Maximal Inequalities and Empirical Central Limit Theorems.

Jérôme Dedecker and Sana Louhichi

Abstract. This work presents some recent developments concerning empirical central limit theorems for dependent sequences. We try to give general conditions under which an adaptive truncation in the chaining procedure works. We show that these conditions are satisfied for a large class of mixing processes satisfying suitable moment inequalities.

1 Introduction

Due to the wide range of its applications in statistics, from parametric to nonparametric estimation, the theory of empirical processes has been widely investigated since the early work of Donsker [9]. For independent observations, Donsker proved in 1952 the weak convergence of the empirical distribution function to a Brownian bridge, which provides as a straightforward consequence the asymptotic behavior of the Kolmogorov-Smirnov statistic.

Dudley [12], [13], extended this result to empirical processes indexed by a class of functions $\mathcal{F}$ and gave a precise definition of weak convegence in the non separable space $\ell^\infty(\mathcal{F})$ of bounded functionals from $\mathcal{F}$ to $\mathbb{R}$. When $\mathcal{F} = \{\mathbf{1}_C, C \in \mathcal{C}\}$ reduces to indicators of sets, he also provided two different kinds of conditions on $\mathcal{F}$ for the empirical process to converge to a $\ell^\infty(\mathcal{F})$-valued Gaussian random variable. The first one involves combinatorial arguments ($\mathcal{C}$ is a V-C class) and is strongly related to the notion of universal entropy. The second one is an integrability condition on the entropy with inclusion of the class $\mathcal{C}$.

In 1982, Pollard [21] proved an empirical central limit theorem under an integrability condition on the universal entropy of $\mathcal{F}$. This condition is optimal in a sense: when replacing universal entropy by the classical metric entropy, it is the minimimal condition (in term of the metric entropy) ensuring the limiting Gaussian process to have continuous sample paths.

Independently, in 1984 Bass [4] obtained the convergence of set-indexed partial sum processes under a minimal condition in terms of entropy with inclusion. His proof is based on a delicate adaptive truncation in the chaining procedure. The main technique is a systematic use of Bernstein's inequality (or, equivalently, of Rosenthal's inequality) which provides the needed truncation level. The same tools and techniques were used in 1987 by Ossiander [18] in the context of empirical processes to obtain optimal conditions on $\mathcal{F}$ in terms of entropy with bracketing (the analogue of entropy with inclusion for classes of functions).

The theory of empirical processes was soon extented to different types of dependent variables. An early result of Billingsley [6] extended Donsker's theorem to (uniformly) ϕ-mixing sequences. For more general classes of functions, see for instance Phillip [19], Doukhan, Léon and Portal [10], Massart [16] or Andrews and Pollard [2].

In terms of entropy with bracketing, the most complete work seems to be that of Doukhan, Massart and Rio [11] in a β-mixing framework. They showed in particular that the right norm to consider in order to measure the size of the brackets is related to the dependence stucture of the variables (see also Rio [25] Chapter 8, for more details on this subject). After a careful reading of this paper, we infer that suitable maximal inequalities are the only essential tools to carry out the chaining procedure.

The paper is organized as follows: Section 2 is devoted to the mathematical background. In Section 3 we state our main result: Theorem 3.3 provides conditions in terms of maximal inequalities for the empirical process indexed by finite sets of functions which are sufficient to derive the asymptotic equicontinuity of the empirical process. In Section 4 we show how to derive such inequalities from Rosenthal-type inequalities. We obtain the tightness of the process under Ossiander's condition in the i.i.d. case and Doukhan, Massart and Rio's in the β-mixing case. The condition we propose for nonuniform ϕ-mixing sequences are new, to our knowledge. In Theorem 5.1 of Section 5, we recall some recent conditions yielding the finite-dimensional convergence of the empirical process, which together with Section 4 imply empirical central limit theorems under various dependence conditions (cf. Theorem 5). The proof of Theorem 3.3 (an adaptation of [11]) is postponed until Section 6. Some technical lemmas concerning the β-mixing case are presented in the Appendix.

2 Empirical process, weak convergence and tightness in $\ell^\infty(\mathcal{F})$.

Let $X = (X_i)_{i\in\mathbb{Z}}$ be a strictly stationary sequence of random elements with values in a Polish space $\mathcal{X}$ with common marginal distribution P. Denote by P_n the empirical probability measure and by Z_n the centered and normalized empirical measure:

$$P_n = \frac{1}{n}\sum_{i=1}^{n}\delta_{X_i}, \quad Z_n = \sqrt{n}(P_n - P)\,.$$

Let $\mathcal{F}$ be a class of measurable functions from $\mathcal{X}$ to $\mathbb{R}$. The space $\ell^\infty(\mathcal{F})$ is the set of all functions z from $\mathcal{F}$ to $\mathbb{R}$ such that $\|z\|_{\mathcal{F}} = \sup_{f\in\mathcal{F}}|z(f)|$ is finite. A random variable X with values in $\ell^\infty(\mathcal{F})$ is tight if for any positive ϵ there exists a compact set K_ϵ of $(\ell^\infty(\mathcal{F}), \|.\|_{\mathcal{F}})$ such that $\mathbb{P}(X \in K_\epsilon) \geq 1-\epsilon$.

Assume that for every x in $\mathcal{X}$, $\sup_{f\in\mathcal{F}}|f(x) - P(f)|$ is finite. Under this minimal condition, the empirical process $\{Z_n(f), f \in \mathcal{F}\}$ can be viewed as a variable with values in $\ell^\infty(\mathcal{F})$, altough it may not be measurable with respect to the Borel σ-algebra generated by $\|.\|_{\mathcal{F}}$. Nevertheless, we say that Z_n converges weakly to a $\ell^\infty(\mathcal{F})$-valued random variable Z (i.e. Borel measurable) if, for every continuous bounded function h from $(\ell^\infty(\mathcal{F}), \|.\|_{\mathcal{F}})$ to $\mathbb{R}$, the outer expectation $\mathbb{E}^*(h(Z_n))$ converges to $\mathbb{E}(h(Z))$ (see for instance [27] p. 4 for the definition of outer expectations and measures, and more details about weak convergence for non-measurable maps).

The variable Z_n with values in $\ell^\infty(\mathcal{F})$ is asymptotically ρ-equicontinuous if there exists a semimetric ρ on $\mathcal{F}$ such that, for every $\epsilon > 0$,

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\mathbb{P}^*\Big(\sup_{\rho(f,g)\leq\delta,\ f,g\in\mathcal{F}}|Z_n(f)-Z_n(g)|>\epsilon\Big)=0\,,$$

where $\mathbb{P}^*$ stands for the outer probability. In the sequel, for short, we shall omit to put the stars. The $\ell^\infty(\mathcal{F})$-valued variable Z_n converges weakly to a tight limit in $\ell^\infty(\mathcal{F})$ if and only if it is asymptotically ρ-equicontinuous for a semimetric ρ such that $(\mathcal{F},\rho)$ is totally bounded, and every one of its finite-dimensional marginals $(Z_n(f_1),\dots Z_n(f_k))$ converges weakly (see for instance [22], Theorem 10.2).

A random variable G with value in $\ell^\infty(\mathcal{F})$ is a Gaussian process if every one of its finite-dimensional marginals $(G(f_1),\dots,G(f_k))$ is normally distributed. Denote by Γ its covariance function $\Gamma(f,g)=\mathrm{Cov}(G(f),G(g))$ and write $\Gamma(f)=\Gamma(f,f)$ for the associated quadratic form. Let K be the set of functions of $\ell^\infty(\mathcal{F})$ which are continuous with respect to $\Gamma(.)$. A zero-mean Gaussian process G is tight if and only if $\ell^\infty(\mathcal{F})$ is relatively compact with respect to $\Gamma(.)$ and G has almost sure continuous sample paths: $\mathbb{P}(G\in K)=1$. If G is a tight Gaussian process then it is also Gaussian as an $\ell^\infty(\mathcal{F})$-valued random variable: for every element d of the dual of $\ell^\infty(\mathcal{F})$, the real valued random variable $d(G)$ is normally distributed.

3 From maximal inequalities to tightness

Before stating the main theorem of this paper, we need some more definitions.

Definition 3.1. *For any $p\geq 1$, let $\mathbb{L}^p(P)$ be the class of real-valued functions on $(\mathcal{X},P)$ such that $\|f\|_p^p=P(|f|^p)$ is finite. We say that a normed space $(L,\|.\|)$ is an $\mathcal{L}$-space if*

1. *There exist two positive numbers a_1 and a_2 such that $\|f\|_1\leq a_1\|f\|$ for any f in L and $\|f\|\leq a_2\|f\|_\infty$ for any f in $\mathbb{L}^\infty(P)$. In particular, this implies that $\|f-P(f)\|\leq(1+a_1a_2)\|f\|$.*
2. *$\|.\|$ is nondecreasing: if f and g are two elements of L such that $|f|\leq|g|$, then $\|f\|\leq\|g\|$.*

Note that $\mathbb{L}^p(P)$ is an $\mathcal{L}$-space as soon as $p\geq 1$.

Definition 3.2. *Entropy with bracketing: let $\mathcal{F}$ be a subset of an $\mathcal{L}$-space $(L,\|.\|)$. Given (f,g) in $L\times L$ with $f\leq g$, the bracket $[f,g]$ is the set of all functions h with $f\leq h\leq g$. An ε-bracket is a bracket $[f,g]$ with $\|f-g\|\leq\varepsilon$. The entropy with bracketing $\mathbb{H}(\mathcal{F},\varepsilon,\|.\|)$ is the logarithm of the smallest number of ε-brackets needed to cover $\mathcal{F}$.*

Theorem 3.3 below gives sufficient conditions for the process Z_n to be asymptotically equicontinuous. They are in fact general conditions under which Ossiander's method (cf. [18]) works. The main point is to find a truncating function for which the maximum of $Z_n(g)$ over a finite set of m-bounded functions is well controlled.

A typical application of Theorem 3.3 is the following: let $\mathcal{G}$ be any finite set of centered functions of $\mathbb{L}^\infty(P)$, and m,δ two positive numbers such that $\|g\|_\infty\leq m$ and $\|g\|_2\leq\delta$

for any g in $\mathcal{G}$. Set $H = \max(1, \log(|\mathcal{G}|))$ and assume that there exists b in $[1, \infty]$ and a constant C depending only on X such that

$$\mathbb{E}\left(\max_{g\in\mathcal{G}} |Z_n(g)|\right) \leq C\left(\delta\sqrt{H} + \frac{m^{\frac{b+1}{b-1}}\delta^{\frac{2}{1-b}}H}{\sqrt{n}}\right). \tag{3.1}$$

Then the $\ell^\infty(\mathcal{F})$-valued variable Z_n is asymptotically equicontinuous as soon as

$$\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, \|.\|_{2b/(b-1)})}\, du < \infty\,.$$

We shall see in Section 4 that Inequality (3.1) holds for i.i.d. sequences with $b = \infty$, and for some non-uniform ϕ-mixing sequences when b equals 1 or 2 (similar bounds may be obtained for b in $]1, 2[$ via interpolation arguments). In the latter case, the value of b is entirely determined by the decay of the ϕ-mixing coefficients.

The fact that the right norm to consider in order to measure the size of the brackets is related to the dependence structure of the variables has been clearly pointed out by Doukhan, Massart and Rio [11] for absolutely regular sequences (see Section 4.3). In this framework, one obtains a bound similar to (3.1) involving the exact rate of mixing of the sequence (cf. equation (4.12)). This inequality leads to the norm $\|.\|_{2,\beta}$ indexed by the whole sequence $\beta = (\beta_i)_{i\geq 0}$ of β-mixing coefficients of X (see equation (4.8) for the definition of $\|.\|_{2,\beta}$). From the inequality $\|g\|_2 \leq \|g\|_{2,\beta}$, we infer that the $\mathcal{L}$-space $\mathcal{L}_{2,\beta}(P) = \{g : \|g\|_{2,\beta} < \infty\}$ is continuously embedded in $\mathbb{L}^2(P)$. Conversely, we shall see that for b in $[1, \infty]$, the space $\mathbb{L}^{2b/(b-1)}(P)$ is continuously embedded in $\mathcal{L}_{2,\beta}(P)$ as soon as the β-mixing coefficients of X satisfy $\sum k^{b-1}\beta(k) < \infty$. This coincides with the i.i.d. case when $b = \infty$, and yields the same rate as for non-uniform ϕ-mixing sequences when b equals 1 or 2.

In each case, the asymptotic equicontinuity of Z_n may be obtained via the following general Theorem:

Theorem 3.3. *Let X be a stationary sequence with marginal distribution P. Let $(L, \|.\|)$ be an $\mathcal{L}$-space and $\mathcal{F}$ a class of functions of L with envelope function F (i.e. $|f| \leq F$ for any f in $\mathcal{F}$).*

Assume that there exists two functions, m_1 from $\mathbb{N}^* \times \mathbb{R}_+^* \times \mathbb{N}^*$ to $\mathbb{R}^+ \cup \{+\infty\}$, nondecreasing in the second variable and nonincreasing in the third, and m_2 from $\mathbb{N}^*\times\mathbb{N}^*$ to $\mathbb{R}^+$ nonincreasing in the second variable, such that both Conditions 1 and 2 hold:

1. *For any (n, δ, k, l) in $\mathbb{N}^* \times \mathbb{R}_+^* \times \mathbb{N}^* \times \mathbb{N}^*$ and any finite subset $\mathcal{G}$ of centered functions of L with cardinality $|\mathcal{G}| \leq k$, whose elements satisfy $\|g\| \leq \delta$ and $\|g\|_\infty \leq \dfrac{m_1(n,\delta,l)}{m_2(n,k)}$, we have*

$$\mathbb{E}\Big(\max_{g\in\mathcal{G}} |Z_n(g)|\Big) \leq C_1\delta\left(\sqrt{\max(1,\log(k))} + \frac{\max(1,\log(k))}{\sqrt{\max(1,\log(l))}}\right)$$

   for some some constant C_1 depending only on X.

2. *For any (n, δ, k) in $\mathbb{N}^* \times \mathbb{R}_+^* \times \mathbb{N}^*$ and any g of L such that $\|g\| \leq \delta$, we have, setting $m(n,\delta,k) = \dfrac{m_1(n,\delta,k)}{m_2(n,k)}$,*

$$\sqrt{n}\|g\mathbf{1}_{4g>m(n,\delta,k)}\|_1 \leq C_2\delta\sqrt{\max(1,\log(k))}\,,$$

   for some constant C_2 depending only on X.

Assume furthermore that there exists a function M from $\mathbb{N}^* \times \mathbb{R}_+^* \times \mathbb{N}^*$ to $\mathbb{R}^+ \cup \{+\infty\}$ such that both Conditions 3 and 4 hold:

3. *For any finite subset $\mathcal{G}$ of centered functions of L with cardinality $|\mathcal{G}| \leq k$, whose elements satisfy $\|g\| \leq \delta$ and $\|g\|_\infty \leq M(n, \delta, k)$, we have*
$$\mathbb{E}\Big(\max_{g \in \mathcal{G}} |Z_n(g)|\Big) \leq C_3 \delta \sqrt{\max(1, \log(k))} + R(n, \delta, k)\,,$$
for some some constant C_3 depending only on X, and $R(n, \delta, k)$ tending to zero as n tends to infinity.

4. *The envelope function F of the class $\mathcal{F}$ satisfies*
$$\text{for any } (\delta, k), \quad \lim_{n \to \infty} \sqrt{n} \|F \mathbf{1}_{4F > M(n,\delta,k)}\|_1 = 0.$$

Then the $\ell^\infty(\mathcal{F})$-valued variable Z_n is asymptotically $\|.\|$-equicontinuous as soon as
$$\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, \|.\|)}\, du < \infty\,. \tag{3.2}$$

Remark 3.4. It is well known that (3.2) implies the existence of an envelope function F belonging to L. In many cases the choice $M \equiv m$ ensures that Condition 4 is satisfied for any F in L (in that case Condition 3 is realized with $C_3 = 2C_1$ and $R \equiv 0$). However, it may happen (see Section 4.3) that the good truncating function M depends on the envelope F.

Remark 3.5. If the function $m_1 \equiv \infty$ satisfies the assumptions of Theorem 3.3 (note that it always satisfies Condition 2), then one can take $M \equiv \infty$ as well, so that Conditions 3 and 4 are automatically realized (see the preceding remark). In such a case, the result can be improved by replacing entropy with bracketing by metric entropy (i.e. the logarithm of the smallest number of balls of radius u with respect to $\|.\|$ necessary to cover $\mathcal{F}$).

Remark 3.6. We have written the theorem with $\sqrt{.}$ because this function appears when considering classical Rosenthal-type inequalities. In fact the result remains true when replacing $\sqrt{.}$ by any increasing subadditive function $\varphi(.)$ in Conditions 1, 2 and 3. The entropy condition (3.2) then becomes
$$\int_0^1 \varphi(\mathbb{H}(\mathcal{F}, u, \|.\|))\, du < \infty\,.$$

Remark 3.7. Let m_1, m_2, M and R be four functions satisfying Conditions 1, 2 and 3 of Theorem 3.3. It follows from Proposition 6.2 that there exists a positive constant K depending only on X such that, for any class $\mathcal{F}$ whose elements satisfy $\|f\| \leq \delta$,
$$\mathbb{E}\Big(\sup_{f \in \mathcal{F}} |Z_n(f)|\Big) \leq K \int_0^\delta \sqrt{1 \vee \mathbb{H}(\mathcal{F}, u, \|.\|)} du + 2\sqrt{n} \|F \mathbf{1}_{2F > M(n,\delta)}\|_1 + 2R(n, \delta)\,,$$
where F is the envelope function of $\mathcal{F}$ and we set $R(n, \delta) = R(n, (1 + a_1 a_2)\delta, \mathbb{N}(\delta))$, $M(n, \delta) = M(n, (1 + a_1 a_2)\delta, \mathbb{N}(\delta))$ and $\mathbb{N}(\delta) = \exp(\mathbb{H}(\mathcal{F}, \delta, \|.\|)$. In particular, if (3.1) holds then $\|.\| = \|.\|_{2b/(b-1)}$, $a_1 = a_2 = 1$, $R_n \equiv 0$ and
$$M(n, \delta) = 2\delta \left(\frac{n}{1 \vee \mathbb{H}(\mathcal{F}, \delta, \|.\|_{2b/(b-1)})} \right)^{\frac{b-1}{2b+2}}\,.$$
In the i.i.d. case (i.e. $b = \infty$), this is the same bound as in Theorem 2.14.2 of [27].

4 From Rosenthal's inequality to maximal inequalities

4.1 The i.i.d. case

In this section, we assume that the sequence $(X_i)_{i\in\mathbb{Z}}$ is i.i.d. If f is a centered function of $\mathbb{L}^p(P)$ for $p \geq 2$, Rosenthal's inequality yields (see for instance Pinelis [20])

$$\|Z_n(f)\|_p \leq K\Big(\sqrt{p}\|f\|_2 + \frac{p}{\sqrt{n}}\Big\|\max_{1\leq i\leq n}|f(X_i)|\Big\|_p\Big)\,, \tag{4.1}$$

for some absolute constant K. Now let $\mathcal{G}$ be any finite set of centered functions of $\mathbb{L}^\infty(P)$, and m, δ two positive numbers such that $\|g\|_\infty \leq m$ and $\|g\|_2 \leq \delta$ for any g in $\mathcal{G}$. From (4.1) we get that

$$\max_{g\in\mathcal{G}}\|Z_n(g)\|_p \leq K\Big(\sqrt{p}\delta + \frac{pm}{\sqrt{n}}\Big)\,. \tag{4.2}$$

Since, for $p \geq 2$,

$$\mathbb{E}\Big(\max_{g\in\mathcal{G}}|Z_n(g)|\Big) \leq \Big\|\max_{g\in\mathcal{G}}|Z_n(g)|\Big\|_p \leq \Big(\sum_{g\in\mathcal{G}}\mathbb{E}|Z_n(g)|^p\Big)^{1/p} \leq |\mathcal{G}|^{1/p}\max_{g\in\mathcal{G}}\|Z_n(g)\|_p\,,$$

we obtain, applying (4.2) with $p = 2H = 2\max(1, \log|\mathcal{G}|)$

$$\mathbb{E}\left(\max_{g\in\mathcal{G}}|Z_n(g)|\right) \leq C\left(\delta\sqrt{H} + \frac{mH}{\sqrt{n}}\right) \tag{4.3}$$

for some absolute constant C.

Set $H(k) = \max(1, \log(k))$. From (4.5) it is clear that the functions defined by

$$m_2 \equiv 1, \quad m(n,\delta,k) = m_1(n,\delta,k) = \delta\sqrt{\frac{n}{H(k)}}$$

satisfy Conditions 1 and 2 of Theorem 3.3 for the space $L = \mathbb{L}^2(P)$ equipped with the norm $\|.\|_2$ (note that Condition 2 is satisfied with $C_2 = 4$). Clearly, the choice $M \equiv m$ ensures that both Conditions 3 and 4 are satisfied, so that the $\ell^\infty(\mathcal{F})$-valued variable Z_n is asymptotically $\|.\|_2$-equicontinuous as soon as

$$\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, \|.\|_2)}du < \infty\,.$$

This result was first obtained by Ossiander [18]. Andersen, Giné, Ossiander and Zinn [1] weakened the bracketing assumption and used majorizing measure instead of metric entropy.

4.2 An extension to martingales

Let $(X_i)_{i\in\mathbb{Z}}$ be a stationary sequence and $\mathcal{M}_i = \sigma(X_k, k \leq i)$. For any f in $\mathbb{L}^1(P)$, define the variable

$$W_n(f) = \frac{1}{\sqrt{n}}\sum_{i=1}^n\Big(f(X_i) - \mathbb{E}(f(X_i)|\mathcal{M}_{i-1})\Big)$$

and the two numbers $d_1(f)$ and $d_2(f)$

$$d_1(f) = \left\|\mathbb{E}\left(|f(X_1)|\middle|\mathcal{M}_0\right)\right\|_\infty \text{ and } d_2(f) = \left\|\mathbb{E}\left((f(X_1))^2\middle|\mathcal{M}_0\right)\right\|_\infty^{1/2} .$$

From [20] again, we have the bounds

$$\|W_n(f)\|_p \leq K\Big(\sqrt{p}d_2(f) + \frac{p}{\sqrt{n}}\Big\|\max_{1\leq i\leq n}|f(X_i) - \mathbb{E}(f(X_i)|\mathcal{M}_{i-1})|\Big\|_p\Big) . \tag{4.4}$$

Arguing as in Section 4.1 we infer that, for any finite set $\mathcal{G}$ of functions of $\mathbb{L}^\infty(P)$, and m, δ two positive numbers such that $\|g\|_\infty \leq m$ and $d_2(g) \leq \delta$ for any g in $\mathcal{G}$,

$$\mathbb{E}\left(\max_{g\in\mathcal{G}}|W_n(g)|\right) \leq C\left(\delta\sqrt{H} + \frac{mH}{\sqrt{n}}\right), \tag{4.5}$$

where $H = \max(1, \log|\mathcal{G}|)$.

Looking at the proof of Theorem 3.3, we see that it may be adapted to the process W_n by replacing the $\mathbb{L}^1$ norm in Conditions 2 and 4 by the norm d_1. Now, since $d_1(f\mathbf{1}_{4f>m}) \leq 4m^{-1}(d_2(f))^2$, the function m of the preceding section works. We infer that the $\ell^\infty(\mathcal{F})$-valued variable W_n is asymptotically d_2-equicontinuous as soon as

$$\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, d_2)}du < \infty . \tag{4.6}$$

If furthermore any function f of $\mathcal{F}$ satisfies $\mathbb{E}(f(X_1)|\mathcal{M}_0) = P(f)$, then the $\ell^\infty(\mathcal{F})$-valued variable W_n coincides with Z_n, and the latter is asymptotically equicontinuous under the entropy condition (4.6).

Let us give two applications of this result

1. Assume that $X_i = (Y_i, Y_{i-1})$ where $(Y_i)_{i\in\mathbb{Z}}$ is a stationary Markov chain with values in an arbitrary space $(E, \mathcal{E})$ and with transition kernel K. Let $\mathcal{F}$ be a class of $\mathbb{L}^2(P)$ such that $\mathbb{E}(f(X_1)|\mathcal{M}_0) = P(f)$ for any f in $\mathcal{F}$. In that case, the $\ell^\infty(\mathcal{F})$-valued variable Z_n is asymptotically d_2-equicontinuous as soon as (4.6) holds, where the norm d_2 is given by

$$(d_2(f))^2 = \sup_{x\in E}\int |f(x,y)|^2 K(x, dy) .$$

2. Assume that $X_i = (\varepsilon_i, Y_i)$ where $\varepsilon = (\varepsilon_i)_{i\in\mathbb{Z}}$ is stationary with marginal distribution P_ε, $Y = (Y_i)_{i\in\mathbb{Z}}$ is i.i.d. with marginal distribution P_Y, and ε is independent of Y. Set $\|g\|_{\varepsilon,p}^p = P_\varepsilon(|g|^p)$ and $\|h\|_{Y,p}^p = P_Y(|h|^p)$. Let $\mathcal{F}_\varepsilon$ be a class of M-bounded functions of $\mathbb{L}^\infty(P_\varepsilon)$, $\mathcal{F}_Y$ be a class of centered functions of $\mathbb{L}^2(P_Y)$ with envelope F_Y and

$$\mathcal{F} = \mathcal{F}_\varepsilon \times \mathcal{F}_Y = \{g \times h, g \in \mathcal{F}_\varepsilon, h \in \mathcal{F}_Y\} .$$

Clearly, $\mathbb{E}(f(X_1)|\mathcal{M}_0) = 0$ for any f in $\mathcal{F}$. Now, if $g_l \leq g \leq g_u$ and $h_l \leq h \leq h_u$, then we have $(g \times h)_l \leq g \times h \leq (g \times h)_u$, where

$$\begin{aligned}
(g_{l+} \times h_{l+}) + (g_{u-} \times h_{u-}) + (g_{l-} \times h_{u+}) + (g_{u+} \times h_{l-}) &= (g \times h)_l \\
(g_{l-} \times h_{l-}) + (g_{u+} \times h_{u+}) + (g_{l+} \times h_{u-}) + (g_{u-} \times h_{l+}) &= (g \times h)_u
\end{aligned}$$

$f_+ = f \vee 0$ and $f_- = f \wedge 0$. If furthermore $|h_u| \vee |h_l| \le F_Y$ and $|g_u| \vee |g_l| \le M$, we easily obtain that

$$d_2\left((g \times h)_u - (g \times h)_l\right) \le 4(\|F_Y\|_{Y,2}\|g_l - g_u\|_{\epsilon,\infty} + M\|h_l - h_u\|_{Y,2}) .$$

From this inequality, we infer that

$$\mathbb{H}(\mathcal{F}, 4u(M + \|F_Y\|_{2,Y}), d_2) \le \mathbb{H}(\mathcal{F}_\varepsilon, u, \|.\|_{\varepsilon,\infty}) + \mathbb{H}(\mathcal{F}_Y, u, \|.\|_{Y,2}) .$$

We conclude that the $\ell^\infty(\mathcal{F})$-valued variable Z_n is asymptotically d_2-equicontinuous as soon as

$$\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}_\varepsilon, u, \|.\|_{\varepsilon,\infty})}du < \infty \quad \text{and} \quad \int_0^1 \sqrt{\mathbb{H}(\mathcal{F}_Y, u, \|.\|_{Y,2})}du < \infty .$$

Setting $\mathcal{M}_{\varepsilon,0} = \sigma(\varepsilon_i, i \le 0)$, the conclusion remains true if we use the norm $(d_{\varepsilon,2}(g))^2 = \|\mathbb{E}(|g(\varepsilon_1)|^2|\mathcal{M}_{\varepsilon,0})\|_\infty$ instead of $\|.\|_{\varepsilon,\infty}$ in the first entropy condition.

The result of this section is a particular case of a remarkable work of Nishiyama [17] (see Theorem 4.2 and Corollary 4.3 therein). In this paper no stationarity assumption is made, the natural random distances

$$d_{2,n}(f) = \Big(\frac{1}{n}\sum_{i=1}^n \mathbb{E}(|f(X_i)|^2|\mathcal{M}_{i-1})\Big)^{1/2}$$

are used instead of d_2, and the condition on $\mathcal{F}$ is expressed in terms of partitioning entropy (in the spirit of Theorem 2.11.9 of [27]). The general context of $\ell^\infty(\mathcal{F})$-valued martingale difference arrays is also considered.

4.3 Absolutely regular sequences

We first recall the definition of the β-mixing coefficient between two σ-algebras (cf. Rozanov and Volkonskii [26]). Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Given two σ-algebras $\mathcal{U}$ and $\mathcal{V}$ of $\mathcal{A}$, the β-mixing coefficient is defined by

$$2\beta(\mathcal{U}, \mathcal{V}) = \sup \sum_{(i,j)\in I\times J} |\mathbb{P}(U_i \cap V_j) - \mathbb{P}(U_i)\mathbb{P}(V_j)| ,$$

where the supremum is taken over all the finite partitions $(A_i)_{i\in I}$ and $(B_j)_{j\in J}$ respectively $\mathcal{U}$ and $\mathcal{V}$-measurable.

Now, let $X = (X_i)_{i\in\mathbb{Z}}$ be a stationary sequence with marginal distribution P and $\mathcal{M}_0$ be the σ-algebra defined by $\mathcal{M}_0 = \sigma(X_i, i \le 0)$. Define the coefficients $\beta_{\infty,k}$ of X by

$$\beta_{\infty,k}(n) = \sup\{\beta\left(\mathcal{M}_0, \sigma(X_{i_1}, \ldots, X_{i_k})\right), 0 < n \le i_1 \le \cdots \le i_k\} \tag{4.7}$$

In this section we shall only consider $\beta_{\infty,\infty}$ coefficients. For the sake of simplicity we omit the indexes, and put

$$\beta(0) = 0 \quad \text{and for } n > 0, \quad \beta(n) = \beta_{\infty,\infty}(n) .$$

The sequences X is said to be absolutely regular as soon as $\beta(n)$ tends to zero as n tends to infinity. Let β^{-1} be the inverse cadlag of the decreasing function $t \to \beta([t])$, and Q_g the inverse cadlag of the tail function $t \to P(|g| > t)$.

From now, assume that the series $\sum \beta(k)$ is finite. This implies that β^{-1} is an integrable function with respect to Lebesgue measure on $[0,1]$. Following Doukhan, Massart and Rio [11], define the set $\mathcal{L}_{2,\beta}(P)$ as the set of functions g for which

$$\|g\|_{2,\beta}^2 = \int_0^1 \beta^{-1}(u) Q_g^2(u)\, du < \infty\,. \tag{4.8}$$

It is proved in Lemma 1 of [11] that $(\mathcal{L}_{2,\beta}(P), \|.\|_{2,\beta})$ is a normed subspace of $\mathbb{L}^2(P)$. Since β^{-1} is greater than 1, we get that $\|g\|_2 \leq \|g\|_{2,\beta}$. On the other hand for any function g in $\mathbb{L}^\infty(P)$, Q_g is bouded by $\|g\|_\infty$, and (4.8) implies that $\|g\|_{2,\beta}^2 \leq \|g\|_\infty^2 \sum \beta(k)$. Collecting the above facts and noting that $|f| \leq |g|$ implies $Q_f \leq Q_g$, we infer that $(\mathcal{L}_{2,\beta}(P), \|.\|_{2,\beta})$ is an $\mathcal{L}$-space.

We now try to find a function m which satisfies the assumptions of Theorem 3.3. To this end, we approximate the original sequence by a sequence of independent random variables (see again [11]). The main tool to perform this approximation is Berbee's coupling lemma (cf. [5]), which we recall hereafter:

Lemma 4.1. *Let X and Y be two random variables taking their values in the Borel spaces S_1 and S_2 respectively, and let U be a random variable with uniform distribution on $[0,1]$, independent of (X,Y). There exists a random variable $Y^* = f(X,Y,U)$, where f is a measurable function from $S_1 \times S_2 \times [0,1]$ into S_2, such that:*

1. *Y^* is independent of X and has the same distribution as Y.*
2. *$\mathbb{P}(Y \neq Y^*) = \beta(\sigma(X), \sigma(Y))$.*

Let q be any positive integer. Starting from Lemma 4.1 we construct by induction a sequence of random variables $(X_i^0)_{i>0}$ such that:

a) For any $i \geq 0$, the random variable $U_i^0 = (X_{iq+1}^0, \ldots, X_{iq+q}^0)$ has the same distribution as $U_i = (X_{iq+1}, \ldots, X_{iq+q})$.

b) The sequence $(U_{2i}^0)_{i\geq 0}$ is i.i.d and so is $(U_{2i+1}^0)_{i\geq 0}$.

c) For any $i \geq 0$, $\mathbb{P}(U_i \neq U_i^0) \leq \beta(q)$.

Let $\mathcal{G}$ be any finite set of centered functions of $\mathbb{L}^\infty(P)$ and m, δ two positive numbers such that $\|g\|_\infty \leq m$ and $\|g\|_{2,\beta} \leq \delta$ for any g in $\mathcal{G}$. Setting $Z_n^0 = n^{-1/2}\sum_{i=1}^n (\delta_{X_i^0} - P)$, we have

$$\mathbb{E}\left(\max_{g\in\mathcal{G}} |Z_n(g)|\right) \leq \mathbb{E}\left(\max_{g\in\mathcal{G}} |Z_n^0(g)|\right) + \mathbb{E}\left(\max_{g\in\mathcal{G}} |Z_n(g) - Z_n^0(g)|\right)\,. \tag{4.9}$$

Clearly

$$|Z_n(g) - Z_n^0(g)| \leq 2\frac{\|g\|_\infty}{\sqrt{n}} \sum_{i=1}^n \mathbf{1}_{X_i \neq X_i^0}\,,$$

and, since $\mathbb{P}(X_i \neq X_i^0) \leq \beta(q)$, we obtain

$$\mathbb{E}\left(\max_{g\in\mathcal{G}} |Z_n(g) - Z_n^0(g)|\right) \leq 2m\beta(q)\sqrt{n} \tag{4.10}$$

It remains to control the first term on right hand in (4.9). Define the random variables $Y_k^0(g)$ by

$$Y_k^0(g) = \sum_{i=kq+1}^{(k+1)q\wedge n} g(X_i^0)\,.$$

For $Y_k^0(g)$, we have the upper bounds

$$\|Y_k^0(g)\|_\infty \le q\|g\|_\infty \le qm \quad \text{and} \quad \|Y_k^0(g)\|_2 \le 4\|g\|_{2,\beta} \le 4\delta\,,$$

the second inequality following from Rio's covariance inequality [23]. Since the random variables $(Y_{2k}^0)_{k\ge 0}$ are independent (as well as $(Y_{2k+1}^0)_{k\ge 0}$), we obtain, applying twice (4.5),

$$\mathbb{E}\left(\max_{g\in\mathcal{G}} |Z_n^0(g)|\right) \le D\Big(\delta\sqrt{H} + \frac{mqH}{\sqrt{n}}\Big)\,,$$

for some absolute constant D. Combining this inequality with (4.10) yields

$$\mathbb{E}\left(\max_{g\in\mathcal{G}} |Z_n(g)|\right) \le D\Big(\delta\sqrt{H} + \frac{mqH}{\sqrt{n}} + \sqrt{n}m\beta(q)\Big)\,. \tag{4.11}$$

We now choose the integer q: define $q(x) = \min\{p \in \mathbb{N}^* : \beta(p) \le px\}$. Taking $q = q(H/n)$ in (4.11), we obtain

$$\mathbb{E}\left(\max_{g\in\mathcal{G}} |Z_n(g)|\right) \le 2D\Big(\delta\sqrt{H} + \frac{mHq(H/n)}{\sqrt{n}}\Big)\,. \tag{4.12}$$

Set $H(k) = \max(1, \log(k))$. From (4.12) it is clear that the functions defined by

$$m_1(n,\delta,k) = 8\delta\sqrt{\frac{n}{H(k)}}, \quad m_2(n,k) = q(H(k)/n), \quad m(n,\delta,k) = \frac{8\delta}{q(H(k)/n)}\sqrt{\frac{n}{H(k)}}$$

satisfy Condition 1 of Theorem 3.3 for the space $(L, \|.\|) = (\mathcal{L}_{2,\beta}(P), \|.\|_{2,\beta})$. In fact it also satisfies Condition 2 (cf. Lemma 4 in [11] and Lemma 7.2 in the Appendix). There seems to be no reason why the function m should be a good truncating function for any envelope F in $\mathcal{L}_{2,\beta}(P)$. We shall see in the Appendix (Lemma 7.3) how to find the second function M from inequality (4.11).

Combining the above facts with the results of the Appendix, we infer that the assumptions of Theorem 3.3 are satisfied for the $\mathcal{L}$-space $(L, \|.\|) = (\mathcal{L}_{2,\beta}(P), \|.\|_{2,\beta})$. This implies that the $\ell^\infty(\mathcal{F})$-valued variable Z_n is asymptotically $\|.\|_{2,\beta}$-equicontinuous as soon as

$$\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, \|.\|_{2,\beta})}du < \infty\,. \tag{4.13}$$

This result is due to [11] and extends on Ossiander's [18] for i.i.d sequences (in which case $\|.\|_{2,\beta} = \|.\|_2$). Let us see how it applies to standard mixing rates, (cf. [11] for a much deeper discussion):

1. If $\sum k^{b-1}\beta(k)$ is finite for some $b \ge 1$, then (4.13) holds as soon as

$$\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, \|.\|_{2b/(b-1)})}du < \infty\,.$$

2. Assume that $\beta(k) = O(b^k)$ for some b in $]0, 1[$, and denote by φ the convex function $\varphi(x) = x^2 \log(1 + |x|)$. The space $\mathbb{L}(\varphi, P)$ of real valued functions f such that

$$\|f\|_\varphi = \inf\Big\{c > 0,\ \mathbb{E}\Big(\varphi\Big(\frac{|f(X_0)|}{c}\Big)\Big) \le 1\Big\} < \infty$$

is continuously embedded in $\mathcal{L}_{2,\beta}(P)$. In particular, (4.13) holds as soon as

$$\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, \|.\|_\varphi)}du < \infty\,.$$

We conclude this section with some bibliographic notes: Rio [25], Chapter 8, considers the minimal case $\sum \beta(k) < \infty$. If $\mathcal{F}$ is a class of bounded functions, he gives entropy conditions on $\mathcal{F}$ which imply asymptotic tightness of Z_n. The proof uses Goldstein's coupling Theorem (see [14]), an analogue of coupling for Harris chains. Comparing to (4.13), the conditions that he obtains are less good for classes of smooth functions but strictly better if $\mathcal{F}$ consists of indicators of sets. As an application, he proves a central limit theorem for the empirical distribution function under the minimal condition $\sum \beta(k) < \infty$, which cannot be obtained via (4.13).

The same author (Rio [24]) uses again Goldstein's theorem together with symmetrization arguments to obtain conditions in terms of uniform entropy of the class $\mathcal{F}$. In particular, Theorem 3 of this paper is an extension to the absolutely regular case of Pollard's empirical central limit theorem (cf. [21]). Results of this type are of a special interest for VC-classes of sets or VC-Major classes (see for instance [27], Chapter 2.6). For these particular classes in a β-mixing context, see also Arcones and Yu [3].

4.4 Nonuniform ϕ-mixing sequences

We first recall the definition of the ϕ-mixing coefficient between two σ-algebras (cf. Ibragimov (1962)). Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Given two σ-algebras $\mathcal{U}$ and $\mathcal{V}$ of $\mathcal{A}$, the ϕ-mixing coefficient is defined by

$$\phi(\mathcal{U}, \mathcal{V}) = \sup\{\|\mathbb{P}(V|\mathcal{U}) - \mathbb{P}(V)\|_\infty\ , V \in \mathcal{V}\}\,.$$

Now, let $X = (X_i)_{i\in\mathbb{Z}}$ be a stationary sequence with marginal distribution P and $\mathcal{M}_0$ be the σ-algebra defined by $\mathcal{M}_0 = \sigma(X_i, i \le 0)$. As in (4.7), define the coefficients $\phi_{\infty,k}$ of X by

$$\phi_{\infty,k}(n) = \sup\{\phi\left(\mathcal{M}_0, \sigma(X_{i_1}, \dots, X_{i_k})\right), 0 < n \le i_1 \le \cdots \le i_k\} \tag{4.14}$$

Recall that $\phi_{\infty,k}$-mixing is more restrictive than $\beta_{\infty,k}$-mixing in the sense that $\beta_{\infty,k}(n) \le \phi_{\infty,k}(n)$. When $k < \infty$, we call these coefficients non-uniform, because they only control the dependence between the past σ-algebra $\mathcal{M}_0$ and any k-tuple $(X_{i_1}, \dots, X_{i_k})$ of the future, while the classical uniform $\phi_{\infty,\infty}$ or $\beta_{\infty,\infty}$ coefficients control the dependence between the past and the whole future $\sigma(X_i, i \ge n)$. Nonuniform coefficients are of a special interest for stationary random fields, for which classical uniform coefficients are much too restrictive. All the results we shall give in this section may be extended to random fields indexed by $\mathbb{Z}^d$ with the appropriate extension of (4.14). Note also that dealing with $\phi_{\infty,k}$ allows to consider nonergodic sequences (or fields), while $\phi_{\infty,\infty}$-mixing sequences as well as absolutely regular sequences are necessarily ergodic (cf. Section 5).

Now from Proposition 1 and Corollary 4(a) in Dedecker [8] (see also Lemma 3 therein) we have, for any centered function f of $\mathbb{L}^\infty(P)$ and $p \ge 3$:

1. Suppose that $\sum k\phi_{\infty,2}(k) < \infty$. Then, for any centered function f of $\mathbb{L}^\infty(P)$ and any two numbers m, δ such that $\|f\|_\infty \leq m$ and $\|f\|_2 \leq \delta$,

$$\|Z_n(f)\|_p \leq K_1(\phi)\left(\delta\sqrt{p} + \frac{pm^3}{\delta^2\sqrt{n}}\right).$$

Let $\mathcal{G}$ be any finite set of centered functions of $\mathbb{L}^\infty(P)$, and m, δ two positive numbers such that $\|g\|_\infty \leq m$ and $\|g\|_4 \leq \delta$ for any g in $\mathcal{G}$. Arguing as in Section (4.1) and setting $H = \max(1, \log(|\mathcal{G}|))$, we obtain

$$\mathbb{E}\left(\max_{g\in\mathcal{G}} |Z_n(g)|\right) \leq C_1(\phi)\left(\delta\sqrt{H} + \frac{m^3 H}{\delta^2\sqrt{n}}\right) \tag{4.15}$$

Set $H(k) = \max(1, \log(k))$. From (4.15) it is clear that the functions defined by

$$m_2 \equiv 1, \quad (m(n,\delta,k))^3 = (m_1(n,\delta,k))^3 = \delta^3\sqrt{\frac{n}{H(k)}}$$

satisfy Conditions 1 and 2 of Theorem 3.3 for the space $L = \mathbb{L}^4(P)$ equipped with the norm $\|.\|_4$ (note that Condition 2 is satisfied with $C_2 = 4^3$). Clearly, the choice $M \equiv m$ ensures that both Conditions 3 and 4 are satisfied, so that the $\ell^\infty(\mathcal{F})$-valued variable Z_n is asymptotically $\|.\|_4$-equicontinuous as soon as

$$\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, \|.\|_4)}du < \infty .$$

2. Assume that $\phi_{\infty,2}(k) = O(k^{-b})$ for some b in $]1,2[$. Then, for any centered function f of $\mathbb{L}^\infty(P)$ and any two numbers m, δ such that $\|f\|_\infty \leq m$ and $\|f\|_2 \leq \delta$,

$$\|Z_n(f)\|_p \leq K_2(\phi)\left(\delta\sqrt{p} + \frac{pm^3\delta^{-2} \vee pm^{\frac{b+1}{b-1}}\delta^{\frac{2}{1-b}}}{\sqrt{n}}\right).$$

Let $\mathcal{G}$ be any finite set of centered functions of $\mathbb{L}^\infty(P)$, and m, δ two positive numbers such that $\|g\|_\infty \leq m$ and $\|g\|_{2b/(b-1)} \leq \delta$ for any g in $\mathcal{G}$. Arguing as above, we obtain

$$\mathbb{E}\left(\max_{g\in\mathcal{G}} |Z_n(g)|\right) \leq C_2(\phi)\left(\delta\sqrt{H} + \frac{Hm^3\delta^{-2} \vee Hm^{\frac{b+1}{b-1}}\delta^{\frac{2}{1-b}}}{\sqrt{n}}\right) \tag{4.16}$$

Define $G_\delta(m) = m^3\delta^{-2} \vee m^{\frac{b+1}{b-1}}\delta^{\frac{2}{1-b}}$ and take

$$m_2 \equiv 1, \quad m(n,\delta,k) = m_1(n,\delta,k) = G_\delta^{-1}\Big(\delta\sqrt{\frac{n}{H(k)}}\Big),$$

so that Condition 1 of Theorem 3.3 is satisfied. Next, write

$$\begin{aligned}\sqrt{n}\|g\mathbf{1}_{4g>m(n,\delta,k)}\|_1 &\leq \frac{\sqrt{H(k)}}{\delta}\|gG_\delta(4g)\|_1 \\ &\leq \sqrt{H(k)}\Big((4/\delta)^3 P(|g|^4) + (4/\delta)^{\frac{b+1}{b-1}} P\big(|g|^{\frac{2b}{b-1}}\big)\Big).\end{aligned}$$

Since $\|g\|_4 \leq \|g\|_{2b/(b-1)}$, the last inequality implies that Condition 2 is satisfied with $C_2 = 4^3 + 4^{(b+1)/(b-1)}$ for the space $L = \mathbb{L}^{2b/(b-1)}(P)$ equipped with the norm $\|.\|_{2b/(b-1)}$. Taking $M \equiv m$, the same arguments apply to Conditions 3 and 4. We conclude that the $\ell^\infty(\mathcal{F})$-valued variable Z_n is asymptotically $\|.\|_{2b/(b-1)}$-equicontinuous as soon as

$$\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, \|.\|_{2b/(b-1)})}du < \infty .$$

3. Assume that $\sum \phi_{\infty,1}(k)$ is finite. Then, for any function f in $\mathbb{L}^\infty(P)$,

$$\|Z_n(f)\|_p \leq K_3(\phi)\sqrt{p}\|f\|_\infty .$$

Arguing as above, we infer that the function $m = \infty$ satisfies the assumptions of Theorem 3.3 for the space $L = \mathbb{L}^\infty(P)$ equipped with the norm $\|.\|_\infty$. Consequently, the $\ell^\infty(\mathcal{F})$-valued variable Z_n is asymptoticaly $\|.\|_\infty$-equicontinuous as soon as

$$\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, \|.\|_\infty)}du < \infty .$$

These results are new, to our knowledge. The comparison between uniform β-mixing (see Section 4.3) and non-uniform ϕ-mixing allows us to make the following conjecture: if, for a positive integer n, the series $\sum k^{n-1}\phi_{\infty,n}(k)$ is finite and the class $\mathcal{F}$ satisfies

$$\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, \|.\|_{2n/(n-1)})}du < \infty ,$$

then the $\ell^\infty(\mathcal{F})$-valued variable Z_n is asymptotically $\|.\|_{2n/(n-1)}$-equicontinuous.

5 Empirical central limit theorems

We first recall a central limit theorem for non uniform mixing sequences (cf. Dedecker [7], Corollary 3):

Theorem 5.1. *Let $X = (X_i)_{i\in\mathbb{Z}}$ be a strictly stationary sequence of random variables with value in a Polish space $\mathcal{X}$ with common marginal distribution P. Let $T : \mathcal{X}^\mathbb{Z} \mapsto \mathcal{X}^\mathbb{Z}$ be the shift operator: $(T(x))_i = x_{i+1}$. Denote by $\mathcal{I}$ the σ-algebra of T-invariant Borel sets of $\mathcal{X}^\mathbb{Z}$. Let f be an element of $\mathbb{L}^2(P)$. If the mixing coefficients of X satisfy either*

$$\sum_{k>0} \int_0^{\beta_{\infty,1}(k)} Q_f^2(u)du < \infty \quad or \quad \sum_{k>0} \phi_{\infty,1}(k) < \infty ,$$

then $Z_n(f)$ converges weakly to $\varepsilon\sqrt{\Gamma_\mathcal{I}(f)}$, where ε is a standard Gaussian random variable independent of $X^{-1}(\mathcal{I})$ and

$$\Gamma_\mathcal{I}(f) = \sum_{k\in\mathbb{Z}} \mathrm{Cov}(f(X_0), f(X_k)|X^{-1}(\mathcal{I})) . \tag{5.1}$$

Applying the non-ergodic version of Ibragimov-Billingsley's central limit theorem for martingales with stationary differences, we infer that the conclusion of Theorem 5.1 remains true if $\mathbb{E}(f(X_1)|\mathcal{M}_0) = P(f)$ (in that case, the sum in (5.1) reduces to the conditional variance term). Note also that, if $\beta_{\infty,2}(n)$ tends to zero as n tends to infinity, then $\sigma(X_0, X_k)$ is independent of $X^{-1}(\mathcal{I})$ and consequently $\Gamma_\mathcal{I}(f)$ is constant. The above facts together with the results of Section 4 yield:

Theorem 5.2. *Let $X = (X_i)_{i \in \mathbb{Z}}$ be a stationary sequence with marginal distribution P and $\mathcal{F}$ be a class of functions of $\mathbb{L}^2(P)$. Consider the following assumptions:*

1. $\mathbb{E}(f(X_1)|\mathcal{M}_0) = P(f)$ *for any f in $\mathcal{F}$, and* $\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, d_2)}du < \infty$.
2. $\sum_{k>0} \beta(k) < \infty$ *and* $\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, \|.\|_{2,\beta})}du < \infty$.
3. $\sum_{k>0} k\phi_{\infty,2}(k) < \infty$ *and* $\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, \|.\|_4)}du < \infty$.
4. $\phi_{\infty,2}(k) = O(k^{-b})$ *for some b in $]1, 2[$, and* $\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, \|.\|_{2b/(b-1)})}du < \infty$.
5. $\sum_{k>0} \phi_{\infty,1}(k) < \infty$ *and* $\int_0^1 \sqrt{\mathbb{H}(\mathcal{F}, u, \|.\|_\infty)}du < \infty$.

If either 2., 3. or 4. holds, then the $\ell^\infty(\mathcal{F})$-valued variable Z_n converges weakly to a zero-mean tight Gaussian process with covariance function

$$\Gamma(f, g) = \sum_{k \in \mathbb{Z}} \mathrm{Cov}(f(X_0), g(X_k))\,.$$

If either 1. or 5. holds then the $\ell^\infty(\mathcal{F})$-valued variable Z_n converges weakly to a tight random variable whose conditional distribution with respect to $X^{-1}(\mathcal{I})$ is that of a zero-mean tight Gaussian process with covariance function

$$\Gamma_{\mathcal{I}}(f, g) = \sum_{k \in \mathbb{Z}} \mathrm{Cov}(f(X_0), g(X_k)|X^{-1}(\mathcal{I}))\,.$$

6 Chaining

The chaining procedure we borrow here is an extension of Ossiander's method due to Doukhan, Massart and Rio [11].

Our purpose is to prove the stochastic equicontinuity of the empirical process over a class of functions $\mathcal{F}$ equipped with the norm $\|.\|$ and fulfilling the entropy condition (3.2). More precisely, we shall prove that

$$\lim_{\sigma \to 0} \limsup_{n \to +\infty} \mathbb{E}\left(\sup_{\|f-g\| \le \sigma} |Z_n(f) - Z_n(g)| \right) = 0. \tag{6.1}$$

Set $\mathcal{F}^\sigma := \{f - g, \quad f, g \in \mathcal{F} \text{ and } \quad \|f - g\| \le \sigma\}$, and consider the familly of bounded functions $\mathcal{F}^\sigma(M) := \{(f - g)\mathbf{1}_{4F \le M}, \quad f, g \in \mathcal{F} \text{ and } \quad \|f - g\| \le \sigma\}$. Clearly, for each positive M,

$$\mathbb{E}\left(\sup_{f \in \mathcal{F}^\sigma} |Z_n(f)| \right) \le \mathbb{E}\left(\sup_{f \in \mathcal{F}^\sigma(M)} |Z_n(f)| \right) + 4\sqrt{n}\|F\mathbf{1}_{4F>M}\|_1. \tag{6.2}$$

Let $\mathbb{N}(\sigma) = \exp(\mathbb{H}(\mathcal{F}^\sigma, \sigma, \|.\|))$ and take

$$M = M_n = M(n, (1 + a_1 a_2)\sigma, \mathbb{N}(\sigma))\,. \tag{6.3}$$

From (6.2) and Condition 4 of Theorem 3.3 we infer that

$$\limsup_{n\to\infty} \mathbb{E}\left(\sup_{f\in\mathcal{F}^\sigma} |Z_n(f)|\right) \leq \limsup_{n\to\infty} \mathbb{E}\left(\sup_{f\in\mathcal{F}^\sigma(M_n)} |Z_n(f)|\right) ,$$

so that (6.1) holds as soon as

$$\lim_{\sigma\to 0}\limsup_{n\to\infty} \mathbb{E}\left(\sup_{f\in\mathcal{F}^\sigma(M_n)} |Z_n(f)|\right) = 0 . \tag{6.4}$$

Note that if f belongs to $\mathcal{F}^\sigma(M_n)$ then $\|f\| \leq \sigma$ and $2\|f\|_\infty \leq M_n$. Note also that $\mathbb{H}(\mathcal{F}^\sigma(M_n), x) \leq \mathbb{H}(\mathcal{F}^\sigma, x) \leq 2\mathbb{H}(\mathcal{F}, x/2)$. From these two elementary remarks, we infer that (6.4) will follow from Proposition 6.2 below.

Before stating this proposition, we introduce as in [11] an appropriate upper bound H for the entropy $\mathbb{H}$:

Lemma 6.1. *For any class of functions $\mathcal{F}$, there exists a nonincreasing upper bound H of $\mathbb{H}(\mathcal{F}, ., \|.\|)$ such that $x \to x^4 H(x)$ is nondecreasing and, for any δ in $[0,1]$,*

$$\psi(\delta) = \int_0^\delta \sqrt{1 \vee H(u)}du \leq 4\int_0^\delta \sqrt{1\vee \mathbb{H}(\mathcal{F}, u, \|.\|)}du$$

See [11] for a proof.

Proposition 6.2. *Let m_1, m_2, M and R be four functions satisfying Conditions 1, 2 and 3 of Theorem 3.3. Let $\mathcal{F}$ be any class of functions of L satisfying (3.2). Define $\mathbb{N}(\sigma) = \exp(\mathbb{H}(\mathcal{F}, \sigma, \|.\|)$ and $M_n = M(n, (1 + a_1a_2)\sigma, \mathbb{N}(\sigma))$. Assume that any function f of $\mathcal{F}$ satisfy the inequalities $2\|f\|_\infty \leq M_n$ and $\|f\| \leq \sigma$. Then there exists a constant K not depending on (n, σ), such that*

$$\mathbb{E}\left(\sup_{f\in\mathcal{F}} |Z_n(f)|\right) \leq K\psi(\sigma) + 2R(n, (1 + a_1a_2)\sigma, \mathbb{N}(\sigma)) .$$

Proof. We start with some preliminary notations.

Definitions 6.3. *Let $\delta_0 = \sigma$ and $\delta_k = 2^{-k}\delta_0$. For each k, we choose a covering of $\mathcal{F}$ by brackets $B_{j,k} := [g_{j,k}, h_{j,k}]$ for $1 \leq j \leq J_k$, with $\|h_{j,k} - g_{j,k}\| \leq \delta_k$ and $J_k \leq \mathbb{N}(\delta_k)$. Since $2|f| \leq M_n$ for all $f \in \mathcal{F}$, we may assume that $|h_{j,k} - g_{j,k}| \leq M_n$. In each bracket $B_{j,k}$, fix a point $f_{j,k}$ belonging to $\mathcal{F}$. Let $\pi_k f := f_{\psi_k f,k}$ where ψ_k is a mapping from $\mathcal{F}$ to $[1, J_k]$ defined by $\psi_k f := \min\{j \in [1, J_k]\ : f \in B_{j,k}\}$. Set $\Delta_k f := h_{\psi_k f,k} - g_{\psi_k f,k}$. Clearly*

$$|f - \pi_k f| \leq \Delta_k f \quad \textit{and} \quad \|\Delta_k f\| \leq \delta_k \tag{6.5}$$

Finally define $\mathbf{H}(\delta) = \sum_{\delta_k\geq\delta} H(\delta_k)$, $\mathbf{N}(\delta) := [\exp(\mathbf{H}(\delta))]$, and $N(\delta_k) := [\exp(H(\delta_k))]$.

Definitions 6.4. *Set $2\alpha = 3(1 + a_1a_2)$. Define the nonincreasing sequences (b_j) and (q_j) and the sequence (m_j) by*

$$b_k = m_1(n, \alpha\delta_k, \mathbf{N}(\delta_{k+1})), \quad q_k = 4m_2(n, \mathbf{N}(\delta_{k+1})), \quad m_k = \frac{b_k}{q_k} . \tag{6.6}$$

For any $f \in \mathcal{F}$, define $\nu := \nu(f) := N \wedge \min\{j \geq 0,\ \ \Delta_j f > m_j\}$, where N is defined by

$$N := \min\left\{k \geq 0\ : \delta_k \leq \frac{\psi(\sigma)}{\sqrt{n}}\right\} \vee 1.$$

Without loss of generality, suppose that $\mathbb{H}(\mathcal{F},\sigma,\|.\|)\geq 1$ (note that the same is necessarily true for $H(\sigma)$ and $\mathbf{H}(\sigma)$). We now list all the inequalities which are useful to prove Proposition 6.2.

Facts 6.5.

$$\text{If}\quad |g|\leq h\quad\text{then}\quad |Z_n(g)|\leq Z_n(h)+2\sqrt{n}\|h\|_1 \tag{6.7}$$

$$\text{For all } k\in\mathbb{N},\quad \mathbf{H}(\delta_{k+1})\leq 17\mathbf{H}(\delta_k) \tag{6.8}$$

There exist positive constants $K_1,\ldots,K_5$ such that, for any integer $1\leq k\leq N-1$, we have

$$\mathbb{E}\left(\sup_{f\in\mathcal{F}}\left|Z_n\left(\Delta_{k-1}f\mathbf{1}_{\Delta_k f>\frac{b_{k-1}}{q_k}}\prod_{i=0}^{k-1}\mathbf{1}_{\Delta_i f\leq m_i}\right)\right|\right)\leq K_1\delta_k\sqrt{\mathbf{H}(\delta_k)} \tag{6.9}$$

$$\mathbb{E}\left(\sup_{f\in\mathcal{F}}\left|Z_n\left(\Delta_k f\mathbf{1}_{m_k<\Delta_k f\leq\frac{b_{k-1}}{q_k}}\prod_{i=0}^{k-1}\mathbf{1}_{\Delta_i f\leq m_i}\right)\right|\right)\leq K_2\delta_k\sqrt{\mathbf{H}(\delta_k)} \tag{6.10}$$

$$\mathbb{E}\left(\sup_{f\in\mathcal{F}}\left|Z_n\left((\pi_k(f)-\pi_{k-1}(f))\mathbf{1}_{\Delta_k f\leq\frac{b_{k-1}}{q_k}}\prod_{i=0}^{k-1}\mathbf{1}_{\Delta_i f\leq m_i}\right)\right|\right)\leq K_3\delta_k\sqrt{\mathbf{H}(\delta_k)} \tag{6.11}$$

$$\mathbb{E}\left(\sup_{f\in\mathcal{F}}\left|Z_n\left(\Delta_{N-1}f\prod_{i=0}^{N-1}\mathbf{1}_{\Delta_i f\leq m_i}\right)\right|\right)\leq K_4\delta_N\sqrt{\mathbf{H}(\delta_N)} \tag{6.12}$$

$$\text{For } 0\leq k\leq N-1,\quad \sqrt{n}\,\|\Delta_k f\mathbf{1}_{\Delta_k f>m_k}\|_1\leq K_5\delta_k\sqrt{\mathbf{H}(\delta_k)} \tag{6.13}$$

Proof of Facts 6.5. Inequality (6.7) is straightforward, and (6.8) follows directly from Lemma 6.1.

We first prove (6.9). Clearly the supremum in (6.9) reduces to a maximum over a finite set of functions with cardinality less than $\mathbf{N}(\delta_k)$. Moreover if g is any function belonging to this set, we have $\|g-P(g)\|\leq(1+a_1a_2)\delta_{k-1}\leq\alpha\delta_{k-1}$ and

$$\|g-P(g)\|_\infty\leq m_{k-1}\leq\frac{m_1(n,\alpha\delta_{k-1},\mathbf{N}(\delta_k))}{m_2(n,\mathbf{N}(\delta_k))}\,,$$

and (6.9) follows from Condition 1 of Theorem 3.3.

We now prove (6.10). As for the proof of (6.9), the supremum in (6.10) reduces to a maximum over a finite set of functions with cardinality less than $\mathbf{N}(\delta_k)\leq\mathbf{N}(\delta_{k+1})$. If g is any function belonging to this set, we have $\|g-P(g)\|\leq(1+a_1a_2)\delta_k\leq\alpha\delta_{k-1}$ and

$$\|g-P(g)\|_\infty\leq\frac{b_{k-1}}{q_k}\leq\frac{m_1(n,\alpha\delta_{k-1},\mathbf{N}(\delta_k))}{m_2(n,\mathbf{N}(\delta_{k+1}))}$$

Condition 1 of Theorem 3.3 yields then

$$\mathbb{E}\left(\sup_{f\in\mathcal{F}}\left|Z_n\left(\Delta_k f\mathbf{1}_{m_k<\Delta_k f\leq\frac{b_{k-1}}{q_k}}\prod_{i=0}^{k-1}\mathbf{1}_{\Delta_i f\leq m_i}\right)\right|\right)\leq\alpha C_1\delta_{k-1}\left(\sqrt{\mathbf{H}(\delta_{k+1})}+\frac{\mathbf{H}(\delta_{k+1})}{\sqrt{\mathbf{H}(\delta_k)}}\right)$$

The last inequality, together with (6.8) yields (6.10).

To prove (6.11), note that the supremum in (6.11) reduces to a maximum over a finite set of functions with cardinality less than $\mathbf{N}(\delta_k)\leq\mathbf{N}(\delta_{k+1})$. If g is any function belonging to this set, we have $\|g-P(g)\|\leq(1+a_1a_2)(\delta_k+\delta_{k-1})=\alpha\delta_{k-1}$ and

$$\|g-P(g)\|_\infty\leq\frac{4b_{k-1}}{q_k}=\frac{m_1(n,\alpha\delta_{k-1},\mathbf{N}(\delta_k))}{m_2(n,\mathbf{N}(\delta_{k+1}))}\,,$$

and we conclude as for (6.10).

To prove (6.12), note that the supremum in (6.12) reduces to a maximum over a finite set of functions with cardinality less than $\mathbf{N}(\delta_{N-1}) \le \mathbf{N}(\delta_N)$. If g is any function belonging to this set, we have $\|g - P(g)\| \le (1 + a_1a_2)\delta_{N-1} \le \alpha\delta_{N-1}$ and

$$\|g - P(g)\|_\infty \le m_{N-1} \le \frac{m_1(n, \alpha\delta_{N-1}, \mathbf{N}(\delta_N))}{m_2(n, \mathbf{N}(\delta_N))},$$

and we conclude as in (6.9).

To prove (6.13) note that $\|\Delta_k f\| \le \delta_k \le \alpha\delta_k$. Since

$$m_k = \frac{m_1(n, \alpha\delta_k, \mathbf{N}(\delta_{k+1}))}{4m_2(n, \mathbf{N}(\delta_{k+1}))},$$

(6.13) follows from (6.8) and Condition 2 of Theorem 3.3. This completes the proof of Facts 6.5.

Chaining.
Arguing as in Doukhan, Massart and Rio [11] we have the decomposition

$$\mathbb{E}\left(\sup_{f\in\mathcal{F}} |Z_n(f)|\right) \le \mathbb{E}_1 + \mathbb{E}_2 + \mathbb{E}_3 + \mathbb{E}_4 + \mathbb{E}_5 + \mathbb{E}_6, \tag{6.14}$$

where

$$\begin{aligned}
\mathbb{E}_1 &= \mathbb{E}\left(\sup_{f\in\mathcal{F}} |Z_n(\pi_0 f)|\right)\\
\mathbb{E}_2 &= \mathbb{E}\left(\sup_{f\in\mathcal{F}} |Z_n(f - \pi_0 f)\mathbf{1}_{\nu(f)=0}|\right)\\
\mathbb{E}_3 &= \sum_{k=1}^{N-1} \mathbb{E}\left(\sup_{f\in\mathcal{F}} |Z_n(f - \pi_{k-1} f)\mathbf{1}_{\nu(f)=k}\mathbf{1}_{\Delta_k f > \frac{b_{k-1}}{q_k}}|\right)\\
\mathbb{E}_4 &= \sum_{k=1}^{N-1} \mathbb{E}\left(\sup_{f\in\mathcal{F}} |Z_n(f - \pi_k f)\mathbf{1}_{\nu(f)=k}\mathbf{1}_{\Delta_k f \le \frac{b_{k-1}}{q_k}}|\right)\\
\mathbb{E}_5 &= \mathbb{E}\left(\sup_{f\in\mathcal{F}} |Z_n(f - \pi_{N-1} f)\mathbf{1}_{\nu(f)=N}|\right)\\
\mathbb{E}_6 &= \sum_{k=1}^{N-1} \mathbb{E}\left(\sup_{f\in\mathcal{F}} \left|Z_n\left((\pi_k(f) - \pi_{k-1}(f))\mathbf{1}_{\Delta_k f \le \frac{b_{k-1}}{q_k}}\mathbf{1}_{\nu(f)\ge k}|\right)\right|\right)
\end{aligned}$$

Control of $\mathbb{E}_1$. When f runs over the set $\mathcal{F}$, $\pi_0 f$ runs over a finite class of functions with cardinality less than $\mathbb{N}(\sigma)$. Moreover, we have the bounds $\|\pi_0 f - P(\pi_0 f)\|_\infty \le M_n$ and $\|\pi_0 f - P(\pi_0 f)\| \le (1 + a_1a_2)\sigma$. Hence according to Condition 3 of Theorem 3.3, we have

$$\mathbb{E}_1 \le C_3(1 + a_1a_2)\sigma\sqrt{H(\sigma)} + R(n, (1 + a_1a_2)\sigma, \mathbb{N}(\sigma)). \tag{6.15}$$

Control of $\mathbb{E}_2$. We deduce from $\{\nu(f) = 0\} = \{\Delta_0 > m_0\}$, (6.5) and (6.7) that

$$\mathbb{E}_2 \le \mathbb{E}\left(\sup_{f\in\mathcal{F}} |Z_n(\Delta_0 f\mathbf{1}_{\Delta_0 f > m_0})|\right) + 2\sqrt{n}\sup_{f\in\mathcal{F}} \|\Delta_0 f\mathbf{1}_{\Delta_0 f > m_0}\|_1. \tag{6.16}$$

Using inequality (6.13), the second term in the right hand side is bounded by

$$\sqrt{n} \sup_{f \in \mathcal{F}} \|\Delta_0 f \mathbf{1}_{\Delta_0 f > m_0}\|_1 \leq K_5 \delta_0 \sqrt{\mathbf{H}(\delta_0)}. \tag{6.17}$$

Note that

$$\|\Delta_0 f \mathbf{1}_{\Delta_0 f > m_0} - \mathbb{E}\left(\Delta_0 f \mathbf{1}_{\Delta_0 f > m_0}\right)\| \leq (1 + a_1 a_2)\sigma \quad \text{using (6.5), and}$$

$$\|\Delta_0 f \mathbf{1}_{\Delta_0 f > m_0} - \mathbb{E}\left(\Delta_0 f \mathbf{1}_{\Delta_0 f > m_0}\right)\|_\infty \leq M_n \, .$$

To control the first term of Inequality (6.16), note that the supremum reduces to a maximum over a finite set of functions with cardinality less than $\mathbb{N}(\sigma)$. Hence according to Condition 3 of Theorem 3.3, we get

$$\mathbb{E}\left(\sup_{f \in \mathcal{F}} |Z_n\left(\Delta_0 f \mathbf{1}_{\Delta_0 f > m_0}\right)|\right) \leq C_3(1+a_1a_2)\sigma\sqrt{H(\sigma)} + R(n, (1+a_1a_2)\sigma, \mathbb{N}(\sigma)) \tag{6.18}$$

Collecting Inequalities (6.16), (6.17), (6.18) we deduce

$$\mathbb{E}_2 \leq C\delta_0\sqrt{\mathbf{H}(\delta_0)} + R(n, (1 + a_1 a_2)\sigma, \mathbb{N}(\sigma)) \tag{6.19}$$

Control of $\mathbb{E}_3$. We deduce from (6.5), (6.7) that

$$\mathbb{E}_3 \leq \sum_{k=1}^{N-1} \mathbb{E}\left(\sup_{f \in \mathcal{F}} |Z_n(\Delta_{k-1}(f)\mathbf{1}_{\nu(f)=k}\mathbf{1}_{\Delta_k f > \frac{b_{k-1}}{q_k}}|\right) + \sum_{k=1}^{N-1} 2\sqrt{n} \sup_{f \in \mathcal{F}} \left\|\Delta_{k-1}(f)\mathbf{1}_{\nu(f)=k}\mathbf{1}_{\Delta_k f > \frac{b_{k-1}}{q_k}}\right\|_1 \tag{6.20}$$

In order to control the first term on right hand in (6.20), note that

$$\{\nu(f) = k\} = \{m_k < \Delta_k f, \ \Delta_i f \leq m_i, \quad i = 0 \ldots, k-1\},$$

Hence Inequality (6.9) yields

$$\sum_{k=1}^{N-1} \mathbb{E}\left(\sup_{f \in \mathcal{F}} |Z_n(\Delta_{k-1}(f)\mathbf{1}_{\nu(f)=k}\mathbf{1}_{\Delta_k f > \frac{b_{k-1}}{q_k}}|\right) \leq K_1 \sum_{k=1}^{N-1} \delta_k \sqrt{\mathbf{H}(\delta_k)}. \tag{6.21}$$

On the other hand, since

$$\mathbf{1}_{\nu(f)=k}\mathbf{1}_{\Delta_k f > \frac{b_{k-1}}{q_k}} \leq \mathbf{1}_{\Delta_k f > \frac{b_{k-1}}{q_k}}\mathbf{1}_{\Delta_{k-1} f \leq \frac{b_{k-1}}{q_{k-1}}} \leq \mathbf{1}_{\Delta_k f > \Delta_{k-1} f}\mathbf{1}_{\Delta_k f > \frac{b_k}{q_k}},$$

we deduce, using Inequality (6.13)

$$\sqrt{n} \sup_{f \in \mathcal{F}} \left\|\Delta_{k-1}(f)\mathbf{1}_{\nu(f)=k}\mathbf{1}_{\Delta_k f > \frac{b_{k-1}}{q_k}}\right\|_1 \leq \sqrt{n} \sup_{f \in \mathcal{F}} \left\|\Delta_k(f)\mathbf{1}_{\Delta_k f > \frac{b_k}{q_k}}\right\|_1 \leq K_5 \delta_k \sqrt{\mathbf{H}(\delta_k)} \, .$$

This last bound together with (6.21) and (6.20) yields

$$\mathbb{E}_3 \leq (K_1 + 2K_5) \sum_{k=1}^{N-1} \delta_k \sqrt{\mathbf{H}(\delta_k)} \, . \tag{6.22}$$

Control of $\mathbb{E}_4$. We deduce from (6.5) and (6.7) that

$$\mathbb{E}_4 \leq \sum_{k=1}^{N-1} \mathbb{E}\left(\sup_{f\in\mathcal{F}} |Z_n(\Delta_k(f)\mathbf{1}_{\nu(f)=k}\mathbf{1}_{\Delta_k f\leq \frac{b_{k-1}}{q_k}}|\right) + \sum_{k=1}^{N-1} 2\sqrt{n}\sup_{f\in\mathcal{F}}\left\|\Delta_k(f)\mathbf{1}_{\nu(f)=k}\mathbf{1}_{\Delta_k f\leq \frac{b_{k-1}}{q_k}}\right\|_1 . \tag{6.23}$$

In order to control the first term on right hand in (6.23), note that

$$\{\nu(f)=k\} = \{m_k < \Delta_k f,\ \Delta_i f \leq m_i, \quad i = 0\dots, k-1\}.$$

Hence Inequality (6.10) yields

$$\sum_{k=1}^{N-1} \mathbb{E}\left(\sup_{f\in\mathcal{F}} |Z_n(\Delta_k(f)\mathbf{1}_{\nu(f)=k}\mathbf{1}_{\Delta_k f\leq \frac{b_{k-1}}{q_k}}|\right) \leq K_2 \sum_{k=1}^{N-1} \delta_k \sqrt{\mathbf{H}(\delta_k)}\,. \tag{6.24}$$

On the other hand since,

$$\mathbf{1}_{\nu(f)=k}\mathbf{1}_{\Delta_k f\leq \frac{b_{k-1}}{q_k}} \leq \mathbf{1}_{\Delta_k f > \frac{b_k}{q_k}}$$

we deduce, using inequality (6.13),

$$\sqrt{n}\sup_{f\in\mathcal{F}}\left\|\Delta_k(f)\mathbf{1}_{\nu(f)=k}\mathbf{1}_{\Delta_k f\leq \frac{b_{k-1}}{q_k}}\right\|_1 \leq \sqrt{n}\sup_{f\in\mathcal{F}}\left\|\Delta_k(f)\mathbf{1}_{\Delta_k f > \frac{b_k}{q_k}}\right\|_1 \leq K_5\delta_k\sqrt{\mathbf{H}(\delta_k)}\,.$$

This last bound together with (6.24) and (6.23) yield

$$\mathbb{E}_4 \leq (K_2 + 2K_5)\sum_{k=1}^{N-1} \delta_k\sqrt{\mathbf{H}(\delta_k)}\,. \tag{6.25}$$

Control of $\mathbb{E}_5$. We deduce from (6.5), (6.7) that

$$\mathbb{E}_5 \leq \mathbb{E}\left(\sup_{f\in\mathcal{F}} |Z_n(\Delta_{N-1}(f)\mathbf{1}_{\nu(f)=N}|\right) + 2\sqrt{n}\sup_{f\in\mathcal{F}}\|\Delta_{N-1}(f)\|_1\,. \tag{6.26}$$

Since $\prod_{i=0}^{N-1}\mathbf{1}_{\Delta_i f\leq m_i} = \mathbf{1}_{\nu(f)=N}$, we infer from Inequality (6.12) that

$$\mathbb{E}\left(\sup_{f\in\mathcal{F}} |Z_n(\Delta_{N-1}(f)\mathbf{1}_{\nu(f)=N}|\right) \leq K_4\delta_N\sqrt{\mathbf{H}(\delta_N)}\,. \tag{6.27}$$

Now

$$\sqrt{n}\sup_{f\in\mathcal{F}}\|\Delta_{N-1}(f)\|_1 \leq a_1\sqrt{n}\sup_{f\in\mathcal{F}}\|\Delta_{N-1}(f)\| \leq a_1\sqrt{n}\delta_{N-1} \leq 2a_1\psi(\sigma),$$

the last inequality following from the definition of N (cf. Definitions 6.4). This last bound together with (6.27) and (6.26) yields

$$\mathbb{E}_5 \leq K_4\delta_N\sqrt{\mathbf{H}(\delta_N)} + 4a_1\psi(\sigma)\,. \tag{6.28}$$

Control of $\mathbb{E}_6$. We deduce from

$$\mathbf{1}_{\Delta_k f \leq \frac{b_{k-1}}{q_k}} \mathbf{1}_{\nu(f)\geq k} = \mathbf{1}_{\Delta_k f \leq \frac{b_{k-1}}{q_k}} \prod_{i=0}^{k-1} \mathbf{1}_{\Delta_i f \leq m_i}$$

and from inequality (6.11) that

$$\mathbb{E}_6 \leq K_3 \sum_{k=1}^{N-1} \delta_k \sqrt{\mathbf{H}(\delta_k)} \tag{6.29}$$

End of the proof of Proposition 6.2.

Collecting Inequalities (6.14), (6.15), (6.19), (6.22), (6.25), (6.28) and (6.29), we obtain

$$\mathbb{E}\left(\sup_{f\in\mathcal{F}} |Z_n(f)|\right) \leq K\Big\{\psi(\sigma) + \sum_{k=1}^{N} \delta_k \sqrt{\mathbf{H}(\delta_k)}\Big\} + 2R(n, (1+a_1a_2)\sigma, \mathbb{N}(\sigma)), \tag{6.30}$$

for a positive constant K. From the definition of $\mathbf{H}$, we get

$$\begin{aligned} \sum_{k=1}^{N} \delta_k \sqrt{\mathbf{H}(\delta_k)} &\leq \sum_{k\geq 0} \delta_k \Big(\sum_{j\leq k} \sqrt{H(\delta_j)} \Big) \\ &\leq 2\sum_{j\geq 0} \delta_j \sqrt{H(\delta_j)} \leq 4\int_0^{\sigma} \sqrt{H(u)}du = 4\psi(\sigma), \end{aligned}$$

which together with Inequality (6.30) completes the proof of Proposition 6.2

7 Appendix

As in [11] define the functions

$$B(t) = \int_0^t \beta^{-1}(u)du \quad \text{and} \quad \delta_h(\epsilon) = \sup_{t\leq\epsilon} Q_h(t)\sqrt{B(t)}\,.$$

The following Lemmas and their proofs are due to Doukhan, Massart and Rio [11].

Lemma 7.1. *Let h be any nonnegative function in $\mathbb{L}^1(P)$. Then, for any ϵ in $]0,1]$, the following inequality holds*

$$\|h\mathbf{1}_{h>Q_h(\epsilon)}\|_1 \leq \frac{2\epsilon\delta_h(\epsilon)}{\sqrt{B(\epsilon)}}\,. \tag{7.1}$$

In particular if $\delta_h(1) \leq \delta$ and $a\sqrt{\epsilon\beta^{-1}(\epsilon)} \geq \delta$, then

$$\|h\mathbf{1}_{h>a}\|_1 \leq 2\delta\sqrt{\frac{\epsilon}{\beta^{-1}(\epsilon)}}\,. \tag{7.2}$$

Proof. Recall that for any function f in $\mathbb{L}^1(P)$, $\|f\|_1 = \int_0^1 Q_f(u)du$. The quantile function of $h\mathbf{1}_{h>a}$ being equal to $Q_h(u)\mathbf{1}_{u\leq Q_h^{-1}(a)}$, we obtain

$$\|h\mathbf{1}_{h>Q_h(\epsilon)}\|_1 = \int_0^{Q_h^{-1}(Q_h(\epsilon))} Q_h(t)dt \leq \int_0^{\epsilon} Q_h(t)dt\,.$$

Bearing in mind the definition of $\delta_h(\epsilon)$ and using the concavity of B, we infer that

$$\|h\mathbf{1}_{h>Q_h(\epsilon)}\|_1 \leq \delta_h(\epsilon)\int_0^{\epsilon}\frac{dt}{\sqrt{B(t)}} \leq \delta_h(\epsilon)\sqrt{\frac{\epsilon}{B(\epsilon)}}\int_0^{\epsilon}\frac{dt}{\sqrt{t}}\,,$$

proving (7.1). Sinces β^{-1} is nonincreasing, $\epsilon\beta^{-1}(\epsilon) \leq B(\epsilon)$ and both the condition on a and the definition of $\delta_h(\epsilon)$ imply that $a \leq Q_h(\epsilon)$. This yields (7.2) and the proof of Lemma 7.1 is complete.

Next, we use Lemma 7.1 to prove the assertions of Section 4.3.

Lemma 7.2. *The function m defined in Section 4.3 by*

$$m(n,\delta,k) = \frac{8\delta}{q(H(k)/n)}\sqrt{\frac{n}{H(k)}}$$

satisfies Condition 2 of Theorem 3.3.

Proof. Apply Lemma 7.1 with

$$4a = m(n,\delta,k) \quad \text{and} \quad \epsilon = \frac{H(k)}{n}\max(1, q(H(k)/n)-1)\,.$$

By definition of $q(.)$, we obtain $\epsilon \leq \beta(q(H(k)/n))$ and therefore

$$\beta^{-1}(\epsilon) \geq q(H(k)/n) \geq \max(1, q(H(k)/n)-1)\,,$$

which implies that $m(n,\delta,k)\sqrt{\epsilon\beta^{-1}(\epsilon)} \geq 4\delta$. Now, take g in $\mathcal{L}_{2,\beta}(P)$ such that $\|g\|_{2,\beta} \leq \delta$. Since $\delta_g(1) \leq \|g\|_{2,\beta} \leq \delta$, inequality (7.2) holds yielding

$$\|g\mathbf{1}_{4g>m(n,\delta,k)}\|_1 \leq 2\delta\sqrt{\frac{\epsilon}{\beta^{-1}(\epsilon)}} \leq 2\delta\sqrt{\frac{H(k)}{n}}\,,$$

and Condition 2 of Theorem 3.3 holds with $C_2 = 2$.

Lemma 7.3. *Let F be any nonnegative function in $\mathcal{L}_{2,\beta}(P)$. Let $\epsilon(n,k)$ be the unique solution of the equation $nx^2 = B(x)H(k)$. Then the function $M(n,k) = 4Q_F(\epsilon(n,k))$ satisfies Conditions 3 and 4 of Theorem 3.3.*

Proof. Let $\mathcal{G}$ be a finite of centered functions of $\mathcal{L}_{2,\beta}(P)$ with cardinality $|\mathcal{G}| \leq k$, whose elements satisfy $\|g\| \leq \delta$ and $\|g\|_\infty \leq 4Q_F(\epsilon)$. According to (4.11), we have

$$\mathbb{E}\Big(\max_{g\in\mathcal{G}}|Z_n(g)|\Big) \leq D\Big(\delta\sqrt{H(k)} + \frac{4Q_F(\epsilon)qH(k)}{\sqrt{n}} + 4\sqrt{n}Q_F(\epsilon)\beta(q)\Big)\,.$$

Choose $q = \beta^{-1}(\epsilon)$. Since $\epsilon\beta^{-1}(\epsilon) \leq B(\epsilon)$, we have $qH(k) \leq n\epsilon$, which leads to

$$\mathbb{E}\Big(\max_{g\in\mathcal{G}}|Z_n(g)|\Big) \leq 4D\Big(\delta\sqrt{H(k)} + 2\sqrt{n}\epsilon Q_F(\epsilon)\Big)\,.$$

To prove that Condition 3 holds, it remains to see that $\sqrt{n}\epsilon Q_F(\epsilon)$ tends to zero as n tends to infinity. Bearing in mind that $Q_F(\epsilon)\sqrt{B(\epsilon)} \le \delta_F(\epsilon)$, we infer that

$$\sqrt{n}\epsilon Q_F(\epsilon) \le \frac{\sqrt{n}\epsilon\delta_F(\epsilon)}{\sqrt{B(\epsilon)}} \le \delta_F(\epsilon)\sqrt{H(k)}\,.$$

Since $\delta_F(\epsilon)$ tends to zero as n tends to infinity, we infer that Condition 3 is satisfied.

It remains to check Condition 4. By inequality (7.1) of Lemma 7.1, we have

$$\sqrt{n}\|F\mathbf{1}_{F>Q_F(\epsilon)}\|_1 \le \frac{2\sqrt{n}\epsilon\delta_F(\epsilon)}{\sqrt{B(\epsilon)}} \le 2\delta_F(\epsilon)\sqrt{H(k)}\,,$$

and we conclude as above. The proof of Lemma 7.3 is complete.

Bibliography

[1] N. T. Andersen, E. Giné, M. Ossiander and J. Zinn: The central limit theorem and the law of the iterated logarithm for empirical processes under local conditions. *Probability Theory and Related Fields* **77** (1988), 271-305.

[2] D. W. K. Andrews and D. Pollard: An introduction to functional central limit theorems for dependent stochastic processes. *International Statistical Review* **62** (1994), 119-132.

[3] M. A. Arcones and B. Yu: Central limit for empirical and U-processes of stationary mixing sequences. *Journal of Theoretical Probability* **7** (1994), 47-71.

[4] R. F. Bass: Law of the iterated logarithm for set-indexed partial sum processes with finite variance. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **70** (1985), 591-608.

[5] H. C. P. Berbee: *Random Walks with Stationary Increments and Renewal Theory.* Mathematical Centre Tracts **112**, Amsterdam, 1979.

[6] P. Billingsley: *Convergence of Probability Measures.* John Wiley, New York, 1968.

[7] J. Dedecker: A central limit theorem for stationary random fields. *Probability Theory and Related Fields* **110** (1998), 397-426.

[8] J. Dedecker: Exponential inequalities and functional central limit theorems for random fields. *ESAIM, Probability and Statistics* **5** (2001), 77-104.

[9] M. D. Donsker: Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems. *Annals of Mathematical Statistics* **23** (1952), 277-281.

[10] P. Doukhan, J. Léon and F. Portal: Principe d'invariance faible pour la mesure empirique d'une suite de variables aléatoires dépendantes. *Probability Theory and Related Fields* **76** (1987), 51-70.

[11] P. Doukhan, P. Massart and E. Rio: Invariance principles for absolutely regular empirical processes. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques* **31** (1995), 393-427.

[12] R. M. Dudley: Weak convergence of measures on nonseparable metric spaces and empirical measures on Euclidean spaces. *Illinois Journal of Mathematics* **10** (1966), 109-126.

[13] R. M. Dudley: Central limit theorems for empirical measures. *Annals of Probability* **6** (1978), 899-929.

[14] S. Goldstein: Maximal coupling. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **46** (1979), 193-204.

[15] I. A. Ibragimov: Some limit theorems for stationary processes. *Theory of Probability and Its Applications* **7** (1962), 349-382.

[16] P. Massart: Invariance principle for empirical processes : the weakly dependent case, Quelques problèmes de vitesse de convergence pour des mesures empiriques. *Thèse d'état, Université de Paris-Sud,* 1987.

[17] Y. Nishiyama: Weak convergence of some classes of martingales with jumps. *Annals of Probability* **28** (2000), 685-712.

[18] M. Ossiander: A central limit theorem under metric entropy with $\mathbb{L}^2$-bracketing. *Annals of Probability* **15** (1987), 897-919.

[19] W. Phillip: Invariance principles for sums of mixing random elements and the multivariate empirical process. *Colloquia Mathematica Societatis János Bolyai* **36** (1982), 843-873.

[20] I. F. Pinelis: Optimum bounds for the distribution of martingales in Banach spaces. *Annals of Probability* **22** (1994), 1679-1706.

[21] D. Pollard: A central limit theorem for empirical processes. *Journal of the Australian Mathematical Society A* **33** (1982), 235-248.

[22] D. Pollard: *Empirical Processes: Theory and Applications. NSF-CBMS Regional Conference Series in Probabity and Statistics* **2**. Institute of Mathematical Statistics and American Statistical Association, 1990.

[23] E. Rio: Covariance inequalities for strongly mixing processes. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques* **29** (1993), 587-597.

[24] E. Rio: Processus empiriques absolument réguliers et entropie universelle. *Probability Theory and Related Fields* **111** (1998), 585-608.

[25] E. Rio: *Théorie asymptotique des processus aléatoires faiblement dépendants. Collection Mathématiques & Applications* **31**. Springer, Berlin, 2000.

[26] Y. A. Rozanov and V. A. Volkonskii: Some limit theorems for random functions I. *Theory of Probability and Its Applications* **4** (1959), 178-197.

[27] A. W. van der Vaart and J. A. Wellner: *Weak Convergence and Empirical Processes.* Springer, Berlin, 1996.

Jérôme Dedecker
L.S.T.A. Université Pierre et Marie Curie,
4, Place Jussieu
F-75252 Paris Cedex 05
France
dedecker@ccr.jussieu.fr

Sana Louhichi
Laboratoire de Probabilités, Statistiques et modélisation,
Université de Paris-Sud, Bât. 425,
F-91405 Orsay Cedex
France
Sana.Louhichi@math.u-psud.fr