Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER Stochastic Processes and their Applications 115 (2005) 339–358
stochastic processes and their applications
www.elsevier.com/locate/spa

Uniform CLT for empirical process

Samir Ben Hariz*,[1]

Département de Mathématiques, Université du Maine, Av. Olivier Messiaen, BP 535, 72017 Le Mans Cedex, France

Received 6 August 2002; received in revised form 23 September 2003; accepted 14 September 2004
Available online 12 October 2004

Abstract

Empirical processes indexed by classes of functions based on dependent observations are considered. Sufficient conditions in order to satisfy stochastic equicontinuity are given. The derived conditions are in terms of bracketing numbers with respect to a norm arising from a Rosenthal type moment inequality satisfied by the process. The application involves mixing sequences and improves on the result of Andrews and Pollard (Int. Statist. Rev. 62 (1) (1994) 119) for strong mixing, Shao and Yu (Ann. Probab. 24 (4) (1996) 2098) for ρ-mixing sequences, and Csörgő and Mielniczuk (Probab. Theory Relat. Fields 104 (1) (1996) 15) for functions of Gaussian sequences.

MSC: primary 60F05; 60F17

Keywords: Bracketing; Chaining; Empirical processes; Functional central limit theorems; Stochastic equicontinuity; Weakly dependent processes

1. Introduction

Let $(X_i)_{i\geqslant 0}$ be a stationary sequence of real random variables defined on a probability space (Ω, A, P) and $\mathscr{F}$ be a class of real valued functions of real variables.

*Tel.: +33 243 83 3082; fax: +33 243 83 3579
E-mail address: Samir.Ben_Hariz@univ-lemans.fr (S. Ben Hariz).
[1]This work is a part of my Ph.D. Thesis in the University of Paris Sud under the supervision of Professor Jean Bretagnolle.

0304-4149/$ - see front matter
doi:10.1016/j.spa.2004.09.006

Let $l^\infty(\mathscr{F})$ denote the space of bounded real functions defined on $\mathscr{F}$. Given a collection $\mathscr{F}$ one can define a map from $\mathscr{F}$ to $\mathbb{R}$ as follows:

$$Z_n : \mathscr{F} \longrightarrow \mathbb{R}$$
$$f \longmapsto Z_n(f) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} [f(X_i) - \mathbb{E}(f(X_i))].$$

If $\sup_{\mathscr{F}} \sup_x |f(x) - \mathbb{E}(f(X_0))|$ exists and is finite, then the map Z_n is an element of $l^\infty(\mathscr{F})$. Consequently, it makes sense to investigate conditions under which the sequences Z_n converge in law in $l^\infty(\mathscr{F})$ endowed with the uniform topology. A class $\mathscr{F}$ for which this is true is called a *Donsker class*. To prove weak convergence in $l^\infty(\mathscr{F})$, according to Pollard [11] (see also Van Der Vaart [15]) we need the following two conditions.

(i) Convergence of marginal: for all $f_1, \ldots, f_k$ elements of $\mathscr{F}$,

$$(Z_n(f_1), \ldots, Z_n(f_k)) \text{ converges in law.}$$

(ii) There exists a pseudo metric ρ such that $(\mathscr{F}, \rho)$ is totally bounded, and for all $\varepsilon > 0$,

$$\lim_{\delta \to 0} \limsup_{n \to +\infty} \mathbb{P}\left(\sup_{\rho(f,g) < \delta} |Z_n(f - g)| > \varepsilon \right) = 0. \tag{1}$$

The second property is known as the stochastic equicontinuity of the family Z_n. It is useful in proving uniform central limit theorems as well as in other contexts (see for example Andrews [2]).

Convergence of the finite dimensional distributions is proved for many classes of processes. Roughly speaking, property (i) is satisfied as soon as the sequence X is sufficiently weak dependent. Dependence between the past and the future of the process is measured either by mixing coefficients such as α-mixing (strong mixing), ρ-, β- and ϕ-mixing, or by the decay of covariances for functions of Gaussian, linear processes and associated sequences. Therefore, in order to conclude the uniform CLT, it remains to prove the stochastic equicontinuity. And this will be the main purpose of the present paper.

Several results exists in the literature. In what follows we recall some of them with an emphasis on those which are close to the spirit of this work. Let $\|\cdot\|_p$ denote the L^p norm for $p < \infty$. In 1987, Ossiander [9] proved that if the variables are i.i.d then (ii) is fulfilled if

$$\int_0^1 \sqrt{\log N_{[\,]}(\varepsilon, \|\cdot\|_2, \mathscr{F})}\, \mathrm{d}\varepsilon < \infty,$$

where $N_{[\,]}(\varepsilon, \|\cdot\|_2, \mathscr{F})$ is the minimal number of ε-brackets sufficient to cover $\mathscr{F}$ (see Definition 1 below).

This result has been generalized by Doukhan et al. [6] to β-mixing sequences under the summability of the sequence of β-mixing and the following condition on the family $\mathscr{F}$:

$$\int_0^1 \sqrt{\log N_{[\,]}(\varepsilon, \|\cdot\|_{2,\beta}, \mathscr{F})}\,d\varepsilon < \infty,$$

where $\|f\|_{2,\beta}^2 = \int_0^1 \beta^{-1}(u) Q_f^2(u)\,du$. The proof in the two cases was based on exponential inequalities for independent random variables.

On the other hand, using a moment inequality of order 2, Arcones [3] showed that the stochastic equicontinuity of $\{Z_n(f),\ f \in \mathscr{F}\}_{n>0}$ holds when the process X is Gaussian with summable covariance function and if the family satisfies the condition.

$$\int_0^1 N_{[\,]}^{1/2}(x, \|\cdot\|_2, \mathscr{F})\,dx < \infty.$$

Andrews and Pollard [2] have concluded the tightness of the empirical process of a strong mixing sequence under the following hypothesis:

$$\sum_{i>0} i^{p-2} \alpha^{\frac{\gamma}{p+\gamma}}(i) < \infty,$$

$$\sup_{\mathscr{F}} |f| \leqslant 1 \quad \text{and} \quad \int_0^1 x^{-\frac{\gamma}{\gamma+2}} N_{[\,]}^{1/p}(x, \|\cdot\|_2, \mathscr{F})\,dx < \infty,$$

where $p \geqslant 2$ and $\gamma > 0$. Here also, the main tool was a moment inequality of order p.

In view of these results we can see that the conditions ensuring the tightness of the empirical process is a kind of balance between the regularity of the process on the one hand, expressed here in term of weak dependence, and the size or the complexity of the family $\mathscr{F}$ on the other hand, measured here by the bracketing numbers with respect to a norm induced by the process.

A goal of this work is to give a general approach to this problem which generalizes and improves on some existing results. The main result asserts that if the process satisfies a Rosenthal type moment inequality of order p and if $\int_0^1 N_{[\,]}^{1/p}(x, \|\cdot\|_{2,X}, \mathscr{F})\,dx < \infty$, where $\mathscr{F}$ is an uniformly bounded class of functions and $\|\cdot\|_{2,X}$ is an appropriate norm induced by the moment inequality then (ii) is satisfied. The paper is structured as follows, in Section 2 we give the main results, several applications are discussed in Section 3 and Section 4 is devoted to the proofs of results.

2. Main results

Before stating the main result we recall the following definition of bracketing numbers.

Definition 1. Given two functions l and u the bracket $[l, u]$ is the set of all functions f with $l \leqslant f \leqslant u$. Given a norm $\|\cdot\|$ on a space containing $\mathscr{F}$, an ε-bracket for $\|\cdot\|$ is a bracket $[l, u]$ with $\|l - u\| < \varepsilon$. The bracketing number $N_{[\,]}(\varepsilon, \|\cdot\|, \mathscr{F})$ is the minimal number of ε-brackets needed to cover $\mathscr{F}$.

For $p \geqslant 2$ we define two kind of conditions, the first one on the process, and the second one on the family $\mathscr{F}$.

H(p, X). There exists constants $a(p)$ and $b(p)$ such that for every measurable f

$$\mathbb{E}|Z_n(f)|^p \leqslant a(p)\|f\|_{2,X}^p + b(p)n^{1-p/2}\|f\|_\infty^{p-2}\|f\|_{2,X}^2, \tag{2}$$

where $\|\cdot\|_{2,X}$ norm is a norm satisfying[1]:

- $\|\cdot\|_1 \leqslant C\|\cdot\|_{2,X}$ for some positive constant C.
- $|f| \leqslant |g| \Rightarrow \|f\|_{2,X} \leqslant \|g\|_{2,X}$.

H($p, \mathscr{F}$). $\mathscr{F}$ is uniformly bounded and

$$\int_0^1 N_{[\,]}^{1/p}(x, \|\cdot\|_{2,X}, \mathscr{F})\,\mathrm{d}x < \infty. \tag{3}$$

We are now able to state our first result.

Theorem 1. *Let $(X_i)_{i \geqslant 0}$ be a strictly stationary sequence of random variables and $\mathscr{F}$ be a class of functions satisfying* H(p, X) *and* H($p, \mathscr{F}$), *then*: $\forall \varepsilon > 0$,

$$\lim_{\delta \to 0} \limsup_{n \to +\infty} \mathbb{P}\left(\sup_{\|f-g\|_{2,X} < \delta} |Z_n(f-g)| > \varepsilon \right) = 0.$$

The condition under which the family $\mathscr{F}$ is uniformly bounded, may be relaxed by strengthening the condition on the covering numbers and imposing further assumptions on the envelope function of the family. This is what is done in the following theorem.

Theorem 2. *Let $(X_i)_{i \geqslant 0}$ be a strictly stationary sequence of random variables and $\mathscr{F}$ be a class of functions satisfying* H(p, X). *Let $F \geqslant \sup_{f \in \mathscr{F}} |f|$ be a measurable function. Assume that $F \in \mathbb{L}^{r+1}$, for some $r > 1$, and*

$$\int_0^1 N_{[\,]}^{v/p}(x, \|\cdot\|_{2,X}, \mathscr{F})\,\mathrm{d}x < \infty, \tag{4}$$

where $1/v = 1 - \frac{1}{r}(1 - \frac{2}{p})$.

Then the conclusion of Theorem 1 *holds.*

In what follows we are aimed to give sufficient conditions for $\mathscr{F}$ in order to satisfy the stochastic equicontinuity property in the case when the α-mixing coefficient decays exponentially. The result is closely related to the work of Massart [8],

[1]The norm $\|\cdot\|_{2,X}$ is simply the L^2 norm in the independent case and is some norm who extends the L^2 norm to the dependent case. In the latter case this norm depends generally on the process X via the measure of dependence used to control the covariance terms. In particular for α-mixing process, this is simply the $\|\cdot\|_{2,\alpha}$ (see Lemma 2 for the definition).

however, the technique's proof is slightly different. The proof of the next result relies on a Rosenthal type moment inequality, with explicit bounds of the coefficients $a(p)$ and $b(p)$, due to Rio and combined, as usual, with a chaining argument.

Theorem 3. *Let* $(X_i)_{i\geqslant 0}$ *be a stationary sequences and* $\mathscr{F}$ *be a family of functions bounded by* 1. *We assume*

(a) $\alpha(i)\leqslant c\exp(-\alpha i)$, *where* $c>0, \alpha>0$.
(b) $\int_0^1 \log^2 N_{[\,]}(\varepsilon, \|\cdot\|_1, \mathscr{F})\,\mathrm{d}\varepsilon<\infty$.

Then, $\forall\varepsilon>0$,

$$\lim_{\delta\to+\infty}\ \limsup_{n\to+\infty}\ \mathbb{P}\left(\sup_{\|f-g\|_1\leqslant\delta}|Z_n(f-g)|>\varepsilon\right)=0.$$

The previous theorem improves on Massart's result. Indeed, under the same hypothesis of mixing, the assumption on $\mathscr{F}$ was

$$\log N_{[\,]}(\varepsilon, \|\cdot\|_1, \mathscr{F})\leqslant C\left(\frac{1}{\varepsilon}\right)^{\xi}, \quad \xi<1/4.$$

We point out however, that Massart shows a rate of convergence for the given weak invariance principle. We note also that Andrews and Pollard [2] conjectured in their paper that the condition implying the stochastic equicontinuity under the same assumption of mixing, may be

$$\int_0^1 \varepsilon^{-\frac{\gamma}{\gamma+2}}\log^2 N_{[\,]}(\varepsilon, \|\cdot\|_1, \mathscr{F})\,\mathrm{d}\varepsilon<\infty,$$

for some positive constant γ.

Remark. In the independent case, the same method of proof shows that the condition is $\int_0^1 \log^{1/2} N_{[\,]}(\varepsilon, \|\cdot\|_1, \mathscr{F})\varepsilon^{-1/2}\,\mathrm{d}\varepsilon<\infty$. This condition is known to be optimal when $\mathscr{F}$ is the class of all subset of $\mathbb{N}$.

3. Examples of application

In this section, we give some examples for which the hypothesis $\mathrm{H}(p, X)$ is fulfilled and we compare with some existing results. For $\mathrm{H}(p,\mathscr{F})$, we refer the reader to [8,15].

First we recall the following measures of dependence. Suppose $(\Omega, \mathscr{K}, \mathscr{P})$ is a probability space. For any two σ-fields $\mathscr{A}$ and $\mathscr{B}$ of $\mathscr{K}$, we define

$$\alpha(\mathscr{A},\mathscr{B})=\sup|\mathbb{P}(A\cap B)-\mathbb{P}(A)\cdot\mathbb{P}(B)|, \quad A\in\mathscr{A}, \quad B\in\mathscr{B}$$

and

$$\rho(\mathscr{A},\mathscr{B})=\sup|\mathrm{corr}(f,g)|, \quad f\in\mathbb{L}^2(\mathscr{A}), \quad g\in\mathbb{L}^2(\mathscr{B}).$$

The following inequality holds:

$$\rho(\mathscr{A},\mathscr{B})\leqslant\alpha(\mathscr{A},\mathscr{B}).$$

If (X_k) is a sequence of random variables we define

$$c(n)=\sup_k c(G_1^k, G_{k+n}^\infty),$$

where $c=\alpha,\rho$ and G_n^m is the σ field generated by $(X_k, n\leqslant k\leqslant m)$.

3.1. Case of α-mixing

Let (X_k) be a stationary sequence, and for u positive real, set $\alpha(u)=\alpha([u])$, where $[x]$ denotes the integer part of x. We denote the quantile function of $|f(X_0)|$ by Q_f, which is the inverse of the tail function $t\longrightarrow\mathbb{P}(|f(X_0)|>t)$. The following corollary is an immediate consequence of Theorem 1.

Corollary 1. *Let* $\|f\|_{2,X}^2=\int_0^1[\alpha^{-1}(u)]Q_f^2(u)\,\mathrm{d}u$. *If* $\mathrm{H}(p,\mathscr{F})$ *is fulfilled, then* $Z_n(f)$ *converges in* $l^\infty(\mathscr{F})$ *to a Gaussian process indexed by* $\mathscr{F}$ *with* $\|f\|_{2,X}$ *continuous sample paths.*

In particular, this convergence holds whenever the following conditions are satisfied:

$$\text{(H1)}\quad \sum_{i\geqslant 0}(i+1)^{(p-1)/(1-\theta)-1}\alpha(i)<\infty$$

$$\text{(H2)}\quad \int_0^1 N_{[\,]}^{1/p}(x,\|\cdot\|_{2/\theta},\mathscr{F})\,\mathrm{d}x<\infty.$$

To compare with the assumptions of Andrews and Pollard we first note that if $N_r(x)=N_{[\,]}(x,\|\cdot\|_r,\mathscr{F})$, and if $\mathscr{F}$ is bounded above by 1, then $N_r(x^{2/r})\leqslant N_2(x)$. By a change of variable we conclude that (H2) is implied by

$$\text{(H}'\text{2)}\quad \int_0^1 N_{[\,]}^{1/p}(x,\|\cdot\|_2,\mathscr{F})x^{-1+\theta}\,\mathrm{d}x<\infty.$$

The assumptions of Andrews and Pollard are

$$\text{(A1)}\quad \sum_{i\geqslant 0}(i+1)^{p-2}\alpha(i)^{(2-2\theta)/(p\theta+2-2\theta)}<\infty$$

$$\text{(A2)}\quad \int_0^1 N_{[\,]}^{1/p}(x,\|\cdot\|_2,\mathscr{F})x^{-1+\theta}\,\mathrm{d}x<\infty.$$

Now (H1) is weaker for $p>2$, (e.g., for a polynomial rate of mixing, say $\alpha(i)\sim ci^{-a}$, (H1) is satisfied if $a>(p-1)/\,(1-\theta)$ while (A1) is fulfilled if $a>(p-1)(p\theta+2-2\theta)/(2-2\theta)$).

3.2. Case of ρ-mixing

The forthcoming corollary considers ρ-mixing sequences. Its proof relies on a moment inequality established by Shao [13] and the CLT for ρ-mixing sequences (see [10] for example).

Corollary 2. *Let (X_k) be a stationary, ρ-mixing sequence. Assume that $F \in L^{2+\delta}$,*

$$\sum_{i=0}^{\infty} \rho(2^i) < \infty \quad \text{and} \quad \int_0^1 N_{[\,]}^{\eta}(x, \|\cdot\|_2, \mathscr{F})\,\mathrm{d}x < \infty,$$

where η, δ are positive reals. Then $Z_n(f)$ converges in $l^{\infty}(\mathscr{F})$ to a Gaussian process indexed by $\mathscr{F}$ with $\|\cdot\|_2$ continuous sample paths.

The corollary applies to the family of quadrants, and generalizes the result of Shao and Yu [14] in the sense that the continuity of the distribution function is not needed here.

3.3. Case of Gaussian sequences

Let $(X_i)_{i \geqslant 0}$ be a stationary Gaussian sequence satisfying: $\mathbb{E}(X_0) = 0$, $\mathbb{E}(X_0^2) = 1$ and let $r(k) = \mathbb{E}(X_0 X_k)$. To apply Theorem 1, we need a Rosenthal type inequality for partial sums of a function of Gaussian sequences. This is the subject of the following lemma. The lemma handles the particular case when $p = 4$. Let $H_k := (-1)^k p^{(k)}/p$ denotes the kth Hermite's polynomial (p is the density of a standard normal distribution). We recall that the *rank* of a real function f is defined by $\operatorname{rank}(f) = \inf\{k > 0 | \mathbb{E}(H_k(X)f(X)) \neq 0\}$.

Lemma 1. *Let f be a real function and assume that*

$$\sum_{k \geqslant 0} |r(k)|^m < \infty,$$

where $m = \inf(\operatorname{rank}(f), rank(f^2))$. Then there exists a constant $K = K(r(\cdot))$ such that

$$\mathbb{E}\left|\sum_{i=1}^{n} f(X_i) - \mathbb{E}f(X_i)\right|^4 \leqslant K(n^2(\mathbb{E}f^2(X_i))^2 + n\|f\|_{\infty}^2 \mathbb{E}f^2(X_i)).$$

As a consequence of the previous lemma and Theorem 1, we deduce that if r belongs to L^1 and if $\int_0^1 N_{[\,]}^{1/4}(x, \|\cdot\|_2, \mathscr{F})\,\mathrm{d}x < \infty$ where $\mathscr{F}$ is a class or family bounded by 1, then (ii) is satisfied. Since the condition that r belongs to L^1 is sufficient for convergence of marginals (see for example [4]), we have then proved the following corollary.

Corollary 3. *Let (X_i) be a stationary Gaussian sequence such that $\mathbb{E}(X_0) = 0$, $\mathbb{E}(X_0^2) = 1$ and let $r(k) = \mathbb{E}(X_0 X_k)$. Let $\mathscr{F}$ be a family of function bounded by 1. If*

$$r \in \mathbb{L}^1 \quad \text{and} \quad \int_0^1 N_{[\,]}^{1/4}(x, \|\cdot\|_2, \mathscr{F})\,\mathrm{d}x < \infty$$

then $\{Z_n(f), f \in \mathscr{F}\}_{n>0}$ converge in $l^\infty(\mathscr{F})$ to a Gaussian, centered process G indexed by $\mathscr{F}$ with covariance function given by

$$\mathrm{Cov}(G(f), G(g)) = \mathbb{E}(G(f)G(g)) = \sum_{i \in Z} \mathrm{Cov}(f(X_0), g(X_i)).$$

In the particular case when $\mathscr{F} = \{1_{G(\cdot) \leqslant x} : x \in R\}$ where G is some measurable function, the condition that r belongs to L^1 can be relaxed to the following one:

$$\sum_{k=1}^{n} |r^m(k)| < \infty,$$

where m is the Hermite rank of the family $\mathscr{F}$. Indeed, in this case the moment inequality of order 4 will be applied to $\mathscr{F} - \mathscr{F} := \{f - g; (f, g) \in (\mathscr{F}, \mathscr{F})\}$. Since for $f \in \mathscr{F}$ we have $\mathrm{rank}(f^2) \geqslant \mathrm{rank}(f)$ it suffices to have $\sum_{k=1}^{n} |r^m(k)| < \infty$. In addition $N_{[\,]}(x, \|\cdot\|_2, \mathscr{F}) \leqslant \frac{C}{x^2}$ for this family. Thus the result applies and this generalizes Theorem 1 in [5] to the case when the distribution function of $G(X)$ is discontinuous.

4. Proof of main results

For any expressions A and B let us write $A \preccurlyeq B$ if $A \leqslant KB$ for some absolute constant K, and let $[x]$ stand for the integer part of x.

4.1. Proof of Theorem 1

By hypothesis (3), for all integers k there exists a finite sequence of pairs of functions $(f_i^k, \Delta_i^k)_{1 \leqslant i \leqslant N(k)}$, where $N(k) = N_{[\,]}(2^{-k}, \|\cdot\|_{2,X}, \mathscr{F})$ such that:

- $\|\Delta_i^k\|_{2,X} \leqslant 2^{-k}$.
- $\forall f \in \mathscr{F}$ there exists i such that $|f - f_i^k| \leqslant \Delta_i^k$.

We set $(\pi_k(f), \Delta_k(f))$ the first pair (f_i^k, Δ_i^k) which satisfies: $|f - f_i^k| \leqslant \Delta_i^k$. Let q_0, k and q be integers verifying $q_0 \leqslant k \leqslant q$. Following a technique used by Arcones [3] we define a map from $\mathscr{F}$ into a finite subset of $\mathscr{F}$ by:

$$T_k(f) = \pi_k \circ \pi_{k+1} \circ \cdots \circ \pi_q(f).$$

For $1 \leqslant i \leqslant N(q_0)$ let us define

$$E_i = \{f \in \mathscr{F} : T_{q_0}(f) = f_i^{q_0}\}$$

then the sets E_i form a partition of $\mathscr{F}$. For $\delta > 0$ we define

$$F_{i,j} = \{(f, g) \in \mathscr{F} \times \mathscr{F} | f \in E_i, g \in E_j \text{ and } \|f - g\|_{2,X} \leqslant \delta\}.$$

Let now $\Lambda = \{(i,j) | F_{i,j} \neq \emptyset\}$. For every pair in Λ we fix an element of $F_{i,j}$ and denote this pair by $(\Phi_{i,j}, \Psi_{i,j})$.

Let (f,g) be a pair satisfying $\|f-g\|_{2,X}\leqslant\delta$, then necessarily $(f,g)\in F_{i,j}$ for some $(i,j)\in\Lambda$. We write

$$f-g=f-T_{q_0}(f)+T_{q_0}(f)-\Phi_{i,j}+\Phi_{i,j}-\Psi_{i,j}+\Psi_{i,j}-T_{q_0}(g)+T_{q_0}(g)-g,$$

but $T_{q_0}(f)=T_{q_0}(\Phi_{i,j})$ and $T_{q_0}(g)=T_{q_0}(\Psi_{i,j})$, since $f,\Phi_{i,j}$ are in E_i and $g,\Psi_{i,j}$ are in E_j. Consequently

$$\sup_{\|f-g\|_{2,X}\leqslant\delta}|Z_n(f-g)|\leqslant 4\sup_{f\in\mathscr{F}}|Z_n(f-T_{q_0}(f))|+\sup_{(i,j)\in\Lambda}|Z_n(\Phi_{i,j}-\Psi_{i,j})|.$$

Take the expectation in the previous inequality to get

$$\mathbb{E}\left[\sup_{\|f-g\|_{2,X}\leqslant\delta}|Z_n(f-g)|\right] \tag{5}$$

$$\leqslant 4\mathbb{E}\left[\sup_{f\in\mathscr{F}}|Z_n(f-T_{q_0}(f))|\right]+\mathbb{E}\left[\sup_{(i,j)\in\Lambda}|Z_n(\Phi_{i,j}-\Psi_{i,j})|\right] \tag{6}$$

$$:=4E_1+E_2.$$

For the shake of brevity we put $\sup_{f\in\mathscr{F}}|Z_n(f)|=\|Z_n(f)\|_{\mathscr{F}}$. In order to control the two terms in Eq. (6) we shall use the following maximal inequality from Pisier, combined with a chaining argument. For all random variables $Z_1,Z_2,\ldots,Z_N$

$$\left(\mathbb{E}\left|\max_{1\leqslant i\leqslant N}|Z_i|\right|\right)\leqslant N^{1/p}\max_{1\leqslant i\leqslant N}(\mathbb{E}|Z_i|^p)^{1/p}. \tag{7}$$

Control of E_1: For f in $\mathscr{F}$ we write

$$\begin{aligned}f-T_{q_0}(f)&=f-T_q(f)+\sum_{k=q_0+1}^{q}T_k(f)-T_{k-1}(f)\\&=f-\pi_q(f)+\sum_{k=q_0+1}^{q}T_k(f)-T_{k-1}(f).\end{aligned}$$

Therefore

$$\begin{aligned}E_1&:=\mathbb{E}\|Z_n(f-T_{q_0}(f))\|_{\mathscr{F}}\\&\leqslant\mathbb{E}\|Z_n(f-\pi_q(f))\|_{\mathscr{F}}+\sum_{k=q_0+1}^{q}\mathbb{E}\|Z_n(T_k(f)-T_{k-1}(f))\|_{\mathscr{F}}\\&\leqslant E_{1,q+1}+2\sqrt{n}\sup_{f\in\mathscr{F}}\mathbb{E}|\Delta_q(f)|+\sum_{k=q_0+1}^{q}E_{1,k},\end{aligned} \tag{8}$$

where

$$\begin{aligned}E_{1,k}&=\mathbb{E}\|Z_n(T_k(f)-T_{k-1}(f))\|_{\mathscr{F}},\qquad q_0+1\leqslant k\leqslant q\\E_{1,q+1}&=\mathbb{E}\|Z_n\Delta_q(f)\|_{\mathscr{F}}.\end{aligned}$$

Now observe that $T_k(f)-T_{k-1}(f)=T_k(f)-\pi_{k-1}(T_k(f))$ and $T_k(f)$ takes values on a finite set with cardinality less than or equal to $N(k)$. Using inequality (7) we can write

$$E_{1,k} \leqslant N(k)^{1/p} \max_{g \in T_k(\mathscr{F})} \|Z_n(g - \pi_{k-1}(g))\|_p. \tag{9}$$

Since by hypothesis $\mathscr{F}$ is uniformly bounded, we may assume that f, $\pi_{k-1}(f)$ and $\Delta_q(f)$ are bounded by 1. Apply hypothesis (2) to $h = g - \pi_{k-1}(g)$ to get

$$\begin{aligned} \|Z_n(h)\|_p &\leqslant 2^{1/p} a^{1/p}(p)\|h\|_{2,X} + 2^{1/p}(b(p)n^{1-p/2}\|h\|_\infty^{p-2}\|h\|_{2,X}^2)^{1/p} \\ &\leqslant 2^{1/p} a^{1/p}(p) 2^{-(k-1)} + 2^{1/p} b^{1/p}(p) n^{1/p-1/2} 2^{-2(k-1)/p}. \end{aligned} \tag{10}$$

Combining (9) and (10) yields

$$E_{1,k} \leqslant 2a^{1/p}(p)N(k)^{1/p}2^{-k} + 2b^{1/p}(p)N(k)^{1/p}2^{-k}(n^{-1/2}2^k)^{1-2/p}.$$

A similar bound holds for $E_{1,q+1}$. Finally, using the fact that $\mathbb{E}|\Delta_q(f)| \leqslant C\|\Delta_q(f)\|_{2,X} \leqslant C2^{-q}$ we obtain

$$\begin{aligned} E_1 &\preccurlyeq \sqrt{n}2^{-q} + \sum_{k=q_0+1}^{q+1} E_{1,k} \\ &\preccurlyeq \sqrt{n}2^{-q} + a^{1/p}(p) \sum_{k=q_0+1}^{q+1} N(k)^{1/p}2^{-k} + b^{1/p}(p) \sum_{k=q_0+1}^{q+1} N(k)^{1/p}2^{-k}(n^{-1/2}2^k)^{1-2/p} \\ &\preccurlyeq \sqrt{n}2^{-q} + a^{1/p}(p) \sum_{k=q_0+1}^{q+1} N(k)^{1/p}2^{-k} + b^{1/p}(p)(n^{-1/2}2^q)^{1-2/p} \sum_{k=q_0+1}^{q+1} N(k)^{1/p}2^{-k}. \end{aligned}$$

Hence

$$E_1 \preccurlyeq \sqrt{n}2^{-q} + c(p) \sum_{k=q_0+1}^{q+1} N(k)^{1/p}2^{-k}(1 + (n^{-1/2}2^q)^{1-2/p}). \tag{11}$$

Control of E_2: Noting that $|\Lambda| \leqslant N^2(q_0)$ and $\|\Phi_{i,j} - \Psi_{i,j}\|_{2,X} \leqslant \delta$ we get

$$\begin{aligned} E_2 &= \mathbb{E}\left[\sup_{(i,j)\in\Lambda} |Z_n(\Phi_{i,j} - \Psi_{i,j})|\right] \\ &\leqslant N^{2/p}(q_0) \max_{(i,j)\in\Lambda} \|Z_n(\Phi_{i,j} - \Psi_{i,j})\|_p. \end{aligned}$$

Again by $\mathrm{H}(p, X)$,

$$\begin{aligned} E_2 &\preccurlyeq N^{2/p}(q_0)\left(a^{1/p}(p)\delta + (b(p)n^{1-p/2}\delta^2)^{1/p}\right) \\ &\preccurlyeq c(p)(N(q_0)\delta)^{2/p}. \end{aligned} \tag{12}$$

Let $W(n,\delta)$ denote $\mathbb{E}(\sup_{\|f-g\|_{2,X}<\delta}|Z_n(f-g)|)$. From (11) and (12) it follows that:

$$\begin{aligned} W(n,\delta) &\leqslant 4E_{1+}E_2 \\ &\preccurlyeq \sqrt{n}2^{-q} + c(p)\sum_{k=q_0+1}^{q+1} N(k)^{1/p}2^{-k}(1+(n^{-1/2}2^q)^{1-2/p}) \\ &\quad + c(p)(N(q_0)\delta)^{2/p}. \end{aligned}$$

Let $q_0 = q_0(\delta)$ be the largest integer satisfying $N(q_0)\leqslant\delta^{-1/2}$. Without loss of generality we may assume that $q_0(\delta)$ goes to infinity as δ goes to zero. Therefore, if we set $\varepsilon(\delta) = \sum_{k=q_0+1}^{\infty} N(k)^{1/p}2^{-k}$ we have by $\mathrm{H}(p,\mathscr{F})$ that $\varepsilon(\delta)\to 0$ when $\delta\to 0$. Take $q = q(n,\delta) = [\frac{1}{2\log 2}\log\frac{n}{\varepsilon(\delta)}]+1$. With this choice $q>q_0$ and $\sqrt{n}2^{-q}<1$ if $n>n(\delta)$ and for $n>n(\delta)$ we have

$$W(n,\delta)\preccurlyeq\sqrt{\varepsilon(\delta)} + c(p)\sqrt{\varepsilon(\delta)} + c(p)\delta^{1/p}.$$

Consequently

$$\lim_{\delta\to 0}\limsup_{n\to+\infty} W(n,\delta)\preccurlyeq \lim_{\delta\to 0}\sqrt{\varepsilon(\delta)} + c(p)\sqrt{\varepsilon(\delta)} + c(p)\delta^{1/p} = 0$$

and Theorem 1 is proved.

4.2. Proof of Theorem 2

We will follow the same lines of the proof of Theorem 1 with small modifications. Therefore notations will also be similar.

Control of E_1:

$$\begin{aligned} E_1 &= \mathbb{E}\|Z_n(f-T_{q_0}(f))\|_{\mathscr{F}} \\ &\leqslant \mathbb{E}\|Z_n((f-T_{q_0}(f))1_{F\leqslant M})\|_{\mathscr{F}} + \mathbb{E}\|Z_n(f-T_{q_0}(f)1_{F>M})\|_{\mathscr{F}} \\ &:= E_{1,M} + E'_{1,M}. \end{aligned} \tag{13}$$

On the one hand, since $F\in\mathbb{L}^{r+1}$ we can write

$$E'_{1,M}\leqslant 2\sqrt{n}\mathbb{E}|F1_{F>M}|\leqslant\frac{\sqrt{n}}{M^r}r(M),$$

where $r(M)$ goes to zero as M goes to $+\infty$. On the other hand

$$E_{1,M}\leqslant\mathbb{E}\left\|Z_n\left((f-\pi_q(f))1_{F\leqslant M}\right)\right\|_{\mathscr{F}} \tag{14}$$

$$+\sum_{k=q_0+1}^{q}\mathbb{E}\|Z_n((T_k(f)-T_{k-1}(f))1_{F\leqslant M})\|_{\mathscr{F}}$$

$$\leqslant E^M_{1,q+1} + 2\sqrt{n}\sup_{f\in\mathscr{F}}\mathbb{E}|\Delta_q(f)| + \sum_{k=q_0+1}^{q} E^M_{1,k}, \tag{15}$$

where

$$E_{1,k}^{M} = \mathbb{E}\|Z_n((T_k(f) - T_{k-1}(f))1_{F\leqslant M})\|_{\mathscr{F}}, \quad q_0+1\leqslant k\leqslant q,$$
$$E_{1,q+1}^{M} = \mathbb{E}\|Z_n\Delta_q(f)1_{F\leqslant M}\|_{\mathscr{F}}.$$

Note that when $F\leqslant M$, we have that $T_k(f)$ and $\Delta_q(f)$ are bounded above by M. Apply hypothesis $\mathrm{H}(p,X)$ to $h := (T_k(f) - T_{k-1}(f))1_{F\leqslant M}$ after applying (7) to obtain

$$\begin{aligned} E_{1,k}^{M} &\leqslant N(k)^{1/p} \max_{f\in\mathscr{F}} \|h\|_p \\ &\leqslant 2a^{1/p}(p)N(k)^{1/p}2^{-k} + 2b^{1/p}(p)N(k)^{1/p}2^{-2k/p}(n^{-1/2}M)^{1-2/p}. \end{aligned} \tag{16}$$

A similar bound holds to $E_{1,q+1}^{M}$ that is

$$E_{1,q+1}^{M} \leqslant 2a^{1/p}(p)N(q)^{1/p}2^{-q} + 2b^{1/p}(p)N(q)^{1/p}2^{-2q/p}(n^{-1/2}M)^{1-2/p}.$$

Therefore

$$\begin{aligned} E_{1,M} &\preccurlyeq \sqrt{n}2^{-q} + \sum_{k=q_0+1}^{q+1} E_{1,k} \\ &\preccurlyeq \sqrt{n}2^{-q} + c(p)\sum_{k=q_0+1}^{q+1} N(k)^{1/p}2^{-k} + c(p)(n^{-1/2}M)^{1-2/p}\sum_{k=q_0+1}^{q+1} N(k)^{1/p}2^{-2k/p}. \end{aligned}$$

Taking $M = n^{1/2r}$, from the estimations of $E_{1,M}$ and $E'_{1,M}$, we deduce that

$$\begin{aligned} E_1 \preccurlyeq{}& \sqrt{n}2^{-q} + c(p)\sum_{k=q_0+1}^{q+1} N(k)^{1/p}2^{-k} \\ &+ c(p)(n^{-1/2}M)^{1-2/p}\sum_{k=q_0+1}^{q+1} N(k)^{1/p}2^{-2k/p} + r'(n), \end{aligned} \tag{17}$$

where $r'(n) \to 0$. Let R denote the third term in the previous equation, then

$$\begin{aligned} R &:= c(p)(n^{-1/2}M)^{1-2/p}\sum_{k=q_0+1}^{q+1} N(k)^{1/p}2^{-2k/p} \\ &\leqslant c(p)(n^{-1/2}n^{1/2r})^{1-2/p}\int_{2^{-q}}^{2^{-q_0}} N^{1/p}(x)x^{2/p-1}\,\mathrm{d}x \\ &\leqslant c(p)n^{(-1/2+1/2r)(1-2/p)}\int_{2^{-q}}^{2^{-q_0}} N^{1/p}(x)x^{2/p-1}\,\mathrm{d}x. \end{aligned}$$

We apply Hölder's inequality to $f = N^{1/p}, g = x^{2/p-1}$ and $1/v = 1 - 1/r(1-2/p)$, to obtain

$$\begin{aligned}
&\int_{2^{-q}}^{2^{-q_0}} N^{1/p}(x)x^{2/p-1}\,\mathrm{d}x \\
&\leqslant \left(\int_{2^{-q}}^{2^{-q_0}} N^{v/p}(x)\,\mathrm{d}x\right)^{1/v} \left(\int_{2^{-q}}^{2^{-q_0}} x^{(2/p-1)\frac{v}{v-1}}\,\mathrm{d}x\right)^{\frac{v-1}{v}} \\
&\leqslant \left(\int_{2^{-q}}^{2^{-q_0}} N^{v/p}(x)\,\mathrm{d}x\right)^{1/v} \left((r-1)^{-1}\left[x^{-r+1}\right]_{2^{-q}}^{2^{-q_0}}\right)^{\frac{v-1}{v}} \\
&\leqslant \left(\int_{2^{-q}}^{2^{-q_0}} N^{v/p}(x)\,\mathrm{d}x\right)^{1/v} (r-1)^{-\frac{v-1}{v}}(2^{-q})^{(1-2/p)(1/r-1)}.
\end{aligned}$$

It follows that

$$R \leqslant c(p,r)(\sqrt{n}2^{-q})^{(1/r-1)(1-2/p)}\left(\int_{2^{-q}}^{2^{-q_0}} N^{v/p}(x)\,\mathrm{d}x\right)^{1/v}. \tag{18}$$

Combining (17) and (18) we obtain

$$\begin{aligned}
E_1 \preccurlyeq\ &\sqrt{n}2^{-q} + c(p)\sum_{k=q_0+1}^{q+1} N(k)^{1/p}2^{-k} \\
&+ c(p,r)(\sqrt{n}2^{-q})^{(1/r-1)(1-2/p)}\left(\int_{2^{-q}}^{2^{-q_0}} N^{v/p}(x)\,\mathrm{d}x\right)^{1/v} + r'(n).
\end{aligned} \tag{19}$$

Control of E_2: Similarly we have

$$E_2 = \mathbb{E}\left[\sup_{(i,j)\in\Lambda}|Z_n(\Phi_{i,j} - \Psi_{i,j})|\right] \leqslant E_{2,M} + E'_{2,M}.$$

Firstly, we write

$$\begin{aligned}
E'_{2,M} &:= \mathbb{E}\left[\sup_{(i,j)\in\Lambda}|Z_n((\Phi_{i,j} - \Psi_{i,j})1_{F>M})|\right] \\
&\leqslant 4\sqrt{n}\mathbb{E}|F1_{F>M}| \leqslant r(n),
\end{aligned} \tag{20}$$

where $r(n)$ goes to zero. Secondly, applying $\mathrm{H}(p, X)$ to $(\Phi_{i,j} - \Psi_{i,j})1_{F\leqslant M}$, which satisfies $\|(\Phi_{i,j} - \Psi_{i,j})1_{F\leqslant M}\|_{2,X} \leqslant \delta$ we obtain

$$\begin{aligned}
E_{2,M} &:= \mathbb{E}\left[\sup_{(i,j)\in\Lambda}|Z_n(\Phi_{i,j} - \Psi_{i,j})1_{F\leqslant M}|\right] \\
&\leqslant N^{2/p}(q_0)\left(a^{1/p}(p)\delta + (b(p)n^{1-p/2}M^{p-2}\delta^2)^{1/p}\right) \qquad (21) \\
&\leqslant c(p)(N(q_0)\delta)^{2/p}. \qquad (22)
\end{aligned}$$

From (20) and (22) we conclude that

$$E_2 \leqslant c(p)(N(q_0)\delta)^{2/p} + r(n). \tag{23}$$

End of the proof: Using that $W(n,\delta)=\mathbb{E}(\sup_{\|f-g\|_{2,X}<\delta}|Z_n(f-g)|)$. Then (19) together with (23) imply

$$W(n,\delta)\leqslant 4E_{1+}E_2$$

$$\preccurlyeq \sqrt{n}2^{-q}+c(p)\sum_{k=q_0+1}^{q+1}N(k)^{1/p}2^{-k}+c(p)(N(q_0)\delta)^{2/p}+r(n)$$

$$+\,c(p,r)(\sqrt{n}2^{-q})^{(1/r-1)(1-2/p)}\left(\int_{2^{-q}}^{2^{-q_0}}N^{v/p}(x)\,\mathrm{d}x\right)^{1/v}+r'(n).$$

Putting $\beta=-(1/r-1)(1-2/p)$ and letting

$$q_0=q_0(\delta)=\max\{k,k\in N,N(k)\leqslant\delta^{-1/2}\}.$$

We may assume that $q_0(\delta)$ goes to infinity as δ goes to zero. Putting

$$\varepsilon(\delta)=\left(\int_0^{2^{-q_0(\delta)}}N^{v/p}(x)\,\mathrm{d}x\right)^{1/v},$$

we have by (4) that $\varepsilon(\delta)\to 0$ when $\delta\to 0$. Now choose $q=q(n,\delta)$ in such a way that $\sqrt{n}2^{-q}$ and $(\sqrt{n}2^{-q})^{(1/r-1)(1-2/p)}\varepsilon(\delta)$ have the same order of magnitude, that is

$$q=q(n,\delta)=\left[\frac{1}{2\log 2}\log\frac{n}{\varepsilon^{1/(1+\beta)}(\delta)}\right]+1.$$

With this choice $q>q_0$ if $n>n(\delta)$, and in this case we have

$$W(n,\delta)\preccurlyeq\varepsilon^{1/(1+\beta)}(\delta)(1+c(p,r))+c(p)\sum_{k=q_0+1}^{q+1}N(k)^{1/p}2^{-k}$$

$$+\,c(p)\delta^{1/p}+r(n)+r'(n).$$

It follows that:

$$\lim_{\delta\to 0}\limsup_{n\to+\infty}\mathbb{P}\left(\sup_{\|f-g\|_{2,X}<\delta}|Z_n(f-g)|>\varepsilon\right)=0$$

and this concludes the proof of Theorem 2.

4.3. Proof of Theorem 3

In the sequel all the inequalities are valid up to a multiplicative constant. First we recall the following moment inequality which is a corollary of Theorem 6.3 in Rio [12].

Lemma 2. *Let $(\alpha_n)_{n\geqslant 0}$ be the sequence of strong mixing coefficients of the process $(X_i)_{i\geqslant 0}$. Let f be a measurable function. Then for all $p\geqslant 2$.*

$$\mathbb{E}|Z_n(f)|^p\leqslant a(p)\|f\|_{2,\alpha}^p+b(p)n^{1-p/2}\sum_{i=1}^{n}(i+1)^{p-2}\alpha_i\|f\|_\infty^p \tag{24}$$

with

- Q_f *is the quantile function of* $|f(X_0)|$.
- $\|f\|_{2,\alpha}^2 = \int_0^1 \alpha^{-1}(u) Q_f^2(u)\,\mathrm{d}u$.
- $a(p) \leqslant (Cp)^{p/2}, b(p) \leqslant (Cp)^p$.

We have assumed that: $\forall f \in \mathscr{F},\ \|f\|_\infty \leqslant 1$. Hence without loss of generality, we may assume that $\forall f \in \mathscr{F},\ \forall k > 0,\ \Delta_k(f) \leqslant 1$. From (30) it follows that:

$$\|Z_n(f)\|_p \leqslant A(p,f) + B(p,f)$$

with

$$A(p,f) \preccurlyeq \sqrt{p}\|f\|_{2,\alpha}$$
$$B(p,f) \preccurlyeq p^2 n^{-1/2+1/p}\|f\|_\infty.$$

Applying Hölder's inequality gives

$$\begin{aligned}
\|f\|_{2,\alpha}^2 &\leqslant \left(\int_0^1 [\alpha^{-1}(u)]^{1/(1-\theta)}\,\mathrm{d}u\right)^{1-\theta} \left(\int_0^1 Q_f^{2/\theta}(u)\,\mathrm{d}u\right)^{\theta} \\
&\leqslant \left(\frac{1}{1-\theta}\sum_{i=1}^n (i+1)^{1/(1-\theta)}\alpha(i)\right)^{1-\theta} \|f\|_{2/\theta}^2 \\
&\leqslant \left(\frac{1}{1-\theta}\sum_{i=1}^n (i+1)^{1/(1-\theta)}\alpha(i)\right)^{1-\theta} \|f\|_1^\theta.
\end{aligned}$$

Therefore

$$A(p,f) \preccurlyeq \sqrt{p}\|f\|_1^{\theta/2},$$
$$B(p,f) \preccurlyeq p^2 n^{-1/2+1/p}\|f\|_\infty.$$

We proceed as in the proof of Theorem 1, and thus we keep the same notation.

Control of E_1: We recall that if $N(k) = N_{[\,]}(2^{-k}, \|\cdot\|_1, \mathscr{F})$ then

$$E_1 \leqslant \mathbb{E}\|Z_n \Delta_q(f)\|_{\mathscr{F}} + 2\sqrt{n}\sup_{f\in\mathscr{F}} \mathbb{E}|\Delta_q(f)| + \sum_{k=q_0+1}^{q} \mathbb{E}\|Z_n T_k(f) - T_{k-1}(f)\|_{\mathscr{F}} \tag{25}$$

we also recall that

$$E_{1,k} = \mathbb{E}\|Z_n T_k(f) - T_{k-1}(f)\|_{\mathscr{F}}, \quad q_0 + 1 \leqslant k \leqslant q,$$
$$E_{1,q+1} = \mathbb{E}\|Z_n \Delta_q(f)\|_{\mathscr{F}}.$$

From the hypothesis and inequality (7) we have

$$E_{1,k} \leqslant N(k)^{1/p} \max_{g\in T_k(\mathscr{F})} \|Z_n(g - \pi_{k-1}(g))\|_p. \tag{26}$$

We now apply the moment inequality to $g-\pi_{k-1}(g)$ which is bounded by $\Delta_{k-1}(g)$ to obtain

$$\begin{aligned} E_{1,k} &\leqslant N(k)^{1/p}\left[\max_{g\in T_k(\mathscr{F})} A(p,g-\pi_{k-1}(g)) + \max_{g\in T_k(\mathscr{F})} B(p,g-\pi_{k-1}(g))\right] \\ &\preccurlyeq N(k)^{1/p}(\sqrt{p}2^{-k\theta/2}+p^2n^{-1/2+1/p}) \\ &\preccurlyeq N(k)^{1/p}(\sqrt{p}2^{k(1-\theta/2)}2^{-k}+p^2(n^{-1/2}2^k)^{1-2/p}2^{2k/p}2^{-k}). \end{aligned} \tag{27}$$

Therefore if $p>2$, $n^{-1/2}2^q\geqslant 1$, we get

$$E_{1,k}\preccurlyeq N(k)^{1/p}\left(\sqrt{p}2^{k(1-\theta/2)}2^{-k}+(n^{-1/2}2^q)p^2 2^{2k/p}2^{-k}\right).$$

Let $p=k+\log N(k)$, then

$$E_{1,k}\preccurlyeq\left(\sqrt{k}+\sqrt{\log N(k)}\right)2^{k(1-\theta/2)}2^{-k}+(n^{-1/2}2^q)(k^2+\log^2 N(k))2^{-k}.$$

A similar bound holds for $E_{1,q+1}$. Hence if we assume that

$$\int_0^1 \log^2 N_{[\,]}(\varepsilon,\|\cdot\|_2,\mathscr{F})\,\mathrm{d}\,\varepsilon<\infty$$

and

$$\int_0^1 \log^{1/2} N_{[\,]}(\varepsilon,\|\cdot\|_2,\mathscr{F})x^{\theta/2-1}\,\mathrm{d}\varepsilon<\infty$$

for some $0<\theta<1$. Then there exists a positive sequence $l(k)$ satisfying $\sum l(k)<\infty$, such that for all $k, q_0\leqslant k\leqslant q+1$, if $n^{-1/2}2^q\geqslant 1$, we have

$$E_{1,k}\preccurlyeq(n^{-1/2}2^q+1)l(k).$$

Since $\int_0^1 \log^2 N_{[\,]}(\varepsilon,\|\cdot\|_2,\mathscr{F})\,\mathrm{d}\varepsilon<\infty$ implies $\int_0^1 \log^{1/2} N_{[\,]}(\varepsilon,\|\cdot\|_2,\mathscr{F})\varepsilon^{\theta/2-1}\,\mathrm{d}\varepsilon<\infty$, for some convenient θ, we conclude that under the hypothesis of the theorem we have: $\forall q\geqslant q_0$ such that $n^{-1/2}2^q\geqslant 1$,

$$E_1\preccurlyeq\sqrt{n}2^{-q}+2n^{-1/2}2^q\sum_{k=q_0+1}^{q+1} l(k). \tag{28}$$

Control of E_2: Recall that $|\Lambda|\leqslant N^2(q_0)$ hence

$$\begin{aligned} E_2 &= \mathbb{E}\left[\sup_{(i,j)\in\Lambda}|Z_n(\Phi_{i,j}-\Psi_{i,j})|\right] \\ &\leqslant N(q_0)\max_{(i,j)\in\Lambda}\|Z_n(\Phi_{i,j}-\Psi_{i,j})\|_2. \end{aligned}$$

Using a moment inequality of order 2,

$$\mathbb{E}|Z_n(f)|^2\leqslant C(\theta',\alpha)\|f\|_1^{\theta'},$$

where $0<\theta'<1/2$. Applying this to $\Phi_{i,j}-\Psi_{i,j}$ which satisfies $\|\Phi_{i,j}-\Psi_{i,j}\|_1\leqslant\delta$ we get

$$E_2\leqslant C(\theta',\alpha)N(q_0)\delta^{\theta'/2}. \tag{29}$$

End of the proof: Let $W(n,\delta)$ denote $\mathbb{E}(\sup_{\|f-g\|_1<\delta}|Z_n(f-g)|)$. Combining (28) and (29) gives

$$W(n,\delta)\preccurlyeq\sqrt{n}2^{-q}+2n^{-1/2}2^q\sum_{k=q_0+1}^{q+1}l(k)+C(\theta',\alpha)N(q_0)\delta^{\theta'/2}.$$

Take $\theta'=1/3$ for example and let $q_0=q_0(\delta)$ the greatest integer satisfying $N(q_0)\leqslant\delta^{-1/12}$. Without loss of generality we may assume that $q_0(\delta)$ tends to infinity as δ tends to zero. Therefore, if we set $\varepsilon(\delta)=\sum_{k=q_0+1}^{\infty}l(k)$ we have that $\varepsilon(\delta)\to 0$ when $\delta\to 0$. Take $q=q(n,\delta)=[\frac{1}{2\log 2}\log\frac{n}{\varepsilon(\delta)}]+1$. Note that $q>q_0$ and $\sqrt{n}2^{-q}<1$ for n sufficiently large, say $n>n(\delta)$ and hence for $n>n(\delta)$

$$W(n,\delta)\preccurlyeq\varepsilon^{1/2}(\delta)+C(\theta',\alpha)\delta^{1/12}.$$

Consequently

$$\lim_{\delta\to 0}\limsup_{n\to+\infty}W(n,\delta)\preccurlyeq\lim_{\delta\to 0}\varepsilon^{1/2}(\delta)+C(\theta',\alpha)\delta^{1/12}=0$$

and this conclude the proof of Theorem 3.

4.4. *Other proofs*

4.4.1. *Proof of Corollary 1*

From Rio [12] Theorem 6.3 we infer that

$$\mathbb{E}|Z_n(f)|^p\leqslant a(p)\left(\int_0^1\alpha^{-1}(u)Q_f^2(u)\,\mathrm{d}u\right)^{p/2}+b(p)n^{1-p/2}\int_0^1[\alpha^{-1}(u)]^{p-1}Q_f^p(u)\,\mathrm{d}u, \tag{30}$$

where Q_f is the quantile function of $|f(X_0)|$. Assume moreover that $\|f\|_\infty\leqslant M$, then (30) can be written

$$\mathbb{E}|Z_n(f)|^p\leqslant a(p)\left(\int_0^1\alpha^{-1}(u)Q_f^2(u)\,\mathrm{d}u\right)^{p/2}+b(p)n^{1-p/2}M^{p-2}\int_0^1[\alpha^{-1}(u)]^{p-1}Q_f^2(u)\,\mathrm{d}u.$$

Therefore, we can apply Theorem 1 with $\|f\|_{2,X}^2=\int_0^1[\alpha^{-1}(u)]^{p-1}Q_f^2(u)\,\mathrm{d}u$. Now $\mathrm{H}(p,\mathscr{F})$ implies that for $f\in\mathscr{F}$, $\int_0^1\alpha^{-1}(u)Q_f^2(u)\,\mathrm{d}u<\infty$, and this implies (i) according to Doukhan et al. (see [7]).

Using Hölder's inequality, we get, for any θ in $(0,1)$,

$$\int_0^1 [\alpha^{-1}(u)]^{p-1} Q_f^2(u)\,\mathrm{d}u \leqslant \left(\int_0^1 [\alpha^{-1}(u)]^{(p-1)/(1-\theta)}\,\mathrm{d}u\right)^{1-\theta} \left(\int_0^1 Q_f^{2/\theta}(u)\,\mathrm{d}u\right)^{\theta}.$$

Since $\int_0^1 [\alpha^{-1}(u)]^q\,\mathrm{d}u \leqslant q\sum_{i\geqslant 0}(i+1)^{q-1}\alpha(i)$ and $Q_f(U) \stackrel{\text{law}}{=} |f(X)|$ if U is uniformly distributed on $[0,1]$, we deduce that

$$\int_0^1 [\alpha^{-1}(u)]^{p-1} Q_f^2(u)\,\mathrm{d}u \leqslant \left(\frac{p-1}{1-\theta}\sum_{i\geqslant 0}(i+1)^{(p-1)/(1-\theta)-1}\alpha(i)\right)^{1-\theta} \left(\|f\|_{2/\theta}\right)^2.$$

Hence the following hypotheses are sufficient to imply (ii),

(H1) $\sum_{i\geqslant 0}(i+1)^{(p-1)/(1-\theta)-1}\alpha(i) < \infty$,

(H2) $\int_0^1 N_{[\,]}^{1/p}(x, \|\cdot\|_{2/\theta}, \mathscr{F})\,\mathrm{d}x < \infty$

and this proves the second part of the corollary.

4.4.2. Proof of Corollary 2

First we recall the following result from Shao [13]. $\forall p \geqslant 2, \exists K = K(\rho(\cdot), p)$ such that for every measurable f,

$$\mathbb{E}|Z_n(f)|^p \leqslant K \exp\left(\sum_{i=0}^{[\log n]} \rho(2^i)\right)\|f(X)\|_2^p + Kn^{1-p/2}\exp\left(K\sum_{i=0}^{[\log n]}\rho^{2/p}(2^i)\right)\|f(X)\|_p^p.$$

In particular, if $\sum_{i=0}^{[\log n]}\rho(2^i) < \infty$, then $\exp(K\sum_{i=0}^{[\log n]}\rho^{2/p}(2^i))$ is a slowly varying function for every p. Hence, $\forall p \geqslant 2$, $\forall \varepsilon > 0 \exists K = K(\rho(\cdot), p, \varepsilon)$ such that for every measurable f,

$$\mathbb{E}|Z_n(f)|^p \leqslant K\|f(X)\|_2^p + Kn^{1+\varepsilon-p/2}\|f(X)\|_p^p. \tag{31}$$

Arguing as in the proof of Theorem 2, it is easy to see that under (31) (ii) is satisfied as soon as F, the envelop function belongs to L^{r+1}, for some $r > 1$ and

$$\int_0^1 N_{[\,]}^{1/p(1-1/r(1-2/p)-2\varepsilon/p)}(x, \|\cdot\|_2, \mathscr{F})\,\mathrm{d}x < \infty.$$

Since p can be chosen arbitrary large and ε arbitrary small, (ii) follows under our hypothesis. The proof of (i) follows from Theorem 1 in [10] for example.

4.4.3. Proof of Lemma 1

We will take back the proof of a similar result given in Csörgő and Mielniczuk [5, inequality 3.2] with small changes. In particular, we recall that for $k = 1, 2, 3, 4$,

$$\mathbb{E}(S_{kn}) = \sum_{1 \leqslant i_1 \neq i_2 \cdots \neq i_k \leqslant n} \mathbb{E}[(f(X_{i_1}) - \mathbb{E}f(X_{i_1})) \cdots (f(X_{i_k}) - \mathbb{E}f(X_{i_k}))].$$

We first assume that $R := \sup_{k \geqslant 1} |r(k)| < 1/3$, then we proceed as in [5] to handle the general case.

$$\mathbb{E}(S_{1n}) = n\mathbb{E}(f(X_0) - \mathbb{E}f(X_0))^4 \preccurlyeq n\|f\|_\infty^2 \mathbb{E}f^2(X_i).$$

$$\mathbb{E}(S_{2n}) = 3 \sum_{1 \leqslant i_1 \neq i_2 \leqslant n} \mathbb{E}\left[(\overline{f}(X_{i_1}))^2 (\overline{f}(X_{i_2}))^2\right] + 4 \sum_{1 \leqslant i_1 \neq i_2 \leqslant n} \mathbb{E}[(\overline{f}(X_{i_1}))^3 \overline{f}(X_{i_2})],$$

where $\overline{f} = f - \mathbb{E}f(X_0)$. The first term is bounded by

$$n^2(\mathbb{E}f^2(X_i))^2 + n \sum_{i=1}^{n} |r^{m(f^2)}(i)| \mathbb{E}|f(X_i)|^4$$

and the second one is bounded by

$$n \sum_{i=1}^{n} |r^{m(f)}(i)| \mathbb{E}^{1/2}|f(X_i)|^2 \mathbb{E}^{1/2}|f(X_i)|^6.$$

Hence

$$\mathbb{E}(S_{2n}) \preccurlyeq n^2(\mathbb{E}f^2(X_i))^2 + n\|f\|_\infty^2 \mathbb{E}f^2(X_i).$$

Using a lemma of Taqqu stated as Lemma 3 in [5], we have

$$\begin{aligned} \mathbb{E}(S_{3n}) &\preccurlyeq n^{3/2} \mathbb{E}^{1/2}|f(X_i)|^2 \mathbb{E}^{1/2}|f(X_i)|^4 \\ &\preccurlyeq n^2(\mathbb{E}f^2(X_i))^2 + n\mathbb{E}f^4(X_i) \\ &\preccurlyeq n^2(\mathbb{E}f^2(X_i))^2 + n\|f\|_\infty^2 \mathbb{E}f^2(X_i). \end{aligned}$$

Again by Lemma 3 we have

$$\mathbb{E}(S_{4n}) \preccurlyeq n^2 \mathbb{E}^2|f(X_i)|^2.$$

This completes the proof.

Acknowledgements

I am deeply grateful to my thesis supervisor, Professor Jean Bretagnolle, for helpful discussions and a careful reading of the manuscript. Also, I sincerely thank the referee for a careful reading of the paper.

References

[1] D.W.K. Andrews, Asymptotics for semiparametric econometric models via stochastic equicontinuity, Econometrica 62 (1) (1994) 43–72.

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

[3] M.A. Arcones, Limit theorems for nonlinear functional of a stationary Gaussian sequence of vectors, Ann. Probab. 22 (4) (1994) 2242–2274.
[4] P. Breuer, P. Major, Central limit theorems for non-linear functional of Gaussian fields, J. Multivariate Anal. 13 (1983) 425–441.
[5] S. Csörgő, J. Mielniczuk, The empirical process of a short-range dependent stationary sequence under Gaussian subordination, Probab. Theory Relat. Fields 104 (1) (1996) 15–25.
[6] P. Doukhan, P. Massart, E. Rio, Invariance principles for absolutely regular empirical processes, Ann. Inst. Henri Poincaré, Probab. Statist. 31 (2) (1995) 393–427.
[7] P. Doukhan, P. Massart, E. Rio, The functional central limit theorem for strongly mixing processes, Ann. Inst. Henri Poincaré, Probab. Statist. 30 (1995) 63–82.
[8] P. Massart, Invariance principles for empirical processes: the weakly dependent case, Ph.D. dissertation, Université de Paris Sud, 1987, Chapter 1B.
[9] M. Ossiander, A central limit theorem under metric entropy with L_2 bracketing, Ann. Probab. 15 (3) (1987) 897–919.
[10] M. Peligrad, On the central limit theorem for ρ-mixing sequences of random variables, Ann. Probab. 15 (4) (1987) 1387–1394.
[11] D. Pollard, Empirical Processes: Theory and Applications, Regional Conference Series in Probability and Statistics, IMS-ASA, Hayward, Alexandria, 1990.
[12] E. Rio, Théorèmes limites pour des variables aléatoires faiblement dépendentes, Preprint, Orsay no. 97/81, 1997.
[13] Q. Shao, Maximal inequalities for partial sums of ρ-mixing sequences, Ann. Probab. 23 (2) (1995) 948–965.
[14] Q. Shao, H. Yu, Weak convergence for weighted empirical processes of dependent sequences, Ann. Probab. 24 (4) (1996) 2098–2127.
[15] A. van der Vaart, J.A. Wellner, Weak Convergence and Empirical Processes, Springer, Berlin, 1996.