On the Coupling of Dependent Random Variables and Applications.

Florence MERLEVÈDE and Magda PELIGRAD*

L.S.T.A. Université Paris 6 and University of Cincinnati

October 31, 2001

Abstract

In this paper we survey some results and further investigate the coupling of a sequence of dependent random variables with an independent one, having the same marginal distributions. The upper bound of the distance between the variables with the same rank is given in terms of mixing coefficients. We shall apply the coupling methods to derive uniform laws of large numbers for the dependent random processes under various types of dependence. We shall also discuss the importance of coupling for obtaining the central limit theorem for strongly mixing sequences.

Mathematics Subject Classifications (1991): 60 F 05, 60 F 15, 60 F 17, 60 F 20, 60 B 12.

Key words: coupling, weak dependent sequences, coefficients of dependence, uniform law of large numbers, central limit theorem.

Short Title: Coupling

---

*Supported in part by a Taft Research Grant and University of Paris X

1 Introduction

One of the most useful techniques in obtaining limit theorems for weakly dependent random variables is the coupling of the initial sequence with an independent one. Generally speaking a coupling theorem enables us to replace the initial dependent sequence by an independent one, having the same marginals. Moreover, a variable in the newly constructed sequence is independent of the past of the initial sequence and it is closed, in some sense, to the variable having the same rank. This approximation has to be carried out in a manner that permits limit theorems for the independent random variables to carry over directly to the dependent ones.

The idea of coupling is going back at least to Doeblin (1937) for the case of the Markov chains with a countable state space and further extended for more general Markov processes by Vasershtein (1969) and Griffeath (1975). The problem of coupling was also considered in ergodic theory, in the context of Bernoulli shifts by Ornstein (1974), and Shields (1973).

This method gained popularity for obtaining limit theorems for dependent random variables since the publication of two articles by Berkes and Philipp (1977, 1979). Their approach is based on estimates of the Prohorov-distance and the Strassen-Dudley Theorem (Strassen (1965), Dudley (1968)). This method was succesfully exploited by many authors including Berkes and Philipp (1977-78), Philipp (1979), Kuelbs and Philipp (1980), Dehling and Philipp (1982), Berger (1982), Dabrowski (1982), Eberlein (1983), Dudley and Philipp (1983), Borovkova, Burton and Dehling (1999).

In parallel, various approaches to coupling have been developed by different authors.

One of the method can be called the "partition method" and is similar to the one used in the ergodic theory setting. It was developed in papers by Goldstein (1979), Berbee (1979), Schwarz (1980), Bryc (1982), Bradley (1983) and Peligrad (2001-a). A variant of this method will be presented in the next section.

Another approach, developed first by the Hungarian school for deriving some approximation theorem for independent random variables, is the quantile transform. Major (1978) extended this construction to the dependent case introducing the conditional quantile transform. The method was further exploited by Rio (1995) and Peligrad (2001-b).

The paper is organized as follows. In Section 2, we shall present some coupling results based on the last two methods. Estimates of the distance between a variable in an original sequence and the corresponding one in the newly constructed one will be given in terms of mixing coefficients. Let us mention that the estimate given at the point (f) of Theorem 2.5 is new. Section 3 is devoted to the application of the coupling results to derive uniform laws of large numbers for dependent processes (some results in this section are new), whereas in Section 4, we show the importance of such coupling results for obtaining the central limit theorem for dependent sequences.

Throughout the paper, the notation $\mathbb{1}(A)$ means the indicator of A and $[\,.\,]$ the integer part as usual.

2 Some results on coupling

In this paper we shall use the following coefficients of dependence. Let $(\Omega, \mathcal{T}, P)$ be a probability space and let $\mathcal{A}$ and $\mathcal{B}$ two σ-algebras of $\mathcal{T}$.

Define the strong mixing coefficient by:

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup_{A\in\mathcal{A}, B\in\mathcal{B}} \left|\mathbb{P}\left(A \cap B\right) - \mathbb{P}(A)\mathbb{P}(B)\right|,$$

the version λ of the maximal coefficient of correlation by:

$$\lambda(\mathcal{A}, \mathcal{B}) = \sup_{A\in\mathcal{A}, B\in\mathcal{B}} \left|\mathbb{P}\left((A \cap B\right) - \mathbb{P}(A)\mathbb{P}(B)\right|/\mathbb{P}^{1/2}(B),$$

the φ mixing coefficient by:

$$\varphi(\mathcal{A}, \mathcal{B}) = \sup_{A\in\mathcal{A}, B\in\mathcal{B}} \left|\mathbb{P}\left(A \cap B\right) - \mathbb{P}(A)\mathbb{P}(B)\right|/\mathbb{P}(B)$$

and the absolute regular coefficient by:

$$\beta(\mathcal{A},\mathcal{B}) = \frac{1}{2}\sup\sum_{i=1}^{I}\sum_{j=1}^{J}|\mathbb{P}(A_i\cap B_j)-\mathbb{P}(A_i)\mathbb{P}(B_j)|\,,$$

where this latter sup is taken over all pairs of partitions $\{A_1,\cdots,A_I\}$ and $\{B_1,\cdots,B_J\}$ of Ω such that $A_i\in\mathcal{A}$ and each $B_j\in\mathcal{B}$.

Obviously $\alpha(\mathcal{A},\mathcal{B})\le\lambda(\mathcal{A},\mathcal{B})\le\varphi(\mathcal{A},\mathcal{B})$ and $\alpha(\mathcal{A},\mathcal{B})\le\beta(\mathcal{A},\mathcal{B})\le\varphi(\mathcal{A},\mathcal{B})$.

Throughout the paper we shall make use of the short notation $\gamma(X,Y)$ to denote $\gamma(\sigma(X),\sigma(Y))$, γ being one of the dependence coefficients.

We shall now recall a well-known coupling result which characterizes the coefficient β of absolutely regularity.

Theorem 2.1. *Let X and Y be random variables defined on $(\Omega,\mathcal{T},\mathbb{P})$ with values in a Polish space S. Let $\sigma(X)$ be a σ-algebra generated by X and U be a random variable uniformly distributed on $[0,1]$ independent of (X,Y). Then there exists a random variable Y^* measurable with respect of $\sigma(X)\vee\sigma(Y)\vee\sigma(U)$, independent of X and distributed as Y, and such that*

$$\mathbb{P}(Y\neq Y^*)=\beta(X,Y)\,.$$

This theorem is frequently named Berbee's Lemma (1979). Related results can be found in works by Goldstein (1979), Berkes and Philipp (1979), Schwarz (1980), Bryc (1982), Dehling and Philipp (1982) and, in the context of ergodic theory in Ornstein (1974) and Shields (1973). In all these papers, the constructions of the variable Y^* are equivalent. We shall sketch here the construction by Bryc (1982) which is versatile and leads to further interesting estimates for the distance between Y and Y^*.

Because the variables are S-valued, by completness, according to the result in Bryc (1982), we can reduce the problem to the case of discrete S-valued random variables. Leaving the technical aspect apart, the main step of the construction is to define the following probabilistic structure :

With the notation $e_{x,y} := \mathbb{P}(X = x, Y = y) - \mathbb{P}(X = x)\mathbb{P}(Y = y)$, we have

$$\mathbb{P}(X = x, Y = y, Y^* = z) = \begin{cases} \dfrac{e^+_{x,y} e^-_{x,z}}{\sum_t e^+_{x,t}} & \text{if } z \neq y \\ \mathbb{P}(X = x)\mathbb{P}(Y = y) & \text{if } z = y \text{ and } e_{x,y} \geq 0 \\ \mathbb{P}(X = x, Y = y) & \text{if } z = y \text{ and } e_{x,y} < 0\,. \end{cases}$$

It is easy to verify that with this construction, we have

1) Y^* distributed as Y,

2) Y^* independent of X,

and for all A, B and C Borel sets of S,

3) $\mathbb{P}(X \in A, Y \in B, Y \neq Y^*) = \sum_{x\in A} \sum_{y \in B} [\mathbb{P}(X = x, Y = y) - \mathbb{P}(X = x)\mathbb{P}(Y = y)]^+$,

4) $\mathbb{P}(X \in A, Y^* \in C, Y \neq Y^*) = \sum_{x\in A} \sum_{y \in C} [\mathbb{P}(X = x, Y = y) - \mathbb{P}(X = x)\mathbb{P}(Y = y)]^-$.

By this construction, one can easily obtain the following proposition.

Proposition 2.1. *Let Y be a random variable defined on $(\Omega, \mathcal{T}, \mathbb{P})$ with values in a Polish space (S, d). Let p be a positive number and s in S. Let Y^* be constructed as in Theorem 2.1. Then, for all $x \geq 0$*

(i) $\mathbb{P}\left(Y \neq Y^* | \sigma(X)\right) \leq \sup_{B\in\sigma(Y)} \left|\mathbb{P}(B|\sigma(X)) - \mathbb{P}(B)\right|$ *a.s.,*

(ii) $\mathbb{P}\left(d(Y, Y^*) \geq x\right) \leq 2\varphi(X, Y)\mathbb{P}\left(d(Y, s) \geq x/2\right)$,

and if we assume that $d^p(Y, s)$ is integrable, then

(iii) $\mathbb{E}\left(d^p(Y, Y^*)\right) \leq 2^p\varphi(X, Y)\mathbb{E}\left(d^p(Y, s)\right)$.

Point *(ii)* comes from Bryc (1982). Point *(i)* is probably well-known (a proof can be found in Peligrad (2001-a)). Point *(iii)* is obtained by integrating with respect to x in *(ii)*.

We shall establish now a new result for estimating the expectation of the distance between Y and Y^* in terms of the absolutely regular coefficient. Before stating it, we need to introduce the following definition : For any nonnegative random variable W, define the "upper tail" quantile function via

$$Q_W(u) = \inf\{t \geq 0 : \mathbb{P}(W > t) \leq u\}\,.$$

With this notation, we have

Proposition 2.2. *Let X and Y be two random variables defined on $(\Omega, \mathcal{T}, \mathbb{P})$ with values in a Polish space S. Let p be a positive number and s in S and assume that $d^p(Y,s)$ is integrable. Let Y^* be constructed as in Theorem 2.1. Then for every positive p*

$$\mathbb{E}\left(d^p(Y,Y^*)\right) \leq 2^{p+2}\int_0^{\beta(X,Y)} Q_{d^p(Y,s)}(u)du\,.$$

Remark 2.1. *In the case when $p \leq 2$, with an other approach the constant is sharper, namely 2^{2p}.*

Remark 2.2. *1) If $\mathbb{E}\left(d^{p+\delta}(Y,s)\right) < +\infty$ for a $\delta > 0$, then*

$$\mathbb{E}\left(d^p(Y,Y^*)\right) \leq 2^{p+2}\beta^{\frac{\delta}{p+\delta}}(X,Y)\Big(\mathbb{E}\left(d^{p+\delta}(Y,s)\right)\Big)^{\frac{p}{p+\delta}}\,.$$

2) Since

$$\begin{aligned}\mathbb{E}\Big(d^p(Y,s)\mathbb{1}\left(d(Y,s) > Q_{d(Y,s)}(\beta(X,Y))\right)\Big) &\leq \int_0^{\beta(X,Y)} Q_{d^p(Y,s)}(u)du \\ &\leq \mathbb{E}\Big(d^p(Y,s)\mathbb{1}\left(d(Y,s) \geq Q_{d(Y,s)}(\beta(X,Y))\right)\Big),\end{aligned}$$

we get

$$\mathbb{E}\left(d^p(Y,Y^*)\right) \leq 2^{p+2}\mathbb{E}\Big(d^p(Y,s)\mathbb{1}\left(d(Y,s) \geq Q_{d(Y,s)}(\beta(X,Y))\right)\Big)\,.$$

Proof of Proposition 2.2

Set $T := Q_{d(Y,s)}\left(\beta(X,Y)\right)$ and notice that the triangle's inequality and the fact that

Y and Y^* have the same distribution, yield for all $s \in S$

$$\begin{aligned}
\mathbb{E}(d^p(Y,Y^*)) &= \mathbb{E}(d^p(Y,Y^*)\mathbb{1}(Y \neq Y^*)) \\
&\leq \mathbb{E}\{(d(Y,s)+d(Y^*,s))^p \mathbb{1}(Y \neq Y^*)\} \\
&\leq 2^p\mathbb{E}(d^p(Y,s)\mathbb{1}(Y \neq Y^*)) + 2^p\mathbb{E}(d^p(Y^*,s)\mathbb{1}(Y \neq Y^*)) \\
&\leq 2^p\mathbb{E}(d^p(Y,s)\mathbb{1}(d(Y,s) \leq T)\mathbb{1}(Y \neq Y^*)) + 2^p\mathbb{E}(d^p(Y,s)\mathbb{1}(d(Y,s) > T)) \\
&\quad +2^p\mathbb{E}(d^p(Y^*,s)\mathbb{1}(d(Y^*,s) \leq T)\mathbb{1}(Y \neq Y^*)) + 2^p\mathbb{E}(d^p(Y^*s)\mathbb{1}(d(Y^*s) > T)) \\
&:= I_1 + I_2 + I_1^* + I_2^* .
\end{aligned}$$

First we use Theorem 2.1 which yields

$$I_1 \leq 2^p T^p \beta(X,Y) .$$

Now by using the equality

$$Q^p_{d(Y,s)}(\beta(X,Y)) = Q_{d^p(Y,s)}(\beta(X,Y)) \tag{2.1}$$

combined with the fact that $Q_{d^p(Y,s)}(u)$ is a nonincreasing function, we get

$$I_1 \leq 2^p \int_0^{\beta(X,Y)} Q_{d^p(Y,s)}(u)du \tag{2.2}$$

In order to treat I_2, let us mention some results. Recall first that if U is a random variable uniformly distributed on the interval $[0,1]$ and if W is a nonnegative random variable, then the r.v. $Q_W(U)$ has the same distribution as the r.v. W, and then

$$\mathbb{E}(W) = \int_0^1 Q_W(u)du .$$

This last equality applied to the r.v. $d^p(Y,s)\mathbb{1}(d(Y,s) > T)$ entails that

$$\mathbb{E}(d^p(Y,s)\mathbb{1}(d(Y,s) > T)) = \int_0^1 Q_{d^p(Y,s)\mathbb{1}(d(Y,s)>T)}(u)du .$$

We also have

$$Q_{d^p(Y,s)\mathbb{1}(d(Y,s)>T)}(u) \leq \begin{cases} Q_{d^p(Y,s)}(u) & \text{if } u < \beta(X,Y) \\ 0 & \text{if } u \geq \beta(X,Y) . \end{cases}$$

From these last considerations we derive that

$$I_2 \leq 2^p \int_0^{\beta(X,Y)} Q_{d^p(Y,s)}(u)du\,. \tag{2.3}$$

Gathering (2.2) and (2.3) and since similar bounds are valid for I_1^* and I_2^*, one gets the result. □

Let us mention now the following result of Bradley (1983) : " For two $\sigma-$fields $\mathcal{A}$ and $\mathcal{B}$ such that one of them (say $\mathcal{A}$) is completely atomic with exactly N atoms, one has $\beta(\mathcal{A},\mathcal{B}) \leq N\alpha(\mathcal{A},\mathcal{B})$." By using this result, we can easily see that some coupling results involving the strong mixing coefficient can be obtained for Y taking its values in S and having a finite numbers of values, and the result will strongly depend on N. For deriving more general strong approximation theorems for random variables taking their values in a Polish space, under the strong mixing condition, we would like to mention the two following results due to Bradley (1983).

Theorem 2.2. *Suppose Y is a random variable, defined on a probability space $(\Omega,\mathcal{T},\mathbb{P})$ and taking its values in a Polish space $(S,\mathcal{S})$. Suppose $\mathcal{A}$ is a σ-field $\subset \mathcal{T}$. Suppose U is a uniform-$[0,1]$ random variable that is independent of the σ-field $\mathcal{A}\vee\sigma(Y)$.*

Suppose N is a positive integer, and $\mathcal{H} := \{H_1,H_2,\ldots,H_N\}$ is a partition of S, with $H_i \in \mathcal{S}$ for all $i=1,2,\ldots,N$.

Then there exists a random variable Y^* which takes its values in $(S,\mathcal{S})$ and is measurable with respect to the σ-field $\mathcal{A}\vee\sigma(Y)\vee\sigma(U)$, such that Y^* has the following three properties:

(a) Y^* is independent of the σ-field $\mathcal{A}$.

(b) Y^* has the same distribution (on $(S,\mathcal{S})$) as the random variable Y.

(c) One has that

$$\mathbb{P}(Y^* \text{ and } Y \text{ are not elements of the same } H_i \in \mathcal{H}) \;\leq\; (8N)^{1/2}\alpha(\mathcal{A},Y)\;.$$

This theorem is taken from Bradley (1983, Theorem 2), and we can mention that it sharpens “codification” of arguments from the proofs of earlier strong approximation theorems in Berkes and Philipp (1977, 1979) and Bryc (1981).

Next theorem and comments that follows are due to Bradley (2001) in a personal communication.

Theorem 2.3. *Suppose (S, d) is a Polish space. Let $\mathcal{S}$ denote the σ-field on S generated by the open balls in the metric d.*

Suppose Y is a random variable, defined on a probability space $(\Omega, \mathcal{T}, \mathbb{P})$ and taking its values in $(S, \mathcal{S})$. Suppose $\mathcal{A}$ is a σ-field $\subset \mathcal{F}$. Suppose U is a uniform-$[0,1]$ random variable which is independent of $\mathcal{A} \vee \sigma(Y)$

Suppose further that $\varepsilon > 0$, $\delta > 0$, N is a positive integer, $D \in \mathcal{S}$, $\mathbb{P}(X \in D) \geq 1 - \delta$, and that there exist points $a_1, a_2, \ldots, a_N \in D$ such that $\forall a \in D, \exists k \in \{1, 2, \ldots, N\}$ satisfying $d(a_k, a) \leq \varepsilon$.

Then there exists a random variable Y^* which takes its values in $(S, \mathcal{S})$ and is measurable with respect to the σ-field $\mathcal{A} \vee \sigma(Y) \vee \sigma(U)$, such that Y^* has the following three properties:

(a) Y^* is independent of the σ-field $\mathcal{A}$.

(b) Y^* has the same distribution as the random variable Y.

(c) $\mathbb{P}(d(Y, Y^*) > 2\varepsilon) \leq \delta + (8(N+1))^{1/2}\alpha(\mathcal{A}, Y)$.

This theorem evolved through several strong approximation theorems:

1. Berkes and Philipp (1977, Theorem 2 (special case)) involving the measure of dependence $\alpha(.,.)$ and random variables taking their values in $\mathbb{Z}^d$ for positive integers d.

2. Berkes and Philipp (1979, Theorem 2), involving the measure of dependence $\varphi(.,.)$ and random variables taking their values in complete separable metric spaces. The “struc-

ture" of the statement and proof of that result is in a certain sense imitated by the statement and proof of Theorem 2.3.

3. Bryc (1981), involving the measure of dependence $\alpha(.,.)$ and real-valued random variables.

4. Bradley (1983, Theorem 3), involving $\alpha(.,.)$ and real-valued random variables.

5. Novak (2000, Lemma 4), involving $\alpha(.,.)$ and random variables taking their values in $\mathbb{Z}^d$ for positive integers d.

In the context when Y is a real random variable, we would like to mention the following useful statement which is a corollary of Theorem 2.3 (see Theorem 3 of Bradley (1983)).

Theorem 2.4. *Suppose (S, d) is a Polish space. Suppose X and Y are random variables taking their values on S and $\mathbb{R}$, respectively. Suppose U is a uniform-$[0,1]$ random variable which is independent of (X, Y); and suppose q and γ are positive numbers such that $q \leq \|Y\|_\gamma < \infty$. Then there exists a random variable Y^* which takes its values in $(S, \mathcal{S})$ and is measurable with respect to the σ-field $\sigma(X) \vee \sigma(Y) \vee \sigma(U)$, such that Y^* has the following three properties:*
(a) Y^* is independent of X.
(b) Y^* has the same distribution as the random variable Y.
(c) $\mathbb{P}(|Y^* - Y| \geq q) \leq 18 \left(\|Y\|_\gamma / q\right)^{\gamma/(2\gamma+1)} \left(\alpha(X, Y)\right)^{2\gamma/(2\gamma+1)}$.

This result has been widely used in the case when Y is bounded (i.e. $\gamma = +\infty$) for instance in works by Bosq (1998) and Guillou and Merlevède (2001).

Next theorem gathers a variety of coupling results for S-valued sequences. The upper bounds for the discrepancy of the variables with the same rank are given in terms of mixing coefficients. The points (a), (b), (d) and (e) are proven in Bryc (1982). The point (c) is proven in Peligrad (2001-a). The point (f) is new.

Theorem 2.5. *Let $(X_n, n \geq 1)$ be a sequence of random variables with values in a Polish space (S,d). Denote by $\mathcal{F}_1^n = \sigma(X_1, ..., X_n)$. Then, we can redefine $\{X_n, n \geq 1\}$ onto a richer probability space (which supports a sequence $(\delta_n, n \geq 1)$ of independent uniformly [0,1]-distributed random variables and independent of $\{X_n, n \geq 1\}$) together with a sequence $\{X_n^*, n \geq 1\}$ of independent random variables such that : for each $1 \leq m \leq n$, we have*

a) X_m^* is independent of $\mathcal{F}_1^{m-1}$,

b) X_m^* has the same distribution as X_m,

c) $\mathbb{P}\left(X_m \neq X_m^* | \mathcal{F}_1^{m-1}\right) \leq \sup_{B \in \sigma(X_m)} \left|\mathbb{P}\left(X_m \in B | \mathcal{F}_1^{m-1}\right) - \mathbb{P}(X_m \in B)\right|$, a.s.

d) $\mathbb{P}(X_m \neq X_m^*) = \beta\left(\mathcal{F}_1^{m-1}, \sigma(X_m)\right)$,

e) $\mathbb{P}(d(X_m, X_m^*) \geq x) \leq 2\varphi(\mathcal{F}_1^{m-1}, \sigma(X_m))\mathbb{P}(d(X_m, s) \geq x/2)$, for all $x \geq 0$

and if $\mathbb{E}(d^p(X_m, s)) < \infty$ for a positive p and $s \in S$,

f) $\mathbb{E}(d^p(X_m, X_m^*)) \leq 2^{p+2} \int_0^{\beta(\mathcal{F}_1^{m-1}, \sigma(X_m))} Q_{d^p(X_m,s)}(u)du$

Remark 2.3. *Of course now the following question arises : " does the point (f) of the above-mentioned approximation theorem also hold in the context of strongly mixing sequences with values in an infinite dimensional space? "*

Following Dehling (1983), it turns out that the answer is negative without any additional assumptions. In his paper he has constructed an example of strongly mixing sequences of ℓ^2-valued random variables X_n with values in an infinite dimensional space, which cannot be approximated by independent random variables X_n^* in such a way that $X_n - X_n^*$ converges to zero in probability.

By integrating (c) of Theorem 2.5 on the set $\bigcap_{i=1}^{m}(X_i \neq X_i^*)$, we easily deduce by an iterative procedure the following useful corollary which estimates the joint distributions of discrepancies between $(X_i)_{i\geq 1}$ and $(X_i^*)_{i\geq 1}$. Before stating it we need to introduce the following coefficient of dependence : let $(X_n, n \geq 1)$ be a sequence of random variables with values in a Polish space S. For $m \geq n$ we denote by $\mathcal{F}_n^m$ the sigma-algebra generated by the variables $(X_n, X_{n+1}, \cdots, X_m)$ and set

$$\tilde{\varphi}_k = \sup_{m\geq 1} \sup \left\{ |\mathbb{P}(B|A) - \mathbb{P}(B)| ; \mathbb{P}(A) \neq 0, A \in \mathcal{F}_1^m, B \in \mathcal{F}_{m+k}^{m+k} \right\} . \tag{2.4}$$

The definition of $\tilde{\varphi}_k$ is using the sigma-algebra generated by only one random variable in the future after k steps. This coefficient has already been used in many papers including Deddens, Peligrad, Yang (1987) and Rio (2000).

Corollary 2.1. (Peligrad (2001-a))

Under the setting of Theorem 2.5, we obtain that : For all $1 \leq i_1 < i_2 < \cdots < i_m$,

$$\mathbb{P}\left(\bigcap_{j=1}^{m} X_{i_j} \neq X_{i_j}^* \right) \leq \tilde{\varphi}_1^m . \tag{2.5}$$

Moreover for $1 \leq \ell \leq n$

$$\mathbb{P}\left(X_i \neq X_i^* \text{ at least } \ell \text{ times for } 1 \leq i \leq n \right) \leq \frac{4}{\sqrt{2\pi}} \frac{2^{n+1}}{\sqrt{n}} \frac{\tilde{\varphi}_1^\ell}{\sqrt{1-\tilde{\varphi}_1^\ell}} . \tag{2.6}$$

Due to Remark 2.3, the case of strong mixing processes has to be analyzed in the real case and with a different approach that is called conditional quantile transform.

In order to construct Y^* we shall use, as in Rio (1995), the method based on the conditional quantile transform, introduced and studied by Major (1978) :

Let us denote: $F(y) = P(Y \leq y)$, $F_{\mathcal{A}}(y) = P(Y \leq y|\mathcal{A})$, and

$$V = F_{\mathcal{A}}(Y-0) + U(F_{\mathcal{A}}(Y) - F_{\mathcal{A}}(Y-0)) ,$$

where U is a random variable uniformly distributed on $[0,1]$ and independent of Y and $\mathcal{A}$.

By Lemma F.1 in Rio (2000), page 161, or by the proof of Bradley (2000), we can see that that $P(V \leq x|\mathcal{A}) = x$, *a.s.*, and, as a consequence, by integrating this expression, we deduce that V is independent on $\mathcal{A}$ and uniformly distributed on [0,1]. In addition:

$$F_{\mathcal{A}}^{-1}(V) = Y \text{ a.s.},$$

where we use the notation of generalized inverse: $h^{-1}(y) = \inf\{t \in R : h(t) \geq y\}$.

We define now:

$$Y^* = F^{-1}(V).$$

According to Major (1978), this choice of X^* minimizes the distance in $\mathbb{L}_1$ between X and any other variable distributed as X and independent of $\mathcal{A}$.

By the above notations and considerations, since V is uniformly distributed on [0,1], clearly:

$$\mathbb{E}|Y^* - Y| = \mathbb{E}|F^{-1}(V) - F_{\mathcal{A}}^{-1}(V)| = \mathbb{E}(\int_0^1 |F^{-1}(u) - F_{\mathcal{A}}^{-1}(u)|du) \tag{2.7}$$

In order to obtain an upper bound for (2.7), we shall introduce the following weaker version of the coefficients of dependence : Let $(\Omega, \mathcal{T}, P)$ be a probability space and let X be a real random variable defined on Ω. Let $\mathcal{A}$ and $\mathcal{B}$ two σ-algebras of $\mathcal{T}$.

Define the strong mixing coefficient by:

$$\overline{\alpha}(\mathcal{A}, X) = \sup_{A \in \mathcal{A}, x \in I\!R} |\mathbb{P}\left((X \geq x) \cap A\right) - \mathbb{P}(X \geq x)\mathbb{P}(A)|\,,$$

the λ-version of the maximal coefficient of correlation by:

$$\overline{\lambda}(\mathcal{A}, X) = \sup_{A \in \mathcal{A}, x \in I\!R} |\mathbb{P}\left(((X \geq x) \cap A\right) - \mathbb{P}(X \geq x)\mathbb{P}(A)| / \mathbb{P}^{1/2}(X \geq x)\,,$$

the φ mixing coefficient by:

$$\overline{\varphi}(\mathcal{A}, X) = \sup_{A \in \mathcal{A}, x \in I\!R} |\mathbb{P}\left((X \geq x) \cap A\right) - \mathbb{P}(X \geq x)\mathbb{P}(A)| / \mathbb{P}(X \geq x)$$

Theorem 2.6. (Peligrad (2001-b)) *Let Y be a real -valued integrable random variable that is defined on a probability space $(\Omega, \mathcal{T}, \mathbb{P})$. Suppose $\mathcal{A}$ is a sub σ-algebra of $\mathcal{T}$. Suppose U is a uniform random variable on $[0, 1]$ that is independent on the σ-algebra generated by $\mathcal{A}$ and Y. Then, there is a random variable Y^*, which is measurable with respect to* *the σ-algebra $\mathcal{A} \cup \sigma(Y) \cup \sigma(U)$ with the following properties:*

$$
\begin{array}{ll}
(a) & Y^* \text{ is independent of the } \sigma\text{-algebra } \mathcal{A} \\
(b) & Y^* \text{ has the same distribution as } X \\
(c) & \mathbb{E}|Y - Y^*| \le 2 \int_0^{\overline{\alpha}(\mathcal{A},Y)} Q_{|X|}(u)du \\
(d) & \mathbb{E}|Y - Y^*| \le 2\overline{\lambda}(\mathcal{A}, Y) \int_0^\infty \mathbb{P}^{1/2}(|Y| > t)dt \\
(d) & \mathbb{E}|Y - Y^*| \le 2\overline{\varphi}(\mathcal{A}, Y)\mathbb{E}|Y|\,.
\end{array}
$$

Remark 2.4. *The point (c) extends a result of Rio (1995) who considered the case where Y is bounded. The point (d) improves on the constant used in the proof of Corollary 2.4 of Bryc (1982).*

The idea of the proof of Theorem 2.6 is the following. Since $Y = Y^+ - Y^-$, we shall assume for simplicity that Y is almost surely positive. (2.7) gives

$$
\mathbb{E}|Y^* - Y| = \mathbb{E}\left(\int_0^\infty |\mathbb{P}(Y > u) - \mathbb{P}(Y > u|\mathcal{A})|du\right)\,.
$$

By classical arguments, we get

$$
\mathbb{E}|\mathbb{P}(Y > u) - \mathbb{P}(Y > u|\mathcal{A})| \le \sup_{A\in\mathcal{A}} |\mathbb{P}\left((Y > u) \cap A\right) - \mathbb{P}(Y > u)\mathbb{P}(A)|\,.
$$

Now, by taking into account our definition of the mixing coefficients we obtain the following upper bound estimates :

$$
\begin{aligned}
\mathbb{E}|\mathbb{P}(Y > u) - \mathbb{P}(Y > u|\mathcal{A})| &\le \min\left(\overline{\alpha}(\mathcal{A}, Y), \mathbb{P}(Y > u)\right)\,, \\
\mathbb{E}|\mathbb{P}(Y > u) - \mathbb{P}(X > u|\mathcal{A})| &\le \overline{\lambda}(\mathcal{A}, Y)\mathbb{P}^{1/2}(Y > u)\,, \\
\mathbb{E}|\mathbb{P}(Y > u) - \mathbb{P}(Y > u|\mathcal{A})| &\le \overline{\varphi}(\mathcal{A}, Y)\mathbb{P}(Y > u)
\end{aligned}
$$

and the results follow by integrating the above relations on $[0, \infty)$.

Concerning the case of a sequence of real random variables, Theorem 2.6 gives the following useful result.

Theorem 2.7. (Peligrad (2001-b)) *Let $(X_n, n \geq 1)$ be a sequence of real-valued random variables. Denote by $\mathcal{F}_1^n = \sigma(X_1, ..., X_n)$. Then, we can redefine $\{X_n, n \geq 1\}$ onto a richer probability space (which supports a sequence $(\delta_n, n \geq 1)$ of independent uniformly [0,1]-distributed random variables and independent of $\{X_n, n \geq 1\}$) together with a sequence $\{X_n^*, n \geq 1\}$ of independent random variables such that : for each $1 \leq m \leq n$, we have*

a) X_m^* *is independent of* $\mathcal{F}_1^{m-1}$,

b) X_m^* *has the same distribution as* X_m,

c) $\mathbb{E}|X_m - X_m^*| \leq 2 \int_0^{\overline{\alpha}(\mathcal{F}_1^{m-1}, X_m)} Q_{|X_m|}(u) du$,

d) $\mathbb{E}|X_m - X_m^*| \leq 2\overline{\lambda}(\mathcal{F}_1^{m-1}, X_m) \int_0^{\infty} \mathbb{P}^{1/2}(|X_m| > t)\, dt$,

e) $\mathbb{E}|X_m - X_m^*| \leq 2\overline{\varphi}(\mathcal{F}_1^{m-1}, X_m)\mathbb{E}|X_m|$.

Let us notice that if $(X_n, n \geq 1)$ is a sequence of real-valued random variables such that for all $n \geq 1$, $\mathbb{P}(a_n \leq X_n \leq b_n) = 1$, then, by using Theorem 2.4 and the relation (5.25) in Rio (2000), we can redefine $\{X_n, n \geq 1\}$ onto a richer probability space on which there exists a sequence $\{X_n^*, n \geq 1\}$ of independent random variables such that, for each $n \geq 1$, X_n and X_n^* have the same distribution and

$$\mathbb{E}|X_n - X_n^*| \leq 3\left(2(b_n - a_n)\right)^{1/3} \left(\overline{\alpha}(\mathcal{F}_1^{n-1}, X_n)\right)^{2/3} . \tag{2.8}$$

Compare to point (c) of Theorem 2.6 which gives in this case

$$\mathbb{E}|X_n - X_n^*| \leq 2(b_n - a_n)\overline{\alpha}(\mathcal{F}_1^{n-1}, X_n),$$

the inequality (2.8) seems to be less powerful. However (2.8) can lead to better upper bound estimates when $(b_n - a_n)$ increases fastly when n tends to infinity.

3 Application to the uniform laws of large numbers for dependent processes

We shall see now how the above-mentioned coupling results are very useful in proving asymptotic results and in this section we focus on the uniform laws of large numbers.

Uniform laws of averages extend the classical laws of large number from a single function to a collection of such functions. The uniform law addresses the question to find a class $\mathcal{F}$ of functions such that

$$\lim_{n\to\infty} \sup_{f\in\mathcal{F}} \frac{1}{n}\left|\sum_{i=1}^{n} f(X_i) - \mathbb{E}f(X_i)\right| = 0\,, \tag{3.1}$$

in some stochastic sense.

If the convergence in (3.1) holds with probability one, $\mathcal{F}$ is said to be a Glivenko-Cantelli class. Alternatively $\mathcal{F}$ is said to satisfy a uniform law of averages with respect to the sequence $(X_i)_{i\geq 1}$.

Vapnik and Cervonenkis (1981) found necessary and sufficient conditions under which a class of functions $\mathcal{F}$ is almost surely uniformly convergent with respect to an i.i.d. process $(X_i)_{i\geq 1}$. The conditions involve the asymptotic behavior of covering numbers associated with the class $\mathcal{F}$ and sample sequences of the process $(X_i)_{i\geq 1}$. For more details, we refer the reader to Vapnik and Cervonenkis (1981) or Pollard (1984).

Concerning weak dependent sequences, the uniform convergence has been studied by many authors. Among them Philipp (1982), Yukich (1986), Massart (1988), Yu (1991), Doukhan, Massart, Rio (1995). All these results used specific rates of convergence to zero of the mixing coefficients.

In this paper we present recent and new results about the uniform law of large numbers for weak dependent sequences. This section is divided in several parts. First we deal with absolutely regular sequences, after with $\tilde{\varphi}$-mixing sequences and finally with strong mixing sequences. In these results no condition is imposed on the rates of convergence to zero of the mixing coefficients.

Since the proofs are mainly based on coupling results, we shall assume in what in follows the probability space to be large enough to accomodate all the variables constructed.

Let $\mathcal{F}$ be a class of real-valued measurable functions on S. A measurable function $F : S \to R$ is said to be an envelope for $\mathcal{F}$, if $|f(x)| \leq F(x)$, for every $f \in \mathcal{F}$. The class $\mathcal{F}$ is said to be uniformly bounded if for some $0 \leq K < \infty$, $F = K$ is an envelope for $\mathcal{F}$. If the class $\mathcal{F}$ is uncountable it must satisfy regularity conditions in order to insure that no problems of measurability will arise when considering quantities such as the supremum. Following Pollard, we use the term *permissible* to indicate that the class $\mathcal{F}$ satisfies such conditions (for more details see Dudley (1978) or Pollard (1984)).

3.1 The case of absolutely regular sequences

Nobel and Dembo (1993) obtained an uniform law for the class of absolute regular sequences without imposing any condition on the rate of convergence to 0 of the dependence coefficients. In fact they obtained the almost sure uniform convergence for absolutely regular strictly stationary sequences as a direct consequence of uniform convergence for independent random variables without imposing any additional conditions on the class of functions. Their result is the following :

Theorem 3.1. *Let $(X_j)_{j\in\mathbb{Z}}$ be a sequence of strictly stationary S-valued variables defined on $(\Omega, \mathcal{T}, \mathbb{P})$ and such that $\beta\left(\mathcal{F}^0_{-\infty}, \mathcal{F}_n^{+\infty}\right) \to 0$ as $n \to \infty$. Let $(X_j^*)_{j\in\mathbb{Z}}$ be an independent sequence with values in S such that for all $j \in \mathbb{Z}$ X_j^* and X_j have the same distribution. Let $\mathcal{F}$ be a permissible class of real-valued functions on S, with an envelope F such that $F \in \mathbb{L}_1(\mathbb{P})$ and such that*

$$\sup_{f\in\mathcal{F}}\left|\frac{1}{n}\sum_{i=1}^{n}\left(f(X_i^*) - \mathbb{E}f(X_i^*)\right)\right| \to 0 \ \ a.s..$$

Then

$$\sup_{f\in\mathcal{F}}\left|\frac{1}{n}\sum_{i=1}^{n}\left(f(X_i) - \mathbb{E}f(X_i)\right)\right| \to 0 \ \ a.s.$$

Let us notice that the coefficient of absolute regularity used by Nobel and Dembo involves all the future of the process which means that the process is ergodic and the proofs of their theorem benefitted by the results in ergodic theory.

3.2 The case of φ-mixing sequences

Concerning the case of $\tilde{\varphi}$-mixing sequences, Peligrad (2001-a) has obtained an uniform strong law without imposing any rates of convergence to zero on $\tilde{\varphi}_n$ and without requiring ergodicity. She has shown that the class of strictly stationary and $\tilde{\varphi}$-mixing sequences satisfies a uniform law of averages if an independent sequence having the same marginals does. Therefore the result available for independent sequences translates to similar results for $\tilde{\varphi}$-mixing sequences. Her result can be formulated as follows :

Theorem 3.2. *Let $\mathcal{F}$ be a permissible class of real-valued functions on S having envelope $F \in \mathbb{L}_1(\mathbb{P})$. Then $\mathcal{F}$ satisfies an uniform strong law of averages with respect to a strictly stationary $\tilde{\varphi}$-mixing sequence of S-valued random variables if it satisfies the same law with respect to an i.i.d.*

The proof of the above theorem can be divided in two steps: the first one involves a class of functions $\mathcal{F}$ which is uniformly bounded and the second one involves a truncation argument. In fact by analyzing these two steps, we can give two new results on the uniform law of convergence for $\tilde{\varphi}$-mixing sequences without imposing stationarity. This is the aim of the next statements.

Proposition 3.1. *Let $(X_j)_{j\geq 1}$ be a sequence of S-valued variables such that $\tilde{\varphi}_n \to 0$ as $n \to \infty$. Let $(X_j^*)_{j\geq 1}$ be an independent sequence such that for all $j \geq 1$, X_j^* and X_j have the same distribution. Let $\mathcal{F}$ be a class of real-valued functions on S which is uniformly bounded and such that for a fixed $p \geq 1$ and each $0 \leq j \leq p-1$ we have*

$$(i) \qquad \lim_{k\to\infty} \sup_{f\in\mathcal{F}} \left| \frac{1}{k} \sum_{i=0}^{k-1} \left(f(X^*_{ip+j}) - \mathbb{E} f(X^*_{ip+j}) \right) \right| = 0 \ \ a.s.,$$

where $X_0^* = 0$.

Then

$$\text{(ii)} \qquad \lim_{n\to\infty} \sup_{f\in\mathcal{F}} \left| \frac{1}{n} \sum_{i=1}^{n} \left(f(X_i) - \mathbb{E}f(X_i)\right) \right| = 0 \ \ a.s.$$

Proof of Proposition 3.1 Without loss of generality we may assume that the sequence $\{f(X_j)\}_{j\geq 1}$ has zero mean.

Let p be a natural number $p < n$. Denote by $k = \left[\frac{n}{p}\right]$, and extend the vector $(X_1, \ldots, X_n)$ to the sequence $(0, X_1, \ldots, X_n, 0, 0, \ldots)$. Define $Y_{i,j} = X_{ip+j}$ for $0 \leq i \leq k-1$, $0 \leq j \leq p-1$. For each $0 \leq j \leq p-1$, consider the vector $(Y^*_{0,j}, \ldots, Y^*_{k-1,j})$ constructed in Theorem 2.5, and notice that

$$\text{(3.2)} \quad \sup_{f\in\mathcal{F}} \frac{1}{n}\left|\sum_{i=1}^{n} f(X_i)\right| \leq \sup_{f\in\mathcal{F}} \frac{1}{n}\left|\sum_{i=1}^{n} f(X_i) - \sum_{i=0}^{k}\sum_{j=0}^{p-1} f(Y^*_{i,j})\right| + \sup_{f\in\mathcal{F}} \frac{1}{kp}\sum_{j=0}^{p-1}\left|\sum_{i=0}^{k-1} f(Y^*_{i,j})\right|.$$

Under assumption (i), the second term in the right-hand side term in the above inequality tends to zero by letting n to tend to infinity and then p. Then it is enought to show that

$$\text{(3.3)} \qquad \lim_{n\to\infty} \sup_{f\in\mathcal{F}} \frac{1}{n}\left|\sum_{i=1}^{n} f(X_i) - \sum_{i=0}^{k-1}\sum_{j=0}^{p-1} f(Y^*_{i,j})\right| = 0\,.$$

By the triangle inequality we have:

$$\text{(3.4)} \qquad \frac{1}{n}\left|\sum_{i=1}^{n} f(X_i) - \sum_{i=0}^{k}\sum_{j=0}^{p-1} f(Y^*_{i,j})\right| \leq \max_{0\leq j\leq p-1} \frac{1}{k}\left|\sum_{i=0}^{k-1}\left(f(Y_{i,j}) - f(Y^*_{i,j})\right)\right|.$$

Fix j, $0 \leq j \leq p-1$ and let η be such that $\tilde{\varphi}_p < \frac{1}{2(1+\eta)}$. Let $\ell_k = \left[k \log \frac{1}{2(1+\eta)} / \log \tilde{\varphi}_p\right]$. We write

$$\text{(3.5)} \qquad \frac{1}{k}\left(\sum_{i=0}^{k-1} f(Y_{i,j}) - f(Y^*_{i,j})\right) =: I_1 + I_2$$

where

$$I_1 = \frac{1}{k}\sum_{i=0}^{k-1}\left(f(Y_{i,j}) - f(Y^*_{i,j})\right) \mathbb{1}\left(\sum_{i=0}^{k-1} \mathbb{1}(Y_{i,j} \neq \tilde{Y}_{i,j}) < \ell_k\right)$$

and

$$I_2 = \frac{1}{k}\sum_{i=0}^{k-1}\left(f(Y_{i,j}) - f(Y^*_{i,j})\right)\mathbb{I}\left(\sum_{i=0}^{k-1}\mathbb{I}(Y_{i,j} \neq Y^*_{i,j}) \geq \ell_k.\right)$$

Since $\mathcal{F}$ is bounded, we can find $C > 0$ such that for all $f \in \mathcal{F}$

$$I_2 \leq CI\left(\sum_{i=0}^{k-1}\mathbb{I}(Y_{i,j} \neq \tilde{Y}_{i,j}) \geq \ell_k\right) \; a.s. \tag{3.6}$$

and also

$$I_1 \leq C\frac{\ell_k}{k} = C\left(\log\frac{1}{2(1+\eta)}\right)/\log\tilde{\varphi}_p \tag{3.7}$$

By combining (3.4) - (3.7) we can find a constant K such that

$$\begin{aligned}\frac{1}{n}\left|\sum_{i=1}^{n} f(X_i) - \sum_{i=0}^{k-1}\sum_{j=0}^{p-1} f(Y^*_{i,j})\right| \;\leq\; & K\Bigg\{\left(\sum_{i=0}^{k-1}\mathbb{I}(Y_{i,j} \neq Y^*_{i,j}) \geq \ell_k\right) + \\ & \left[\log\frac{1}{2(1+\eta)}\right]/\log\tilde{\varphi}_p\Bigg\} \qquad \text{a.s..}\end{aligned}$$

By applying now (2.6) of Corollary 2.1, it is easy to obtain via Borel-Cantelli lemma that

$$\mathbb{P}\left(\sum_{i=0}^{k-1}\mathbb{I}(Y_{i,j} \neq Y^*_{i,j}) \geq \ell_k \; i.o.\right) = 0$$

which yields

$$\limsup_{n\to\infty}\sup_{f\in\mathcal{F}}\frac{1}{n}\left|\sum_{i=1}^{n} f(X_i) - \sum_{i=0}^{k-1}\sum_{j=0}^{p-1} f(Y^*_{i,j})\right| \leq K\left[\log\frac{1}{2(1+\eta)}\right]/\log\tilde{\varphi}_p \quad \text{a.s..}$$

The result follows by letting $p \to \infty$. □

Remark 3.1. *In view of the proof, Theorem 3.1 holds with convergence in probability remplacing almost sure convergence everywhere.*

In the next theorem we remove the condition of boundedness of the class of function $\mathcal{F}$ and replace it with an integrability condition imposed on the envelope of the class. However our result is now a convergence in probability instead of an almost sure result.

Theorem 3.3. *Assume that $(X_j)_{j\geq 1}$ and $(X_j^*)_{j\geq 1}$ are as in Theorem 3.1. Let $\mathcal{F}$ be a class of functions with an envelope F such that $\{F(X_i)\}_{i\geq 1}$ is an uniformly integrable family and*

$$(i) \qquad \sup_{f\in\mathcal{F}} \left| \frac{1}{n} \sum_{i=1}^{n} (f(X_i^*) - \mathbb{E}f(X_i^*)) \right| \xrightarrow{\mathbb{P}} 0\,.$$

Then

$$(ii) \qquad \sup_{f\in\mathcal{F}} \left| \frac{1}{n} \sum_{i=1}^{n} (f(X_i) - \mathbb{E}f(X_i)) \right| \xrightarrow{\mathbb{P}} 0$$

Proof of Theorem 3.3

Once again we may assume that the sequence $\{f(X_j)\}_{j\geq 1}$ has zero mean.

Let C be a positive constant and let $f \in \mathcal{F}$. Denote by $f_1 = f\mathbb{1}(F \leq C)$ and by $f_2 = f\mathbb{1}(F > C)$. We obviously have:

(3.8)

$$\sup_{f\in\mathcal{F}} \frac{1}{n} \left| \sum_{i=1}^{n} f(X_i) \right| = \sup_{f_1, f\in\mathcal{F}} \frac{1}{n} \left| \sum_{i=1}^{n} (f_1(X_i) - \mathbb{E}f_1(X_i)) \right| + \sup_{f_2, f\in\mathcal{F}} \frac{1}{n} \left| \sum_{i=1}^{n} (f_2(X_i) - \mathbb{E}f_2(X_i)) \right| .$$

By Proposition 3.1 and Remark 3.1, since the family $\{f_1, f \in \mathcal{F}\}$ has a bounded envelope, we will obtain that the first term in the right-hand side of equality (3.8) converges to zero in probability as soon as assumption (i) of Theorem 3.1 holds in probability. The following statement together with assumption (ii) of Theorem 3.3 prove this point.

Claim 1: *Let $(Z_j)_{j\geq 1}$ be a sequence of independent S-valued random variables defined on $(\Omega, \mathcal{T}, \mathbb{P})$ and let $\mathcal{F}$ be a permissible class of real-valued functions on S such that*

$$\sup_{f\in\mathcal{F}} \left| \frac{1}{n} \sum_{i=1}^{n} (f(Z_i) - \mathbb{E}f(Z_i)) \right| \xrightarrow{\mathbb{P}} 0\,.$$

Then for each p, $p \geq 1$ and all $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, where $k = [\frac{n}{p}]$ we have

$$\sup_{f\in\mathcal{F}} \left| \frac{1}{k} \sum_{j=1}^{k} \left(f(Z_{i_j}) - \mathbb{E}f(Z_{i_j})\right) \right| \xrightarrow{\mathbb{P}} 0\,.$$

Proof of the claim : The proof is based on the fact that if X and Y are two S-valued random variables independent such that for all $f \in \mathcal{F}$ we have $\mathbb{E}f(Y) = 0$, then for all

$\varepsilon > 0$

$$\mathbb{P}\left(\sup_{f\in\mathcal{F}} |f(X)| > \varepsilon\right) \le \mathbb{P}\left(\sup_{f\in\mathcal{F}} |f(X)+f(Y)| > \varepsilon\right) . \tag{3.9}$$

We give here a short proof of (3.9) based on an suggestion of Pollard (personal communication). Notice first that since $\mathbb{E}f(Y) = 0$ and X and Y are independent

$$f(X) = \mathbb{E}(f(X) + f(Y)|X) .$$

Then

$$\begin{aligned}
\mathbb{P}\left(\sup_{f\in\mathcal{F}} |f(X)| > \varepsilon\right) &= \mathbb{P}\left(\sup_{f\in\mathcal{F}} |\mathbb{E}\left(f(X)+f(Y)\big|X\right)| > \varepsilon\right) \\
&\le \mathbb{P}\left(\mathbb{E}\left(\sup_{f\in\mathcal{F}} |f(X)+f(Y)|\big|X\right) > \varepsilon\right) \\
&= \mathbb{P}\left(\mathbb{E}\Big(\sup_{f\in\mathcal{F}} |f(X)+f(Y)|\big|X\Big) > \varepsilon\, ,\, \sup_{f\in\mathcal{F}} |f(X)+f(Y)| \le \varepsilon\right) \\
&\quad + \mathbb{P}\left(\mathbb{E}\Big(\sup_{f\in\mathcal{F}} |f(X)+f(Y)|\big|X\Big) > \varepsilon\, ,\, \sup_{f\in\mathcal{F}} |f(X)+f(Y)| > \varepsilon\right) \\
&\le \mathbb{P}\left(\sup_{f\in\mathcal{F}} |f(X)+f(Y)| > \varepsilon\right) \square
\end{aligned}$$

Now in order to treat the last term in the right-hand side of equality (3.8), it suffices to notice that

$$\mathbb{E} \sup_{f_2, f\in\mathcal{F}} \frac{1}{n}\left|\sum_{i=1}^{n} (f_2(X_i) - \mathbb{E}f_2(X_i))\right| \le \frac{2}{n}\sum_{i=1}^{n} \mathbb{E}|F(X_i)|\mathbb{I}(|F(X_i)| > C)$$

The result follows by the uniform integrability of $\{F(X_i)\}_{i\ge 1}$, by letting $n \to \infty$ and then $C \to \infty$. $\square$

3.3 The case of strong mixing sequences

Concerning the case of strong mixing sequences of real-valued random variables, some new results have been established recently by Peligrad (2001-b). The proofs are mainly based on the coupling result (c) of Theorem 2.7 and on a decomposition in "p-sequences" as in the proof of Proposition 3.1. The results are again obtained without imposing rates on the strong mixing coefficients and without requiring ergodicity. She has shown that this class of strong mixing sequences satisfies an uniform weak law of averages if an independent sequence having the same marginal does. Therefore the results available for independent sequences imply similar results for strong mixing sequences.

Let us also notice that since, as we mentioned above, the proofs are mainly based on the coupling result for strongly mixing sequences. Due to Remark 2.3, results for strong mixing sequences taking their values in a Polish space cannot be derived by a similar approach.

For a sake of simplicity, we set $\overline{\alpha}_n = \overline{\alpha}(\mathcal{F}_1^{n-1}, X_n)$, where $\overline{\alpha}$ is defined in the previous section.

Theorem 3.4. (Peligrad (2001-b) *Assume* $\{X_n\}_{n\geq 1}$ *is a sequence of identically distributed real-valued random variables in* $\mathbb{L}_1$ *such that* $\overline{\alpha}_n \to 0$ *as* $n \to \infty$. *Let* $\{X_n^*\}_{n\geq 1}$ *be a sequence of independent identically distributed real-valued random variables such that* X_1^* *has the same distribution as* X_1. *Let* $\mathcal{F}$ *be a class of Lipschitz functions with the Lipschitz constants bounded by* C $(C > 0)$ *such that*

(i)
$$\lim_{n\to\infty} \mathbb{E}\sup_{f\in\mathcal{F}} \frac{1}{n}\left|\sum_{i=1}^{n} (f(X_i^*) - \mathbb{E}f(X_i^*))\right| = 0\,.$$

Then

(ii)
$$\lim_{n\to\infty} \mathbb{E}\sup_{f\in\mathcal{F}} \frac{1}{n}\left|\sum_{i=1}^{n} (f(X_i) - \mathbb{E}f(X_i))\right| = 0\,.$$

In the statement of Theorem 3.4, let us notice that the assumption that the variables have the same distribution can be relaxed by imposing some weaker forms of stationarity. In this direction, we would like to mention the following result :

Proposition 3.2. *If we assume in Theorem 3.4 that* $\{X_n\}_{n\geq 1}$ *is uniformly integrable instead of stationarity then Theorem 3.4 holds.*

Proof of Proposition 3.2

To prove the result, we consider the "p-sequences" as in the proof of Proposition 3.1, for each $1 \leq j \leq p-1$, the vector $(Y^*_{0,j}, \dots, Y^*_{k-1,j})$ being this time constructed as in Theorem 2.7. Assuming that $\mathbb{E}f(X_i) = 0$ for all i, we notice that by the triangle's inequality, we have

$$\begin{aligned} \mathbb{E}\sup_{f\in\mathcal{F}} \frac{1}{n}\left|\sum_{i=1}^{n} f(X_i)\right| &\leq \mathbb{E}\sup_{f\in\mathcal{F}} \frac{1}{n}\sum_{j=0}^{p-1} \Big|\sum_{i=0}^{k-1} \big(f(Y_{i,j}) - f(Y^*_{i,j})\big)\Big| \\ &\quad + \mathbb{E}\sup_{f\in\mathcal{F}} \frac{1}{kp}\sum_{j=0}^{p-1}\left|\sum_{i=0}^{k-1} f(Y^*_{i,j})\right| . \end{aligned} \tag{3.10}$$

First let us notice that it is easy to see that conclusion of Claim 1 (used for the proof of Theorem 3.3) holds if we replace everywhere the convergence in probability by the $\mathbb{L}_1$-convergence. Then under assumption (i), the last term in the right-hand side of inequality (3.10) tends to zero by letting n to tend to infinity. Now, to treat the first term, we first notice that since f is Lipschitz, by Theorem 2.7, we derive that

$$\begin{aligned} \mathbb{E}\sup_{f\in\mathcal{F}} \frac{1}{n}\sum_{j=0}^{p-1} \Big|\sum_{i=0}^{k-1} \big(f(Y_{i,j}) - f(Y^*_{i,j})\big)\Big| &\leq \frac{2}{n} C \sum_{j=0}^{p-1}\sum_{i=0}^{k-1} \int_0^{\overline{\alpha}_p} Q_{|X_{ip+j}|}(u)du \\ &\leq 2C \sup_{1\leq i\leq n} \int_0^{\overline{\alpha}_p} Q_{|X_i|}(u)du\,. \end{aligned}$$

To obtain the desired result we let first n tend to infinity and then we apply the following lemma :

Lemma 3.1. *Suppose that $\{X_n\}_{n\geq 1}$ is an uniformly integrable sequence of random variables, then*

$$\lim_{\varepsilon\to 0}\sup_{i\geq 1} \int_0^{\varepsilon} Q_{|X_i|}(u)du = 0 \tag{3.11}$$

Proof of Lemma 3.1 : Let suppose that $\operatorname{ess\,sup}\sup_{i\geq 1}|X_i| = +\infty$, otherwise (3.11) is trivial. Now notice that

$$\sup_{i\geq 1}\int_0^{\varepsilon} Q_{|X_i|}(u)du \leq \sup_{i\geq 1}\mathbb{E}\big\{\,|X_i|\,\mathbb{I}(|X_i| \geq Q_{|X_i|}(\varepsilon))\big\}\ ,$$

which converges to zero when $\varepsilon \to 0$ by using the fact that $\lim_{\varepsilon\to 0} Q_{|X_i|}(\varepsilon) = +\infty$ together with the uniform integrability condition. □

We can also formulate the following new result using strong mixing coefficients.

Theorem 3.5. *Assume $\{X_n\}_{n\geq 1}$ is a sequence of uniformly integrable, centered real-valued random variables and such that $\overline{\alpha}_n \to 0$ as $n \to \infty$. Then*

$$\lim_{n\to\infty} \mathbb{E} \max_{1\leq j\leq n} \frac{1}{n}\Big|\sum_{i=1}^{j} X_i\Big| = 0 .$$

Proof of Theorem 3.5

Let us first truncate the random variables in the following way : Take A be a positive number and set

$$X_i' = X_i \mathbb{1}\left(|X_i| \leq A\right) - \mathbb{E}X_i \mathbb{1}\left(|X_i| \leq A\right)$$

and

$$X_i'' = X_i \mathbb{1}\left(|X_i| > A\right) - \mathbb{E}X_i \mathbb{1}\left(|X_i| > A\right) .$$

Now we set $S_k = \sum_{i=1}^k X_i$ and $S_k' = \sum_{i=1}^k X_i'$ and we notice that

$$\frac{1}{n}\mathbb{E}\max_{1\leq j\leq n} |S_j| \leq \frac{\sum_{i=1}^n \mathbb{E}|X_i''|}{n} + \frac{1}{n}\mathbb{E}\max_{1\leq j\leq n} |S_j'| . \tag{3.12}$$

Under the condition of uniform integrability of the sequence $\{X_i\}$, it is clear that the first term in the right-hand side of the above inequality tends to zero by letting n to tend to infinity and after A. Then it remains to treat the second term of inequality (3.12).

Now let p be a natural number $p < n$. Denote by $k = \left[\dfrac{n}{p}\right]$, and extend the vector $(X_1', \ldots, X_n')$ to the sequence $(0, X_1', \ldots, X_n', 0, 0, \ldots)$. Define $Y_{i,j}' = X_{ip+j}'$ for $0 \leq i \leq k-1$, $0 \leq j \leq p-1$. For each $0 \leq j \leq p-1$, consider the vector $({Y'}^*_{0,j}, \ldots, {Y'}^*_{k-1,j})$ constructed in Theorem 2.7, and notice that

$$\begin{aligned}
\frac{1}{n}\mathbb{E}\max_{1\leq j\leq n} |S_j'| &= \frac{1}{n}\mathbb{E}\max_{1\leq m\leq p}\max_{1\leq \ell\leq k} \Big|\sum_{j=0}^{m-1}\sum_{i=0}^{\ell-1} X_{ip+j}'\Big| \\
&\leq \frac{1}{n}\mathbb{E}\max_{1\leq m\leq p}\max_{1\leq \ell\leq k} \Big|\sum_{j=0}^{m-1}\sum_{i=0}^{\ell-1} {Y'}^*_{i,j}\Big| + \frac{1}{n}\mathbb{E}\max_{1\leq m\leq p}\max_{1\leq \ell\leq k} \Big|\sum_{j=0}^{m-1}\sum_{i=0}^{\ell-1} \left({Y'}_{i,j} - {Y'}^*_{i,j}\right)\Big| . \qquad (3.13)
\end{aligned}$$

First we treat the last term in the right-hand side of the above inequality. By using Theorem 2.7, we derive that

$$\begin{aligned}\frac{1}{n}\mathbb{E}\max_{1\le m\le p}\max_{1\le \ell\le k}\Big|\sum_{j=0}^{m-1}\sum_{i=0}^{\ell-1}\left(Y'_{i,j}-Y'^{*}_{i,j}\right)\Big| &\le \frac{1}{n}\sum_{j=0}^{p-1}\sum_{i=0}^{k-1}\mathbb{E}|Y'_{i,j}-Y'^{*}_{i,j}|\\ &\le 2\sup_{1\le i\le n}\int_0^{\overline{\alpha}_p} Q_{|X'_i|}(u)du\\ &\le 2A\overline{\alpha}_p\end{aligned}$$

which converges to zero by letting respectively n and after p to tend to infinity. To treat now the first term in the right-hand side of inequality (3.13), first we write

$$\frac{1}{n}\mathbb{E}\max_{1\le m\le p}\max_{1\le \ell\le k}\Big|\sum_{j=0}^{m-1}\sum_{i=0}^{\ell-1}Y'^{*}_{i,j}\Big| \le \frac{1}{n}\sum_{j=0}^{p-1}\mathbb{E}\max_{1\le \ell\le k}\Big|\sum_{i=0}^{\ell-1}Y'^{*}_{i,j}\Big|.$$

By using Rosenthal's inequality applied to the independent sequence (see e.g. Hall and Heyde (1980)-Theorem 2.11), there exists a constant C such that

$$\mathbb{E}\max_{1\le \ell\le k}\Big|\sum_{i=0}^{\ell-1}Y'^{*}_{i,j}\Big| \le C\left\{\left(\sum_{i=0}^{k-1}\mathbb{E}(Y'^{*}_{i,j})^2\right)^{1/2}+\mathbb{E}\max_{1\le i\le k-1}|Y'^{*}_{i,j}|\right\}.$$

It follows that

$$\frac{1}{n}\mathbb{E}\max_{1\le m\le p}\max_{1\le \ell\le k}\Big|\sum_{j=0}^{m-1}\sum_{i=0}^{\ell-1}Y'^{*}_{i,j}\Big| \le AC\left(\sqrt{\frac{p}{n}}+\frac{1}{n}\right).$$

which converges to zero by letting n to tend to infinity.

Gathering all these considerations, we obtain the desired result. □

4 Application to the Central Limit Theorem

In this section we shall see how the coupling results of Theorem 2.7 can be used to solve the central limit theorem question for stationary strongly mixing sequences, say $X := \{X_k, k\in\mathbb{Z}\}$. In this section, we mean by strongly mixing sequences, sequences such that

$$\alpha_n := \alpha\left(\mathcal{F}_{-\infty}^{0},\mathcal{F}_{n}^{+\infty}\right) \to 0 \text{ as } n\to\infty$$

where $\mathcal{F}_m^p := \sigma(X_j, m \leq j \leq p)$.

Since 1956, a vaste body of work has been devoted to study the limiting behaviour of strongly mixing sequences. This is a large class of random variables which contains both weakly dependent sequences and sequences with long range dependence. Examples include time series, Gaussian processes and Markov processes. These processes appear in other branches of mathematics, as well as statistics and mathematical physics, giving rise to a great deal of interest in their asymptotic properties.

The question concerning the central limit theorem in this setting is the following: under what assumptions, besides that of strong mixing, do there exist real numbers $a_1, a_2, a_3, \ldots$ and positive numbers $b_1, b_2, b_3, \ldots$ with $b_n \to \infty$ as $n \to \infty$, such that

$$\frac{S_n - a_n}{b_n} \xrightarrow{\mathcal{D}} N \sim \mathcal{N}(0,1) \ \text{, as } n \to \infty, \tag{4.14}$$

where $S_n = \sum_{i=1}^n X_i$.

By answering a conjecture of Bradley (1997), Merlevède and Peligrad (2000) obtained the following sharp central limit theorem that extends previous results by Ibragimov and Linnik (1971) and Doukhan, Massart and Rio (1994).

Theorem 4.1. *Suppose that $\{X_k, k \in \mathbb{Z}\}$ is a strictly stationary, centered, strong mixing sequence with finite second moment. Assume that*

$$\liminf_{n\to\infty} \frac{\mathbb{E}(S_n^2)}{n} > 0\,. \tag{4.15}$$

and

$$\int_0^{\alpha_n} Q^2_{|X_0|}(u)du = o(\frac{1}{n}) \ \text{as } n \to \infty. \tag{4.16}$$

Then

$$\frac{S_n}{\sqrt{\frac{\pi}{2}}\mathbb{E}|S_n|} \xrightarrow{\mathcal{D}} N \sim \mathcal{N}(0,1) \ \text{, as } n \to \infty. \tag{4.17}$$

The result of the above theorem was predicted by Bradley (1997). In his paper he formulated the following conjecture: Under the assumptions of Theorem 4.1, there exists a sequence $b_n \to \infty$ as $n \to \infty$, such that $\dfrac{S_n}{b_n}$ converges in distribution to a centered Gaussian random variable with variance 1.

The proof of Theorem 4.1 in Merlevède and Peligrad (2000) is based on the Bernstein-type blocking arguments and the coupling results. In an earlier version of this paper, the coupling was used only for identifying the normalizing sequence b_n. In a recent paper of Merlevède (2001), the proof does not require coupling. we would like to mention here the idea of proof based on coupling.

The proof Theorem 4.1 involves several steps.

<u>*Step 1*</u> *:* As a consequence of (4.15) and (4.16), we proved that $\sigma_n^2 := \operatorname{Var}(S_n)$ has the representation :

$$\sigma_n^2 = nh(n), \tag{4.18}$$

where $h(n)$ is a slowly varying function of n.

<u>*Step 2*</u> *:* For preparing the decomposition in "big" and "small" blocks, we constructed two sequences p_n and q_n which converge to infinity and such that $q_n = o(p_n)$. In fact the construction of p_n and q_n is the crucial part of the proof. The sequences are somehow implicit solutions of an equation involving the mixing coefficients and a continuous approximation of the quantile function.

<u>*Step 3*</u> *:* The proof continues with the following truncation : Set $T_n = Q_{|X_0|}(\alpha_{q_n})$ and

$$X_i' = \begin{cases} X_i \mathbb{1}(|X_i| \leq T_n) - \mathbb{E}X_i \mathbb{1}(|X_i| \leq T_n) & \text{if } X \text{ is an unbounded sequence} \\ X_i & \text{if ess.sup} X = A \text{ a.s.} \end{cases}$$

(4.19) and

$$X_i'' = \begin{cases} X_i \mathbb{1}(|X_i| > T_n) - \mathbb{E}X_i \mathbb{1}(|X_i| > T_n) & \text{if } X \text{ is an unbounded sequence} \\ 0 & \text{if ess.sup} X = A \text{ a.s.} \end{cases}$$

and let $S_n' = \sum_{i=1}^n X_i'$ and $S_n'' = \sum_{i=1}^n X_i''$. Then $S_n = S_n' + S_n''$.

Step 4 : At this step we reduce the proof to a central limit theorem for S'_n properly normalized by showing that $\dfrac{S''_n}{\sqrt{n}}$ converges to zero in $\mathbb{L}_1$.

Step 5 : Here we partion the index set $1, 2, 3, \cdots$ into an alternating sequence of "big" blocks of size p_n and "small" blocks of size q_n (in fact we will obtain k_n "big" blocks and $k_n + 1$ "small" blocks, where $k_n = \left[\dfrac{n}{p_n + q_n}\right]$).

Let $Y'_1, Y'_2, Y'_3, \cdots$ denote the respective sums of the X'_k's in the "big" blocks each having p_n summands, and let $Z'_1, Z'_2, Z'_3, \cdots$ denote the respective sums of the X'_k's in the "small" blocks each having q_n summands.

Moreover we use the following selection of b_n :

$$b_n^2 = k_n \sigma_{p_n}^2 \,,$$

and prove that the asymptotic behaviour of $\left\{\dfrac{\sum_{j=1}^{k_n+1} Z'_j}{b_n}\right\}_{n\geq 1}$ is negligible for the convergence in distribution.

Step 6 : At this step, we use coupling to study the asymptotic behaviour of $\left\{\dfrac{\sum_{j=1}^{k_n} Y'_j}{b_n}\right\}_{n\geq 1}$. To this aim, we consider a sequence $\left\{Y_j'^*\right\}_{1\leq i\leq k_n}$ of independent real random variables each respectively distributed as Y'_j and constructed as in Theorem 2.7. Since for all $1 \leq j \leq k_n$, $|Y'_j| \leq 2p_n Q_{|X_0|}(\alpha_{q_n})$, point (c) of Theorem 2.7 yields

$$\frac{E\left|\sum_{j=1}^{k_n}\left(Y'_j - Y_j'^*\right)\right|}{b_n} \leq \frac{8p_n k_n \alpha_{q_n} Q_{|X_0|}(\alpha_{q_n})}{b_n}\,, \tag{4.20}$$

which converges to 0 under (4.15) and our selection of p_n and q_n. This result implies that the proof of Theorem 4.1 is reduced to study the limiting behaviour of $\left\{\dfrac{\sum_{j=1}^{k_n} Y_j'^*}{b_n}\right\}_{n\geq 1}$.

Step 7 : We show that $\dfrac{\sum_{j=1}^{k_n} Y_j'^*}{b_n}$ converges in distribution to a standard normal random variable by checking the Lindeberg's condition.

Step 8 : To end the proof we identify the normalizing sequence b_n. To this aim we prove that $\left\{\dfrac{S_n}{b_n}\right\}_{n\geq 1}$ is an uniformly integrable family. Notice that (4.20) entails that $\left\{\dfrac{\sum_{j=1}^{k_n}\left(Y'_j - Y_j'^*\right)}{b_n}\right\}_{n\geq 1}$ is an uniformly integrable family. This consideration together with Steps 4, 5 and 7 imply the desired result. We finish the proof by using Theorem 5.4 in Billingsley (1968) which yields

$$\lim_{n\to\infty} \frac{b_n}{\sqrt{\frac{\pi}{2}}E|S_n|} = 1\,.$$

For the sake of applications, we give the following corollary of Theorem 4.1 in term of conditions imposed to mixing rates and to moments of individual summands. It extends the corresponding results of Ibragimov (1962).

Corollary 4.1. *Suppose that* $X := \{X_k, k \in \mathbb{Z}\}$ *is a strictly stationary, centered, strong mixing sequence which satisfies (4.15). In addition if*
1) X *has moments of order* $2+\delta$ *finite, for a* $\delta > 0$ *and*

$$n\alpha_n^{\frac{\delta}{2+\delta}} \to 0 \text{ as } n \to \infty \tag{4.21}$$

or if
2) X *is bounded and*

$$n\alpha_n \to 0 \text{ as } n \to \infty, \tag{4.22}$$

then (4.17) holds.

Let us notice that under the additional assumption $\lim_{n\to\infty} \frac{\mathbb{E}(S_n^2)}{n} = \sigma^2 > 0$, all the results of type (4.17) hold with the normalization $\sigma\sqrt{n}$ instead of $\sqrt{\frac{\pi}{2}}\mathbb{E}|S_n|$.

Moreover we can mentioned that the functional version of Theorem 4.1 has also been established in the paper of Merlevède and Peligrad (2000), and it can be formulated as follows :

Theorem 4.2. *Assume that the conditions of Theorem 4.1 are satisfied. Then* W_n *converges in distribution to* W *in the Skorohod space* $D([0,1])$*, where* $W_n(t) = \frac{\sum_{i=1}^{[nt]} X_i}{\sqrt{\frac{\pi}{2}}\mathbb{E}|S_n|}$*, and* W *is the standard Brownian motion on* $[0,1]$.

Acknoledgement

The authors would like to thank Richard Bradley for providing us with useful informations regarding estimates for the coupling error under strong mixing conditions for multivariate random variables. This paper was written when Magda Peligrad visited the University of Paris X.

References

[1] Arcones, M.A. and Yu, B.,1994. Central limit theorems for empirical and U-processes of stationary mixing sequences. *J.Theoret.Probab.***7**, 47-71.

[2] Berbee, H.P.C., 1979. Random walks with stationary increments and renewal theory, *Math. Center Tracts* **112**. Amsterdam.

[3] Berger, E., 1982. Fastsichere Approximation von Partialsummen unabhängiger und stationärer ergodischer Folgen von Zufallsvektoren, Diss. Univ. Gottingen.

[4] Berkes, I., Philipp, W., 1977/78. An almost sure invariance principle for the empirical distribution function of mixing random variables, *Z. Wahr. verw. Gebiete* **41**, 115-137.

[5] Berkes, I., Philipp, W., 1979. Approximation theorems for independent and weakly dependent random variables, *Ann. Probability* **7**, 29-54.

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

[7] Borovkova, S., Burton,R., Dehling, H.G., 1999. Limit theorems for functionals of mixing processes with applications to U-Statistics and dimension estimation. *University of Groningen IWI-preprint* 99-5-10.

[8] Bosq, D., 1998. Nonparametric statistics for stochastic processes, estimation and prediction. Second Edition. *Springer.*

[9] Bradley, R.C., 1983. Approximation theorems for strongly mixing random variables, *Michigan Math. J.* **30**, 69-81.

[10] Bradley, R.C., 1997. On quantiles and the central limit theorem question for strongly mixing sequences. *J. Theor. Probab.* **10** 507-555.

[11] Bradley, R.C., 2000. A strong approximation theorem of Rio. *Personal Communication* .

[12] Bryc, W., 1981. Sequences of weakly dependent random variables, *Ph. D. Dissertation (In Polish).*

[13] Bryc, W., 1982. On the approximation theorem of Berkes and Philipp, *Demonstratio Matematica* **15, 3**, 807-816.

[14] Dabrowski, A. R., 1982. A note on a theorem of Berkes and Philipp for dependent sequences, *Statistic and Probability Letters* **1**, 53-55.

[15] Deddens, J., Yang, T. and Peligrad, M., 1987. On strong consistency of kernel estimators under dependence assumptions. *Proceedings of the Sixth Pannonian Symposium of Mathematical Statistics*, D. Reidel Publishing Company, 33-43.

[16] Dehling, H., and Philipp, W., 1982. Almost sure invariance principles for weakly dependent vector valued random variables, *Ann. Probability* **10**, 689-701.

[17] Dehling, H., 1983. A note on a theorem of Berkes and Philipp, *Z. Wahr. verw. Gebiete* **62**, 39-42.

[18] Doeblin, W., 1937. Exposé de la théorie des chaînes simples constantes de Markov à un nombre fini d'états, *Rev. Math. de l'Union Interbalkanique* **2**, 77-105.

[19] Doukhan, P., Massart, P., Rio, E., 1994. The functional central limit theorem for strongly mixing processes, *Ann. Inst. H. Poincaré Probab. Statist.* **30**, 63-82.

[20] Doukhan, P., Massart, P., Rio, E., 1995. Invariance principles for absolutely regular empirical processes, *Ann. Inst. H. Poincaré Probab. Statist.* **31**, 393-427.

[21] Dudley, R.M., 1968. Distances of probability measures and random variables, *Ann. Math. Stat.* **39**, 1563-1572.

[22] Dudley, R.M., 1978. Central limit theorems for empirical measures, *Ann. Probab.* **6**, 899-929.

[23] Dudley, R.M., and Philipp, W., 1983. Invariance principles for sums of Banach space valued random elements and empirical processes, *Z. Wahr. verw. Gebiete* **62**, 509-552.

[24] Eberlein, E., 1983. Strong approximation of very weak Bernoulli processes, *Z. Wahr. verw. Gebiete* **62**, 17-37.

[25] Goldstein, S., 1979. Maximal coupling, *Z. Wahr. verw. Gebiete* **46**, 193-204.

[26] Griffeath, D., 1975. A maximal coupling for Markov chains, *Z. Wahr. verw. Gebiete* **31**, 95-106.

[27] Guillou, A., and Merlevède, F., 2001. Second-order properties of the blocks of blocks bootstrap for density estimators for continuous time processes, *submitted.*

[28] Hall, P. and Heyde, C. C. 1980. *Martingale limit theory and its application.* North-Holland, New-York.

[29] Ibragimov, I.A., 1962. Some limit theorems for stationary processes. *Teor. Verojatnost. i Primenen.* **7** 361-392.

[30] Ibragimov, I.A. and Linnik, Yu. V. 1971. *Independent and Stationary Sequences of Random Variables.* Wolters-Noorhoff, Groningen.

[31] Kuelbs, J., and Philipp, W., 1980. Almost sure invariance principles for partial sums of mixing B valued random variables. *Ann. Probab.* **8**, 1003-1036.

[32] Major, P., 1978. On the invariance principle for sums of identically distributed random variables. *J. Multivariate Anal.* **8**, 487-517.

[33] Massart, P., 1988. Invariance principles for empirical processes: the weakly dependent case, *PhD Thesis, Univ. of Paris* (Paris).

[34] Merlevède, F., and Peligrad, M., 2000. The functional central limit theorem for strong mixing sequences of random variables. *Ann. Probab.* **28** 3, 1336-1352.

[35] Merlevède, F., 2001. On the central limit theorem for strongly mixing sequences with values in a Hilbert space. *preliminary version.*

[36] Nobel, A. Dembo, A., 1993. A note on uniform laws of averages for dependent processes, *Stat. and Prob. Letters* **17**, 169-172.

[37] Novak, S., 2000. A note on uniform laws of averages for dependent processes, *Stat. and Prob. Letters* **17**, 169-172.

[38] Ornstein, D., 1974. *Ergodic theory, randomness, and dynamical systems.* New Haven and London, Yale University Press.

[39] Peligrad, M., 2001-a. A note on the uniform laws for dependent processes via coupling. *J. Theor. Probab.*(to appear).

[40] Peligrad, M., 2001-b. Some remarks on coupling of dependent random variables. *submitted.*

[41] Philipp, W., 1979. Almost sure invariance principles for sums of B-valued random variables, Probability in Banach spaces (Proc. Conf. Oberwolfach 1978), *Lecture Notes In Math.* **709**, 171-193. Springer, New-York, Heidelberg, Berlin.

[42] Philipp, W., 1982. Invariance principles for sums of mixing random elements and the multivariate empirical process, in: *Colloq. MathSoc. János Bolyai No. 36. limit Theorems in Probability and Statistics, Veszprém, Hungary* (North-Holland, Amsterdam), 843-873.

[43] Pollard, D., 1984. *Convergence of Stochastic Processes* (Springer, New York).

[44] Rio, E., 1995. The functional law of the iterated logarithm for stationary strongly mixing sequences, *Ann. Probability* **23**, 1188-1203.

[45] Rio, E., 2000. Théorie asymtotique des processus aléatoires faiblement dépendants, *Mathematiques et Applications de la SMAI* **31**. Springer-Verlag Berlin Heidelberg .

[46] Schwarz, G., 1980. Finitely determined processes–an indiscrete approach. *J. Math. Anal. Appl.* **76**, 146-158.

[47] Shields, P., 1973. *The Theory of Bernoulli Shifts* (University of Chicago Press, Chicago).

[48] Strassen, V., 1965. The existence of probability measures with given marginals. *Ann. Math. Statist.* **36**, 423-439.

[49] Vaprik, V.N. and Cervonenkis ,A. Ya., 1981. Necessary and sufficient conditions for the uniform convergence of means to their expectations. *Theory Probab. Appl.* **26**, 532-553.

[50] Vasershtein, L.N., 1969. Markov processes on countable product spaces describing large systems of automata. *Problemy Peredači Informacii* **3**, 64-72.

[51] Yu, B., 1991. Rates of convergence and central limit theorems for empirical processes of stationary mixing sequences, *Tech. Rept. 260, Dept. of Statistics*, University of California (Berkeley, CA).

[52] Yukich, J.E., 1986. Rates of convergence for classes of functions: the non-i.i.d. case, *J. Multivariate Anal.* **20**, 175-189.

LSTA, Université Paris VI
Tour 45-55, 4 place Jussieu
75252 Paris Cedex 5
France.
E-mail : merleve@ccr.jussieu.fr

Department of Mathematical Sciences
University of Cincinnati
PO Box 210025
Cincinnati, OH 45221-0025
USA.
E-mail : peligrm@math.uc.edu