INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2005-51

An Invariance Principle for New Weakly Dependent Stationary Models using Sharp Moment Assumptions

P. DOUKHAN[1]
O. WINTENBERGER[2]

[1] CREST, Laboratoire de Statistique, 3 avenue Pierre Larousse, 92245 Malakoff Cedex, France and SAMOS, Statistique Appliquée et MOdélisation Stochastique, Université Paris 1, Centre Pierre Mendès France, 90 rue de Tolbiac, 75634 Cedex 13, France. *Author for correspondence* : doukhan@ensae.fr
[2] SAMOS.

An invariance principle for new weakly dependent stationary models using sharp moment assumptions

Paul Doukhan and Olivier Wintenberger

3rd April 2006

Résumé

Ce travail a pour objet de préciser et d'améliorer un principe d'invariance prouvé pour des suites stationnaires par Doukhan & Louhichi (1999). Notre condition est simultanément plus faible que celles de mélange ou de dépendance faible causale considérée par Dedecker et Doukhan (2003);ces auteurs obtenaient un énoncé de caractère optimal améliorant celui prouvé dans Doukhan *et alii* (1995) sous une condition de mélange fort. Notre résultat, accompagné de bornes pour la vitesse de convergence suppose l'existence de moments d'ordre > 2 alors que Doukhan & Louhichi (1999) supposaient un moment d'ordre > 4. Comme ces auteurs nous utilisons une condition de dépendance faible non causale permettant de traiter de classes très larges de modèles. En plus de conditions déjà introduites de η- et de κ-dépendance faible, nous introduisons une condition mixte, λ, adaptée pour considérer des schémas de Bernoulli à entrées dépendantes..

Abstract

This paper is aimed at sharpen a weak invariance principle for stationary sequences in Doukhan & Louhichi (1999). Our assumption is both beyond mixing and the causal θ-weak dependence in Dedecker and Doukhan (2003); those authors obtained a sharp result which improves on an optimal one in Doukhan *et alii* (1995) under strong mixing. We prove this result and we also precise convergence rates under existence of moments with order > 2 while Doukhan & Louhichi (1999) assume a moment of order > 4. Analogously to those authors, we use a non-causal condition to deal with some general classes of stationary and weakly dependent sequences. Besides the previously used η- and κ-weak dependence conditions, we introduce a mixed condition, λ, adapted to consider Bernoulli shifts with dependent inputs.

1 Introduction and main results

Working with times series provides a huge amount of applications. Several ways of modeling the weak dependence have already been worked out. One of the most popular is the notion of mixing, see Doukhan (1994) for bibliography; this allows a very nice asymptotic theory (see Rio, 2000). However, using mixing presents lots of restrictions. For example, Andrews (1984) exhibits the simple counter-example of an auto-regressive process which does not satisfy a mixing condition. Doukhan and Louhichi (1999) introduced several new weak dependence conditions to solve those problems. The present work aims to prove a sharp weak invariance principle under such conditions.

An usual way to derive a sharp weak invariance principle for stationary sequences is based on the approximation with martingales differences. This approach was first explored by Gordin (1969) under a moment condition of order 2 and was recently improved in Dedecker & Rio (2000). Such approximations are very natural in causal cases, when the process $(X_t)_t$ is adapted with a filtration $\sigma(\xi_i, i \le t)$ where $(\xi_i)_i$ is an iid (independent and identically distributed) sequence. Weak invariance principle holds as soon as a specific decay of the dependence coefficients holds. This is made for example in the causal case of θ-dependence in Dedecker & Doukhan (2003).

In non causal context, the approximation technique of Gordin (1969) also works, as proved in Heyde (1975). This approach has been improved in Dedecker & Rio (2000):

Theorem 1 (Dedecker & Rio (2000)) *Here we denote by $\mathbb{E}_0(Y)$ the conditional expectation of a random variable Y with respect to some fixed σ-algebra $\mathcal{M}_0$.*
Let $(X_i)_{i\in\mathbb{Z}}$ be a 0-mean stationary sequence satisfying the moment assumption $\mathbb{E}|X_0|^m < \infty$ (see (5) below). We set $m' = m/(m-1)$ for the conjugate exponent of m. Assume that the series $\sum_{n=0}^{\infty} \mathbb{E}_0(X_n)$ converge in $\mathbb{L}^{m'}$ and:

$$\sum_{n=0}^{\infty} \|X_{-n} - \mathbb{E}_0(X_{-n})\|_{m'} < \infty.$$

If moreover the series $\sigma^2 = \sum_{k\in\mathbb{Z}} Cov(X_0, X_k) = \sum_{k\in\mathbb{Z}} \mathbb{E}X_0X_k \ge 0$ (see eqn. (6) below) converges, then:

$$W_n(t) \to_{n\to\infty} \sigma W(t), \quad \textit{in distribution, in the Skorohod space } D([0,1]).$$

This result improves on Heyde (1975) who states this result for $m = m' = 2$.

Another way to derive weak invariance principle in a non-causal context is the result of Newman & Wright (1981). Those authors work under association which is not of a causal nature. In our paper, we will exhibit various examples which are neither associated nor approximable by increments of martingales. For this we explore the case of non-causal two-sided linear models in section 3.1:

$$X_t = \sum_{i=-\infty}^{\infty} b_i\, Y_{t+i}, \tag{1}$$

where summations run from $-\infty$ to $+\infty$, and the inputs $(Y_t)_{t\in\mathbb{Z}}$ are non-linear and dependent sequences.

If $(Y_t)_{t\in\mathbb{Z}}$ is a sequence adapted to the filtration $\mathcal{M}_i = \sigma(Y_j; j \le i)$, then Peligrad & Utev (2006) provide conditions on the inputs to derive weak invariance principles, see [26]. Other conditions may be found in Dedecker & Rio (2000) and in Louhichi & Soulier (2000). It seems however quite uneasy to verify effectively such conditions in various cases; for example:

1. consider the equation (1) with non causal sequence and non linear inputs $(Y_t)_{t\in\mathbb{Z}}$ given by:
$$Y_t = \xi_t \sum_{i\neq 0} a_i\, \xi_{t+i}, \tag{2}$$
for some iid sequence $(\xi_t)_{t\in\mathbb{Z}}$. For such models, the conditions of Theorem 1 do not seem to hold.

2. other linear models (1) may be written with associated inputs $(Y_t)_{t\in\mathbb{Z}}$. In this case, the projective criteria (see also Merlevède and Peligrad (2004)) do not seem possible to use directly. However, X_t writes here as a linear combination of associated variables; indeed, selecting non-negative and non-positive coefficients b_i allows to write such models as the difference of two associated sequences (this remark follows from Louhichi and Soulier (2000)). An example of such inputs is:
$$Y_t = \xi_t \sum_{i\neq 0} a_i\, \exp(\xi_{t+i}), \tag{3}$$
where $(\xi_t)_{t\in\mathbb{Z}}$ denotes an iid sequence.

3. we now recall that both associated and Gaussian sequences are $\kappa-$weakly dependent (see the definition below). Let $(\xi'_t)_{t\in\mathbb{Z}}$ be a Gaussian process, independent of some associated random process $(\xi_t)_{t\in\mathbb{Z}}$, we set
$$Y_t = \xi_t \exp(\xi'_t), \tag{4}$$
then $(Y_t)_{t\in\mathbb{Z}}$ is still $\kappa-$weakly dependent. The process (1) is now λ-weakly dependent; this yields a simple way to derive a weak invariance principle. Morevover the previous trick for association is not more valid here.

Hence for a non causal linear model (1) with dependent inputs, various criteria may be used to prove a weak invariance principle.

Section 3.2 is devoted to the more general case of Bernoulli shifts $X_n = H(Y_{n-j}, j\in\mathbb{Z})$ with weakly dependent input process $(Y_t)_{t\in\mathbb{Z}}$. Analogue models with dependent inputs are already considered by Borovkova *et al.* (2001) in [4]. The new notion of λ-weak dependence yields a weak invariance principle for those models, see theorem 4.

Our idea in this work also drives the idea of weak dependence: *a general criterion for limit theorems implies more robustness relatively to the model.* We really believe that it is important for practitioners to dispose of criteria simple to use. In special cases when such general results are not good enough, one needs to specify the models in order to use a sharp criterion.

Our proofs use the Bernstein blocks technique and Lindeberg method. We relax Doukhan and Louhichi (1999)'s moment condition on the observations from an order $m > 4$ to an order $m = 2+\zeta > 2$. Their result in [12] is clearly not optimal and the reason of such assumptions is that their proof, based on the Bernstein blocks technique, relies on a combinatorial moment inequality of order 4; this excludes lower order moment assumptions. We propose an alternative proof to bypass this problem: for this, we use a moment inequality of order $2+\delta$, where $0<\delta<\zeta$. It writes $\|X_1+\cdots+X_n\|_{2+\delta} \le c\sqrt{n}$ and implies both the tightness (see Billingsley, 1968, [3]) and a rate in the central limit theorem (CLT). The method used here, originated from Ibragimov (1975) was recently used by Bulinski and Sashkin (2005) for κ'-weakly dependent random fields (see the definition in section 2).

The paper is organized as follows. In the forthcoming section 2, we introduce various weak dependent coefficients and give our results. Then, in section 3, we give examples of weak dependent models and we also discuss the validity of our weak invariance principle in each case. Besides more standard examples, we shall focus on some examples of λ-weakly dependent sequences. We compare our assumptions with those in Dedecker & Rio (2000). The other sections are devoted to the proofs. We first derive conditions ensuring the convergence of the series σ^2 and we then obtain a bound for the moment of order $(2+\delta)$-th of a sum (of an independent interest), in section 4. The proofs are collected in section 5. The standard Lindeberg method with Bernstein blocks is developed in § 5.1 and yield our versions of Donsker theorem. Rates of convergence in the CLT are obtained in § 5.2.

2 Definitions and main results

2.1 Weak dependence assumptions

Definition 1 (see [12]) *A vector valued process $(X_n)_{n\in\mathbb{Z}}$ with values in $\mathbb{R}^d$, endowed with some norm $\|\cdot\|$, is said to be (ϵ,ψ)-weakly dependent if there exist a sequence $\epsilon_r \downarrow 0$ (as $r \uparrow \infty$) and a function $\psi : \mathbb{N}^2 \times (\mathbb{R}^+)^2 \to \mathbb{R}^+$ such that:*

$$|Cov(f(X_{s_1},\dots,X_{s_u}),g(X_{t_1},\dots,X_{t_v})| \le \psi(u,v,Lip\,f,Lip\,g)\epsilon_r,$$

for any $r \ge 0$ and any $(u+v)$-tuples such that $s_1 \le \cdots \le s_u \le s_u + r \le t_1 \le \cdots \le t_v$, where the real valued functions f,g are defined respectively on $\left(\mathbb{R}^d\right)^u$ and $\left(\mathbb{R}^d\right)^v$, and they satisfy $\|f\|_\infty, \|g\|_\infty \le 1$ and $Lip\,f + Lip\,g < \infty$ where we set

$$Lip\,f = \sup_{(x_1,\dots,x_u)\neq(y_1,\dots,y_u)} \frac{|f(x_1,\dots,x_u) - f(y_1,\dots,y_u)|}{\|x_1-y_1\| + \cdots + \|x_u - y_u\|}$$

Specific functions ψ yield notions of weak dependence appropriate to describe various examples of models:

- *κ-weak dependence for which $\psi(u,v,a,b) = uvab$, in this case we simply denote ϵ_r as κ_r,*
- *κ' (causal) weak dependence for which $\psi(u,v,a,b) = vab$, in this case we simply denote ϵ_r as κ'_r; this is the causal counterpart of κ coefficients which we recall only for completeness,*
- *η-weak dependence, $\psi(u,v,a,b) = ua + vb$, in this case we write $\epsilon_r = \eta_r$ for short,*
- *θ-weak dependence is a causal dependence which refers to $\psi(u,v,a,b) = vb$, in this case we simply denote $\epsilon_r = \theta_r$ (see Dedecker & Doukhan, 2003); this is the causal counterpart of η coefficients which we recall only for completeness,*
- *λ-weak dependence $\psi(u,v,a,b) = uvab + ua + vb$, in this case we write $\epsilon_r = \lambda_r$.*

Remarks.

- Besides the fact that it includes η- and κ-weak dependence, this new notion of λ-weak dependence will be proved to be convenient, for example, for Bernoulli shifts with associated inputs (see Theorem 3 below).

- If now the functions f, g take their values in $\mathbb{C}$, the previous inequalities remain true by replacing ϵ_r by $\epsilon_r/2$. A useful case of such complex valued functions is $f(x_1, \ldots, x_u) = \exp\left(it \cdot (x_1 + \cdots + x_u)\right)$ for each $t \in \mathbb{R}^d$, $u \in \mathbb{N}^*$ and $x_1, \ldots, x_u \in \mathbb{R}^d$ (see section 5.1). This indeed corresponds to the characteristic function adapted to derive the convergence in distribution.

2.2 Main results

We consider a stationary, 0-mean, and real valued sequence $(X_n)_{n\in\mathbb{Z}}$ such that

$$\mu = \mathbb{E}|X_0|^m < \infty, \qquad \text{for a real number } m = 2 + \zeta > 2. \tag{5}$$

We also set

$$\sigma^2 = \sum_{k\in\mathbb{Z}} \operatorname{Cov}(X_0, X_k) = \sum_{k\in\mathbb{Z}} \mathbb{E}X_0X_k, \tag{6}$$

W denotes the standard Brownian motion and

$$W_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{nt} X_i, \quad \text{for } t \in [0,1],\ n \geq 1. \tag{7}$$

We now present our main results, they are new versions of the Donsker weak invariance principle.

Theorem 2 (κ-dependence) *Assume that the κ-weakly dependent stationary process satisfies (5) and $\kappa_r = \mathcal{O}(r^{-\kappa})$ (as $r \uparrow \infty$) for $\kappa > (2m-3)/(m-2)$ then the previous expression $\sigma^2 \geq 0$ is well defined and, moreover:*

$$W_n(t) \to_{n\to\infty} \sigma W(t), \qquad \textit{in distribution in the Skorohod space } D([0,1]).$$

Remark. Under the more restrictive κ' condition, Bulinski & Shashkin (2005) obtain invariance principles with the sharper assumption $\kappa' > (2m-3)/(m-2)-1$. The difference of both conditions is natural. Our loss is explained by the fact that κ'-weakly dependent sequences satisfies $\kappa'_r \geq \sum_{s\geq r} \kappa_s$. This simple bound directly follows from the definitions.

The following result relaxes the previous dependence assumptions to the price of a faster decay for the dependence coefficients.

Theorem 3 (λ-dependence) *Assume that the λ-weakly dependent stationary inputs satisfies (5) and $\lambda_r = \mathcal{O}(r^{-\lambda})$ (as $r \uparrow \infty$) for $\lambda > (4m-6)/(m-2)$ then $\sigma^2 \geq 0$ is well defined and, moreover:*

$$W_n(t) \to_{n\to\infty} \sigma W(t), \qquad \textit{in distribution in the Skorohod space } D([0,1]).$$

Remarks.

- We do not achieve better results for η or θ-weak dependence cases than the one for $\lambda-$dependence. Comparatively with the result obtained by Dedecker & Doukhan (2003), our results are thus not good under θ-weak dependence.

  We work under more restrictive moment conditions than those authors; the same remark applies for all projective measures of dependence; we refer here to results by Heyde (1975) and Dedecker & Rio (2000) already mentioned as Theorem 1.

- However we mention here that the example in subsection 3.1.2 stresses the fact that such results do not improve systematically on theorem 3; in fact for this example we even conjecture that theorem 1 does not apply.

  The above mentioned idea of robustness with respect to the model is another strong justification of this result.

- The technique of the proofs is based on Lindeberg method and we prove in fact that $|\mathbb{E}\left(\phi(S_n/\sqrt{n}) - \phi(\sigma N)\right)| \le Cn^{-c^*}$ for constants $c^*, C > 0$ and (ϕ denotes here the characteristic function) depending only on the parameters ζ and κ or λ respectively and where $c^* < \frac{1}{2}$ (see Proposition 2 in section 5.2 for more details). As κ or λ tends to infinity, $c^* \to (m-2)/(m+2)$. If now ζ and κ (or λ) both tend to infinity, we notice that $c^* \to \frac{1}{3}$.

- Using a smoothing lemma also yields an analogue bound for the Levy distance:

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}\left(\frac{1}{\sqrt{n}}S_n \le t\right) - \mathbb{P}\left(\sigma N \le t\right)\right| \le Cn^{-c'}.$$

  A first and easy way to control c' is to set $c' = c^*/4$ but the corresponding rate is really a bad one (see e.g. in [14]). Petrov (1995) obtains the optimal exponent $\frac{1}{2}$ in the iid case and Rio (2000) reaches the exponent $\frac{1}{3}$ for strongly mixing sequences. In proposition 3 section 5.2, we achieve $c' = c^*/3$. Analogue suboptimal convergence rates have been settled in the case of weakly dependent random fields in [10]. Previous results by Heyde (1975) and Dedecker & Rio (2000) as well as Dedecker & Doukhan (2003) or Peligrad & Utev (2006) do not derive such convergence rates for the Levy distance.

The λ-coefficient is very useful to study Bernoulli shifts $X_n = H(Y_{n-j}, j \in \mathbb{Z})$ with weakly dependent input process $(Y_i)_i$ (see section 3.2 for more details). We shall restrict to functions $H : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}$ such that, for each $s \in \mathbb{Z}$, if $x, y \in \mathbb{R}^{\mathbb{Z}}$ satisfy $x_i = y_i$ for each index $i \neq s$ then:

$$|H(x) - H(y)| \le b_s(\|z\|^l \vee 1)|x_s - y_s| \tag{8}$$

where z is defined by $z_s = 0$ and $z_i = x_i = y_i$ for each $i \neq s$. Here $\|x\| = \sup_{i\in\mathbb{Z}} |x_i|$.

Theorem 4 *Let $(Y_i)_i$ be a stationary $\lambda_{Y,r}$-weakly dependent process and $H : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}$ satisfies the condition (8) for some $l \ge 0$ and some sequence $b_j \ge 0$ such that $\sum_j |j|b_j < \infty$. Let assume that there exist real numbers m, m' with $\mathbb{E}|Y_0|^{m'} < \infty$ such that $m > 2$ and $m' \ge (l+1)m$.*

Then $X_n = H(Y_{n-i}, i \in \mathbb{Z})$ satisfies the weak invariance principle in the following cases:

- ***Geometric case:*** *$b_r = \mathcal{O}\left(e^{-|r|b}\right)$ and $\lambda_{Y,r} = \mathcal{O}\left(e^{-ra}\right)$ as $|r| \uparrow \infty$.*

- ***Mixed cases:***
  - $b_r = \mathcal{O}\left(e^{-|r|b}\right)$ *and* $\lambda_{Y,r} = \mathcal{O}\left(r^{-a}\right)$ *as* $|r| \uparrow \infty$ *with*
$$a > \frac{(4m-6)(m'-1)}{(m-2)(m'-1-l)}.$$
  - $b_r = \mathcal{O}\left(|r|^{-b}\right)$ *and* $\lambda_{Y,r} = \mathcal{O}\left(e^{-ar}\right)$ *as* $|r| \uparrow \infty$ *with*
$$b > \frac{6m-10}{m-2}.$$
- ***Riemanian case:*** *If* $b_r = \mathcal{O}\left(|r|^{-b}\right)$ *for some* $b > 2$ *and* $\lambda_{Y,r} = \mathcal{O}\left(r^{-a}\right)$ *as* $|r| \uparrow \infty$ *with*
$$a > \frac{(4m-6)b(m'-1)}{(m-2)(b-2)(m'-1-l)}.$$

The constants $b > 0$ *and* $a > 0$ *are different for each case, they only rely on the involved parameters.*

3 Examples

Theorem 4 is useful to derive the weak invariance principle in various cases. This section is aimed to precise the case of Bernoulli shifts with dependent inputs. The important class of linear models with dependent inputs is presented in a separate section since the importance of our results is enlighten by linear models of the first subsection. Non linear models are reported to a second subsection.

3.1 Linear processes with dependent inputs

We now focus on specific examples of two sided linear sequences (1) with dependent inputs. As quoted in the introduction, we compare the four different conditions from Dedecker & Rio (2000), Louhichi & Soulier (2000) (see also Louhichi (2001)), Peligrad & Utev (2006) with our weak invariance principle in four distincts examples of inputs $(Y_t)_{t\in\mathbb{Z}}$.

3.1.1 LARCH(∞) inputs

A vast literature is devoted to the study of conditionally heteroscedastic models. A simple equation in terms of a vector valued process allows a unified treatment of those models, see [15]. Let $(\xi_t)_{t\in\mathbb{Z}}$ be an iid centered real valued sequence , and $a, a_j, j \in \mathbb{N}^*$ be real numbers. A LARCH(∞) model is a solution of the recurrence equation

$$Y_t = \xi_t \left(a + \sum_{j=1}^{\infty} a_j Y_{t-j} \right). \tag{9}$$

We provide below sufficient conditions for the following chaotic expansion

$$Y_t = \xi_t \left(a + \sum_{k=1}^{\infty} \sum_{j_1,\ldots,j_k \geq 1} a_{j_1}\xi_{t-j_1} a_{j_2} \cdots a_{j_k} \xi_{t-j_1-\cdots-j_k} a \right). \tag{10}$$

Assume that $\Lambda = \|\xi_0\|_{m''} \sum_{j\geq 1} |a_j| < 1$ then one (essentially unique) stationary solution of eqn. (9) in $\mathbb{L}^{m''}$ is given by (10). In this case (thanks to a referee) we note that theorem 1 applies. More precisely (see Peligrad & Utev (2006), theorem 1 in [26]) the weak invariance principle holds for the process $(X_t)_{t\in\mathbb{Z}}$ as soon as

$$\sum_{k\geq 1} \sqrt{\sum_{i\geq k} b_i^2} < +\infty,$$

and conditions on $(a_i)_i$ ensuring that the chaotic expansion (10) holds in $\mathbb{L}^2$. We then derive necessary conditions for weak invariance principle, assuming $|a_j| \leq Cj^{-a}$ for some $a > 1$, $|b_j| \leq C'(|j|+1)^{-b}$, for some $b > 3/2$, $\mathbb{E}\xi_0^2 < +\infty$, and $\|\xi_0\|_2 \sum_{j\geq 1} |a_j| < 1$.

The solution (10) of equation (9) is $\theta-$weakly dependent with

$$\theta_r = \left(\mathbb{E}|\xi_0| \sum_{k=1}^{r-1} k\Lambda^{k-1} A\left(\frac{r}{k}\right) + \frac{\Lambda^r}{1-\Lambda} \right) \mathbb{E}|\xi_0||a|,$$

where $A(x) = \sum_{j\geq x} |a_j|$. There exists some constant $K > 0$ and $b, C > 0$ such that

$$\theta_r \leq \begin{cases} K\frac{(\log(r))^{(a-1)_+}}{r^{a-1}}, \text{ under Riemanian decay } A(x) \leq Cx^{1-a}, \\ K(a \vee \Lambda)^{\sqrt{r}}, \text{ under geometric decay } A(x) \leq Ca^x. \end{cases}$$

We also may use Theorem 4 to derive the weak invariance principle; it provides the stronger conditions $\mathbb{E}|\xi_0|^m < +\infty$ for $m > 2$, $b > 2$, and $a > \dfrac{b(4m-6)}{(b-2)(m-2)} + 1$. Our conditions are thus not at all optimal compared to those of Dedecker & Rio (2000) (see also Peligrad & Utev (2006)).

3.1.2 Non-causal, non-linear inputs

Consider first the two sided linear sequence (1) with non-causal and non-linear inputs $(Y_i)_{i\in\mathbb{Z}}$ as in (2) where the process $(\xi_i)_{i\in\mathbb{Z}}$ is iid and $\mathbb{E}\xi_0 = 0$. This example of inputs is a sub-case of the general one called Bernoulli shifts. We recall here some weak dependence properties of such models (see Doukhan & Louhichi (1999) for more details). Let $H : (\mathbb{R}^d)^{\mathbb{Z}} \to \mathbb{R}$ be a measurable function. If the sequence $(\xi_n)_{n\in\mathbb{Z}}$ is independent and identically distributed on $\mathbb{R}^d$, a Bernoulli shift with input process $(\xi_n)_{n\in\mathbb{Z}}$ is defined as

$$Y_n = H\left((\xi_{n-i})_{i\in\mathbb{Z}}\right), \qquad n \in \mathbb{Z}.$$

The most simple case of infinitely dependent Bernoulli shift is the infinite moving average process (1), with now independent inputs. Such Bernoulli shifts are $\eta-$weakly dependent (see Doukhan & Louhichi (1999)) with $\eta_r \leq 2\delta_{[r/2]}$ if

$$\mathbb{E}\left|H\left(\xi_j, j \in \mathbb{Z}\right) - H\left(\xi_j \mathbb{1}_{|j|\leq r}, j \in \mathbb{Z}\right)\right| \leq \delta_r. \tag{11}$$

In our case of dependent inputs (2), we set

$$\mathbb{E}\left|\xi_0 \left(\sum_{|j|>r} a_j \xi_j\right)\right| = \delta_r.$$

Under a condition of order 2 on ξ_0, we derive the η-weak dependence of the process $(Y_t)_{t\in\mathbb{Z}}$ with $\eta_r \le 2\sqrt{\sum_{|j|>[r/2]} a_j^2}$. Under Riemanian decays ($a_r = \mathcal{O}(r^{-a})$ and $b_r = \mathcal{O}((|r|+1)^{-b})$ for some $a, b > 0$), the weak invariance principle on the process $(X_t)_t$ follows from Theorem 4 with $\mathbb{E}|\xi_0|^m < \infty$, for $m > 2$, $b > 2$, $a > \dfrac{b(4m-6)}{(b-2)(m-2)} + \dfrac{1}{2}$.
We now try to derive the assumptions of Theorem 1 in the special case $m = m' = 2$.
- We fix as usual (but perhaps this is a bad choice) the σ-algebra $\mathcal{M}_0$ as $\sigma(\xi_i, i \le 0)$. Then,

$$\mathbb{E}_0 X_{-n} = \mathbb{E}_0\left(\sum_i \sum_{j\neq 0} b_i a_j \xi_{-n+i}\xi_{-n+i+j}\right) = \sum_i \sum_{j\neq 0} b_i a_j \mathbb{E}_0\left(\xi_{-n+i}\xi_{-n+i+j}\right).$$

Thanks to the conditional expectation, the terms in the sums are null for $i > n$ or $j > n-i$. Otherwise, the expectation is equal to the product $\xi_{-n+i}\xi_{-n+i+j}$. Then, we compute the difference:

$$X_{-n} - \mathbb{E}_0 X_{-n} = \sum_i \sum_{\substack{j > n-i \\ j\neq 0}} b_i a_j \xi_{-n+i}\xi_{-n+i+j} + \sum_{i>n} \sum_{\substack{j \le n-i \\ j\neq 0}} b_i a_j \xi_{-n+i}\xi_{-n+i+j}.$$

The expectation of the square of the difference is equal to:

$$\mathbb{E}(X_{-n} - \mathbb{E}_0 X_{-n})^2 \quad = \quad \mathbb{E}(\xi_0^2)^2 \sum_i \sum_{\substack{j > n-i \\ j\neq 0}} b_i^2 a_j^2 + C,$$

where $C > 0$ depends on n, on the sequences $(a_i)_{i\in\mathbb{Z}}$, $(b_i)_{i\neq 0}$ and on $\|\xi_0\|_2$. Thus the assumptions of Theorem 1 do not hold with this standard choice of $\mathcal{M}_0$.

Other choices of $\mathcal{M}_0$ are possible.
- If we consider the σ-algebra $\mathcal{M}_0 = \sigma(\xi_i, i \ge 0)$, the same conclusion remains true, we have:

$$\mathbb{E}(X_{-n} - \mathbb{E}_0 X_{-n})^2 \ge \mathbb{E}(\xi_0^2)^2 \sum_i \sum_{\substack{j < n-i \\ j\neq 0}} b_i^2 a_j^2.$$

- Consider now the σ-algebra $\mathcal{M}_0 = \sigma(\xi_{2i}, i \in \mathbb{Z})$, then the quantities $\mathbb{E}_0(X_n^2)$ are equal for all $n \ge 0$.

Perhaps a better choice of $\mathcal{M}_0$ would allow to conclude but we did not find it out... Peligrad & Utev (2006) assume causal assumptions (Y_i writes as a function of the paste $g(\xi_j, j \le i)$), see [26]. This condition seems to be necessary to approximate X_n by increments of martingales, it is clearly not fulfilled here.

3.1.3 Non-Causal LARCH(∞) inputs

The previous approach extends for the case of Non-Causal LARCH(∞) inputs:

$$Y_t = \xi_t\left(a + \sum_{j\neq 0} a_j Y_{t-j}\right).$$

Doukhan, Teyssière and Winant (2005) prove the same results of existence as for the previous causal case (only replace summation for $j > 0$ by summation for $j \neq 0$) and the dependence is now of the η type with

$$\eta_r = \left(\|\xi_0\|_\infty \sum_{0 \le 2k < r} k\Lambda^{k-1} A\left(\frac{r}{2k}\right) + \frac{\Lambda^{r/2}}{1-\Lambda} \right) \mathbb{E}|\xi_0||a|$$

where now $A(x) = \sum_{|j| \ge x} |a_j|$, $\Lambda = \|\xi_0\|_\infty \sum_{j \ge 1} |a_j| < 1$. Notice that a very restrictive new assumption is that inputs need to be uniformly bounded in this non-causal case.

Under Riemanian decays ($a_r = \mathcal{O}(|r|^{-a})$ and $b_r = \mathcal{O}((|r|+1)^{-b})$ for some $a, b > 0$), our weak invariance principle, theorem 4, holds for $(X_t)_{t \in \mathbb{Z}}$ if $a > \dfrac{b(4m-6)}{(b-2)(m-2)} + 1$, $\|\xi_0\|_\infty < \infty$ and $b > 2$. Here again, this is possible to prove that assumptions of Theorem 1 do not hold.

Remark. We now mention that Doukhan, Teyssière and Winant (2005) provide a vector valued version of those models. Bilinear models $X_t = \zeta_t \left(\alpha + \sum_{j=1}^{\infty} \alpha_j X_{t-j} \right) + \beta + \sum_{j=1}^{\infty} \beta_j X_{t-j}$ where the variables are real valued and ζ_t is the input take this form. For this, we set $\xi_t = \begin{pmatrix} \zeta_t \\ 1 \end{pmatrix}$, $a = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ and $a_j = \begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix}$ (expansion (10) coincides with the chaotic expansion in [19]). For example, the classical GARCH(p, q) models,

$$\begin{cases} r_t = \sigma_t \varepsilon_t \\ \sigma_t^2 = \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + \gamma + \sum_{j=1}^{q} \gamma_j r_{t-j}^2 \end{cases}$$

where $\gamma > 0$, $\gamma_i \ge 0$, $\beta_i \ge 0$ (and the variables ε_t are centered at expectation) are bilinear. For this, Giraitis & Surgailis (2002) set in [19]: $\alpha_0 = \frac{\gamma_0}{1 - \sum \beta_i}$ et $\sum \alpha_i z^i = \frac{\sum \gamma_i z^i}{1 - \sum \beta_i z^i}$.

3.1.4 Associated inputs

The κ-weak dependence condition is known to hold for associated or Gaussian sequences. Recall that a process is associated if $\mathrm{Cov}(f(Y^{(n)}), g(Y^{(n)})) \ge 0$ for any coordinatewise non-decreasing function $f, g : \mathbb{R}^n \to \mathbb{R}$ such that the previous covariance makes sense with $Y^{(n)} = (Y_1, \ldots, Y_n)$. In both cases this condition will hold with

$$\kappa_r = \sup_{j \ge r} |\mathrm{Cov}(Y_0, Y_j)|$$

Notice the absolute values are needed only in the second case since for associated processes these covariances are nonnegative if they are finite. Independent sequences as well are associated and Pitt (1982) proves that a Gaussian process with nonnegative covariances is also associated. Finally, we recall that non-decreasing functions of associated sequences remain associated. This is a standard way to construct associated models from iid sequences (see e.g. Louhichi, 2001). Hence, it is clear that inputs as in equation (3) is an associated process.

Suppose that the inputs $(Y_t)_{t\in\mathbb{Z}}$ are such that $\kappa_r = \mathcal{O}(r^{-a})$ (for some $a > 0$). Then, under $b_r = \mathcal{O}((|r|+1)^{-b})$ (for some $b > 0$), the invariant principle of Newman & Wright (1981) follow from the remark of Louhichi & Soulier (2000) as soon as $\mathbb{E}Y^2 < +\infty$, $a > 1$ and $b > 1$. Those conditions are optimal, they correspond to $\sum_j \mathrm{Cov}(X_0, X_j) < \infty$. Such strong conditions are due to the fact that uncorrelation implies independence for associated processes. Our conditions from theorem 4 are much stronger: $\mathbb{E}Y^m < +\infty$ with $m > 2$, $b > 2$ and $a > \dfrac{(4m-6)b}{(m-2)(b-2)}$. For the special case of associated innovation given by equation (3), analogue difficulties for martingale approximation techniques occur as in subsection 3.1.2.

3.1.5 κ-weak dependent inputs

Our weak invariance principle theorem 4 remains however valid for the case of κ-weak dependent inputs $(Y_j)_{j\in\mathbb{Z}}$. Suppose that $\kappa_r = \mathcal{O}(r^{-a})$ (for some $a > 0$), then our conditions remains the same as for the associated case $\mathbb{E}Y^m < +\infty$ with $m > 2$, $b > 2$ and $a > \dfrac{(4m-6)b}{(m-2)(b-2)}$.

In the special case of κ-weak dependent inputs that are not associated as in (4), the optimal weak invariance principle of Newman & Wright (1981) is no valid anymore. Let us remark that in equation (4), the Gaussian process $(\xi'_t)_t$ is not necessarily an iid sequence.

3.2 Bernoulli shifts with dependent inputs

Now we omit the linearity condition of model (1). We still denote by $(Y_i)_i$ the weakly dependent input process. We restrict us in this section to the case $d = 1$. Let $H : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}$ be a measurable function and $X_n = H(Y_{n-i}, i \in \mathbb{Z})$. Such models are proved to exhibit either λ- or η-weak dependence properties. Because Bernoulli shifts of κ-weak dependent inputs are neither κ- nor η-weakly dependent, the κ case is here included in the λ one. In order to derive weak dependence properties of such processes, we assume that H satisfies the condition (8), which is a stronger assumption than for the case of independent inputs (see eqn. (11)). The following lemma proves both the existence and the weak dependence properties of such models:

Lemma 1 *Let $(Y_i)_i$ be a stationary $\lambda_{Y,r}$-weakly dependent process and $H : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}$ satisfies the condition (8) for some $l \geq 0$ and some sequence $b_j \geq 0$ such that $\sum_j |j| b_j < \infty$. Let assume that there exist a couple of real numbers (m, m') with $\mathbb{E}|Y_0|^{m'} < \infty$ such that $m > 2$ and $m' \geq (l+1)m$. Then,*

- *the process $X_n = H(Y_{n-i}, i \in \mathbb{Z})$ is well defined in $\mathbb{L}^m$: this is a strongly stationary process.*
- *if the input process $(Y_i)_{i\in\mathbb{Z}}$ is λ-weakly dependent (the weak dependence coefficients are denoted $\lambda_{Y,r}$), then X_n is λ-weakly dependent with*

$$\lambda_k = c \inf_{r \leq [k/2]} \left(\sum_{|j| \geq r} |j| b_j \right) \vee \left((2r+1)^2 \lambda_{Y,k-2r}^{\frac{m'-1-l}{m'-1}} \right).$$

- *if the input process $(Y_i)_{i\in\mathbb{Z}}$ is η-weakly dependent (the weak dependence coefficients are denoted $\eta_{Y,r}$) then X_n is η-weakly dependent and there exists a constant $c>0$ such that*
$$\eta_k = c \inf_{r\le[k/2]}\left(\sum_{|j|\ge r}|j|b_j\right)\vee\left((2r+1)^{1+\frac{l}{m'-1}}\eta_{Y,k-2r}^{\frac{m'-2}{m'-1}}\right).$$

Such models where already mentioned by Billingsley in [3] and Borovkova *et alii* in [4]. The proof of our results is deferred to section 5.3.

3.2.1 Explicit dependence rates

We now specify the decay rates from lemma 1. For standard decays of the previous sequences, this is easy to get the following explicit bounds. Here $\lambda>0$ and $\eta>0$ are constants which may differ in each case.

- If $b_j=\mathcal{O}\left((|j|+1)^{-b}\right)$ for some $b>2$ and $\lambda_{Y,j}=\mathcal{O}\left(j^{-\lambda}\right)$, resp. $\eta_{Y,j}=\mathcal{O}\left(j^{-\eta}\right)$ (as $|j|\uparrow\infty$) then from a simple calculation, we optimize both terms in order to prove that $\lambda_k=\mathcal{O}\left(k^{-\lambda\left(1-\frac{2}{b}\right)\frac{m'-1-l}{m'-1}}\right)$, resp. $\eta_k=\mathcal{O}\left(k^{-\eta\frac{(b-2)(m'-2)}{(b-1)(m'-1)-l}}\right)$.
  Note that in the case $m'=\infty$ this exponent may be arbitrarily close to λ for large values of $b>0$. This exponent may thus take all possible values between 0 and λ.
- If $b_j=\mathcal{O}\left(e^{-|j|b}\right)$ for some $b>0$ and $\lambda_{Y,j}=\mathcal{O}\left(e^{-j\lambda}\right)$, resp. $\eta_{Y,j}=\mathcal{O}\left(e^{-j\eta}\right)$ (as $|j|\uparrow\infty$) we have $\lambda_k=\mathcal{O}\left(k^2e^{-\lambda k\frac{b(m'-1-l)}{b(m'-1)+2\eta(m'-1-l)}}\right)$, resp. $\eta_k=\mathcal{O}\left(k^{\frac{m'-1-l}{m'-1}}e^{-\eta k\frac{b(m'-2)}{b(m'-1)+2\eta(m'-2)}}\right)$.
  The geometric decay of both $(b_j)_j$ and the weak dependence coefficients of the inputs ensures the geometric decay of the weak dependence coefficients of the Bernoulli shift.
- If we assume that the coefficients $(b_j)_j$ associated to the the Bernoulli shift have a geometric decay, say $b_j=\mathcal{O}\left(e^{-|j|b}\right)$ and that $\lambda_{Y,j}=\mathcal{O}\left(j^{-\lambda}\right)$ as $|j|\uparrow\infty$ (resp. $\eta_{Y,j}=\mathcal{O}\left(j^{-\eta}\right)$) we find $\lambda_k=\mathcal{O}\left((\log k)^2k^{-\lambda\frac{m'-1-l}{m'-1}}\right)$, resp. $\eta_k=\mathcal{O}\left((\log k)^{1+\frac{l}{m'-1}}k^{-\eta\frac{m'-2}{m'-1}}\right)$.
  If $m'=\infty$ this thus means that we only loose at most a factor $\log^2 k$ with respect to the dependence coefficients of the input dependent series $(Y_i)_i$.
- If we assume that the coefficients $(b_j)_j$ associated to the the Bernoulli shift have a Riemaniann decay, say $b_j=\mathcal{O}\left((|j|+1)^{-b}\right)$ and that $\lambda_{Y,j}=\mathcal{O}\left(e^{-j\lambda}\right)$ as $|j|\uparrow\infty$ (resp. $\eta_{Y,j}=\mathcal{O}\left(e^{-j\eta}\right)$) we find $\lambda_k=\mathcal{O}\left(k^{2-b}\right)$, resp. $\eta_k=\mathcal{O}\left(k^{2-b}\right)$.

After the simplest examples quoted in the previous section 3.1, we precise below some more examples of Bernoulli shifts with dependent inputs for which our results apply.

3.2.2 Volterra models with dependent inputs

Consider the function H defined by:
$$H(x)=\sum_{k=0}^{K}\sum_{j_1,\dots,j_k}a^{(k)}_{j_1,\dots,j_k}x_{j_1}\cdots x_{j_k},$$

then if x, y are as in eqn. (8):

$$H(x) - H(y) = \sum_{k=1}^{K} \sum_{u=1}^{k} \sum_{\substack{j_1, \ldots, j_{u-1} \\ j_{u+1}, \ldots, j_k}} a^{(k)}_{j_1, \ldots, j_{u-1}, s, j_{u+1}, \ldots, j_k} x_{j_1} \cdots x_{j_{u-1}} (x_s - y_s) x_{j_{u+1}} \cdots x_{j_k}.$$

From the triangular inequality we thus derive that suitable constants may be written as $l = K - 1$ and

$$b_s = \sum_{k=1}^{K} \sum\nolimits^{(k,s)} |a^{(k)}_{j_1, \ldots, j_k}|$$

where $\sum^{(k,s)}$ stands for the sums over all indices in $\mathbb{Z}^k$ where one of the indices $j_1, \ldots, j_k$ takes the value s and

$$L \equiv \sum_{k=0}^{K} \sum_{j_1, \ldots, j_k} \left| a^{(k)}_{j_1, \ldots, j_k} \right|.$$

E.g. $|a^{(k)}_{j_1, \ldots, j_k}| \leq C\,(j_1 \vee \cdots \vee j_k)^{-\alpha}$ or $\leq C \exp\left(-\alpha(j_1 \vee \cdots \vee j_k)\right)$ respectively yield $b_s \leq C' s^{d-1-\alpha}$ or $b_s \leq C' e^{-\alpha s}$ for some constant $C' > 0$.

3.2.3 Uniform Lipschitz Bernoulli shifts

Assume that condition (8) holds with $l = 0$, then the previous result still hold. An example of such a situation is the case of LARCH(∞) non-causal processes of § 3.1.3 with bounded ($m' = +\infty$) and dependent stationary inputs. In this cas, it is simple to check that there exists a constant $C >$ such that lemma 1 applies with $b_s \leq C|a_s|$ (in the vector valued LARCH models, $|a_s|$ is simply replaced by $\|a_s\|$ in this bound).

3.2.4 Markov stationary inputs

Markov stationary sequences satisfy a recurrence equation

$$Z_n = F(Z_{n-1}, \ldots, Z_{n-d}, \xi_n)$$

where the sequence (ξ_n) is iid. In this case $Y_n = (Z_n, \ldots, Z_{n-d+1})$ is a Markov chain such that $Y_n = M(Y_{n-1}, \xi_n)$ with

$$M(x_1, \ldots, x_d, \xi) = (F(x_1, \ldots, x_d, \xi), x_1, \ldots, x_{d-1}).$$

Duflo (1996) proves in Theorem 1.IV.24, [16], that there exists a stationary solution $(Z_n)_n$ of this equation in $\mathbb{L}^m$ for $m > 2$ as soon as $\|F(0, \xi)\|_m < \infty$ and there exist a norm $\|\cdot\|$ on $\mathbb{R}^d$ and a real $0 \leq a < 1$ such that $\|F(x, \xi) - F(y, \xi)\|_m \leq a\|x - y\|$. In this setting θ-dependence holds with $\theta_{Z,r} = \mathcal{O}(a^{r/d})$ (as $r \uparrow \infty$). We shall not detail more the significative examples provided in [11]. Indeed, we already mentioned that our results are sub-optimal in such causal cases; such dependent sequences may however also be used as inputs for a Bernoulli shift.

4 Moments inequalities

Our proof for central limit theorems is based on a truncation method. For a truncation level $T \geq 1$ we shall denote $\overline{X}_k = f_T(X_k) - \mathbb{E} f_T(X_k)$ with $f_T(X) = X \vee (-T) \wedge T$. Let us simply remark that $\overline{X}_k$ has moments of any orders because it is bounded. Furthermore, for any $a \leq m$, we control the moment $\mathbb{E}|f_T(X_0) - X_0|^a$ with Markov inequality:

$$\mathbb{E}|f_T(X_0) - X_0|^a \leq \mathbb{E}|X_0|^a \mathbb{1}_{\{|X_0| \geq T\}} \leq \mu T^{a-m},$$

thus using Jensen inequality yields

$$\|\overline{X}_0 - X_0\|_a \leq 2\mu^{\frac{1}{a}} T^{1-\frac{m}{a}}. \tag{12}$$

Deriving from this truncation, we are now able to control the limiting variance as well as the higher order moments.

4.1 Variances

Lemma 2 (Variances) *If one of the following conditions holds then the series σ^2 is convergent*

$$\sum_{k=0}^{\infty} \kappa_k \quad < \quad \infty \tag{13}$$

$$\sum_{k=0}^{\infty} \lambda_k^{\frac{m-2}{m-1}} \quad < \quad \infty \tag{14}$$

Proof. Using the fact that $\overline{X}_0 = g_T(X_0)$ is a function of X_0 with $\mathrm{Lip}\, g_T = 1$, $\|g_T\|_\infty \leq 2T$ we derive

$$|\mathrm{Cov}(\overline{X}_0, \overline{X}_k)| \leq \kappa_k \text{ or } 4T\lambda_k, \text{ respectively} \tag{15}$$

In the κ dependent case, the truncation may thus be omitted and

$$|\mathrm{Cov}(X_0, X_k)| \leq \kappa_k. \tag{16}$$

We shall only consider λ dependence, below. Now we develop

$$\mathrm{Cov}(X_0, X_k) = \mathrm{Cov}(\overline{X}_0, \overline{X}_k) + \mathrm{Cov}(X_0 - \overline{X}_0, X_k) + \mathrm{Cov}(\overline{X}_0, X_k - \overline{X}_k)$$

and using a truncation T to be determined we use the two previous bounds (12) and (15) with Hölder inequality with the exponents $\frac{1}{a} + \frac{4}{m} = 1$ to derive

$$\begin{aligned} |\mathrm{Cov}(\overline{X}_0, \overline{X}_k)| &\leq 4T\lambda_k + 2\|X_0\|_m \|\overline{X}_0 - X_0\|_a \\ &\leq 4T\lambda_k + 4\mu^{1/a+1/m} T^{1-m/a} \\ &\leq 4(T\lambda_k + \mu T^{2-m}). \end{aligned}$$

Note that we used the relation $1 - m/a = 2 - m$. Thus using the truncation such that $T^{1+\zeta} = \frac{\mu}{\lambda_k}$ yields the bound

$$|\mathrm{Cov}(X_0, X_k)| \leq 8\mu^{\frac{1}{\zeta+1}} \lambda_k^{\frac{\zeta}{\zeta+1}} = 8\mu^{\frac{1}{m-1}} \lambda_k^{\frac{m-2}{m-1}}. \tag{17}$$

4.2 A $(2+\delta)$-order moment bound

Lemma 3 *Assume that the stationay and centered process* $(X_i)_{i\in\mathbb{Z}}$ *satisfies* $\mathbb{E}|X_0|^{2+\zeta}<\infty$*, and it is either* κ*-weakly dependent with* $\kappa_r=\mathcal{O}\left(r^{-\kappa}\right)$ *or* λ*-weakly dependent with* $\lambda_r=\mathcal{O}\left(r^{-\lambda}\right)$*. Then if* $\kappa>2+\frac{1}{\zeta}$*, or* $\lambda>4+\frac{2}{\zeta}$*, then for all* $\delta\in]0,A\wedge B\wedge 1[$ *(where* A *and* B *are constants smaller than* ζ *and only depending of* ζ *and respectively* κ *or* λ*, see 21 and 22), there exist* $C>0$ *such that:*

$$\|S_n\|_\Delta\le C\sqrt{n},\qquad \textit{where}\qquad \Delta=2+\delta.$$

Remarks.

- The constant C satisfies $C>\left(\dfrac{5}{2^{\delta/2}-1}\right)^{1/\Delta}\sum_{k\in\mathbb{Z}}|\mathrm{Cov}(X_0,X_k)|$. Under the conditions of this lemma, the lemma 2 entails
$$c\equiv\sum_{k\in\mathbb{Z}}|\mathrm{Cov}(X_0,X_k)|<\infty.$$
- The result is sketched from Bulinski and Sashkin (2005); notice, however that their condition of dependence is of a causal nature while our is not which explains a loss with respect to the exponents λ and κ. In their $\kappa'-$weak dependence setting the best possible value of the exponent is 1 while this 2 for our non causal dependence.

Proof of lemma 3. As Bulinski and Sashkin (2005) who adapted the ideas in Ibragimov (1979) to weak dependence, we proceed by recurrence on k for $n\le 2^k$ to prove that:

$$\|1+|S_n|\|_\Delta\le C\sqrt{n}. \tag{18}$$

We thus assume (18) for all $n\le 2^{K-1}$. Setting $N=2^K$ we have to bound $\|1+|S_N|\|_\Delta$. We always can divide the sum S_N into three blocks, the two first block have the same size $n\le 2^{K-1}$ are denoted by A and B, and the third block V, located between A and B, has cardinality $q<n$. We then have $\|1+|S_N|\|_\Delta\le\|1+|A|+|B|\|_\Delta+\|V\|_\Delta$. The term $\|V\|_\Delta$ is directly bound with $\|1+|V|\|_\Delta\le C\sqrt{q}$ from the recurrence assumption. Writing $q=N^b$ with $b<1$, then this term is of order strictly smaller than $\sqrt{N}$. For $\|1+|A|+|B|\|_\Delta$, we have:

$$\begin{aligned}\mathbb{E}(1+|A|+|B|)^\Delta &\le \mathbb{E}(1+|A|+|B|)^2(1+|A|+|B|)^\delta,\\ &\le \mathbb{E}(1+2|A|+2|B|+(|A|+|B|)^2(1+|A|+|B|)^\delta.\end{aligned}$$

We expand the RHS of this expression; the following terms appear:

- $\mathbb{E}(1+|A|+|B|)^\delta\le 1+|A|_2^\delta+|B|_2^\delta\le 1+2c^\delta(\sqrt{n})^\delta$,
- $\mathbb{E}|A|(1+|A|+|B|)^\delta\le\mathbb{E}|A|((1+|B|)^\delta+|A|^\delta)\le\mathbb{E}|A|(1+|B|)^\delta+\mathbb{E}|A|^{1+\delta}$. The term $\mathbb{E}|A|^{1+\delta}$ is bounded with $\|A\|_2^{1+\delta}$ and then $c^{1+\delta}(\sqrt{n})^{1+\delta}$. The term $\mathbb{E}|A|(1+|B|)^\delta$ is bounded using Hölder $\|A\|_{1+\delta/2}\|1+|B|\|_\Delta^\delta$ and then is at least of order $cC^\delta(\sqrt{n})^{1+\delta}$ (a term analogue to this one exchanges the roles of B and A).

- $\mathbb{E}(|A|+|B|)^2(1+|A|+|B|)^\delta$. For this term, we use an inequality from Bulinski:

  $$\mathbb{E}(|A|+|B|)^2(1+|A|+|B|)^\delta \le \mathbb{E}|A|^\Delta + \mathbb{E}|B|^\Delta + 5(\mathbb{E}A^2(1+|B|)^\delta + \mathbb{E}B^2(1+|A|)^\delta).$$

  Now $\mathbb{E}|A|^\Delta \le C^\Delta(\sqrt{n})^\Delta$ is bounded by using (18). The second term is its analogue with B in place of A. The third term has to be handled with a particular care, as follows.

We now control $\mathbb{E}A^2(1+|B|)^\delta$ and its counterpart relative to B. For this, we use the weak dependence. We then have to truncate the variables. Denote by $\overline{X}$ the variable $X \vee T \wedge (-T)$ for a real $T > 0$ to be determined later. We then note by extension $\overline{A}$ and $\overline{B}$ the sums of the truncated variables $\overline{X}_i$. Remarking that $|B| - |\overline{B}| \ge 0$, we have:

$$\mathbb{E}|A|^2(1+|B|)^\delta \le \mathbb{E}A^2(|B|-|\overline{B}|)^\delta + \mathbb{E}(A^2-\overline{A}^2)(1+|\overline{B}|)^\delta + \mathbb{E}\overline{A}^2(1+|\overline{B}|)^\delta.$$

We begin with a control of $\mathbb{E}A^2(|B|-|\overline{B}|)^\delta$. Set $m = 2+\zeta$, then using Hölder inequality with $2/m + 1/m' = 1$ yields:

$$\mathbb{E}A^2(|B|-|\overline{B}|)^\delta \le \|A\|_m^2 \|(|B| - \|\overline{B}|)^\delta\|_{m'}$$

$\|A\|_\Delta$ is bounded using (18) and we remark that:

$$(|B|-|\overline{B}|)^{\delta m'} \le (|B| - |B|1\!\!1_{\{\forall i, |X_i| \le T\}})^{\delta m'} \le |B|^{\delta m'} 1\!\!1_{\{\exists i, |X_i| > T\}} \le |B|^{\delta m'} 1\!\!1_{|B|>T}.$$

We then bound $1\!\!1_{|B|>T} \le (|B|/T)^\alpha$ with $\alpha = m - \delta m'$, hence:

$$\mathbb{E}||B|-|\overline{B}||^{\delta m'} \le E|B|^m T^{\delta m' - m}.$$

By convexity and stationarity, we have $\mathbb{E}|B|^m \le n^m \mathbb{E}|X_0|^m$, so that:

$$\mathbb{E}A^2(|B|-|\overline{B}|)^\delta \preceq n^{2+m/m'} T^{\delta - m/m'}.$$

Finally, remarking that $m/m' = m-2$, we obtain:

$$\mathbb{E}A^2(|B|-|\overline{B}|)^\delta \preceq n^m T^{\Delta - m}.$$

We obtain the same bound for the second term:

$$\mathbb{E}(A^2-\overline{A}^2)(1+|\overline{B}|)^\delta \preceq n^m T^{\Delta-m}.$$

For the third term, we introduce a covariance term:

$$\mathbb{E}\overline{A}^2(1+|\overline{B}|)^\delta \le \mathrm{Cov}(\overline{A}^2, (1+|\overline{B}|)^\delta) + \mathbb{E}\overline{A}^2\mathbb{E}(1+|\overline{B}|)^\delta.$$

The last term is bounded with $|A|_2^2|B|_2^\delta \le c^\Delta \sqrt{n}^\Delta$. The covariance is controlled with weak-dependence:

- in the κ-dependent case: $n^2 T \kappa_q$,
- in the λ-dependent case: $n^3 T^2 \lambda_q$.

We then choose either the truncation $T^{m-\delta-1} = n^{m-2}/\kappa_q$ or $T^{m-\delta} = n^{m-3}/\lambda_q$. Now the three terms of the decomposition have the same order:

$$\begin{aligned}
\mathbb{E}|A|^2(1+|B|)^\delta &\preceq \left(n^{3m-2\Delta}\kappa_q^{m-\Delta}\right)^{1/(m-\delta-1)}, \text{ under } \kappa\text{-dependence,}\\
\mathbb{E}|A|^2(1+|B|)^\delta &\preceq \left(n^{5m-3\Delta}\lambda_q^{m-\Delta}\right)^{1/(m-\delta)}, \text{ under } \lambda\text{-dependence.}
\end{aligned}$$

Set $q = N^b$, we note that $n \le N/2$ and this term has order $N^{\frac{3m-2\Delta+b\kappa(\Delta-m)}{m-\delta-1}}$ under κ-weak dependence and the order $N^{\frac{5m-3\Delta+b\lambda(\Delta-m)}{m-\delta}}$ under λ-weak dependence. Those terms are thus negligible with respect to $N^{\Delta/2}$ if:

$$\kappa > \frac{3m-2\Delta-\Delta/2(m-\delta-1)}{b(m-\Delta)}, \text{ under } \kappa\text{-dependence,} \tag{19}$$

$$\lambda > \frac{5m-3\Delta-\Delta/2(m-\delta)}{b(m-\Delta)}, \text{ under } \lambda\text{-dependence.} \tag{20}$$

Finally, using this assumption, $b < 1$ and $n \le N/2$ we derive the bound for some suitable constants $a_1, a_2 > 0$:

$$\mathbb{E}(1+|S_N|)^\Delta \le \left(2^{-\delta/2}C^\Delta + 5\cdot 2^{-\delta/2}c^\Delta + a_1 N^{-a_2}\right)\left(\sqrt{N}\right)^\Delta.$$

Using the relation between C and c, we conclude that (18) is also true at the step N if the constant C satisfies $2^{-\delta/2}C^\Delta + 5\cdot 2^{-\delta/2}c^\Delta + a_1 N^{-a_2} \le C^\Delta$. Choose $C > \left(\frac{5c^\Delta + a_1 2^{\delta/2}}{2^{\delta/2}-1}\right)^{1/\Delta}$ with $c = \sum_{k\in\mathbb{Z}} |\mathrm{Cov}(X_0, X_k)|$, then the previous relation holds. Finally, we use eqns. (19) and (20) to find a condition on δ.
In the case of κ-weak dependence, we rewrite inequality (19) as:

$$0 > \delta^2 + \delta(2\kappa - 3 - \zeta) - \kappa\zeta + 2\zeta + 1.$$

It leads to the following condition on λ:

$$\delta < \frac{\sqrt{(2\kappa-3-\zeta)^2 + 4(\kappa\zeta - 2\zeta - 1)} + \zeta + 3 - 2\kappa}{2} = A. \tag{21}$$

We do the same in the case of the λ-weak dependence:

$$\delta < \frac{\sqrt{(2\lambda-6-\zeta)^2 + 4(\lambda\zeta - 4\zeta - 2)} + \zeta + 6 - 2\lambda}{2} = B. \quad \square \tag{22}$$

Remark. The bounds A and B are always smaller than ζ.

5 Proof of the main results

In this section we first prove the central limit corresponding to the convergence $W_n(1) \to W(1)$ in the theorems 2 and 3, then we shall provide rates in those central limit results. The weak invariance principle is obtained in a standard way from such central limit theorems and tightness which follows from lemma 1, by using the classical Kolmogorov-Centsov tightness criterion, see Billingsley (1968). In the last subsection, we prove the lemma 3 which states the properties of our (new) Bernoulli shifts with dependent inputs.

5.1 Proof of Theorems 2 & 3

Set $S = \frac{1}{\sqrt{n}} S_n$ and consider $p = p(n)$ and $q = q(n)$ in such a way that

$$\lim_{n\to\infty} \frac{1}{q(n)} = \lim_{n\to\infty} \frac{q(n)}{p(n)} = \lim_{n\to\infty} \frac{p(n)}{n} = 0$$

and $k = k(n) = \left[\frac{n}{p(n)+q(n)}\right]$

$$Z = \frac{1}{\sqrt{n}} \left(U_1 + \cdots + U_k\right), \qquad \text{with } U_j = \sum_{i\in B_j} X_i$$

where $B_j =](p+q)(j-1), (p+q)(j-1)+p] \cap \mathbb{N}$ is a subset of p successive integers from $\{1, \ldots, n\}$ such that, for $j \neq j'$, B_j and $B_{j'}$ are at least distant of $q = q(n)$. We note B'_j the block between B_j and B_{j+1} and $V_j = \sum_{i\in B_j} X_i$. V_k is the last block of X_i between the end of B_k and n. Furthermore, set $\sigma_p^2 = \mathrm{Var}\,(U_1)/p$, we let

$$Y = \frac{V'_1 + \cdots + V'_k}{\sqrt{n}}, \qquad V'_j \sim \mathcal{N}(0, p\sigma_p^2)$$

where the Gaussian variables V_j are independent and independent of the sequence $(X_n)_{n\in\mathbb{Z}}$. We also consider a sequence $U_1^*, \ldots, U_k^*$ of independent random variables with the same distribution as U_1 and we set $Z^* = \left(U_1^* + \cdots + U_k^*\right)/\sqrt{n}$. We fix $t \in \mathbb{R}^d$ and we define $f : \mathbb{R}^d \to \mathbb{C}$ by $f(x) = \exp\{it \cdot x\}$. Then:

$$\mathbb{E}f(S) - f(\sigma N) = \mathbb{E}f(S) - f(Z) + \mathbb{E}f(Z) - f(Z^*) + f(Z^*) - f(Y) + \mathbb{E}f(Y) - f(\sigma N)$$

Lindeberg method is devoted to prove that this expression converges to 0 as $n \to \infty$. The first and the last term in this inequality are referred to as the auxiliary terms in this Bernstein-Lindeberg method. They come from the replacement of the individual initial - non-Gaussian and Gaussian respectively - random variables by their block counterparts. The second term is analogue to that obtained with decoupling and turns the proof of the central limit theorem to the independent case. The third term is referred to as the main term and following the proof under independence it will be bounded above by using a Taylor expansion. Because of the dependence structure, in the corresponding bounds, some additional covariance terms will appear.

The following subsections are organized as follows: we first consider the auxiliary terms and the main terms are then decomposed by the usual Lindeberg method and the corresponding terms coming from the dependence or the usual remainder terms (standard for the independent case) are considered in separated subsections. In a last subsection we recollect those calculations to derive the central limit theorem.

5.1.1 Auxiliary terms

Using Taylor expansions up to the second order, we obtain:

$$|\mathbb{E}f(S) - f(Z)| \le \frac{\|f''\|_\infty^2}{2} \mathbb{E}|S - Z|^2 \text{ and } |\mathbb{E}f(Y) - f(\sigma N)| \le \frac{\|f''\|_\infty^2}{2} \mathbb{E}|Y - \sigma N|^2.$$

We note that $\mathbb{E}|Z-S|^2 \preceq \frac{\mathbb{E}(V_1+\ldots+V_k)^2}{n}$. Quoting that $V_1+\ldots+V_k$ is a sum of X_is with cardinality $\le (k+1)q+p$, then (17) and (16), under conditions (14) or (13), respectively entail: $\mathbb{E}|Z-S|^2 \preceq \frac{(k+1)q+p}{n}$. Now $Y \sim \sqrt{\frac{kp}{n}}\sigma_p N$, thus:

$$\mathbb{E}|Y-\sigma N|^2 \quad \le \quad \left|\frac{kp}{n}-1\right|\sigma_p^2 + \left|\sigma_p^2-\sigma^2\right|$$

Remarking that $kp/n-1 \preceq q/p$ we need to bound

$$|\sigma_p^2-\sigma^2| \quad \le \quad \sum_{|i|<p}\frac{|i|}{p}|\mathbb{E}X_0X_i| + \sum_{|i|>p}|\mathbb{E}X_0X_i|$$

Set $a_i = |\mathbb{E}X_0X_i|$, under conditions (14) or (13) (respectively), the series $\sum_{i=0}^{\infty} a_i$ converge thus $s_j = \sum_{i=j}^{\infty} a_i \to_{j\to\infty} 0$ and

$$|\sigma_p^2-\sigma^2| \le 2\sum_{i=0}^{p-1}\frac{i}{p}\cdot a_i + 2s_p \le \frac{2}{p}\sum_{i=0}^{p-1}s_i + 2s_p,$$

Cesaro lemma entails that the first term converges to 0. Hence $|\mathbb{E}f(S)-f(Z)|+|\mathbb{E}f(Y)-f(\sigma N)|$ tends to 0 as $n\uparrow\infty$.
To precise a convergence rate, we assume that $a_i = \mathcal{O}(i^{-\alpha})$ for some $\alpha>1$; then

$$|\sigma_p^2-\sigma^2| \preceq p^{1-\alpha}.$$

The convergence rate is thus given by $\frac{q}{p}+\frac{p}{n}+p^{1-\alpha}$ if $\mathbb{E}X_0X_i = \mathcal{O}(i^{-\alpha})$. Remarking that $\mathbb{E}X_0X_i = \mathrm{Cov}(X_0,X_i)$, we then use equations (16) and (17) and we find $\alpha=\kappa$ or $\alpha = \lambda(m-2)/(m-1)$ depending of the weak-dependence setting.

With $p=n^a$, $q=n^b$, those bounds become:

$n^{b-a}+n^{a-1}+n^{a(1-\kappa)}$, in the κ-weak dependence setting,

$n^{b-a}+n^{a-1}+n^{a(1-\lambda(m-2)/(m-1))}$, under λ-weak dependence.

5.1.2 Main terms

It remains to control the second and the third terms of the sum. They are bounded as usual by:

$$|\mathbb{E}f(Z)-f(Z^*)| \le \sum_{j=1}^{k}|\mathbb{E}\Delta_j|, \quad |\mathbb{E}f(Z^*)-f(Y)| \le \sum_{j=1}^{k}|\mathbb{E}\Delta_j'|,$$

where $\Delta_j = f(W_j+x_j)-f(W_j+x_j^*)$, for $j=1,\ldots,k$ with $x_j = \frac{1}{\sqrt{n}}U_j$, $x_j^* = \frac{1}{\sqrt{n}}U_j^*$, $W_j = w_j+\sum_{i>j}x_i^*$, $w_j = \sum_{i<j}x_i$ and $\Delta_j' = f(W_j'+x_j^*)-f(W_j'+x_j')$, for $j=1,\ldots,k$ with $x_j' = \frac{1}{\sqrt{n}}V_j'$, $W_j' = \sum_{i<j}x_i^*+\sum_{i>j}x_i'$.
Now we use the special form of f and the independence properties of the variables U_i^* and V_i' to write:

$$\begin{aligned}\mathbb{E}\Delta_j &= \left(\mathbb{E}f(w_j)f(x_j)-\mathbb{E}f(w_j)\mathbb{E}f(x_j^*)\right)\mathbb{E}f\left(\sum_{i>j}x_i^*\right),\\ \mathbb{E}\Delta_j' &= \left(\mathbb{E}f(x_j^*)-\mathbb{E}f(x_j')\right)\mathbb{E}f\left(W_j'\right).\end{aligned}$$

We then control the terms $\mathbb{E}f\left(\sum_{i>j} x_i^*\right)$ and $\mathbb{E}f\left(W_j'\right)$ by the fact that $\|f\|_\infty \le 1$ and we use the coupling to introduce a covariance term:

$$\begin{aligned}
|\mathbb{E}\Delta_j| &\le \left|\mathrm{Cov}\left(f\left(\sum_{i<j} x_i\right), f(x_j)\right)\right|, \\
|\mathbb{E}\Delta_j'| &= \left|\mathbb{E}f(x_j^*) - \mathbb{E}f(x_j')\right|.
\end{aligned}$$

- For Δ_j, we use weak dependence to control it.
  Write for this $|\mathrm{Cov}(F(X_m, m \in B_i, i<j), G(X_m, m \in B_j))|$, with $\|F\|_\infty \le 1$ and the control of $\mathrm{Lip}\,F$ for $F(z_1,\dots,z_{kp}) = f\left(\frac{1}{\sqrt{n}}\sum_{i<j} u_i\right)$ where $u_i = \sum_{\ell\in B_i} z_\ell$ and:

$$\begin{aligned}
\left|f\left(\frac{1}{\sqrt{n}}\sum_{i<j}\sum_{\ell\in B_i} z_\ell\right) - f\left(\frac{1}{\sqrt{n}}\sum_{i<j}\sum_{\ell\in B_i} z_\ell'\right)\right| &\le \left|1 - \exp it\cdot\left(\frac{1}{\sqrt{n}}\sum_{i<j}\sum_{\ell\in B_i}(z_\ell - z_\ell')\right)\right| \\
&\le \frac{\|t\|_2}{\sqrt{n}}\sum_{\ell=1}^{kp}\|z_\ell - z_\ell'\|_2.
\end{aligned}$$

  For $G(z_1,\dots,z_p) = f\left(\sum_{i=1}^{p} z_i/\sqrt{n}\right)$, we have $\|G\|_\infty = 1$ and $\mathrm{Lip}\,G \preceq 1/\sqrt{n}$. We then distinguish the two cases, remarking the gap between the left and the right terms in the covariance is at least q:

  - In the κ dependent setting: $|\mathbb{E}\Delta_j| \preceq kp\cdot p\cdot\frac{1}{\sqrt{n}}\cdot\frac{1}{\sqrt{n}}\cdot\kappa_q$.
  - Under λ dependence: $|\mathbb{E}\Delta_j| \preceq \left(kp\cdot p\cdot\frac{1}{\sqrt{n}}\cdot\frac{1}{\sqrt{n}} + kp\cdot\frac{1}{\sqrt{n}} + p\cdot\frac{1}{\sqrt{n}}\right)\cdot\lambda_q$.

  Remark that those bounds do not depend on j, the contribution of the second term is:

$$\begin{aligned}
|\mathbb{E}f(Z) - f(Z^*)| &\preceq kp\cdot\kappa_q, && \text{under } \kappa, \\
&\preceq kp(1+\sqrt{k/p})\cdot\lambda_q, && \text{under } \lambda.
\end{aligned}$$

  Reminding that $p = n^a$, $q = n^b$, $\kappa_r = \mathcal{O}\left(r^{-\kappa}\right)$ or $\lambda_r = \mathcal{O}\left(r^{-\lambda}\right)$, those bounds become $n^{1-\kappa b}$ or $n^{1+(1/2-a)_+-\lambda b}$ in respectively the κ or the λ dependence context.

- For Δ_j', writing Taylor formula up to order 2 or 3 respectively:

$$\begin{aligned}
|f(x_j^*) - f(x_j')| &\le |x_j^* - x_j'|\|f'\|_\infty + \frac{1}{2}(x_j^* - x_j')^2\|f''\|_\infty + r_j \\
r_j &\le \frac{1}{2}\|f''\|_\infty(x_j^* - x_j')^2, \text{ or} \\
&\le \frac{1}{6}\|f'''\|_\infty|x_j^* - x_j'|^3,
\end{aligned}$$

We then have for an arbitrary $\delta \in [0,1]$:

$$\begin{aligned}\mathbb{E}r_j &\preceq \mathbb{E}((|x_j^*|^2 + |x_j'|^2) \wedge (|x_j^*|^3 + |x_j'|^3)) \\ &\preceq \mathbb{E}(|x_j^*|^2 \wedge |x_j^*|^3) + \mathbb{E}(|x_j'|^2 \wedge |x_j'|^3) \\ &\preceq \mathbb{E}|x_j^*|^{2+\delta} + \mathbb{E}|x_j'|^{2+\delta}.\end{aligned}$$

Then, using the stationarity of the sequence X_n we obtain:

$$|\mathbb{E}\Delta_j'| \quad \preceq \quad n^{-1-\frac{\delta}{2}} \left(\mathbb{E}|S_p|^{2+\delta} \vee p^{1+\frac{\delta}{2}} \right).$$

We then use lemma 3 to bound the moment $\mathbb{E}|S_p|^{2+\delta}$. If $\kappa > 2 + \frac{1}{\zeta}$, or $\lambda > 4 + \frac{2}{\zeta}$, where $\kappa_r = \mathcal{O}\left(r^{-\kappa}\right)$ or $\lambda_r = \mathcal{O}\left(r^{-\lambda}\right)$ then there exists $\delta \in]0, \zeta \wedge 1[$ and $C > 0$ such that:

$$\mathbb{E}|S_p|^{2+\delta} \leq C p^{1+\delta/2}.$$

We then obtain:

$$|\mathbb{E}f(Z^*) - f(Y)| \preceq k(p/n)^{1+\delta/2}.$$

Reminding that $p = n^a$, this bound has order $n^{(a-1)\delta/2}$ in both the κ or the λ-weak dependence setting.

We now collect the previous bounds to conclude that a multidimensional CLT holds under both assumptions of theorems 2 and 3. Tightness follows from Kolmogorov-Chentsov criterion (see [3]) and lemma 3; thus theorems 2 and 3 both follow from a repeated use of the previous CLT. □

5.2 Rates of convergence

We now present two propositions of independent interest. Rates of convergence in the clt are explicited in both the Dudley metric as well as a Berry-Essen bound.

Proposition 1 *Assume that the weakly dependent stationary process $(X_n)_n$ satisfies (5) then the difference between the characteristic functions is bounded by ($C > 0$ is a constant):*

$$\left|\mathbb{E}\left(f(S_n/\sqrt{n}) - f(\sigma N)\right)\right| \leq C n^{-c^*},$$

where c^* depends of the weak dependent coefficients:

- *under λ-dependence, if $\lambda_r = \mathcal{O}(r^{-\lambda})$ for $\lambda > 4 + \frac{2}{\zeta}$, then* $c^* = \dfrac{A}{2}\dfrac{2\lambda - 1}{(2+A)(\lambda+1)}$ *where*
$$A = \frac{\sqrt{(2\lambda - 6 - \zeta)^2 + 4(\lambda\zeta - 4\zeta - 2)} + \zeta + 6 - 2\lambda}{2} \wedge 1,$$
- *under κ-dependence, if $\kappa_r = \mathcal{O}(r^{-\kappa})$ for $\kappa > 2 + \frac{1}{\zeta}$, then* $c^* = \dfrac{(\kappa - 1)B}{\kappa(2+B)}$ *where*
$$B = \frac{\sqrt{(2\kappa - 3 - \zeta)^2 + 4(\kappa\zeta - 2\zeta - 1)} + \zeta + 3 - 2\kappa}{2} \wedge 1.$$

We use Theorem 5.1 of Petrov (1995) to obtain:

Proposition 2 (A rate in the Berry Essen bounds) *Assume that the real weakly dependent stationary process $(X_n)_n$ satisfies the assumptions from Proposition 1. We obtain:*

$$\sup_x |F_n(x) - \Phi(x)| = \mathcal{O}\left(n^{-c^*/3}\right).$$

where c^* is defined in Proposition 2.

Proof of proposition 1. In the previous section, the different terms have already be bounded as follows:

- In the λ-dependent case, we only need to consider the three largest exponents of n which appear in those rates: $(a-1)\delta/2$, $1 + (1/2 - a)_+ - \lambda b$ and $b - a$. The previous optimal choice of a^* is smaller than 1/2, then we have to consider the rate $3/2 - a - \lambda b$ and not $1 - \lambda b$. Thus:

  $$a^* = \frac{(1+\lambda)\delta + 3}{(2+\delta)(\lambda+1)} \in \left]0, \frac{1}{2}\right[, \qquad b^* = a^* \frac{3}{2(\lambda+1)} \in]0, a^*[$$

  Finally, we obtain the rate n^{-c^*}.

- In the κ-dependence case, the exponents of n write as:
  - in auxiliary terms: $b - a$, $a - 1$ and $a(1-\kappa)$,
  - in main terms: $1 - \kappa b$ and $(a-1)\delta/2$.

  The idea is to choose carefully a^* and $b^* \in]0, 1[$ such that the main rates are equal. Because $\delta < 1$, $a > b$, we directly see that $(a-1)\delta/2 > a - 1$ and $1 - \kappa b > a(1-\kappa)$, so that the only rate of the auxiliary term it remains to consider is $b - a$. Finally, we obtain

  $$a^* = 1 - \frac{2\kappa - 2}{(2+\delta)\kappa + \delta} \in]0, 1[, \quad b^* = a^* \frac{2 + 2\delta}{2 + \delta + \delta\kappa} \in]0, a^*[.$$

  We conclude with standard calculations. □

Proof of proposition 2. For a fixed t, we control the $\mathbb{L}^1$ distance between the characteristic functions of S and σN by a term proportional to $t^2 n^{-c^*}$. Here the factor t^2 appears from the relation $\|f^{(j)}\|_\infty \leq |t|^j$. Let Φ be the distribution function of σN and F_n the one of S. Theorem 5.1 p. 142 in Petrov (1995) gives, for every $T > 0$:

$$\sup_x |F_n(x) - \Phi(x)| \preceq n^{-c^*} T^2 + 1/T.$$

We optimize T to get the proposed rate of convergence in the central limit theorem. □

5.3 Proof of lemma 1

The proof of the properties of Bernoulli shifts with dependent inputs is divided in two sections devoted respectively to the definition of such models and to their weak dependence properties.

5.3.1 Existence

We first prove the existence in $\mathbb{L}^1$ of a Bernoulli shift with dependent inputs. Here we set $Y^{(s)} = (Y_{-i}1\!\!1_{|i|<s})_{i\in\mathbb{Z}}$ and $Y_+^{(s)} = (Y_{-i}1\!\!1_{-s<i\le s})_{i\in\mathbb{Z}}$ for $i \in \mathbb{Z}\cup\{\infty\}$. In order to prove the existence of Bernoulli shift with dependent inputs, we show that $H(Y^{(\infty)})$ is the sums of a normally convergent series in $\mathbb{L}^1$. Then formally

$$\begin{aligned} X_0 = H(Y^{(\infty)}) &= H(0) + (H(Y^{(1)}) - H(0)) \\ &\quad + \sum_{s=1}^{\infty}\Big((H(Y^{(s+1)}) - H(Y_+^{(s)})) + (H(Y_+^{(s)}) - H(Y^{(s)}))\Big) \end{aligned}$$

From (8) we obtain

$$\begin{aligned} |H(Y^{(1)}) - H(0)| &\le b_0|Y_0| \\ |H(Y^{(s+1)}) - H(Y_+^{(s)})| &\le b_{-s}(\|Y_+^{(s)}\|_\infty^l \vee 1)|Y_{-s}| \\ |H(Y_+^{(s)}) - H(Y^{(s)}| &\le b_s(\|Y^{(s)}\|_\infty^l \vee 1)|Y_s| \end{aligned}$$

Using Hölder inequality yields

$$\begin{aligned} \mathbb{E}\left|H(Y^{(1)}) - H(0)\right| + \sum_{s=1}^{\infty}\mathbb{E}\left|H(Y^{(s+1)}) - H(Y_+^{(s)})\right| + \mathbb{E}\left|H(Y_+^{(s)}) - H(Y^{(s)})\right| \\ \le \sum_{i\in\mathbb{Z}} 2|i|b_i(\|Y_0\|_1 + \|Y_0\|_{l+1}^{l+1}) \quad (23) \end{aligned}$$

Hence assumptions $l+1 \le m'$ and $\sum_{i\in\mathbb{Z}}|i|b_i < \infty$ together imply that the variable $H(Y)$ is well defined. The same way proves that the process $X_n = H(Y_{n-i}, i \in \mathbb{Z})$ is a well defined process in $\mathbb{L}^1$ and it is strongly stationary. The proof extends in $\mathbb{L}^m$ if $m \ge 1$ is such that $(l+1)m \le m'$.

5.3.2 Weak dependence properties

Here, we exhibit some Lipschitz function and we then truncate. We write $\overline{Y} = Y\vee(-T)\wedge T$ for a truncation T set below. Denote $X_n^{(r)} = H(Y^{(r)})$ and $\overline{X}_n^{(r)} = H(\overline{Y}^{(r)})$. Furthermore, for any $k \ge 0$ and any $(u+v)$-tuples such that $s_1 < \cdots < s_u \le s_u + k \le t_1 < \cdots < t_v$, we set $X_{\mathbf{s}} = (X_{s_1}, \ldots, X_{s_u})$, $X_{\mathbf{t}} = (X_{t_1}, \ldots, X_{t_v})$ and $\overline{X}_{\mathbf{s}}^{(r)} = (\overline{X}_{s_1}^{(r)}, \ldots, \overline{X}_{s_u}^{(r)})$, $\overline{X}_{\mathbf{t}}^{(r)} = (\overline{X}_{t_1}^{(r)}, \ldots, \overline{X}_{t_v}^{(r)})$. Then we have for all f, g satisfying $\|f\|_\infty, \|g\|_\infty \le 1$ and $\mathrm{Lip}\, f + \mathrm{Lip}\, g < \infty$:

$$\begin{aligned} |\mathrm{Cov}(f(X_{\mathbf{s}}), g(X_{\mathbf{t}}))| &\le |\mathrm{Cov}(f(X_{\mathbf{s}}) - f(\overline{X}_{\mathbf{s}}^{(r)}), g(X_{\mathbf{t}}))| && (24) \\ &+ |\mathrm{Cov}(f(\overline{X}_{\mathbf{s}}^{(r)}), g(X_{\mathbf{t}}) - g(\overline{X}_{\mathbf{t}}^{(r)}))| && (25) \\ &+ |\mathrm{Cov}(f(\overline{X}_{\mathbf{s}}^{(r)}), g(\overline{X}_{\mathbf{t}}^{(r)}))|. && (26) \end{aligned}$$

Using $\|g\|_\infty \le 1$, the term (24) is bounded by:

$$2\mathrm{Lip}\, f \cdot \mathbb{E}\left|\sum_{i=1}^{u}(X_{s_i} - \overline{X}_{s_i}^{(r)})\right| \le 2u\mathrm{Lip}\, f\left(\max_{1\le i\le u}\mathbb{E}\left|X_{s_i} - X_{s_i}^{(r)}\right| + \max_{1\le i\le u}\mathbb{E}\left|X_{s_i}^{(r)} - \overline{X}_{s_i}^{(r)}\right|\right).$$

With the same arguments that for the proof of the existence of $H(Y^{(\infty)})$ (see equation (23)), the first term in the right sight is bounded by $\sum_{i\geq s} 2|i|b_i(\|Y_0\|_1 + \|Y_0\|_{l+1}^{l+1})$. Notice now that if x, y are sequences with $x_i = y_i = 0$ if $|i| \geq r$ then a repeated application of the previous inequality (8) yields

$$|H(x) - H(y)| \leq L(\|x\|_\infty^l \vee \|y\|_\infty^l \vee 1)\|x - y\| \tag{27}$$

where $L = \sum_{i\in\mathbb{Z}} b_i < \infty$ because $\sum_{i\in\mathbb{Z}} |i|b_i < \infty$. Then the second term is bounded by using equation (27):

$$\begin{aligned}
\mathbb{E}\left|X_{s_i}^{(r)} - \overline{X}_{s_i}^{(r)}\right| &= \mathbb{E}\left|H\left(Y^{(r)}\right) - H\left(\overline{Y}^{(r)}\right)\right| \\
&\leq L\mathbb{E}\left(\left(\max_{-r\leq i\leq r} |Y_i|\right)^l \sum_{-r\leq j\leq r} \{|Y_j|\mathbb{1}_{Y_j\geq T}\}\right) \\
&\leq L(2r+1)^2\mathbb{E}\left(\max_{-r\leq i,j\leq r} |Y_i|^l\{|Y_j|\mathbb{1}_{|Y_j|\geq T}\}\right) \\
&\leq L(2r+1)^2\|Y_0\|_{m'}^{m'}T^{l+1-m'}
\end{aligned}$$

The second term (25) of the sum is analogously bounded. The last term (26) writes

$$\left|\mathrm{Cov}(\overline{F}^{(r)}(Y_{s_i+j}, 1 \leq i \leq u, |j| \leq r), \overline{G}^{(r)}(Y_{t_i+j}, 1 \leq i \leq v, |j| \leq r)\right|,$$

where $\overline{F}^{(r)} : \mathbb{R}^{u(2r+1)} \to \mathbb{R}$ and $\overline{G}^{(r)} : \mathbb{R}^{u(2r+1)} \to \mathbb{R}$. Under the assumption $r \leq [k/2]$, we use the $\epsilon = \eta$ or λ-weak dependence of Y in order to bound this covariance term by $\psi(\mathrm{Lip}\,\overline{F}^{(r)}, \mathrm{Lip}\,\overline{G}^{(r)}, u(2r+1), v(2r+1))\epsilon_{k-2r}$, with respectively $\psi(u,v,a,b) = uvab$ or $\psi(u,v,a,b) = uvab + ua + vb$.

$$\mathrm{Lip}\,\overline{F}^{(r)} = \sup \frac{|f(H(\overline{x}_{s_i+l}, 1 \leq i \leq u, |l| \leq r) - f(H(\overline{y}_{s_i+l}, 1 \leq i \leq u, |l| \leq r)|}{\sum_{j=1}^u \|x_j - y_j\|},$$

where the sup extends to $(x_1, \ldots, x_u) \neq (y_1, \ldots, y_u)$ where $x_i, y_i \in \mathbb{R}^{2r+1}$. Using (27):

$$\begin{aligned}
|\overline{F}^{(r)}(x) - \overline{F}^{(r)}(y)| &\leq \mathrm{Lip}\,f L \sum_{i=1}^u \left(\|\overline{x}_{s_i}\|_\infty \vee \|\overline{y}_{s_i}\|_\infty \vee 1\right)^l \|\overline{x}_{s_i} - \overline{y}_{s_i}\| \\
&\leq \mathrm{Lip}\,f LT^l \sum_{i=1}^u \sum_{-r\leq l\leq r} |x_{s_i+l} - y_{s_i+l}|.
\end{aligned}$$

We thus obtain $\mathrm{Lip}\,F^{(r)} \leq \mathrm{Lip}\,f \cdot L \cdot T^l$. Similarly $\mathrm{Lip}\,G^{(r)} \leq \mathrm{Lip}\,g \cdot L \cdot T^l$.
Under η-weak dependent inputs, we bound the covariance:

$$\begin{aligned}
&|\mathrm{Cov}(f(X_{\mathbf{s}}), g(X_{\mathbf{t}}))| \leq (u\mathrm{Lip}\,f + v\mathrm{Lip}\,g)\times \\
&\times \left[4\sum_{|i|\geq r} |i|b_i(\|Y_0\|_1 + \|Y_0\|_{l+1}^{l+1}) + (2r+1)L\left((2r+1)2\|Y_0\|_{m'}^{m'}T^{l+1-m'} + T^l\eta_{Y,k-2r}\right)\right]
\end{aligned}$$

We then fix the truncation $T^{m'-1} = \dfrac{2(2r+1)\|Y_0\|_{m'}^{m'}}{\eta_{Y,k-2r}}$ to obtain the result of the lemma 1 in the η-weak dependent case.
Under λ-weak dependent inputs:

$$|\mathrm{Cov}(f(X_{\mathbf{s}}), g(X_{\mathbf{t}}))| \le (u\mathrm{Lip}\, f + v\mathrm{Lip}\, g + uv\mathrm{Lip}\, f\mathrm{Lip}\, g)\times$$
$$\times \left(\left\{ 4 \sum_{|i|\ge r} |i| b_i (\|Y_0\|_1 + \|Y_0\|_{l+1}^{l+1}) + (2r+1)L \left(2(2r+1)T^{l+1-m'}\|Y_0\|_{m'}^{m'} + T^l \lambda_{Y,k-2r} \right) \right\} \right.$$
$$\left. \vee \ \left\{ (2r+1)^2 L^2 T^{2l} \lambda_{Y,k-2r} \right\} \right)$$

We then set a truncation such that $T^{l+m'-1} = \dfrac{2\|Y_0\|_{m'}^{m'}}{L\lambda_{Y,k-2r}}$ to obtain the result of the lemma 1 in the η-weak dependent case. □

Acknowledgement. We wish to thank an anonymous referee for his comments; he pointed out the reference of Heyde (1975). After this, we exhibited examples for which this method based on martingale arguments does not seem to apply.

References

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

[2] Bardet, J. M., Doukhan, P., Leon, J. R. (2005) *A uniform central limit theorem for the periodogram and its applications to Whittle parametric estimation for weakly dependent time series.* submitted.

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

[4] Borovkova, S.A. Burton, R.M. Dehling, H.G. (2001) *Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation.* T. Am. Math. Soc. 353, 4261-4318.

[5] Bulinski, A., Shashkin, A. (2005) *Strong invariance principle for dependent random fields.* Preprint.

[6] Dedecker, J. and Rio, E. (2000) *On the functional central limit theorem for stationary processes.* Ann. Inst. H. Poincaré 34, 134.

[7] Dedecker, J., Doukhan, P. (2003) *A new covariance inequality and applications.* Stoch. Proc. Appl. 63-80, Vol. 106, n 1.

[8] Doukhan, P. (1994) *Mixing: Properties and Examples.* Lecture Notes in Statistics 85. Springer Verlag.

[9] Doukhan, P. (2002) *Models inequalities and limit theorems for stationary sequences.* in Theory and applications of long range dependence (Doukhan *et alii* ed.) Birkhäuser, 43-101.

[10] Doukhan, P., Lang, G. (2002) *Rates of convergence in the weak invariance principle for the empirical repartition process of weakly dependent sequences.* Statist. Inf. for Stoch. Proc. 5, 199-228.

[11] Doukhan, P., Lang, G., Surgailis, D., Viano, M. C. (2005) *Functional limit theorem for the empirical process shifts with long memory.* J. Theor. Probab. 18, 1, 109-134.

[12] Doukhan, P., Louhichi, S. (1999) *A new weak dependence condition and applications to moment inequalities.* Stoch. Proc. Appl. 84, 313-342.

[13] Doukhan, P., Massart, P., Rio, E. (1995) *Invariance principle for the empirical measure of a weakly dependent process.* Ann. I.H.P. 31-2, 393-427.

[14] Doukhan, P., Portal, F. (1987) *Principe d'invariance faible pour la fonction de répartition empirique dans un cadre multidimensionnel et mélangeant.* Probab. Math. Stat. 8.2, 117-132.

[15] Doukhan, P. Teyssière, G., Winant, P. (2005) *Vector valued LARCH infinity processes.* To appear in a special issue of LNS, *Statistics of times series*, edited by P. Bertail *et alii.*

[16] Duflo, M. (1996) *Algorithmes stochastiques.* Math. & Appl. 23, Springer.

[17] Gordin, M. I. (1969). *The central limit theorem for stationary processes.* Soviet. Math. Dokl. 10, n.5, 11741176.

[18] Esary, J., Proschan, F., Walkup, D. (1967) *Association of random variables with applications.* Ann. Math. Statist. 38, 1466-1476.

[19] Giraitis, L., Surgailis, D. (2002) *ARCH-type bilinear models with double long memory.* Stochastic Processes and their Applications, 100, 275-300.

[20] Heyde, C. C. (1975) *On the central limit theorem and iterated logarithm law for stationary processes.* Bull. Austral. Math. Soc. 12, 18.

[21] Ibragimov, I. A. (1975) *A note on the central limit theorem for dependent random variables.* Theory Probab. Appl. 20, 135-141.

[22] Louhichi, S. (2001) *Rates of convergence in the CLT theorem for some weakly dependent random variables.* Theory Probab. Appl. 46-2, 297-315, and Teor. Veroyatn. Primen. 46-2, 345-364.

[23] Louhichi, S., Soulier, P. (2000) *Marcinkiewicz-Zygmund strong laws for infinite variance time series.* Statist. Inf. for Stoch. Proc., 3(1-2), p. 31-40.

[24] Merlevède, F. and Peligrad, M. (2004) *On the weak invariance principle for stationary sequences under projective criteria.* To appear in J. Theor. Probab. http://www.geocities.com/irina8/proj.pdf.

[25] Newman, C.M. and Wright, A.L. (1981) *An invariance principle for certain dependent sequences.* Ann. Probab.,9 , 671-765.

[26] Peligrad, M. and Utev, S. (2006) *Central limit theorem for stationary linear processes.* To appear in The Annals of Probability. http://arxiv.org/abs/math.PR/0509682

[27] Petrov, V. (1995) *Limit theorems of probability theory.* Oxford University Press.

[28] Pitt, L. (1982) *Positively Correlated Normal Variables are Associated.* Ann. Probab. 10, 496-499.

[29] Rio, E. (2000) *Théorie asymptotique pour des processus aléatoires faiblement dépendants.* SMAI, Mathématiques et Applications 31, Springer.

Paul Doukhan, LS-CREST, Laboratoire de Statistique, CREST. Timbre J340, 3, avenue Pierre Larousse, 92240 Malakoff, France and SAMOS, Statistique Appliquée et MOdélisation Stochastique, Université Paris 1, Centre Pierre Mendès France, 90 rue de Tolbiac, F-75634 Paris Cedex 13, France.
Olivier Wintenberger, SAMOS.