Coupling for τ-dependent sequences and applications

Jérôme DEDECKER*and Clémentine PRIEUR†

Abstract

Let X be a real-valued random variable and $\mathcal{M}$ a σ-algebra. We show that the minimum $\mathbb{L}^1$-distance between X and a random variable distributed as X and independant of $\mathcal{M}$ can be viewed as a dependence coefficient $\tau(\mathcal{M}, X)$ whose definition is comparable (but different) to that of the usual β-mixing coefficient between $\mathcal{M}$ and $\sigma(X)$. We compare this new coefficient to other well known measures of dependence, and we show that it can be easily computed in various situations, such as causal Bernoulli shifts or stable Markov chains defined *via* iterative random maps. Next, we use coupling techniques to obtain Bennett and Rosenthal-type inequalities for partial sums of τ-dependent sequences. The former is used to prove a strong invariance principle for partial sums.

Mathematics Subject Classifications (2000): 60 F 05, 60 F 17, 60 J 10.
Key words: Coupling, dependent sequences, Bernoulli shifts, Markov chains, exponential inequalities, strong invariance principle, law of the iterated logarithm.
Short title: Coupling for dependent sequences

1 Introduction

To define the dependence coefficient τ between a real-valued random variable X and a σ-algebra $\mathcal{M}$, we need the following classical result about conditional probability (see for instance Billingsley (1995) Theorem 33.3).

Lemma 1 *Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space,* $\mathcal{M}$ *a* σ*-algebra of* $\mathcal{A}$ *and* X *a real-valued random variable with distribution* P_X*. There exists a function* $P_{X|\mathcal{M}}$ *from* $\mathcal{B}(\mathbb{R}) \times \Omega$ *to* $[0, 1]$ *such that*

*Laboratoire de Statistique Théorique et Appliquée, Université Paris 6, Site Chevaleret, 13 rue Clisson, 75013 Paris, France. E-mail: dedecker@ccr.jussieu.fr

†Laboratoire de Statistique et Probabilités, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cedex 4, France. E-mail: prieur@cict.fr

1. *For any ω in Ω, $P_{X|\mathcal{M}}(.,\omega)$ is a probability measure on $\mathcal{B}(\mathbb{R})$.*

2. *For any $A \in \mathcal{B}(\mathbb{R})$, $P_{X|\mathcal{M}}(A,.)$ is a version of $\mathbb{E}(\mathbb{1}_{X\in A}|\mathcal{M})$.*

The function $P_{X|\mathcal{M}}$ is a conditional distribution of X given $\mathcal{M}$. We denote by $P_{X|\mathcal{M}}(A)$ the random variable $P_{X|\mathcal{M}}(A,.)$.

According to Proposition (3.22) (III) in Bradley (2002) the function

$$V(P_{X|\mathcal{M}}) = \sup\Big\{\Big|\int f(x)P_{X|\mathcal{M}}(dx) - \int f(x)P_X(dx)\Big|, f \text{ s. t. } \|f\|_\infty \le 1\Big\} \quad (1.1)$$

is a $\mathcal{M}$-measurable random variable and the usual β-mixing coefficient between $\mathcal{M}$ and $\sigma(X)$ may be defined as

$$\beta(\mathcal{M}, \sigma(X)) = \frac{1}{2}\|V(P_{X|\mathcal{M}})\|_1 . \quad (1.2)$$

One of the most important properties of that coefficient is Berbee's coupling lemma (1979): if Ω is rich enough, there exists a random variable X^* independent of $\mathcal{M}$ and distributed as X such that $\mathbb{P}(X \neq X^*) = \beta(\mathcal{M}, \sigma(X))$.

Unfortunately, many simple Markov chains are not β-mixing (which means that $\beta(\sigma(X_0), \sigma(X_n))$ does not tend to zero as n tends to infinity). For instance, let X_n be the stationary solution of

$$X_n = f(X_{n-1}) + \epsilon_n \quad (1.3)$$

where f is k-lipschitz with $k < 1$ and the innovations are i.i.d. and integrable. It is well known that the chain is geometrically β-mixing if the distribution of ϵ_i has an absolutely continuous component which is bounded away from zero in a neighborhood of the origin. However if we omit the assumption on the innovations, this may be no longer true (see for instance the counter-example of Andrews (1984)).

In this paper we introduce a new dependence coefficient wich is easier to compute than β. The definition is similar to that of β except that the supremum in (1.1) is taken over the class $\Lambda_1(\mathbb{R})$ of 1-Lipschitz functions from $\mathbb{R}$ to $\mathbb{R}$. If the real-valued random variable X is integrable, we shall see in Lemma 2 that the function

$$W(P_{X|\mathcal{M}}) = \sup\Big\{\Big|\int f(x)P_{X|\mathcal{M}}(dx) - \int f(x)P_X(dx)\Big|, f \in \Lambda_1(\mathbb{R})\Big\}, \quad (1.4)$$

is a $\mathcal{M}$-measurable random variable. The coefficient τ is now defined by

$$\tau(\mathcal{M}, X) = \|W(P_{X|\mathcal{M}})\|_1 . \quad (1.5)$$

As for β, this definition does not depend on the choice of the conditional distribution. For the model described in (1.3) we obtain the bound $\tau(\sigma(X_0), X_n) \le Ck^n$ without any additional assumption on the distribution of the innovations (the constant C is smaller than $2\|X_0\|_1$).

The main result of Section 2 is the following coupling lemma: if Ω is rich enough, the coefficient $\tau(\mathcal{M}, X)$ is the infimum of $\|X-Y\|_1$ where Y is independent of $\mathcal{M}$ and distributed as X, and this infimum is reached for some particular random variable X^* introduced by Major (1978). We also give a less precise result when the random variable X takes its values in a Banach space E. In Section 3 we compare τ to the strong mixing coefficient of Rosenblatt (1956) and to the s-dependence coefficient introduced in Coulon-Prieur and Doukhan (2000). To conclude this section we give three large classes of examples for which the coefficient τ can be easily computed. In Section 4 we establish some deviation inequalities for the maximum of partial sums. Theorem 1 (resp. Corollary 1) extends Bennett's inequality (resp. Rosenthal's inequality) to τ-dependent sequences. Using the comparison between τ and α, we obtain the same bounds as those given in Rio (2000) for strongly mixing sequences. These inequalities are the main tools to prove a strong invariance principle for partial sums of τ-dependent sequences (Section 6).

2 Coupling

The following lemma ensures that the coefficient τ defined in (1.5) does not depend on the choice of the conditional probability.

Lemma 2 *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $\mathcal{M}$ a σ-algebra of $\mathcal{A}$ and X an integrable real-valued random variable with distribution P_X. The function $W(P_{X|\mathcal{M}})$ defined in (1.4) is a $\mathcal{M}$-measurable random variable. Furthermore, if $P'_{X|\mathcal{M}}$ is another conditional distribution of X given $\mathcal{M}$, then $W(P_{X|\mathcal{M}}) = W(P'_{X|\mathcal{M}})$ $\mathbb{P}$-almost surely.*

Proof. It is easy to see that there exists a countable subset $\Lambda_{1,0}(\mathbb{R})$ of $\Lambda_1(\mathbb{R})$ such that: for any f in $\Lambda_1(\mathbb{R})$, there exists a sequence f_n in $\Lambda_{1,0}(\mathbb{R})$ such that $f_n(x)$ converges to $f(x)$ for any real x. Since X is integrable, we infer from the dominated convergence theorem that

$$W(P_{X|\mathcal{M}}) = \sup\Big\{\Big|\int f(x)P_{X|\mathcal{M}}(dx) - \int f(x)P_X(dx)\Big|, f \in \Lambda_{1,0}(\mathbb{R})\Big\}. \qquad (2.1)$$

It follows that $W(P_{X|\mathcal{M}})$ is a $\mathcal{M}$-measurable random variable.

Now, if $P'_{X|\mathcal{M}}$ is another conditional probability of X given $\mathcal{M}$, we see that for any f in $\Lambda_1(\mathbb{R})$, $\mathbb{P}$-almost surely

$$\Big|\int f(x)P_{X|\mathcal{M}}(dx) - \int f(x)P_X(dx)\Big| = \Big|\int f(x)P'_{X|\mathcal{M}}(dx) - \int f(x)P_X(dx)\Big|.$$

Since $\Lambda_{1,0}(\mathbb{R})$ is countable, we infer from (2.1) that $W(P_{X|\mathcal{M}}) = W(P'_{X|\mathcal{M}})$ $\mathbb{P}$-almost surely.

In Definitions 1 below, we extend the definition of τ to Banach spaces and we introduce some well known coefficients.

Definitions 1. Given a Banach space $(E, |.|)$, let $\Lambda_1(E)$ be the set of 1-Lipschitz functions from E to $\mathbb{R}$. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. We say that a random variable X with values in E is integrable if the variable $|X|$ is integrable, and we write $\|X\|_1 = \mathbb{E}(|X|)$. For any σ-algebra $\mathcal{M}$ of $\mathcal{A}$ and any E-valued integrable variable X, define

$$\tau(\mathcal{M}, X) = \sup\{\tau(\mathcal{M}, f(X)), f \in \Lambda_1(E)\}\,. \tag{2.2}$$

The s-dependence coefficient of Coulon-Prieur and Doukhan (2000) is defined as

$$\theta(\mathcal{M}, X) = \sup\{\|\mathbb{E}(f(X)|\mathcal{M}) - \mathbb{E}(f(X))\|_1, f \in \Lambda_1(E)\}\,. \tag{2.3}$$

If X is real-valued, define the strong mixing coefficient

$$\alpha(\mathcal{M}, X) = \sup\{|\mathbb{P}(A \cap X \le t) - \mathbb{P}(A)\mathbb{P}(X \le t)|\,, A \in \mathcal{M}, t \in \mathbb{R}\}\,. \tag{2.4}$$

If $\mathcal{M}$ and $\mathcal{U}$ are two σ-algebra of $\mathcal{A}$, the strong mixing coefficient of Rosenblatt (1956) is defined by

$$\alpha(\mathcal{M}, \mathcal{U}) = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|\,, A \in \mathcal{M}, B \in \mathcal{U}\}\,. \tag{2.5}$$

The following elementary lemma will be very usefull to obtain upper bounds for the coefficient τ (see Section 3.1).

Lemma 3 *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, X an integrable random variable with values in a Banach space $(E, |.|)$, and $\mathcal{M}$ a σ-algebra of $\mathcal{A}$. If Y is a random variable distributed as X and independent of $\mathcal{M}$, then*

$$\tau(\mathcal{M}, X) \le \|X - Y\|_1\,. \tag{2.6}$$

Proof. If Y is a random variable independent of $\mathcal{M}$ and distributed as X, then for any function f of $\Lambda_1(E)$ and any function g of $\Lambda_1(\mathbb{R})$ the random variable

$$T(g) = \Big|\int g(x) P_{f(X)|\mathcal{M}}(dx) - \int g(x) P_{f(X)}(dx)\Big|$$

is a version of $|\mathbb{E}(g \circ f(X)|\mathcal{M}) - \mathbb{E}(g \circ f(Y)|\mathcal{M})|$. Consequently $T(g)$ is $\mathbb{P}$-almost surely smaller than any version of $\mathbb{E}(|X - Y||\mathcal{M})$. From (2.1), we infer that $W(P_{f(X)|\mathcal{M}}) = \sup\{T(g), g \in \Lambda_{1,0}(\mathbb{R})\}$ is $\mathbb{P}$-almost surely smaller than any version of $\mathbb{E}(|X - Y||\mathcal{M})$. By definition $\tau(\mathcal{M}, f(X)) = \|W(P_{f(X)|\mathcal{M}})\|_1$, which implies that $\tau(\mathcal{M}, f(X)) \le \|X - Y\|_1$. This being true for any function f of $\Lambda_1(E)$, the result follows.

The main result of this section is that if X is real-valued and Ω is rich enough, the equality can be reached in (2.6). This result is based on Major's quantile transformation (1978) which we recall in Lemma 4 below (see also Rio (2000), page 161).

Notation 1. For any distribution function F, define the generalized inverse as follows: for any u in $[0,1]$, $F^{-1}(u) = \inf\{t \in \mathbb{R} : F(t) \geq u\}$. It is clear that $F(t) \geq u$ if and only if $t \geq F^{-1}(u)$.

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $\mathcal{M}$ a σ-algebra of $\mathcal{A}$ and X a real-valued random variable. Let $F_{\mathcal{M}}(t,\omega) = P_{X|\mathcal{M}}(]-\infty, t], \omega)$. For any ω, $F_{\mathcal{M}}(.,\omega)$ is a distribution function, and for any t, $F_{\mathcal{M}}(t,.)$ is a $\mathcal{M}$-measurable random variable. For any ω, define the generalized inverse $F_{\mathcal{M}}^{-1}(u,\omega)$ as in Notation 1. From the equality $\{\omega : t \geq F_{\mathcal{M}}^{-1}(u,\omega)\} = \{\omega : F_{\mathcal{M}}(t,\omega) \geq u\}$, we infer that $F_{\mathcal{M}}^{-1}(u,.)$ is $\mathcal{M}$-measurable. In the same way, $\{(t,\omega) : F_{\mathcal{M}}(t,\omega) \geq u\} = \{(t,\omega) : t \geq F_{\mathcal{M}}^{-1}(u,\omega)\}$, which implies that the mapping $(t,\omega) \to F_{\mathcal{M}}(t,\omega)$ is measurable with respect to $\mathcal{B}(\mathbb{R}) \otimes \mathcal{M}$. The same arguments imply that the mapping $(u,\omega) \to F_{\mathcal{M}}^{-1}(u,\omega)$ is measurable with respect to $\mathcal{B}([0,1]) \otimes \mathcal{M}$. Denote by $F_{\mathcal{M}}(t)$ (resp.$F_{\mathcal{M}}^{-1}(u)$) the random variable $F_{\mathcal{M}}(t,.)$ (resp. $F_{\mathcal{M}}^{-1}(u,.)$), and let $F_{\mathcal{M}}(t-0) = \sup_{s<t} F_{\mathcal{M}}(s)$.

Lemma 4 *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $\mathcal{M}$ a σ-algebra of $\mathcal{A}$ and X an integrable real-valued random variable. Assume that there exists a random variable δ uniformly distributed over $[0,1]$, independent of the σ-algebra generated by X and $\mathcal{M}$. Define*

$$U = F_{\mathcal{M}}(X-0) + \delta(F_{\mathcal{M}}(X) - F_{\mathcal{M}}(X-0))\,.$$

The random variable U is uniformly distributed over $[0,1]$ and independent of $\mathcal{M}$. Moreover $F_{\mathcal{M}}^{-1}(U) = X$ $\mathbb{P}$-almost surely.

With the help of Lemma 4, we can now establish our coupling result

Lemma 5 *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, X an integrable real-valued random variable, and $\mathcal{M}$ a σ-algebra of $\mathcal{A}$. Assume that there exists a random variable δ uniformly distributed over $[0,1]$, independent of the σ-algebra generated by X and $\mathcal{M}$. Then there exists a random variable X^*, measurable with respect to $\mathcal{M} \vee \sigma(X) \vee \sigma(\delta)$, independent of $\mathcal{M}$ and distributed as X, such that*

$$\|X - X^*\|_1 = \tau(\mathcal{M}, X)\,. \tag{2.7}$$

Remark 1. From Lemma 5, we infer that $\|X - X^*\|_1$ is the infimum of $\|X - Y\|_1$ where Y is independent of $\mathcal{M}$ and distributed as X. This result is due to Major (1978).

Proof. Let U be the random variable defined in Lemma 4 and F be the distribution function of X. The random variable $X^* = F^{-1}(U)$ is measurable with respect to

$\mathcal{M} \vee \sigma(X) \vee \sigma(\delta)$, independent of $\mathcal{M}$ and distributed as X. Since $X = F_{\mathcal{M}}^{-1}(U)$ $\mathbb{P}$-almost surely, we have

$$\|X - X^*\|_1 = \mathbb{E}\Big(\int_0^1 |F_{\mathcal{M}}^{-1}(u) - F^{-1}(u)|du\Big)\,. \tag{2.8}$$

For two distribution functions F and G, denote by $M(F,G)$ the set of all probability measures on $\mathbb{R} \times \mathbb{R}$ with marginals F and G. Define

$$d(F,G) = \inf\left\{\int |x-y|\mu(dx,dy) : \mu \in M(F,G)\right\}\,,$$

and recall that (see Dudley (1989), Section 11.8, Problems 1 and 2 page 333)

$$d(F,G) = \int_{\mathbb{R}} |F(t) - G(t)|dt = \int_0^1 |F^{-1}(u) - G^{-1}(u)|du\,. \tag{2.9}$$

On the other hand, Kantorovich and Rubinstein (Theorem 11.8.2 in Dudley (1989)) have proved that

$$d(F,G) = \sup\Big\{\Big|\int f dF - \int f dG\Big| : f \in \Lambda_1(\mathbb{R})\Big\}\,. \tag{2.10}$$

Combining (2.8), (2.9) and (2.10), we have that

$$\|X - X^*\|_1 = \mathbb{E}\Big(\sup\Big\{\Big|\int f dF_{\mathcal{M}} - \int f dF\Big| : f \in \Lambda_1(\mathbb{R})\Big\}\Big)\,,$$

and the proof is complete.

3 Comparison of coefficients and examples

Starting from Lemma 5, we can compare the coefficients $\theta(\mathcal{M},X)$, $\tau(\mathcal{M},X)$ and $\alpha(\mathcal{M},X)$ when X is some real-valued random variable.

Notation 2. Let X be some random variable with values in some Banach space $(E, |.|)$. Let $H_{|X|}(x) = \mathbb{P}(|X| > x)$, and
– $Q_{|X|}$ the generalized inverse of $H_{|X|}$: if $u \in [0,1]$, $Q_{|X|}(u) = \inf\{t \in \mathbb{R} : H_{|X|}(t) \leq u\}$.
– $G_{|X|}$ the inverse of $x \to \int_0^x Q_{|X|}(u)du$.
– $L_{|X|}$ the inverse of $x \to xG_{|X|}(x)$.

Lemma 6 *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, X an integrable real-valued random variable, and $\mathcal{M}$ a σ-algebra of $\mathcal{A}$. The following inequalities hold*

$$\tau(\mathcal{M},X) \leq 2\int_0^{2\alpha(\mathcal{M},X)} Q_{|X|}(u)du \tag{3.1}$$

$$\theta(\mathcal{M},X) \leq \tau(\mathcal{M},X) \leq 4L_{|X|}(\theta(\mathcal{M},X))\,. \tag{3.2}$$

Remark 2. In particular, if p and q are two conjugate exponents, (3.1) yields the upper bound $\tau(\mathcal{M}, X) \le 2\|X\|_p(2\alpha(\mathcal{M}, X))^{1/q}$.

Proof. To obtain (3.1), we use a recent result of Peligrad (2002). In (c) of Theorem 1 of her paper, she proves that if X^* is the random variable of Lemma 5,

$$\|X - X^*\|_1 \le 4\int_0^{\alpha(\mathcal{M},X)} Q_{|X|}(u)du\,,$$

which is not exactly the required inequality. After a carefull reading of the proof, we see that Peligrad establishes that if $X_+ = X \vee 0$ and $X_- = (-X) \vee 0$,

$$\|X - X^*\|_1 \le 2\int_0^{\infty} (\alpha(\mathcal{M}, X) \wedge \mathbb{P}(X_+ > u) + \alpha(\mathcal{M}, X) \wedge \mathbb{P}(X_- > u))\, du\,.$$

Since $a \wedge b + c \wedge d \le (a + c) \wedge (b + d)$, we obtain that

$$\|X - X^*\|_1 \le 2\int_0^{\infty} (2\alpha(\mathcal{M}, X)) \wedge H_{|X|}(u)du \le 2\int_0^{\infty}\int_0^{2\alpha(\mathcal{M},X)} \mathbb{1}_{t<H_{|X|}(u)}dt\, du\,,$$

and the result follows by applying Fubini and by noting that $t < H_{|X|}(u)$ if and only if $u < Q_{|X|}(t)$.

To obtain (3.2), note that from (2.8) and (2.9),

$$\begin{aligned}\|X - X^*\|_1 &= \int_{\mathbb{R}} \|\mathbb{E}(\mathbb{1}_{X\le t}|\mathcal{M}) - \mathbb{P}(X \le t)\|_1\, dt\\ &= \int_0^{\infty} \|\mathbb{E}(\mathbb{1}_{X_+\le t}|\mathcal{M}) - \mathbb{P}(X_+ \le t)\|_1\, dt\\ &\quad + \int_0^{\infty} \|\mathbb{E}(\mathbb{1}_{X_-<t}|\mathcal{M}) - \mathbb{P}(X_- < t)\|_1\, dt\,. \end{aligned} \tag{3.3}$$

For any positive ϵ, let $f_\epsilon^t(x) = \mathbb{1}_{x\le t} + \epsilon^{-1}(t + \epsilon - x)\mathbb{1}_{t<x\le t+\epsilon}$. Clearly

$$\begin{aligned}\|\mathbb{E}(\mathbb{1}_{X_+\le t}|\mathcal{M}) - \mathbb{P}(X_+ \le t)\|_1 &\le \|\mathbb{E}(f_\epsilon^t(X_+)|\mathcal{M}) - \mathbb{E}(f_\epsilon^t(X_+))\|_1 + 2\mathbb{P}(X_+ \in [t, t + \epsilon])\\ &\le \epsilon^{-1}\theta(\mathcal{M}, X) \wedge 2\mathbb{P}(X_+ > t) + 2\mathbb{P}(X_+ \in [t, t + \epsilon])\end{aligned}$$

and the same is true for X_-. Integrating (3.3) and applying Fubini, we have that

$$\begin{aligned}\|X - X^*\|_1 &\le \int_0^{\infty} \epsilon^{-1}\theta(\mathcal{M}, X) \wedge 2\mathbb{P}(X_+ > t) + \epsilon^{-1}\theta(\mathcal{M}, X) \wedge 2\mathbb{P}(X_- > t)\, dt + 2\epsilon\\ &\le 2\int_0^{\infty} \epsilon^{-1}\theta(\mathcal{M}, X) \wedge \mathbb{P}(|X| > t)\, dt + 2\epsilon\,,\end{aligned}$$

and consequently

$$\|X - X^*\|_1 \le 2\int_0^{\epsilon^{-1}\theta(\mathcal{M},X)} Q_{|X|}(u)du + 2\epsilon\,.$$

The two terms on right hand are equal for $\epsilon = G_{|X|}^{-1}(\theta(\mathcal{M}, X)/\epsilon)$ which means that $\epsilon = L_{|X|}(\theta(\mathcal{M}, X))$. This completes the proof.

We now define the coefficients θ, τ and α for a sequence $(X_i)_{i>0}$ of Banach-valued random variables.

Definitions 2. Let $(E, |.|)$ be some Banach space. On E^k we put the norm $|.|_1$ defined by $|x-y|_1 = |x_1 - y_1| + \cdots + |x_k - y_k|$. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $(X_i)_{i>0}$ a sequence of E-valued random variables and $\mathcal{M}_i$ a sequence of σ-algebra of $\mathcal{A}$. For any positive integer k, define

$$\theta_k(i) = \max_{1\le l\le k} \frac{1}{l} \sup\{\theta(\mathcal{M}_p, (X_{j_1}, \ldots, X_{j_l})), p+i \le j_1 < \cdots < j_l\} \text{ and } \theta(i) = \sup_{k\ge 0} \theta_k(i).$$

Define $\tau_k(i)$ and $\tau(i)$ in the same way. As usual, the coefficient $\alpha(i)$ is defined by

$$\alpha_k(i) = \sup\{\alpha(\mathcal{M}_p, \sigma(X_{j_1}, \ldots, X_{j_k})), p+i \le j_1 < \cdots < j_k\} \text{ and } \alpha(i) = \sup_{k\ge 0} \alpha_k(i).$$

With this definition, it is clear that if $(X_i)_{i>0}$ is some sequence with coefficients $\theta(i)$ (resp. $\tau(i)$) and if g is some L-lipschitz function, then the coefficients of the sequence $(g(X_i))_{i>0}$ are smaller than $L\theta(i)$ (resp. $L\tau(i)$).

The following lemma allows to compare the coefficients $\theta_k(i)$, $\tau_k(i)$ and $\alpha_k(i)$.

Lemma 7 *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $(X_i)_{i>0}$ a sequence of random variables with values in a Banach space $(E, |.|)$ and $\mathcal{M}_i$ a sequence of σ-algebra of $\mathcal{A}$. Let X be some nonnegative random variable such that $Q_X \ge \sup_{k\ge 1} Q_{|X_k|}$. The following inequalities hold*

$$\tau_k(i) \le 2\int_0^{2\alpha_k(i)} Q_X(u)du, \quad \textit{and} \quad \theta_k(i) \le \tau_k(i) \le 4L_X(\theta_k(i))\,.$$

Proof. We first compare τ and α. Without loss of generality, we can assume that $f(0, \ldots, 0) = 0$ in the definition of $\tau(\mathcal{M}, X)$. For $p+i \le j_1 < \cdots < j_l$ and f in $\Lambda_1(E^l)$, we infer from Lemma 6 that

$$\tau(\mathcal{M}_p, f(X_{j_1}, \ldots, X_{j_l})) \le 2\int_0^{2\alpha_k(i)} Q_{|f(X_{j_1},\ldots,X_{j_l})|}(u)du \le 2\int_0^{2\alpha_k(i)} Q_{|X_{j_1}|+\cdots+|X_{j_l}|}du\,. \tag{3.4}$$

From Lemma 2.1 in Rio (2000) we know that if Z_1, Z_2 et Z_3 are three nonnegative random variables, then

$$\int_0^1 Q_{Z_1+Z_2}(u)Q_{Z_3}(u)du \le \int_0^1 (Q_{Z_1}(u) + Q_{Z_2}(u))Q_{Z_3}(u)du\,.$$

Applying this result with $Q_{Z_3}(u) = \mathbb{1}_{u\le 2\alpha_k(i)}$, we infer from (3.4) that

$$\tau(\mathcal{M}_p, f(X_{j_1}, \ldots, X_{j_l})) \le 2\int_0^{2\alpha_k(i)} Q_{|X_{j_1}|}(u) + \cdots + Q_{|X_{j_l}|}(u)du \le 2l\int_0^{2\alpha_k(i)} Q_X(u)du\,,$$

and the result follows from the definition of $\tau_k(i)$.

Let us prove that $\tau_k(i) \leq 4L_X(\theta_k(i))$. Applying once again Lemma 2.1 in Rio (2000), we obtain that

$$\int_0^x Q_{|X_{j_1}|+\cdots+|X_{j_l}|}(u)du \leq l \int_0^x Q_X(u)du\,.$$

Hence $G_{|f(X_{j_1},\ldots,X_{j_l})|}(u) \geq G_X(u/l)$ and therefore $L_{|f(X_{j_1},\cdots,X_{j_l})|}(u) \leq lL_X(u/l)$. Consequently, for $p+i \leq j_1 < \cdots < j_l$,

$$\tau(\mathcal{M}_p, f(X_{j_1},\ldots,X_{j_l})) \leq 4\,lL_X(\theta(\mathcal{M}_p, f(X_{j_1},\ldots,X_{j_l}))/l) \leq 4\,lL_X(\theta_k(i))\,,$$

and the result follows from the definition of $\tau_k(i)$.

3.1 Examples.

We can use Lemma 7 to obtain upper bounds for the coefficients $(\tau(i))_{i>0}$ of an α-mixing sequence (or a s-dependent sequence). We refer to the book of Doukhan (1994) for examples of α-mixing processes and to the paper by Coulon-Prieur and Doukhan (2000) for examples of s-dependent sequences (see also the paper by Doukhan and Louhichi (1999)).

Concerning the examples given by Coulon-Prieur and Doukhan, it is easy to see that we can obtain the same bounds for $\tau(i)$ as those obtained for $\theta(i)$. In each case, the result follows by applying Lemma 3.

In this section, we show how to compute upper bounds for the coefficient $\tau(i)$ for three large classes of examples. In Examples 1 and 2 we apply Lemma 3 to obtain an upper bound for $\tau(i)$, while in Example 3 we start from the definition of $\tau(i)$.

Example 1: causal Bernoulli shifts. Let $(\xi_i)_{i\in\mathbb{Z}}$ be a sequence of i.i.d. random variables with values in a measurable space $\mathcal{X}$. Assume that there exists a function H defined on a subset of $\mathcal{X}^{\mathbb{N}}$, with values in a Banach space $(E, |.|)$, and such that $H(\xi_0, \xi_{-1}, \xi_{-2}, \ldots,)$ is defined almost surely. The stationary sequence $(X_n)_{n>0}$ defined by $X_n = H(\xi_n, \xi_{n-1}, \xi_{n-2}, \ldots)$ is called a causal Bernoulli shift.

Let $(\xi'_i)_{i\in\mathbb{Z}}$ be a sequence of i.i.d. random variables, independent of $(\xi_i)_{i\in\mathbb{Z}}$ and distributed as $(\xi_i)_{i\in\mathbb{Z}}$. Let $(\delta_i)_{i>0}$ be a decreasing sequence such that

$$\|H(\xi_i, \xi_{i-1}, \xi_{i-2}, \ldots) - H(\xi_i, \xi_{i-1}, \xi_{i-2}, \ldots, \xi_1, \xi'_0, \xi'_{-1}, \ldots)\|_1 \leq \delta_i\,.$$

If $\mathcal{M}_i = \sigma(X_j, j \leq i)$, the coefficient $\tau_k(i)$ of $(X_n)_{n>0}$ satisfies $\tau_k(i) \leq \delta_i$.

Proof. Define $X'_n = H(\xi_n, \ldots, \xi_1, \xi'_0, \xi'_{-1}, \ldots)$. The sequence $(X'_n)_{n>0}$ is distributed as $(X_n)_{n>0}$ and is independent of the σ-algebra $\mathcal{M}_0$. From Lemma 3 we have that, for $j_k > \cdots > j_1 \geq i$,

$$\tau(\mathcal{M}_0, (X_{j_1}, \ldots, X_{j_k})) \leq \sum_{l=1}^{k} \|X_{j_l} - X'_{j_l}\|_1 \leq k\delta_i\,.$$

The result follows by using the stationarity of $(X_n)_{n>0}$ and the definition of $\tau(i)$.

Application 1: causal linear processes. In that case $\mathcal{X} = \mathbb{R}$ and X_n is defined by $X_n = \sum_{j\geq 0} a_j \xi_{n-j}$. If $|\xi_0|$ is integrable, we can take $\delta_i = \sum_{j\geq i} |a_j| \|\xi_0 - \xi_0'\|_1$. Let $\Delta(\xi_0) = \inf_{a\in\mathbb{R}} \|\xi_0 - a\|_1$. Since $\|\xi_0 - \xi_0'\|_1 = 2\Delta(\xi_0)$, we see that $\delta_i = 2\Delta(\xi_0)\sum_{j\geq i}|a_j|$. If ξ_0^2 is integrable, we can take $\delta_i = (2\mathrm{Var}(\xi_0)\sum_{j\geq i} a_j^2)^{1/2}$. For instance, if $a_i = 2^{-i-1}$ and $\xi_0 \sim \mathcal{B}(1/2)$, $\delta_i = 2^{-i}\sqrt{1/6}$. Recall that in that case, $\alpha_i = 1/4$ for any positive integer i.

Example 2: iterative random functions. Let $(X_n)_{n\geq 0}$ be a stationary Markov chain, with values in a Banach space $(E, |.|)$ and such that $X_n = F(X_{n-1}, \xi_n)$ for some measurable function F and some i.i.d. sequence $(\xi_i)_{i\geq 0}$. Let X_0' be a random variable independent of X_0 and distributed as X_0, and define $X_n' = F(X_{n-1}', \xi_n)$. The sequence $(X_n')_{n\geq 0}$ is distributed as $(X_n)_{n\geq 0}$ and independent of X_0. We infer from Lemma 3 that, for $j_k > \cdots > j_1 \geq i$,

$$\tau(\sigma(X_0), (X_{j_1}, \ldots, X_{j_k})) \leq \sum_{l=1}^{k} \|X_{j_l} - X_{j_l}'\|_1 .$$

Let μ be the distribution of X_0 and $(X_n^x)_{n\geq 0}$ the chain starting from $X_0^x = x$. With these notations, we have that

$$\tau(\sigma(X_0), (X_{j_1}, \ldots, X_{j_k})) \leq \sum_{l=1}^{k} \iint \|X_{j_l}^x - X_{j_l}^y\|_1 \mu(dx)\mu(dy) . \tag{3.5}$$

For instance, assume that there exists a decreasing sequence $(\delta_i)_{i\geq 0}$ of positive numbers such that $\|X_i^x - X_i^y\|_1 \leq \delta_i |x-y|$. In that case

$$\tau(\sigma(X_0), (X_{j_1}, \ldots, X_{j_k})) \leq k\delta_i \|X_0 - X_0'\|_1 ,$$

and consequently $\tau(i) \leq \delta_i \|X_0 - X_0'\|_1$. For instance in the usual case where $\|F(x,\xi_0) - F(y,\xi_0)\|_1 \leq \kappa|x-y|$ for some $\kappa < 1$, we can take $\delta_i = \kappa^i$. An important example is $X_n = f(X_{n-1}) + \xi_n$ for some κ-lipschitz function f.

Application 2: functional autoregressive processes. We give a simple example of a non contractive function for which the coefficient $\tau(i)$ decreases arithmetically. Given $\delta \in [0,1[$, $C \in]0,1]$ and $S \geq 1$, let $\mathcal{L}(C,\delta)$ be the class of 1-Lipschitz functions f such that

$$f(0) = 0 \quad \text{and} \quad |f'(t)| \leq 1 - C(1+|t|)^{-\delta} \quad \text{almost everywhere.}$$

Let $ARL(C,\delta,S)$ be the class of real-valued Markov chains $(X_n)_{n>0}$ solutions of the equation $X_n = f(X_{n-1}) + \xi_n$ where $f \in \mathcal{L}(C,\delta)$ and $\|\xi_0\|_S < \infty$. Dedecker and Rio (2000) have proved that for any Markov chain belonging to $ARL(C,\delta,S)$, there exists an unique invariant probability μ, and that $\mu(|x|^{S-\delta}) < \infty$. Starting from (3.5) and arguing as in Dedecker and Rio, we can prove that if $S > 1+\delta$ the coefficients of the stationary chain satisfy $\tau(n) = O(n^{(\delta+1-S)/\delta})$.

Example 3: other Markov chains. Let P be a Markov kernel defined on a measurable subset $\mathcal{X}$ of a Banach space $(E, |.|)$. For any continuous bounded function f from $\mathcal{X}$ to $\mathbb{R}$ we have

$$P(f)(x) = \int_{\mathcal{X}} f(z)P(x, dz)\,.$$

We make the following assumptions on P

H For some $0 < \kappa < 1$, P maps $\Lambda_1(\mathcal{X})$ to $\Lambda_\kappa(\mathcal{X})$.

Let $(X_n)_{n\in\mathbb{N}}$ be a stationary Markov chain with values in $\mathcal{X}$, with marginal distribution μ and transition kernel P satisfying H. Then for $j_k > \cdots > j_1 \geq i$ and f in $\Lambda_1(\mathcal{X}^k)$, the function $\mathbb{E}(f(X_{j_1}, \ldots, X_{j_k})|X_{j_1=x})$ belongs to $\Lambda_{1+\kappa+\cdots+\kappa^{k-1}}(\mathcal{X})$ and consequently $f_{j_1,\ldots,j_k}(x) = \mathbb{E}(f(X_{j_1}, \ldots, X_{j_k})|X_0 = x)$ belongs to $\Lambda_{k\kappa^i}(\mathcal{X})$. Since

$$\tau(\sigma(X_0), (X_{j_1}, \ldots, X_{j_k})) \leq \iint \sup_{f\in\Lambda_1(\mathcal{X}^k)} |f_{j_1,\ldots,j_k}(x) - f_{j_1,\ldots,j_k}(y)|\mu(dx)\mu(dy),$$

we infer that

$$\tau(\sigma(X_0), (X_{j_1}, \ldots, X_{j_k})) \leq k\kappa^i \|X_0 - X_0'\|_1 \tag{3.6}$$

where X_0' is independent of X_0 and distributed as X_0. By definition of τ, we infer from (3.6) that $\tau(i) \leq \kappa^i \|X_0 - X_0'\|_1$ and the chain is geometrically τ-dependent.

In the case of iterated random maps (Example 2 above) the map F is a measurable function from $\mathcal{X} \times \mathcal{Y}$ to $\mathcal{X}$, and the transition kernel P has the form

$$P(f)(x) = \int_{\mathcal{Y}} f(F(x, z))\nu(dz) \tag{3.7}$$

for some probability measure ν on $\mathcal{Y}$. Assumption H is satisfied as soon as

$$\int |F(x, z) - F(y, z)|\nu(dz) \leq \kappa|x - y|\,,$$

which was the condition previously found.

We now consider the more general situation

$$P(f)(x) = \int_{\mathcal{Y}} b(x, z)f(F(x, z))\nu(dz)\,, \tag{3.8}$$

where ν is a measure on $\mathcal{Y}$ and $b(x, y)\nu(dy)$ is a probability measure for each x in $\mathcal{X}$. For simplicity, we assume that $E = \mathbb{R}$ and that $\mathcal{X} = I$ is either $\mathbb{R}$, $[a, b]$, $[a, \infty[$ or $]-\infty, b]$. According to property 34.5 in McShane (1947) a function g from I to $\mathbb{R}$ is M-lipshitz if and only if

$$\sup_{x\in I} \limsup_{h\to 0} \frac{|f(x+h) - f(x)|}{|h|} \leq M\,.$$

Starting from this property, we infer that Condition H is satisfied as soon as, for any x in I,

$$\limsup_{h\to 0} \frac{1}{|h|}\Big(\int b(x+h,z)|F(x+h,z)-F(x,z)|\nu(dz) \\ + \int |b(x+h,z)-b(x,z)||F(x,z)|\nu(dz)\Big) \leq \kappa\,. \quad (3.9)$$

In particular, if for ν-almost every z the functions $x \to b(x,z)$ and $x \to F(x,z)$ are derivable and the functions $b(x,z)|F'(x,z)|$ and $|b'(x,z)F(x,z)|$ are each bounded by integrable functions not depending on z, (3.9) writes in fact: for any x in I,

$$\int b(x,z)|F'(x,z)| + |b'(x,z)F(x,z)|\nu(dz) \leq \kappa\,. \quad (3.10)$$

Application 3: Markov kernel associated to expanding maps. In this example, $\mathcal{X} = I = [0,1]$. We consider a transformation T from $[0,1]$ to $[0,1]$. Assume that there exist N in $\mathbb{N}^* \cup +\infty$ and a partition $([a_j, a_{j+1}[)_{1\leq j\leq N}$ of $[0,1[$ such that:

- For each $1 \leq j \leq N$, the restriction T_j of T to $]a_j, a_{j+1}[$ is strictly monotonic and can be extended to a function $\overline{T}_j$ belonging to $C^1([a_j, a_{j+1}])$. Moreover $\overline{T}_j([a_j, a_{j+1}]) = [0,1]$.

Denote by λ the Lebesgue measure on $[0,1]$, and define the operator Φ from $\mathbb{L}^1(I,\lambda)$ to $\mathbb{L}^1(I,\lambda)$ *via* the equality

$$\int_0^1 \Phi(f)(x)g(x)\lambda(dx) = \int_0^1 f(x)(g\circ T)(x)\lambda(dx) \quad \text{where } f \in \mathbb{L}^1(I,\lambda) \text{ and } g \in \mathbb{L}^\infty(I,\lambda).$$

By definition of T, we infer that for any continuous bounded function f,

$$\Phi(f)(x) = \sum_{j=1}^{N} f(\sigma_j(x))|\sigma_j'(x)| \quad \text{where } \sigma_j = \overline{T}_j^{-1}\,.$$

(see Broise (1996) for more details on the operator Φ). Assume that there exists a positive density h such that T preserves the probability measure $h\lambda$ (or equivalently $\Phi(h) = h$). On the probability space $(I, h\lambda)$, the sequence $(g \circ T^j)_{j\geq 0}$ is strictly stationary. Moreover the vector $(g, g\circ T, \ldots, g\circ T^n)$ has the same distribution as $(g(X_n), g(X_{n-1}), \ldots, g(X_0))$ where $(X_i)_{i\geq 0}$ is a stationary Markov chain with invariant distribution $h\lambda$ and transition kernel P given by

$$P(f)(x) = \frac{\Phi(hf)(x)}{h(x)} = \frac{\sum_{j=1}^N h(\sigma_j(x))f(\sigma_j(x))|\sigma_j'(x)|}{\sum_{j=1}^N h(\sigma_j(x))|\sigma_j'(x)|}\,.$$

In particular, P has the form (3.8) with $\nu = \sum_{i=1}^N \delta_i$, $F(x,i) = \sigma_i(x)$ and $b(x,i) = h(\sigma_i(x))|\sigma_i'(x)|/h(x)$. Consequently, the chain is geometrically τ-dependent as soon as (3.9) holds. Let us chek this condition on some examples.

- If T_j is linear, then $h \equiv 1$, σ_j' is constant, and $\sum_{j=1}^N |\sigma_j'| = 1$. Hence the chain is geometrically τ-dependent as soon as $N \geq 2$, with $\kappa = \sup_{1 \leq i \leq N} |\sigma_i'|$. In fact, this is an example of iterated random maps, since the kernel P has the form (3.7) with $F(x, i) = \sigma_i(x)$ and $\nu = \sum_{j=1}^N |\sigma_j'| \delta_i$. In particular, if $T(x) = Kx \bmod 1$ for some integer $K \geq 2$, then $\kappa = 1/K$.

- Denote by $\{x\} = x - [x]$, $[x]$ being the integer part of x. If $T(x) = \{a(x^{-1} - 1)\}$ for some positive real a, then $h(x) = 1/((x + a)\ln(1 + 1/a))$ and

$$P(f)(x) = (x+a) \sum_{n=0}^{\infty} f\Big(\frac{a}{x+n+a}\Big)\Big(\frac{1}{x+n+a} - \frac{1}{x+n+1+a}\Big).$$

  If $a \geq 1$, one can easily see that (3.10) holds (for $a = 1$, $\kappa = 421/432$ works).

4 Exponential and moment inequalities

The first theorem of this section extends Bennett's inequality for independent sequences to the case of τ-dependent sequences. For any positive integer q, we obtain an upper bound involving two terms: the first one is the classical Bennett's bound at level λ for a sum $\sum_n$ of independent variables ξ_i such that $\mathrm{Var}(\sum_n) = v_q$ and $\|\xi_i\|_\infty \leq qM$, and the second one is equal to $n\lambda^{-1}\tau_q(q+1)$. Using lemma 7, we obtain the same inequalities as those established by Rio (2000) for strongly mixing sequences. This is not surprising, for we follow the proof of Rio and we use Lemma 5 instead of Rio's coupling lemma.

Theorem 1 *Let $(X_i)_{i>0}$ be a sequence of real-valued random variables sucht that $\|X_i\|_\infty \leq M$, and $\mathcal{M}_i = \sigma(X_k, 1 \leq k \leq i)$. Let $S_k = \sum_{i=1}^k (X_i - \mathbb{E}(X_i))$ and $\overline{S}_n = \max_{1 \leq k \leq n} |S_k|$. Let q be some positive integer, v_q some nonnegative number such that*

$$v_q \geq \|X_{q[n/q]+1} + \cdots + X_n\|_2^2 + \sum_{i=1}^{[n/q]} \|X_{(i-1)q+1} + \cdots + X_{iq}\|_2^2 .$$

and h the function defined by $h(x) = (1+x)\ln(1+x) - x$.

1. *For any positive λ, we have*

$$\mathbb{P}(|S_n| \geq 3\lambda) \leq 4\exp\Big(-\frac{v_q}{(qM)^2} h\Big(\frac{\lambda q M}{v_q}\Big)\Big) + \frac{n}{\lambda}\tau_q(q+1).$$

2. *For any $\lambda \geq Mq$, we have*

$$\mathbb{P}(\overline{S}_n \geq (\mathbb{1}_{q>1} + 3)\lambda) \leq 4\exp\Big(-\frac{v_q}{(qM)^2} h\Big(\frac{\lambda q M}{v_q}\Big)\Big) + \frac{n}{\lambda}\tau_q(q+1).$$

Proof. We proceed as in Rio (2000) page 83. For $1 \leq i \leq [n/q]$, define the variables $U_i = S_{iq} - S_{iq-q}$ and $U_{[n/q]+1} = S_n - S_{q[n/q]}$. Let $(\delta_j)_{1\leq j\leq [n/q]+1}$ be independent random variables uniformly distributed over $[0,1]$ and independent of $(U_i)_{1\leq j\leq [n/q]+1}$. We now apply Lemma 5: For any $1 \leq i \leq [n/q]+1$, there exists a measurable function F_i such that $U_i^* = F_i(U_1, \ldots, U_{i-2}, U_i, \delta_i)$ satisfies the conclusions of Lemma 5, with $\mathcal{M} = \sigma(U_l, l \leq i-2)$. The sequence $(U_i^*)_{1\leq j\leq [n/q]+1}$ has the following properties:

a. For any $1 \leq i \leq [n/q]+1$, the random variable U_i^* is distributed as U_i.

b. The random variables $(U_{2i}^*)_{2\leq 2i\leq [n/q]+1}$ are independent and so are the variables $(U_{2i-1}^*)_{1\leq 2i-1\leq [n/q]+1}$.

c. Moreover $\|U_i - U_i^*\|_1 \leq \tau(\sigma(U_l, l \leq i-2), U_i)$.

Since for $1 \leq i \leq [n/q]$ we have $\tau(\sigma(U_l, l \leq i-2), U_i) \leq q\tau_q(q+1)$, we infer that

$$\begin{aligned} \text{for } 1 \leq i \leq [n/q], \quad \|U_i - U_i^*\|_1 &\leq q\tau_q(q+1) \\ \text{and} \quad \|U_{[n/q]+1} - U_{[n/q]+1}^*\|_1 &\leq (n - q[n/q])\tau_{n-q[n/q]}(q+1)\,. \end{aligned} \tag{4.1}$$

Proof of 1. Clearly

$$|S_n| \leq \sum_{i=1}^{[n/q]+1} |U_i - U_i^*| + \Big| \sum_{i=1}^{([n/q]+1)/2} U_{2i}^* \Big| + \Big| \sum_{i=1}^{[n/q]/2+1} U_{2i-1}^* \Big|. \tag{4.2}$$

Combining (4.1) with the fact that $\tau_{n-q[n/q]}(q+1) \leq \tau_q(q+1)$, we obtain

$$\mathbb{P}\Big(\sum_{i=1}^{[n/q]+1} |U_i - U_i^*| \geq \lambda \Big) \leq \frac{n}{\lambda}\, \tau_q(q+1)\,. \tag{4.3}$$

The result follows by applying Bennett's inequality to the two other sums in (4.2).

Proof of 2. Any integer j being distant from at most $[q/2]$ of an element of $q\mathbb{N}$, we have that

$$\max_{1\leq k\leq n} |S_k| \leq 2[q/2]M + \max_{1\leq j\leq [n/q]+1} \Big| \sum_{i=1}^{j} U_i \Big|\,.$$

Hence Theorem 1 follows from the bound

$$\mathbb{P}\Big(\max_{1\leq j\leq [n/q]+1} \Big| \sum_{i=1}^{j} U_i \Big| \geq 3\lambda \Big) \leq 4\exp\Big(-\frac{v_q}{(qM)^2} h\Big(\frac{\lambda q M}{v_q} \Big) \Big) + \frac{n}{\lambda}\, \tau_q(q+1)\,. \tag{4.4}$$

Using again the variables U_i^* we find:

$$\max_{1\leq j\leq [n/q]+1} \Big| \sum_{i=1}^{j} U_i \Big| \leq \sum_{i=1}^{[n/q]+1} |U_i - U_i^*| + \max_{2\leq 2j\leq [n/q]+1} \Big| \sum_{i=1}^{j} U_{2i}^* \Big| + \max_{1\leq 2j-1\leq [n/q]+1} \Big| \sum_{i=1}^{j} U_{2i-1}^* \Big|. \tag{4.5}$$

Inequality (4.4) follows by applying (4.3) to the first term on right hand in (4.5) and Bennett's inequality to the two others terms.

Proceeding as in Theorem 1, we establish Fuk-Nagaev type inequalities (see Fuk and Nagaev (1971)) for sums of τ-dependent sequences. Applying Lemma 7, we obtain the same inequalities (up to some numerical constant) as those established by Rio (2000) for strongly mixing sequences.

Notation 3. For any non-increasing sequence $(\delta_i)_{i\geq 0}$ of nonnegative numbers, define $\delta^{-1}(u) = \sum_{i\geq 0} \mathbb{1}_{u<\delta_i} = \inf\{k \in \mathbb{N} : \delta_k \leq u\}$. For any non-increasing cadlag function f define the generalized inverse $f^{-1}(u) = \inf\{t : f(t) \leq u\}$. Note that δ^{-1} is the generalized inverse of the cadlag function $x \to \delta_{[x]}$, $[.]$ denoting the integer part.

Theorem 2 *Let $(X_i)_{i>0}$ be a sequence of centered and square integrable random variables, and define $(\mathcal{M}_i)_{i>0}$ and $\overline{S}_n$ as in Theorem 1. Let X be some positive random variable sucht that $Q_X \geq \sup_{k\geq 1} Q_{|X_k|}$ and*

$$s_n^2 = \sum_{i=1}^n \sum_{j=1}^n |\mathrm{Cov}(X_i, X_j)|\,.$$

Let $R = ((\tau/2)^{-1} \circ G_X^{-1})Q_X$ and $S = R^{-1}$. For any $\lambda > 0$ and $r \geq 1$,

$$\mathbb{P}(\overline{S}_n \geq 5\lambda) \leq 4\Big(1 + \frac{\lambda^2}{rs_n^2}\Big)^{-r/2} + \frac{4n}{\lambda}\int_0^{S(\lambda/r)} Q_X(u)du\,. \tag{4.6}$$

Proof. We proceed as in Rio (2000) and we use the notations of the proof of Theorem 1. Define $\overline{U}_i = (U_i \wedge qM) \vee (-qM)$ et $\varphi_M(x) = (|x| - M)_+$. Arguing as in Rio we can show that

$$\overline{S}_n \leq \max_{1\leq j\leq [n/q]} \Big|\sum_{i=1}^j (\overline{U}_i - \mathbb{E}(\overline{U}_i))\Big| + qM + \sum_{k=1}^n (\mathbb{E}(\varphi_M(X_k)) + \varphi_M(X_k))\,. \tag{4.7}$$

Choose $v = S(\lambda/r)$, $q = (\tau/2)^{-1} \circ G_X^{-1}(v)$ and $M = Q_X(v)$. We have that

$$qM = R(v) = R(S(\lambda/r)) \leq \lambda/r\,.$$

We use the same arguments as in the proof of inequality (4.4). Since $\overline{U}_i$ a 1-Lipschitz function of U_i, we have that $\tau(\sigma(U_l, l \leq i-2), \overline{U}_i) \leq q\tau_q(q+1)$. This fact together with the inequality $s_n^2 \geq \|\overline{U}_1\|_2^2 + \cdots + \|\overline{U}_{[n/q]}\|_2^2$ yield

$$\mathbb{P}\Big(\max_{1\leq j\leq [n/q]} \Big|\sum_{i=1}^j (\overline{U}_i - \mathbb{E}(\overline{U}_i))\Big| \geq 3\lambda\Big) \leq 4\Big(1 + \frac{\lambda^2}{rs_n^2}\Big)^{-r/2} + \frac{n}{\lambda}\tau(q+1)\,. \tag{4.8}$$

On the other hand, since $M = Q_X(v)$,

$$\mathbb{P}\Big(\sum_{k=1}^n (\mathbb{E}(\varphi_M(X_k)) + \varphi_M(X_k)) \geq \lambda\Big) \leq \frac{2n}{\lambda}\int_0^v Q_X(u)du\,. \tag{4.9}$$

The choice of q implies that $\tau(q) \leq 2\int_0^v Q_X(u)du$. Since $qM \leq \lambda$, the result follows from (4.7), (4.8) and (4.9).

Corollary 1 *Let* $(X_i)_{i>0}$ *be a sequence of centered random variables belonging to* $\mathbb{L}^p$ *for some* $p \geq 2$*. Define* $(\mathcal{M}_i)_{i>0}$, $\overline{S}_n$, Q_X *and* s_n *as in Theorem 2. The following inequalities hold*

$$\|\overline{S}_n\|_p^p \leq a_p s_n^p + n b_p \int_0^{\|X\|_1} ((\tau/2)^{-1}(u))^{p-1} Q_X^{p-1} \circ G_X(u) du \,,$$

where $a_p = 4p5^p(p+1)^{p/2}$ *and* $(p-1)b_p = 4p5^p(p+1)^{p-1}$*. Moreover we have that*

$$s_n^2 \leq 4n \int_0^{\|X\|_1} (\tau/2)^{-1}(u) Q_X \circ G_X(u) du \,.$$

Proof. The result follows by integrating the inequality of Theorem 2 (as done in Rio (2000) page 88) and by noting that

$$\int_0^1 Q(u)(R(u))^{p-1}(u)du = \int_0^{\|X\|_1} ((\tau/2)^{-1}(u))^{p-1} Q_X^{p-1} \circ G_X(u) du \,.$$

The upper bound for s_n^2 holds with θ instead of τ (cf. Dedecker and Doukhan (2002), Proposition 2).

5 Strong invariance principle

The main result of this section is a strong invariance principle for partial sums of τ-dependent sequences. Let $(X_i)_{i\in\mathbb{Z}}$ be a stationary sequence of zero-mean square integrable random variables, and $\mathcal{M}_i = \sigma(X_j, j \leq i)$. Define

$$S_n = X_1 + \cdots + X_n \quad \text{and} \quad S_n(t) = S_{[nt]} + (nt - [nt])X_{[nt]+1} \,.$$

Assume that $n^{-1}\mathrm{Var}(S_n)$ converges to some constant σ^2 as n tends to infinity (this will always be true for any of the conditions we shall use hereafter). For $\sigma > 0$, we study the almost sure behavior of the partial sum process

$$\left\{\sigma^{-1}\left(2n \ln\ln n\right)^{-1/2} S_n(t) \,:\, t \in [0,1]\right\}. \tag{5.1}$$

Let $\mathcal{S}$ be the subset of $C([0,1])$ consisting of all absolutely continuous functions with respect to the Lebesgue measure such that $h(0) = 0$ and $\int_0^1 (h'(t))^2 dt \leq 1$.

In 1964, Strassen proved that if the sequence $(X_i)_{i\in\mathbb{Z}}$ is i.i.d. then the process defined in (5.1) is relatively compact with a.s. limit set $\mathcal{S}$. This result is known as the functional law of the iterated logarithm (FLIL for short). Heyde and Scott (1973) extended the FLIL to the case where $\mathbb{E}(X_1|\mathcal{M}_0) = 0$ and the sequence is ergodic. Starting from this result and from a coboundary decomposition due to Gordin (1969), Heyde (1975) proved that the FLIL holds if $\mathbb{E}\left(S_n|\mathcal{M}_0\right)$ converges in

$\mathbb{L}_2$ and the sequence is ergodic. Applying Proposition 1 in Dedecker and Doukhan (2002), we see that Heyde's condition holds as soon as

$$\sum_{k=1}^{\infty} k \int_0^{\gamma(k)/2} Q \circ G(u)\, du < \infty, \tag{5.2}$$

where the functions $Q = Q_{|X_0|}$ and $G = G_{|X_0|}$ have been defined in Notation 2 and $\gamma(k) = \|\mathbb{E}\,(X_k|\mathcal{M}_0)\,\|_1$ is the coefficient introduced by Gordin (1973).

For strongly mixing sequences, Rio (1995) proved the FLIL (and even a strong invariance principle) for the process defined in (5.1) as soon as the DMR (Doukhan, Massart and Rio, 1994) condition (5.3) is satisfied

$$\sum_{k=1}^{\infty} \int_0^{2\alpha(k)} Q^2(u)\, du < \infty. \tag{5.3}$$

From Lemma 7, we easily infer that

$$\int_0^{\gamma(k)/2} Q \circ G(u)\, du \le \int_0^{\tau(k)/2} Q \circ G(u)\, du \le \int_0^{2\alpha(k)} Q^2(u)\, du. \tag{5.4}$$

Hence a reasonable conjecture for the FLIL is that condition (5.2) holds without the k in front of the integral. Actually, we can only prove this conjecture with $\tau(k)$ instead of $\gamma(k)$, that is the FLIL holds as soon as

$$\sum_{k=1}^{\infty} \int_0^{\tau(k)/2} Q \circ G(u)\, du < \infty. \tag{5.5}$$

More precisely, we shall prove that

Theorem 3 *Let $(X_n)_{n\in\mathbb{Z}}$ be a strictly stationary sequence of centered and square integrable random variables satisfying (5.5). Then $n^{-1}\mathrm{Var}(S_n)$ converges to σ^2, and there exists a sequence $(Y_n)_{n\in\mathbb{N}}$ of independent $\mathcal{N}(0,\sigma^2)$-distributed random variables (possibly degenerate) such that*

$$\sum_{i=1}^{n} (X_i - Y_i) = o\left(\sqrt{n \ln\ln n}\right) \ a.s.$$

Such a result is known as a strong invariance principle. If $\sigma > 0$, Theorem 3 and Strassen's FLIL for the Brownian motion yield the FLIL for the process (5.1).

Starting from (5.5) and applying Lemma 2 of Dedecker and Doukhan (2002), we obtain some simple sufficient conditions for the FLIL to hold.

Corollary 2 *Let $(X_i)_{i\in\mathbb{Z}}$ be a strictly stationary sequence of centered and square integrable random variables. Any of the following conditions implies (5.5) and hence the FLIL.*

1. $\mathbb{P}(|X_0| > x) \le (c/x)^r$ *for some* $r > 2$*, and* $\sum_{i\ge 0} (\tau(i))^{(r-2)/(r-1)} < \infty$.

2. $\|X_0\|_r < \infty$ *for some* $r > 2$*, and* $\sum_{i\ge 0} i^{1/(r-2)}\tau(i) < \infty$.

3. $\mathbb{E}(|X_0|^2 \ln(1+|X_0|)) < \infty$ *and* $\tau(i) = O(a^i)$ *for some* $a < 1$.

Now, according to Lemma 7 and to the examples given in Doukhan, Massart and Rio (1994), we can see that Condition (5.5) is essentially optimal. For instance, Corollary 3 below follows easily from Proposition 3 in Doukhan, Massart and Rio.

Corollary 3 *For any* $r > 2$*, there exists a strictly stationary Markov chain* $(X_i)_{i\in\mathbb{Z}}$ *such that* $\mathbb{E}(X_0) = 0$ *and*

1. *For any nonnegative real* x*,* $\mathbb{P}(|X_0| > x) = \min(1, x^{-r})$.

2. *The sequence* $(\tau_i)_{i\ge 0}$ *satisfies* $\sup_{i\ge 0} i^{(r-1)/(r-2)}\tau(i) < \infty$.

3. $\displaystyle\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{n\ln\ln n}} = +\infty$ *almost surely.*

5.1 Proofs

Notations 4. Define the set

$$\Psi = \left\{ \psi : \mathbb{N} \to \mathbb{N}\,,\ \psi \text{ increasing}\,,\ \frac{\psi(n)}{n} \xrightarrow[n\to+\infty]{} +\infty\,,\ \psi(n) = o\left(n\sqrt{LLn}\,\right) \right\}.$$

If ψ is some function of Ψ, let $M_1 = 0$ and $M_n = \sum_{k=1}^{n-1}(\psi(k)+k)$ for $n \ge 2$. For $n \ge 1$, define the random variables

$$U_n = \sum_{i=M_n+1}^{M_n+\psi(n)} X_i\,, \quad V_n = \sum_{i=M_{n+1}+1-n}^{M_{n+1}} X_i\,, \quad \text{and} \quad U_n' = \sum_{i=M_n+1}^{M_{n+1}} |X_i|.$$

If $Lx = \max(1, \ln x)$, define the truncated random variables

$$\overline{U}_n = \max\left(\min\left(U_n, \frac{n}{\sqrt{LLn}}\right), \frac{-n}{\sqrt{LLn}}\right).$$

Theorem 3 is a consequence of the following Proposition

Proposition 1 *Let* $(X_n)_{n\in\mathbb{Z}}$ *be a strictly stationary sequence of centered and square integrable random variables satisfying condition (5.5). Then* $n^{-1}\mathrm{Var}(S_n)$ *converges to* σ^2 *and there exist a function* $\psi \in \Psi$ *and a sequence* $(W_n)_{n\in\mathbb{N}}$ *of independent* $\mathcal{N}(0, \psi(n)\sigma^2)$*-distributed random variables (possibly degenerate) such that*

(a) $\displaystyle\sum_{i=1}^{n}(W_i - \overline{U}_i) = o\left(\sqrt{M_n LLn}\right)$ *a.s.*

(b) $\displaystyle\sum_{n=1}^{\infty} \frac{\mathbb{E}(|U_n - \overline{U}_n|)}{n\sqrt{LLn}} < \infty$

(c) $U_n' = o\left(n\sqrt{LLn}\right)$ *a.s.*

Proof of Proposition 1. It is adapted from the proof of Proposition 2 in Rio (1995).

Proof of (b). Note first that

$$\mathbb{E}|U_n - \overline{U}_n| = \mathbb{E}\left(\left(|U_n| - \frac{n}{\sqrt{LLn}}\right)_+\right) \text{ so that } \mathbb{E}|U_n - \overline{U}_n| = \int_{\frac{n}{\sqrt{LLn}}}^{+\infty} \mathbb{P}(|U_n| > t)dt. \tag{5.6}$$

In the following we write Q instead of $Q_{|X_0|}$. Since U_n is distributed as $S_{\psi(n)}$, we infer from Theorem 2 that

$$\mathbb{P}(|U_n| > t) \leq 4\Big(1 + \frac{t^2}{25\, r\, s^2_{\psi(n)}}\Big)^{-\frac{r}{2}} + \frac{20\,\psi(n)}{t} \int_0^{S\left(\frac{t}{5r}\right)} Q(u)du. \tag{5.7}$$

Consider the two terms

$$A_{1,n} = \frac{4}{n\sqrt{LLn}} \int_{\frac{n}{\sqrt{LLn}}}^{+\infty} \Big(1+\frac{t^2}{25\, r\, s^2_{\psi(n)}}\Big)^{-\frac{r}{2}} dt\,,\ A_{2,n} = \frac{20\,\psi(n)}{n\sqrt{LLn}} \int_{\frac{n}{\sqrt{LLn}}}^{+\infty} \frac{1}{t} \int_0^{S\left(\frac{t}{5r}\right)} Q(u)du\, dt.$$

From (5.6) and (5.7), we infer that

$$\frac{\mathbb{E}|U_n - \overline{U}_n|}{n\sqrt{LLn}} \leq A_{1,n} + A_{2,n}\,. \tag{5.8}$$

Study of $A_{1,n}$. Since the sequence $(X_n)_{n\in\mathbb{N}}$ satisfies (5.5), $s^2_{\psi(n)}/\psi(n)$ converges to some positive constant. Let C_r denote some constant depending only on r which may vary from line to line. We have that

$$A_{1,n} \leq \frac{4}{n\sqrt{LLn}} \int_{\frac{n}{\sqrt{LLn}}}^{+\infty} \frac{t^{-r}}{C_r\, s^{-r}_{\psi(n)}} dt \leq C_r\, s^r_{\psi(n)}\, \frac{n^{-r}}{LLn^{1-\frac{r}{2}}}.$$

We infer that $A_{1,n} = O(\psi(n)^{r/2} n^{-r} LLn^{(r-2)/2})$ as n tends to infinity. Since $\psi \in \Psi$ and $r > 2$, we infer that $\sum_{n\geq 1} A_{1,n}$ is finite.

Study of $A_{2,n}$. We use the elementary result: if $(a_i)_{i\geq 1}$ is a sequence of positive numbers, then there exists a sequence of positive numbers $(b_i)_{i\geq 1}$ such that $b_i \to \infty$ and $\sum_{i\geq 1} a_i b_i < \infty$ if and only if $\sum_{i\geq 1} a_i < \infty$ (note that $b_n^2 = (\sum_{i=n}^{\infty} a_i)^{-1}$ works). Consequently $\sum_{n\geq 1} A_{2,n}$ is finite for some $\psi \in \Psi$ if and only if

$$\sum_{n\geq 1} \frac{1}{\sqrt{LLn}} \int_{\frac{n}{\sqrt{LLn}}}^{+\infty} \frac{1}{t} \int_0^{S\left(\frac{t}{5r}\right)} Q(u)du\, dt < +\infty. \tag{5.9}$$

Recall that $S = R^{-1}$, with the notations of Theorem 2. To prove (5.9), write

$$\begin{aligned}\int_{\frac{n}{\sqrt{LLn}}}^{+\infty} \frac{1}{t} \int_0^{S\left(\frac{t}{5r}\right)} Q(u)du\, dt &= \int_{\frac{n}{\sqrt{LLn}}}^{+\infty} \frac{1}{t} \int_0^1 1_{R(u)\geq \frac{t}{5r}} Q(u)du\, dt \\ &= \int_0^1 Q(u) \int_{\frac{n}{\sqrt{LLn}}}^{5\,r\,R(u)} \frac{1}{t} dt du \\ &= \int_0^1 Q(u) \ln \frac{5\,r\,R(u)}{\frac{n}{\sqrt{LLn}}} \, 1_{R(u)\geq \frac{n}{5\,r\,\sqrt{LLn}}} du.\end{aligned}$$

Consequently (5.9) holds if and only if

$$\int_0^1 Q(u) \sum_{n\geq 1} \frac{1}{\sqrt{LLn}} \ln \frac{5\,r\,R(u)}{\frac{n}{\sqrt{LLn}}} \, 1_{R(u)\geq \frac{n}{5\,r\,\sqrt{LLn}}} du < +\infty. \tag{5.10}$$

To see that (5.10) holds, we shall prove the following result: if f is any increasing function such that $f(0) = 0$ and $f(1) = 1$, then for any positive R we have that

$$\sum_{n\geq 1} \ln\left(\frac{R}{f(n)}\right) (f(n) - f(n-1)) \, 1_{f(n)\leq R} \leq \max(R-1, 0) \leq R\,. \tag{5.11}$$

Applying this result to $f(x) = x(LLx)^{-1/2}$ and $R = 5rR(u)$, and noting that $(LLn)^{-1/2} \leq C\,(f(n) - f(n-1))$ for some constant $C > 1$, we infer that

$$\int_0^1 Q(u) \sum_{n\geq 1} \frac{1}{\sqrt{LLn}} \ln \frac{5\,r\,R(u)}{\frac{n}{\sqrt{LLn}}} \, 1_{R(u)\geq \frac{n}{5\,r\,\sqrt{LLn}}} du \leq \; 5Cr \int_0^1 Q(u)R(u)du\,,$$

which is finite as soon as (5.5) holds.

It remains to prove (5.11). If $R \leq 1$, the result is clear. Now, for $R > 1$, let x_R be the greatest integer such that $f(x_R) \leq R$ and write $R^* = f(x_R)$. Note first that

$$\sum_{n\geq 1} \ln(R)\,(f(n) - f(n-1)) \, 1_{f(n)\leq R} \leq R^* \ln(R). \tag{5.12}$$

On the other hand, we have that

$$\sum_{n\geq 1} \ln(f(n))\,(f(n) - f(n-1)) \, 1_{f(n)\leq R} = \sum_{n=1}^{x_R} \ln(f(n))\,(f(n) - f(n-1))\,.$$

It follows that

$$\sum_{n\geq 1} \ln(f(n))\,(f(n) - f(n-1)) \geq \int_1^{R^*} \ln(x)dx = R^* \ln(R^*) - R^* + 1. \tag{5.13}$$

Using (5.12) and (5.13) we get that

$$\sum_{n\geq 1} \ln\left(\frac{R}{f(n)}\right) (f(n) - f(n-1)) \, 1_{f(n)\leq R} \leq R^* - 1 + R^*(\ln(R) - \ln(R^*)). \tag{5.14}$$

Using Taylor's inequality, we have that $R^*(\ln(R) - \ln(R^*)) \le R - R^*$ and (5.11) follows. The proof of (b) is complete.

Proof of (c). Let $T_n = \sum_{i=M_n+1}^{M_{n+1}} (|X_i| - \mathbb{E}|X_i|)$. We easily see that

$$U_n^{'} = (\psi(n) + n)\,\mathbb{E}(|X_1|) + T_n. \tag{5.15}$$

By definition of Ψ, we have $\psi(n) = o\left(n\sqrt{LLn}\right)$. Here note that

$$T_n \le \frac{n}{\sqrt{LLn}} + \sup\left(0, T_n - \frac{n}{\sqrt{LLn}}\right). \tag{5.16}$$

Using same arguments as for the proof of (b), we obtain that

$$\sum_{n\ge 1} \frac{\mathbb{E}\left(\sup\left(0, T_n - \frac{n}{\sqrt{LLn}}\right)\right)}{n\sqrt{LLn}} < +\infty, \text{ so that } \sum_{n\ge 1} \frac{\left(\sup\left(0, T_n - \frac{n}{\sqrt{LLn}}\right)\right)}{n\sqrt{LLn}} < +\infty \text{ a.s.}$$

Consequently $\max(0, T_n - n(LLn)^{-1/2}) = o(n\sqrt{LLn})$ almost surely, and the result follows from (5.15) and (5.16).

Proof of (a). In the following, $(\delta_n)_{n\ge 1}$ and $(\eta_n)_{n\ge 1}$ denote independent sequences of independent random variables with uniform distribution over $[0, 1]$, independent of $(X_n)_{n\ge 1}$. Since $\overline{U}_n$ is a 1-Lipschitz function of U_i, $\tau(\sigma(U_i, i \le n-1), \overline{U}_n) \le \psi(n)\tau(n)$. Using Lemma 5 and arguing as in the proof of Theorem 1, we get the existence of a sequence $(\overline{U}_n^*)_{n\ge 1}$ of independent random variables with the same distribution as the random variables $\overline{U}_n$ such that $\overline{U}_n^*$ is a measurable function of $\left(\overline{U}_l, \delta_l\right)_{l\le n}$ and

$$\mathbb{E}\left(|\overline{U}_n - \overline{U}_n^*|\right) \le \psi(n)\,\tau(n).$$

Since (5.5) holds, we have that

$$\sum_{n\ge 1} \frac{\mathbb{E}\left(|\overline{U}_n - \overline{U}_n^*|\right)}{\sqrt{M_n\,LLn}} < +\infty \quad \text{so that} \quad \sum_{n\ge 1} \frac{|\overline{U}_n - \overline{U}_n^*|}{\sqrt{M_n\,LLn}} < +\infty \text{ a.s.}$$

Applying Kronecker's lemma, we obtain that

$$\sum_{i=1}^{n} (\overline{U}_i - \overline{U}_i^*) = o\left(\sqrt{M_n\,LLn}\right) \text{ a.s.} \tag{5.17}$$

We infer from (5.5) and from Dedecker and Doukhan (2002) that

$$(\psi(n))^{-1}\,\mathrm{Var}\,U_n \xrightarrow[n\to+\infty]{} \sigma^2 \quad \text{and} \quad (\psi(n))^{-1/2}\,U_n \xrightarrow[n\to+\infty]{\mathcal{D}} \mathcal{N}\left(0, \sigma^2\right).$$

Hence the sequence $(U_n^2/\psi(n))_{n\ge 1}$ is uniformly integrable (Theorem 5.4. in Billingsley (1968)). Consequently, since the random variables $\overline{U}_n^*$ have the same distribution as the random variables $\overline{U}_n$, we deduce from the above limit results, from Strassen's representation theorem (see Dudley (1968)), and from Skorohod's lemma (1976) that one can construct some sequence $(W_n)_{n\geq 1}$ of $\sigma(\overline{U}_n^*, \eta_n)$-measurable random variables with respective distribution $\mathcal{N}(0, \psi(n)\,\sigma^2)$ such that

$$\mathbb{E}\left(\left(\overline{U}_n^* - W_n\right)^2\right) = o\,(\psi(n)) \text{ as } n \to +\infty, \tag{5.18}$$

which is exactly equation (5.17) of the proof of Proposition 2(c) in Rio (1995). The end of the proof is the same as that of Rio.

Proof of Theorem 3. By Skohorod's lemma (1976), there exists a sequence $(Y_i)_{i\geq 1}$ of independent $\mathcal{N}(0, \sigma^2)$-distributed random variables satisfying for all positive n $W_n = \sum_{i=M_n+1}^{M_n+\psi(n)} Y_i$. Define the random variable $V_n' = \sum_{i=M_{n+1}+1-n}^{M_{n+1}} Y_i$.

Let $n(k) := \sup\{n \geq 0 : M_n \leq k\}$, and note that by definition of M_n we have $n(k) = o(\sqrt{k})$. Applying Proposition 1(c) we see that

$$\Big|\sum_{i=1}^{k} X_i - \sum_{i=1}^{n(k)}(U_i + V_i)\Big| \leq U'_{n(k)} = o\left(\sqrt{k\,LLk}\right) \text{ a.s.} \tag{5.19}$$

From (5.26) in Rio (1995), we infer that

$$\sum_{i=1}^{n(k)} V_i = o\left(\sqrt{k\,LLk}\right) \text{ a.s.} \quad \text{ and } \quad \sum_{i=1}^{n(k)} V_i' = o\left(\sqrt{k\,LLk}\right) \text{ a.s.} \tag{5.20}$$

Gathering (5.19), (5.20) and Proposition 1(a) and (b), we obtain that

$$\sum_{i=1}^{k} X_i - \sum_{i=1}^{n(k)}(W_i + V_i') = o\left(\sqrt{k\,LLk}\right) \text{ a.s.} \tag{5.21}$$

Clearly $\sum_{i=1}^{k} Y_i - \sum_{i=1}^{n(k)}(W_i + V_i')$ is normally distributed with variance smaller than $\psi(n(k)) + n(k)$. Since $n(k) = o(\sqrt{k})$ we have that $\psi(n(k)) + n(k) = o(\sqrt{kLLk})$ by definition of ψ. An elementary calculation on Gaussian random variables shows that

$$\sum_{i=1}^{k} Y_i - \sum_{i=1}^{n(k)}(W_i + V_i') = o\left(\sqrt{k\,LLk}\right) \text{ a.s.} \tag{5.22}$$

Theorem 3 follows from (5.21) and (5.22).

Acknowledgments. We thank Paul Doukhan, Magda Peligrad and Emmanuel Rio for helpfull discussions. Magda Peligrad has told one of the authors that there should be an appropriate coefficient adapted to the coupling in $\mathbb{L}^1$.

References

[1] D. W. K. Andrews, (1984). Nonstrong mixing autoregressive processes. *J. Appl. Probab.* **21** 930-934.

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

[3] P. Billingsley, (1968). *Convergence of probability measures.* Wiley, New-York.

[4] P. Billingsley, (1995). *Probability and Measure, third edition.* Wiley, New-York.

[5] R. C. Bradley, (2002). *Introduction to Strong Mixing Conditions, Volume 1.* Technical Report, Department of Mathematics, I. U. Bloomington.

[6] A. Broise, (1996). Transformations dilatantes de l'intervalle et théorèmes limites. Études spectrales d'opérateurs de transfert et applications. *Astérisque* **238** 1-109.

[7] C. Coulon-Prieur and P. Doukhan, (2000). A triangular CLT for weakly dependent sequences. *Statist. Probab. Lett.* **47** 61-68.

[8] J. Dedecker and P. Doukhan, (2002). A new covariance inequality and applications. To appear in *Stochastic Process. Appl.*

[9] J. Dedecker and E. Rio, (2000). On the functional central limit theorem for stationary processes, *Ann. Inst. H. Poincaré Probab. Statist.* **36** 1-34.

[10] P. Doukhan, (1994). *Mixing: properties and examples.* Lecture Notes in Statist. **85**, Springer-Verlag.

[11] P. Doukhan and S. Louhichi (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Process. Appl.* **84** 313-342.

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

[13] R. M. Dudley, (1968). Distances of probability measures and random variables. *Ann. Math. Statist.* **39** 1563-1572.

[14] R. M. Dudley, (1989). *Real analysis and probability.* Wadworsth Inc., Belmont, California.

[15] D. K. Fuk and S. V. Nagaev, (1971). Probability inequalities for sums of independent random variables. *Theory Probab. Appl.* **16** 643-660.

[16] M. I. Gordin, (1969). The central limit theorem for stationary processes, *Dokl. Akad. Nauk SSSR* **188** 739-741.

[17] M. I. Gordin, (1973). Abstracts of Communication, T.1: A-K, International Conference on Probability Theory, Vilnius.

[18] C. C. Heyde and D. J. Scott, (1973). Invariance principles for the law of the iterated logarithm for martingales and processes with stationary increments. *Ann. Probab.* **1** 428-436.

[19] C. C. Heyde, (1975). On the central limit theorem and iterated logarithm law for stationary processes. *Bull. Amer. Math. Soc.* **12** 1-8.

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

[21] E. J. McShane, (1947). *Integration.* Princeton university press, London.

[22] M. Peligrad, (2002). Some remarks on coupling of dependent random variables. *Stat. Prob. Letters* **60** 201-209.

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

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

[25] M. Rosenblatt, (1956). A central limit theorem and a strong mixing condition, *Proc. Nat. Acad. Sci. U. S. A.* **42** 43-47.

[26] A. V. Skorohod, (1976). On a representation of random variables. *Theory Probab. Appl.* **21** 628-632.

[27] V. Strassen (1964). An invariance principle for the law of the iterated logarithm. *Z. Wahrsch. Verw. Gebiete* **3** 211-226.