of a limiting variance for $Z_n(f)$ does not imply the CLT. Secondly it does not allow bracketing. We will provide a norm depending on P and on the mixing structure of the sequence of observations, which coincides with the usual $\mathcal{L}_2(P)$-norm in the independent case. Denoting by $\mathcal{L}_{2,\beta}(P)$ the so-defined normed space, the finite dimensional convergence of Z_n to some Gaussian vector with covariance function Γ holds on $\mathcal{L}_{2,\beta}(P)$. Γ is the limiting covariance of Z_n and is majorized by the square of the $\mathcal{L}_{2,\beta}(P)$-norm. Moreover, this norm allows bracketing and we will obtain a generalization of Ossiander's theorem by simply measuring the size of the brackets with the $\mathcal{L}_{2,\beta}(P)$-norm instead of the $\mathcal{L}_2(P)$-norm.

2. STATEMENT OF RESULTS

β-mixing

Throughout the sequel, the underlying probability space $(\Omega, \mathcal{T}, \mathbb{P})$ is assumed to be rich enough in the following sense: there exists a random variable U with uniform distribution over $[0,1]$, independent of the sequence $(\xi_i)_{i\in\mathbb{Z}}$.

For any numerical integrable function f, we set $E_P(f) = \int_{\mathcal{X}} f(x)P(dx)$. For any $r \geq 1$, let $\mathcal{L}_r(P)$ denote the class of numerical functions on $(\mathcal{X}, P)$ such that $\|f\|_r = [E_P(|f|^r)]^{1/r} < +\infty$.

Since $(\xi_i)_{i\in\mathbb{Z}}$ is a strictly stationary sequence, the mixing coefficients $(\beta_n)_{n>0}$ of the sequence $(\xi_i)_{i\in\mathbb{Z}}$ are defined by $\beta_n = \beta(\mathcal{F}_0, \mathcal{G}_n)$, where $\mathcal{F}_0 = \sigma(\xi_i : i \leq 0)$ and $\mathcal{G}_n = \sigma(\xi_i : i \geq n)$. $(X_i)_{i\in\mathbb{Z}}$ is called a β-mixing sequence if $\lim_{n\to+\infty} \beta_n = 0$. Examples of such sequences may be found in Davydov (1973), Bradley (1986) and Doukhan (1994).

$\beta(A,B) = \sup \sum_{(i,j)\in I\times J} |P(A_i \cap B_j) - P(A_i)P(B_j)|$

We now need to recall the definitions of entropy and entropy with bracketing.

Entropy. – Given a metric set (V,d), let $\mathcal{N}(\delta, V, d)$ be the minimum of

covering

$$\inf\{n \in \mathbb{N} : \exists S_n = \{x_1, \dots, x_n\} \subset S, \text{ s.t. } \forall x \in S, d(x, S_n) \leq \delta\}.$$

The entropy function $H(\delta, V, d)$ is the logarithm of $\mathcal{N}(\delta, V, d)$.

Bracketing. – Let V be some linear subspace of the space of numerical functions on $(\mathcal{X}, P)$. Assume that there exists some application $\Lambda : V \to \mathbb{R}^+$ such that, for any f and any g in V,

$$|f| \leq |g| \text{ implies } \Lambda(f) \leq \Lambda(g). \tag{2.1}$$

Assume that $\mathcal{F} \subset V$. A pair $[f, g]$ of elements of V such that $f \leq g$ is called a bracket of V. $\mathcal{F}$ is said to satisfy (A.1) if, for any $\delta > 0$, there exists a finite collection $S(\delta)$ of brackets of V such that

(2.2) for all $f \in \mathcal{F}$, there exists $[g, h]$ in $S(\delta)$ such that $g \leq f \leq h$ and $\Lambda(h - g) \leq \delta$.

log

The bracketing number $\mathcal{N}_{[\,]}(\delta, \mathcal{F})$ of $\mathcal{F}$ with respect to (V, Λ) is the minimal cardinality of such collections $S(\delta)$. The entropy with bracketing $H_{[\,]}(\delta, \mathcal{F}, \Lambda)$ is the logarithm of $\mathcal{N}_{[\,]}(\delta, \mathcal{F})$. When Λ is a norm, denoting by d_Λ the corresponding metric, it follows from (2.1) and (2.2) that

$x_0 > 0.$

$\Lambda(x_0) = 6$

$|x_0 - f| \geq x_0.$

$$H(\delta, \mathcal{F}, d_\Lambda) \leq H_{[\,]}(\delta, \mathcal{F}, \Lambda). \tag{2.3}$$

$|x_1 - x_y| < x_0$

We now define a new norm, which emerges from a covariance inequality due to Rio (1993).

The $\mathcal{L}_{2,\beta}(P)$-spaces

Notations. – Throughout the sequel, we make the convention that $\beta_0 = 1$. If $(u_n)_{n \geq 0}$ is a nonincreasing sequence of nonnegative real numbers, we denote by $u(.)$ the rate function defined by $u(t) = u_{[t]}$. For any nonincreasing function ψ, let ψ^{-1} denote the inverse function of ψ,

$x_1 \cdots x_n$

$n \quad t$

$$\psi^{-1}(u) = \inf\, \{t : \psi(t) \leq u\}.$$

For any f in $\mathcal{L}_1(P)$, we denote by Q_f the quantile function of $|f(\xi_0)|$, which is the inverse of the tail function $t \to \mathbb{P}(|f(\xi_0)| > t)$.

Let $(\alpha_n)_{n \geq 0}$ denote the sequence of strong mixing coefficients of $(\xi_i)_{i \in \mathbb{Z}}$. It follows from (1.5) that $\alpha^{-1}(u) \leq \beta^{-1}(2u)$. Hence, by Theorem 1.2 in Rio (1993), the following result holds.

$2\alpha(A, B) \leq \beta(A, B).$

PROPOSITION 1 [Rio, 1993]. – *Assume that the β-mixing coefficients of $(\xi_i)_{i \in \mathbb{Z}}$ satisfy the summability condition*

$$\sum_{n \geq 0} \beta_n < +\infty. \tag{2.4}$$

Let $\mathcal{L}_{2,\beta}(P)$ denote the class of numerical functions f such that

$$\|f\|_{2,\beta} = \sqrt{\int_0^1 \beta^{-1}(u)[Q_f(u)]^2\, du} < +\infty. \tag{2.5}$$

Then, for any f *in* $\mathcal{L}_{2,\beta}(P)$,

$$\sum_{t \in \mathbb{Z}} |\mathrm{Cov}\,(f(\xi_0), f(\xi_t))| \leq 4\|f\|_{2,\beta}^2,$$

and denoting by $\Gamma(f, f)$ *the sum of the series* $\sum_{t \in \mathbb{Z}} \mathrm{Cov}\,(f(\xi_0), f(\xi_t))$, *we have:*

$$(2.6) \qquad \begin{cases} \mathrm{Var}\, Z_n(f) \leq 4\|f\|_{2,\beta}^2 \\ \text{and} \\ \lim_{n \to +\infty} \mathrm{Var}\, Z_n(f) = \Gamma(f, f) \leq 4\|f\|_{2,\beta}^2. \end{cases}$$

Remark. – Note that $\sum_{n \geq 0} \beta_n = \int_0^1 \beta^{-1}(u)du$. So, under condition (2.4), $\mathcal{L}_{2,\beta}(P)$ contains the space $\mathcal{L}_\infty(P)$ of bounded functions. Moreover, we will prove in section 6 the following basic lemma.

LEMMA 1. – *Assume that the sequence of* β*-mixing coefficients of* $(\xi_i)_{i \in \mathbb{Z}}$ *satisfies condition* (2.4). *Then,* $\mathcal{L}_{2,\beta}(P)$*, equipped with* $\|.\|_{2,\beta}$ *is a normed subspace of* $\mathcal{L}_2(P)$ *satisfying* (2.1), *and for any* f *in* $\mathcal{L}_{2,\beta}(P)$, $\|f\|_2 \leq \|f\|_{2,\beta}$.

We now define the corresponding weak space as follows. Let $B(t) = \int_0^t \beta^{-1}(u)du$. For any measurable numerical function f, we set

$$(2.7) \qquad \Lambda_{2,\beta}(f) = \sup_{t \in]0,1]} Q_f(t)\sqrt{B(t)}.$$

Throughout the sequel, $\Lambda_{2,\beta}(P)$ denotes the space of measurable functions f on $\mathcal{X}$ such that $\Lambda_{2,\beta}(f) < +\infty$. Clearly $\mathcal{L}_{2,\beta}(P) \subset \Lambda_{2,\beta}(P)$ and

$$(2.8) \qquad \Lambda_{2,\beta}(f) \leq \|f\|_{2,\beta} \quad \text{for any } f \in \mathcal{L}_{2,\beta}(P).$$

Now we compare the so defined spaces with the Orlicz spaces of functions, and with the so-called weak Orlicz spaces of functions.

Comparison with Orlicz spaces

Let Φ be the class of increasing functions

$$\Phi = \Big\{\phi : \mathbb{R}^+ \to \mathbb{R}^+ : \phi \nearrow \text{ convex and differentiable,}$$
$$\phi(0) = 0,\ \lim_{+\infty} \frac{\phi(x)}{x} = \infty\Big\}.$$

For any $\phi \in \Phi$, we denote by $\mathcal{L}_{\phi,2}(P)$ the space of functions

$$\mathcal{L}_{\phi,2}(P) = \{h : \mathcal{X} \to \mathbb{R} \text{ s.t. } E_P(\phi(th^2)) < +\infty \text{ for some } t > 0\}.$$

$\mathcal{L}_{\phi,2}(P)$ is equipped with the following norm :

$$\|h\|_{\phi,2} = \inf \left\{ c > 0 : E_P\left(\phi\left(\left|\frac{h}{c}\right|^2\right)\right) \leq 1 \right\}.$$

This norm is the usual Orlicz norm associated with the function $x \to \phi(x^2)$.

Let us now define the corresponding weak Orlicz spaces. For any ϕ in Φ, let $\Lambda_\phi(P)$ be the space of measurable functions f on $\mathcal{X}$ such that

$$\sup_{u>0} [uP(\phi(f^2) > u)] = \sup_{t>0} [\phi(t^2)P(|f| > t)] < +\infty,$$

or equivalently such that

$$\Lambda_\phi(f) = \sup_{u\in]0,1]} ([\phi^{-1}(1/u)]^{-1/2} Q_f(u)) < +\infty.$$

Of course, by Markov's inequality, $\Lambda_\phi(h) \leq \|h\|_{\phi,2}$.

The purpose of the following lemma is to relate the $\mathcal{L}_{2,\beta}$-norm with $\|.\|_{\phi,2}$ and $\Lambda_\phi(.)$.

LEMMA 2. – *For any element ϕ of Φ, let define the dual function ϕ^* by $\phi^*(y) = \sup_{x>0} [xy - \phi(x)]$. If*

$$\int_0^1 \phi^*(\beta^{-1}(u))\, du < +\infty, \tag{2.9a}$$

then

$$\|f\|_{2,\beta} \leq \|f\|_{\phi,2} \sqrt{1 + \int_0^1 \phi^*(\beta^{-1}(u))\, du}\,, \tag{a}$$

and so $\mathcal{L}_{\phi,2}(P) \subset \mathcal{L}_{2,\beta}(P)$. If

$$\int_0^1 \phi^{-1}(1/u)\beta^{-1}(u) du < +\infty\,, \tag{2.9b}$$

then

$$\|f\|_{2,\beta} \leq \Lambda_\phi(f) \sqrt{\int_0^1 \phi^{-1}(1/u)\beta^{-1}(u) du}\,, \tag{b}$$

and so $\Lambda_\phi(P) \subset \mathcal{L}_{2,\beta}(P)$.

The main result

In a recent paper, Doukhan, Massart and Rio (1994) prove that, for any finite subset $\mathcal{G} \subset \mathcal{L}_{2,\beta}(P)$, the convergence of $\{Z_n(g) : g \in \mathcal{G}\}$ to a Gaussian random vector with covariance function Γ holds true. Moreover, a counterexample shows that fidi convergence may fail to hold if $\mathcal{G} \not\subset \mathcal{L}_{2,\beta}(P)$. So, it is quite natural to assume that $\mathcal{F} \subset \mathcal{L}_{2,\beta}(P)$. Now the question raises whether an integrability condition on the metric entropy with bracketing in $\mathcal{L}_{2,\beta}(P)$ is sufficient to imply the uniform CLT. In fact the answer is positive, as shown by the following theorem.

THEOREM 1. – *Let $(\xi_i)_{i\in\mathbb{Z}}$ be a strictly stationary and β-mixing sequence of random variables with common marginal distribution P. Assume that the sequence $(\beta_n)_{n\geq 0}$ satisfies* (2.4). *Let $\mathcal{F}$ be a class of functions f, $\mathcal{F} \subset \mathcal{L}_{2,\beta}(P)$. Assume that the entropy with bracketing with respect to $\|.\|_{2,\beta}$, which we denote by $H_\beta(\delta, \mathcal{F})$, satisfies the integrability condition*

$$\int_0^1 \sqrt{H_\beta(u, \mathcal{F})}\, du < +\infty. \tag{2.10}$$

Then,

(i) *the series* $\sum_{t\in\mathbb{Z}} \operatorname{Cov}(f(\xi_0), f(\xi_t))$ *is absolutely convergent over $\mathcal{F}$ to a nonnegative quadratic form $\Gamma(f, f)$, and*

$$(\Gamma(f, f))^{1/2} = \|f\|_\Gamma \leq 2\|f\|_{2,\beta}.$$

(ii) *There exists a sequence $(Z^{(n)})_{n>0}$ of Gaussian processes indexed by $\mathcal{F}$ with covariance function Γ and a.s. uniformly continuous sample paths such that*

Z_t $\quad Z_n = n^{-1/2} \sum_{i=1}^{n} (\delta_{\xi_i} - P)$

$$\sup_{f\in\mathcal{F}} |Z_n(f) - Z^{(n)}(f)| \to_P 0 \quad \text{as } n \to +\infty.$$

Applications of Theorem 1

1. *Orlicz spaces.* – Let ϕ be some element of Φ such that the mixing coefficients β_n satisfy

$$\sum_{n>0} (\phi')^{-1}(n)\beta_n < +\infty \qquad \lim_{x\to\infty} \frac{\phi(x)}{x} = \infty \tag{S.1}$$

and suppose that $\mathcal{F} \subset \mathcal{L}_{\phi,2}(P)$. Then (2.9*a*) holds [*see* Rio, 1993]. Hence it follows from Lemma 2 that $\mathcal{F}$ satisfies the assumptions of Theorem 1 if the metric entropy with bracketing of $\mathcal{F}$ in $\mathcal{L}_{\phi,2}(P)$ fulfills the integrability condition

$$\int_0^1 \sqrt{H_{[\,]}(t,\mathcal{F},\|.\|_{\phi,2})}dt < +\infty. \tag{2.11}$$

For example, if $\phi(x) = x^r$, $\mathcal{L}_{\phi,2}(P) = \mathcal{L}_{2r}(P)$, $\|.\|_{\phi,2}$ is the usual norm $\|.\|_{2r}$ and (S.1) means that the series $\sum_{n>0} n^{1/(r-1)}\beta_n$ is convergent.

As an application of (2.11), we can derive the following striking result. Assume that the mixing coefficients satisfy $\beta_k = O(b^k)$ for some b in $]0,1[$. There exists some $s > 0$ such that (S.1) holds with

$$\phi(x) = h(sx) = (1+sx)\log(1+sx) - sx. \tag{2.12}$$

$(\phi^{-1})(n)\,\beta_n$

We notice that $\mathcal{L}_{\phi,2}(P)$ is the space $L_{2\log^+}(P)$ of numerical functions f such that $E_P(f^2\log^+|f|) < \infty$ and it is equipped with a norm which is equivalent to the usual Orlicz norm in this space. Hence, by Lemma 2 and Theorem 1, the uniform CLT for the empirical process holds as soon as the entropy with bracketing of $\mathcal{F}$ in $L_{2\log^+}(P)$ satisfies the usual integrability condition.

2. *Weak Orlicz spaces.* – Let ϕ be some element of Φ such that $x \to x^{-r}\phi(x)$ is nondecreasing for some $r > 1$. Then, (2.9*b*) is equivalent to the summability condition of Herrndorf (1985) on the mixing coefficients:

$$\sum_{n>0} \phi^{-1}(1/\beta_n)\beta_n < +\infty \tag{S.2}$$

$\phi(x)$

β_n^{-1} $\beta_n^{-\frac{1}{r}}$

It follows from Lemma 2 that $\mathcal{F} \subset \Lambda_\phi(P)$ satisfies the assumptions of Theorem 1 if the entropy with bracketing of $\mathcal{F}$ with respect to $\Lambda_\phi(.)$ verifies

$$\int_0^1 \sqrt{H_{[\,]}(t,\mathcal{F},\Lambda_\phi)}\,dt < +\infty. \tag{2.13}$$

Some calculations (*cf.* Rio, 1993) show that (S.2) is stronger than (S.1). For example, if $\phi(x) = x^r$ for some $r > 1$, $\Lambda_\phi(P)$ is the usual weak $\mathcal{L}_{2r}(P)$-space $\Lambda_{2r}(P)$, equipped with the usual weak norm $\Lambda_{2r}(.)$ and (S.2) is equivalent to the convergence of the series $\sum_{n>0} \beta_n^{1-1/r}$, while (S.1) holds iff the series $\sum_{n>0} n^{1/(r-1)}\beta_n$ is convergent, which is a weaker condition. We refer the reader to application 3 of Theorem 1 in Doukhan, Massart and Rio (1994), for more about comparisons with the previous conditions of Ibragimov (1962) and Herrndorf (1985).

3. *Conditions involving the envelope function of $\mathcal{F}$.* – Let $\mathcal{G} \subset \mathcal{L}_\infty(P)$ be a class of functions satisfying the entropy condition

$$\int_0^1 \sqrt{H(t, \mathcal{G}, \|.\|_\infty)}\, dt < +\infty.$$

$\ell_\infty(P) \subseteq \ell_{2,\beta}(P)$

Let F be some element of $\mathcal{L}_{2,\beta}(P)$ satisfying $F \geq 1$ and $\mathcal{F} = \{gF : g \in \mathcal{G}\}$. Both Theorem 1 and the elementary inequality $\|gF\|_{2,\beta} \leq \|g\|_\infty \|F\|_{2,\beta}$ imply the uniform CLT.

4. *Conditions involving the $\mathcal{L}_2(P)$-entropy of $\mathcal{F}$.* – In this subsection, we consider the following problem: given a class $\mathcal{F} \subset \mathcal{L}_2(P)$, with metric entropy with bracketing $H_2(.)$ with respect to the metric induced by $\|.\|_2$, we want to find a condition on the mixing coefficients and on the $\mathcal{L}_2(P)$-entropy implying (2.10). Throughout application 4, we assume that the envelope function of $\mathcal{F}$ is in $\Lambda_{2r}(P)$, for some $r \in]1, +\infty]$, where we make the convention that $\Lambda_\infty(P) = \mathcal{L}_\infty(P)$. We shall prove in appendix C that condition (2.10) is satisfied if any of the three following conditions is fulfilled:

$$(2.15)\quad \begin{cases} \beta_n = O(n^{-b}) \quad \text{with } b > r/(r-1), \\ \text{and} \\ \displaystyle\int_0^1 v^{-r/(b(r-1))} H_2^{1/2}(v)\,dv < +\infty, \end{cases}$$

$1 + \theta \frac{1}{r-1}$

$$(2.16)\quad \begin{cases} H_2(u) = O(u^{-2\zeta}) \quad \text{with } \zeta \in]0, 1[, \\ \text{and} \\ \displaystyle\sum_{n>0} n^{-1/2} \beta_n^{(1-\zeta)(r-1)/(2r)} < +\infty, \end{cases}$$

$u \in [0,1]$

$$(2.17)\quad H_2(u) = O(|\log u|) \quad \text{and} \quad \sum_{n>0} \frac{1}{n} \sqrt{\frac{\sum_{k \geq n} \beta_k^{1-1/r}}{\log n}} < +\infty.$$

In particular, (2.15) and (2.16) are satisfied if $\beta_n = O(n^{-b})$, $H_2(u) = O(u^{-2\zeta})$ with $b(1-\zeta) > r/(r-1)$. Also (2.17) holds whenever $\beta_n = O(n^{-1}(\log n)^{-2-\delta})$ and $\mathcal{F}$ is the class of quadrants, which improves in this special case on Corollary 3 of Arcones and Yu (1994).

The outline of the paper is as follows: in section 3, the technique of blocking for mixing processes is applied to establish an upper bound on the mean of the supremum of Z_n over a finite class of bounded functions. Next, in section 4, we show how both the upper bound of section 3 and a generalization of Ossiander's method for proving tightness of the empirical process yield the stochastic equicontinuity of $\{Z_n(f) : f \in \mathcal{F}\}$ under the assumptions of Theorem 1. In section 5, we weaken the bracketing conditions, in the spirit of Andersen *et al.* (1988). The stochastic equicontinuity of Z_n is ensured in the bounded case by the following result, which is in fact the crucial technical part of the paper. In what follows, for any positive (non necessarily measurable) random element X, $\mathbb{E}(X)$ denotes the smallest expectation of the (measurable) random variables majorizing X.

THEOREM 2. – *Let σ be a positive number and let $\mathcal{F}_\sigma \subset \mathcal{L}_{2,\beta}$ be a class of functions satisfying the condition $\|f\|_{2,\beta} \leq \sigma$ for any f in $\mathcal{F}_\sigma$. Suppose that $\mathcal{F}_\sigma$ fulfills* (2.10) *and that for some $M \geq 1$, $|f| \leq M$ for any function f in $\mathcal{F}_\sigma$. Then there exists a constant K, depending only on $\theta = (\sum_{n \geq 0} \beta_n)^{1/2}$, such that, for any positive integer q,*

$$\mathbb{E}(\sup_{f \in \mathcal{F}} |Z_n(f)|) \leq K\left(\varphi(\sigma) + \frac{Mq\varphi^2(\sigma)}{\sigma^2\sqrt{n}} + \sqrt{n}M\beta_q\right). \tag{2.18}$$

3. A MAXIMAL INEQUALITY FOR β-MIXING PROCESSES INDEXED BY FINITE CLASSES OF FUNCTIONS

In order to establish the stochastic equicontinuity of the empirical process $\{Z_n(f) : f \in \mathcal{F}\}$, we need to control the mean of the supremum of the empirical process $Z_n(.)$ over a finite class of bounded functions. This control is performed via the approximation of the original process by conveniently defined independent random variables. The main argument is Berbee's coupling lemma.

LEMMA [Berbee (1979)]. – *Let X and Y be two r.v.'s taking their values in Borel spaces S_1 and S_2 respectively and let U be a r.v. with uniform distribution over* $[0,1]$, *independent of* (X, Y). *Then, 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:*

- *Y^* is independent of X and has the same distribution as Y.*
- *$\mathbb{P}(Y \neq Y^*) = \beta$, where β denotes the β-mixing coefficient between the σ-fields generated by X and Y respectively.*