^2 \leqslant \mathrm{Cov}(X_1, X_2)$, *then* $\mathrm{Cov}(f_{a,h}(X_1), f_{a,h}(X_2)) \leqslant 5h$.
- *If* $\mathrm{Cov}(X_1, X_2) \leqslant h^4$, *then*

$$\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_2)\big) \leqslant 10h\mathrm{Cov}^{1/4}(X_1, X_2).$$

- *If* $h^4 \leqslant \mathrm{Cov}(X_1, X_2) \leqslant h^2$, *then*

$$\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_2)\big) \leqslant 10\mathrm{Cov}^{1/2}(X_1, X_2).$$

Proof of Lemma 3. – The proof of this lemma uses repeatedly the following covariance inequality. For a stationary associated vector (X_1, X_2)

$$\big|\mathrm{Cov}\big(f(X_1), f(X_2)\big)\big| \leqslant \min\big(2\mathbb{E}f^2(X_1), \|f'\|_\infty^2\mathrm{Cov}(X_1, X_2)\big). \quad (27)$$

We suppose first that $h \neq 0$, $0 \leqslant a \leqslant h$ and we consider the following cases.

Case 1. Suppose that $a \geqslant \min(\mathrm{Cov}^{1/4}(X_1, X_2), (\mathrm{Cov}(X_1, X_2)/h)^{1/3})$, then (27) yields:

(1) If $a \geqslant \mathrm{Cov}^{1/4}(X_1, X_2)$, then

$$\begin{aligned}\big|\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_2)\big)\big| &\leqslant \min\big(\mathrm{Cov}^{1/2}(X_1, X_2), 2\rho^2(f_{a,h})\big)\\ &\leqslant \sqrt{2}\mathrm{Cov}^{1/4}(X_1, X_2)\rho(f_{a,h}).\end{aligned}$$

(2) If $a \geqslant (\mathrm{Cov}(X_1, X_2)/h)^{1/3}$, then

$$\big|\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_2)\big)\big| \leqslant h^{2/3}\mathrm{Cov}^{1/3}(X_1, X_2).$$

Hence (25) of Lemma 3 follows, by noting that

$$h^{2/3}\mathrm{Cov}^{1/3}(X_1, X_2) \leqslant h^{1/2}\mathrm{Cov}^{1/4}(X_1, X_2) \quad \text{and}$$
$$\rho^2(f_{a,h}) \leqslant 2a + h \leqslant 3h, \quad \text{since } a \leqslant h \leqslant 1.$$

Case 2. Suppose now that

$$a < \min\big(\mathrm{Cov}^{1/4}(X_1, X_2), (\mathrm{Cov}(X_1, X_2)/h)^{1/3}\big).$$

A bound for $\mathrm{Cov}(f_{a,h}(X_1), f_{a,h}(X_2))$

If $b = \mathrm{Cov}^{1/4}(X_1, X_2)$, then the condition on a yields $a < b$. Hence $0 \leqslant f_{a,h} \leqslant f_{b,h}$, and

$$\begin{aligned}&\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_2)\big)\\ &\quad\leqslant \mathrm{Cov}\big(f_{b,h}(X_1), f_{b,h}(X_2)\big) + \mathbb{E}f_{b,h}(X_1)\mathbb{E}f_{b,h}(X_2)\end{aligned}$$

$$
\begin{aligned}
&- \mathbb{E} f_{a,h}(X_1)\mathbb{E} f_{a,h}(X_2) \\
&\leqslant \frac{1}{b^2}\mathrm{Cov}(X_1, X_2) + (h+2b)^2 - h^2 \\
&\leqslant 5\mathrm{Cov}^{1/2}(X_1, X_2) + 4h\mathrm{Cov}^{1/4}(X_1, X_2).
\end{aligned}
$$

(The last inequalities are obtained using the fact that $\mathbb{E} f_{a,h}(X_1) \geqslant h$ and that $\mathbb{E} f_{b,h}(X_1) \leqslant h + 2b$.) Hence we obtain (recall that $a \leqslant h$):

$$
\begin{aligned}
&\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_2)\big) \\
&\quad \leqslant \min\big(5\mathrm{Cov}^{1/2}(X_1, X_2) + 4h\mathrm{Cov}^{1/4}(X_1, X_2), \rho^2(f_{a,h})\big) \\
&\quad \leqslant 5\min\big(\mathrm{Cov}^{1/2}(X_1, X_2) + h\mathrm{Cov}^{1/4}(X_1, X_2), h\big).
\end{aligned}
$$

Now, we bound $\min(\mathrm{Cov}^{1/2}(X_1, X_2) + h\mathrm{Cov}^{1/4}(X_1, X_2), h)$ by

- h if $h^2 \leqslant \mathrm{Cov}(X_1, X_2)$,
- $2h\mathrm{Cov}^{1/4}(X_1, X_2)$ if $\mathrm{Cov}(X_1, X_2) \leqslant h^4$,
- $2\mathrm{Cov}^{1/2}(X_1, X_2)$ otherwise.

With these estimates, the last part of Lemma 3 follows.

A bound for $-\mathrm{Cov}(f_{a,h}(X_1), f_{a,h}(X_2))$

If $\mathrm{Cov}(X_1, X_2) \geqslant h^4/8$, then the covariance inequality (26) is obviously satisfied:

$$
\begin{aligned}
-\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_2)\big) &\leqslant \big(\mathbb{E} f_{a,h}(X_1)\big)^2 \leqslant 9h^2 \\
&\leqslant 81 h^{2/3}\mathrm{Cov}^{1/3}(X_1, X_2).
\end{aligned}
$$

We suppose now $\mathrm{Cov}(X_1, X_2) < h^4/8$ and we consider, for some $b \in [a, h/2[$, the function $g_b := f_{s+b,b,h-2b}$. Clearly, $0 \leqslant g_b \leqslant f_{s,a,h}$, $h - 2b \leqslant \mathbb{E} g_b(X_1)$ and $\|g'\|_\infty \leqslant 1/b$. Hence:

$$
\begin{aligned}
&-\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_2)\big) \\
&\quad \leqslant -\mathrm{Cov}\big(g_b(X_1), g_b(X_2)\big) - \mathbb{E} g_b(X_1)\mathbb{E} g_b(X_2) + (2a+h)^2 \\
&\quad \leqslant \frac{1}{b^2}\mathrm{Cov}(X_1, X_2) + (h+2a)^2 - (h-2b)^2 \\
&\quad \leqslant \frac{1}{b^2}\mathrm{Cov}(X_1, X_2) + 8bh \quad \text{(recall that } a \leqslant b < h/2\text{)}.
\end{aligned}
$$

Now let $b := (\mathrm{Cov}(X_1, X_2)/h)^{1/3}$. This b satisfies $a \leqslant b < h/2$. Hence the last inequality yields:

$$
-\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_2)\big) \leqslant 9h^{2/3}\mathrm{Cov}^{1/3}(X_1, X_2).
$$

The proof of Lemma 3 is complete if we bound $|\mathrm{Cov}(f_{a,0}(X_1), f_{a,0}(X_2))|$. To this end we consider again two cases.

(1) If $a > \mathrm{Cov}^{1/4}(X_1, X_2)$, then obviously

$$\begin{aligned} \left|\mathrm{Cov}\big(f_{a,0}(X_1), f_{a,0}(X_2)\big)\right| &\leqslant \min\left(2\rho^2(f_{a,0}), \frac{\mathrm{Cov}(X_1, X_2)}{a^2}\right) \\ &\leqslant \min\left(2\rho^2(f_{a,0}), \mathrm{Cov}^{1/2}(X_1, X_2)\right) \\ &\leqslant \sqrt{2}\rho(f_{a,0})\mathrm{Cov}^{1/4}(X_1, X_2). \end{aligned}$$

(2) Suppose now that $a \leqslant \mathrm{Cov}^{1/4}(X_1, X_2) =: b$, then $0 \leqslant f_{a,0} \leqslant f_{b,0}$ and

$$\begin{aligned} &\mathrm{Cov}\big(f_{a,0}(X_1), f_{a,0}(X_2)\big) \\ &\quad\leqslant \mathbb{E} f_{a,0}(X_1) f_{a,0}(X_2) \\ &\quad\leqslant \mathrm{Cov}\big(f_{b,0}(X_1), f_{b,0}(X_2)\big) + \mathbb{E} f_{b,0}(X_1)\mathbb{E} f_{b,0}(X_2) \\ &\quad\leqslant \frac{1}{b^2}\mathrm{Cov}(X_1, X_2) + 4b^2 = 5\mathrm{Cov}^{1/2}(X_1, X_2). \end{aligned}$$

Hence

$$\begin{aligned} \mathrm{Cov}\big(f_{a,0}(X_1), f_{a,0}(X_2)\big) &\leqslant \min\left(5\mathrm{Cov}^{1/2}(X_1, X_2), 2\rho^2(f_{a,0})\right) \\ &\leqslant 5\mathrm{Cov}^{1/4}(X_1, X_2)\rho(f_{a,0}) \end{aligned}$$

and

$$\begin{aligned} -\mathrm{Cov}\big(f_{a,0}(X_1), f_{a,0}(X_2)\big) &\leqslant \min\left(4a^2, \rho^2(f_{a,0})\right) \\ &\leqslant \min\left(4\mathrm{Cov}^{1/2}(X_1, X_2), \rho^2(f_{a,0})\right) \\ &\leqslant 2\rho(f_{a,0})\mathrm{Cov}^{1/4}(X_1, X_2). \end{aligned}$$

Let us now conclude. We denote by I the set of all positive integers $i \leqslant n$ for which

$$a < \min\left(\mathrm{Cov}^{1/4}(X_1, X_i), \left(\frac{\mathrm{Cov}(X_1, X_i)}{h}\right)^{1/3}\right).$$

Let J be its complement in $\{1, \ldots, n\}$. Clearly

$$\begin{aligned} \sum_{j=1}^{n} \left|\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_j)\big)\right| &= \sum_{j\in I} \left|\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_j)\big)\right| \\ &\quad + \sum_{j\in J} \left|\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_j)\big)\right|. \end{aligned}$$

The last sum in the last inequality is bounded, using Lemma 3, by:

$$\sum_{j\in J}\left|\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_j)\big)\right| \leqslant \sqrt{6}h^{1/2}\sum_j \mathrm{Cov}^{1/4}(X_1, X_j).$$

In order to bound the first sum, we denote by:

$$\begin{aligned} I_1 &:= \{i \in I,\ \mathrm{Cov}(X_1, X_i) \leqslant h^4\},\\ I_2 &:= \{i \in I,\ h^4 < \mathrm{Cov}(X_1, X_i) \leqslant h^2\},\\ I_3 &:= \{i \in I,\ h^2 < \mathrm{Cov}(X_1, X_i)\}. \end{aligned}$$

Using Lemma 3 and assumption (22), we obtain:

$$\begin{aligned} &\sum_{j\in I}\left|\mathrm{Cov}\big(f_{a,h}(X_1), f_{a,h}(X_j)\big)\right|\\ &\quad\leqslant 81\Big\{h^{2/3}\sum_j \mathrm{Cov}^{1/3}(X_1, X_j) + \sum_{j\in I_3} h + \sum_{j\in I_2}\mathrm{Cov}^{1/2}(X_1, X_j)\\ &\qquad + \sum_{j\in I_1} h\mathrm{Cov}^{1/4}(X_1, X_j)\Big\}\\ &\quad\leqslant Ch^{1/2} \leqslant C\rho(f_{a,h}). \end{aligned}$$

Lemma 2 is so proved.

ACKNOWLEDGEMENTS

I would like to thank Professors Paul Doukhan and Emmanuel Rio for all their comments and helpful discussions.

APPENDIX A

A.1. Proof of Proposition 1

Throughout the proof we use the following maximal inequality due to Pisier (cf. Andrews et al. [1]): for random variables $Z_1, \ldots, Z_N$,

$$\Big\|\max_{i\leqslant N}|Z_i|\Big\|_r \leqslant N^{1/r}\max_{i\leqslant N}\|Z_i\|_r. \qquad \text{(A.1)}$$

The proof is done in three steps.

Step 1. We first show that there exists a sequence $f_{k(n)}$ in $\mathcal{F}_k$ such that $G_n(f)$ is approximated by $G_n(f_{k(n)})$ for a suitable choice of the sequence $k := k(n)$, i.e.

CLAIM 1. – *There exists a sequence $k(n)$ that diverges at infinity for which*

$$\lim_{n\to+\infty}\left\|\sup_{f\in\mathcal{F}}\left|G_n(f)-G_n(f_k)\right|\right\|_r=0.$$

In order to prove Claim 1, let f be in $\mathcal{F}$. Invoking the definition of bracketing numbers, one may find some functions f_k, g_k in $\mathcal{F}_k$ for which $f_k \leqslant f \leqslant g_k$, $\rho^2(g_k - f_k) \leqslant C2^{-k}$. On the one hand

$$\begin{aligned}
&G_n(f)-G_n(f_k)\\
&\quad\leqslant \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big((g_k-f_k)(X_i)-\mathbb{E}(g_k-f_k)(X_i)+\mathbb{E}(g_k-f)(X_i)\big)\\
&\quad\leqslant \left|G_n(g_k)-G_n(f_k)\right|+C\sqrt{n}2^{-k},
\end{aligned}$$

(the two last inequalities are obtained since $f \leqslant g_k$ and $\mathbb{E}(g_k - f_k) \leqslant C2^{-k}$). On the other hand

$$\begin{aligned}
G_n(f_k)-G_n(f)&=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big((f_k-f)(X_i)-\mathbb{E}(f_k-f)(X_i)\big)\\
&\leqslant C\sqrt{n}2^{-k},\quad \text{since } f_k-f\leqslant 0.
\end{aligned}$$

Hence: $|G_n(f) - G_n(f_k)| \leqslant |G_n(g_k) - G_n(f_k)| + C\sqrt{n}2^{-k}$, this yields

$$\sup_{f\in\mathcal{F}}\left|G_n(f)-G_n(f_k)\right|\leqslant \max_{f_k\in\mathcal{F}_k}\left|G_n(g_k)-G_n(f_k)\right|+C\sqrt{n}2^{-k}. \tag{A.2}$$

The maximum in the last inequality, is considered over the functions f_k in $\mathcal{F}_k$ (not over the functions (g_k, f_k) in $\mathcal{F}_k \times \mathcal{F}_k$). Indeed g_k is well defined given f_k and f: assume $\mathcal{F}_k$ to be enumerated, then g_k is the first function satisfying

$$f_k\leqslant f\leqslant g_k,\qquad \rho^2(g_k-f_k)\leqslant C2^{-k}. \tag{A.3}$$

Hence inequality (A.2), together with (A.1), yields

$$\begin{aligned}
&\left\|\sup_{f\in\mathcal{F}}\left|G_n(f)-G_n(f_k)\right|\right\|_r\\
&\qquad\leqslant \mathcal{N}_k^{1/r}\max_{f_k\in\mathcal{F}_k}\left\|G_n(g_k)-G_n(f_k)\right\|_r+C\sqrt{n}2^{-k}.
\end{aligned}\tag{A.4}$$

So that, Assumption $\mathcal{A}(r)$ and inequality (A.4) imply

$$\left\| \sup_{f\in\mathcal{F}} \left| G_n(f) - G_n(f_k) \right| \right\|_r \leqslant C_r \left(\mathcal{N}_k^{1/r} 2^{-k/4} + \mathcal{N}_k^{1/r} 2^{(2k)/r} n^{(1+\mu)/r-1/2} + \sqrt{n} 2^{-k} \right), \quad \text{(A.5)}$$

(here and in the sequel C_r is a positive constant independent of the family $\mathcal{F}$, that may be different from line to line). The second term on the right hand side of (A.5) tends to 0 at infinity if $k = k(n)$ satisfies $2^k = \sqrt{n}/a_n$ and if conditions (17) of Proposition 1 are fulfilled.

Step 2. The second step is to prove, via a chaining argument, that for fixed m and n large enough, $G_n(f_{k(n)})$ is uniformly approximated by $G_n(f_m)$. That is:

CLAIM 2. – *For a fixed $\varepsilon > 0$, there exists some $m \in \mathbb{N}$ such that for n large enough, one has*

$$\left\| \max_{f\in\mathcal{F}} \left| G_n(f_{k(n)}) - G_n(f_m) \right| \right\|_r \leqslant \varepsilon.$$

Since $k(n)$ diverges at infinity, we deduce that $m < k(n)$ for n large enough. Hence

$$\left\| \max_{f\in\mathcal{F}} \left| G_n(f_{k(n)}) - G_n(f_m) \right| \right\|_r \leqslant \sum_{i=m+1}^{k(n)} \left\| \max_{f\in\mathcal{F}} \left| G_n(f_i) - G_n(f_{i-1}) \right| \right\|_r. \quad \text{(A.6)}$$

The functions f_i appearing in the last sum are defined inductively as follows. Given f_i in $\mathcal{F}_i$, we define the function f_{i-1} from $\mathcal{F}_{i-1}$ that approximate the function f_i in the sense:

$$\rho^2(f_i - f_{i-1}) \leqslant C 2^{-i+1}.$$

From the above construction, we deduce that

$$\left\| \max_{f\in\mathcal{F}} \left| G_n(f_i) - G_n(f_{i-1}) \right| \right\|_r \leqslant \left\| \max_{f_i\in\mathcal{F}_i} \left| G_n(f_i) - G_n(f_{i-1}) \right| \right\|_r. \quad \text{(A.7)}$$

Inequalities (A.6), (A.7) and (A.1), together with Assumption $\mathcal{A}(r)$, yield:

$$\left\| \max_{f\in\mathcal{F}} \left| G_n(f_{k(n)}) - G_n(f_m) \right| \right\|_r \leqslant C_r \sum_{i=m+1}^{k(n)} \left(\mathcal{N}_i^{1/r} 2^{-i/4} + \mathcal{N}_i^{1/r} 2^{(2i)/r} n^{(1+\mu)/r-1/2} \right)$$

$$\leqslant C_r \sum_{i=m+1}^{k(n)} \left(\mathcal{N}_i^{1/r} 2^{-i/4} + \mathcal{N}_{k(n)}^{1/r} n^{(1+\mu)/r-1/2} 2^{2k(n)/r}\right),$$

the last inequality is obtained since the sequence $(\mathcal{N}_k)_k$ is nondecreasing.

Now, using conditions (17), the second term on the right hand side of the last inequality is small than $\varepsilon/2$ for n large enough (since $2^{k(n)} = \sqrt{n}/a_n$). The first term on the right hand side is bounded by $C \int_0^{2^{-m}} x^{-3/4} \mathcal{N}^{1/r}(x, \mathcal{F})\, \mathrm{d}x$, which is smaller than $\varepsilon/2$ for an appropriate fixed m.

Step 3. Now, starting from Claims 1 and 2 and arguing exactly as in Andrews and Pollard [1] (see their paragraph 'comparison of pairs'), we obtain

$$\left\| \sup_{\rho(f-g)\leqslant\delta} \left|G_n(f) - G_n(g)\right| \right\|_r \leqslant 8\varepsilon + \mathcal{N}_m{}^{2/r} \sup_{\rho(f-g)\leqslant 2\delta} \left\|G_n(f) - G_n(g)\right\|_r.$$

REFERENCES

[1] Andrews D.W.K., Pollard D., An introduction to functional central limit theorems for dependent stochastic processes, Int. Stat. Rev. 62 (1994) 119–132.

[2] Billingsley P., Probability and Measure, Wiley, New York, 1968.

[3] Bradley R.C., Basic properties of strong mixing conditions, in: Eberlein E., Taqqu M.S. (Eds.), Dependence in Probability and Statistics, Birkhäuser, Boston, 1986, pp. 165–192.

[4] Burton R.M., Dabrowski A.R., Dehling H., An invariance principle for weakly associated random variables, Stoch. Proc. Appl. 23 (1986) 301–306.

[5] Doukhan P., Mixing: Properties and Examples, Lecture Notes in Statist., Vol. 85, Springer-Verlag, Berlin, 1994.

[6] Doukhan P., Massart P., Rio E., Invariance principle for the empirical measure of a weakly dependent process, Ann. I.H.P. 31 (2) (1995) 393–427.

[7] Esary J., Proschan F., Walkup D., Association of random variables with applications, Ann. Math. Stat. 38 (1967) 1466–1476.

[8] Giraitis L., Surgailis D., A central limit theorem for the empirical process of a long memory linear sequence, Preprint, 1994.

[9] Newman C.M., Asymptotic independence and limit theorems for positively and negatively dependent random variables, in: Tong Y.L. (Ed.), Inequalities in Statistics and Probability, I.M.S. Lecture Notes-Monograph Series, Vol. 5, 1984, pp. 127–140.

[10] Petrov V.V., Limit Theorems of Probability Theory: Sequences of Independent Random Variables, Clarendon Press, Oxford, 1995.

[11] Pitt L., Positively correlated normal variables are associated, Ann. Probab. 10 (1982) 496–499.

[12] Pham T.D., Tran L.T., Some mixing properties of time series models, Stoch. Proc. Appl. 19 (1985) 297–303.

[13] Shao Q.M., Yu H., Weak convergence for weighted empirical processes of dependent sequences, Ann. Probab. 24 (4) (1996) 2052–2078.

[14] Volkonski V.A., Rozanov Y.A., Some limit theorems for random functions, Theory Probab. Appl. 4 (1959) 178–197.

[15] Yu H., A Glivenko–Cantelli lemma and weak convergence for empirical processes of associated sequences, Probab. Theory Related Fields 95 (1993) 357–370.