PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 27, Fasc. 1 (2007), pp. 45–73

AN INVARIANCE PRINCIPLE FOR WEAKLY DEPENDENT STATIONARY GENERAL MODELS

BY

PAUL DOUKHAN (MALAKOFF) AND OLIVIER WINTENBERGER (PARIS)

Abstract. The aim of this paper is to refine a weak invariance principle for stationary sequences given by Doukhan and Louhichi [10]. Since our conditions are not causal, our assumptions need to be stronger than the mixing and causal θ-weak dependence assumptions used in Dedecker and Doukhan [5]. Here, if moments of order greater than 2 exist, a weak invariance principle and convergence rates in the CLT are obtained; Doukhan and Louhichi [10] assumed the existence of moments with order greater than 4. Besides the η- and κ-weak dependence conditions used previously, we introduce a weaker one, λ, which fits the Bernoulli shifts with dependent inputs.

2000 AMS Mathematics Subject Classification: Primary: 60F17.

Key words and phrases: Invariance principle, weak dependence, the Bernoulli shifts.

1. INTRODUCTION

Let $(X_t)_{t\in Z}$ be a real-valued stationary process. A huge amount of applications make use of such times series.

Several ways of modeling weak dependence have already been proposed. One of the most popular is the notion of mixing (see [7] for bibliography); this notion leads to a very nice asymptotic theory, in particular a weak invariance principle under very sharp conditions (see [26] for the strong mixing case). Such mixing conditions entail restrictions on the model. For example, Andrews exhibits in [1] the simple counterexample of an autoregressive process which does not satisfy any mixing condition and innovations need much regularity in both MA(∞) and Markov models. Doukhan and Louhichi introduced in [10] new weak dependence conditions in order to solve those problems. We intend to sharpen their assumptions leading to a weak invariance principle. A common approach to derive a weak invariance principle for stationary sequences is based on a martingale difference approximation. This approach was first explored by Gordin in [14]; necessary and sufficient conditions were found by

Heyde in [15]. Let $\mathcal{M}_t$ be a filtration. Heyde's martingale difference approximation is equivalent to the existence of moments of order 2 and

$$\sum_{t=0}^{\infty} E(X_t \mid \mathcal{M}_0) \text{ and } \sum_{t=0}^{\infty} (X_{-t} - E(X_{-t} \mid \mathcal{M}_0)) \text{ converge in } \mathbb{L}^2. \tag{1.1}$$

Martingale theory leads directly to invariance principles (see also [27]). In the following, the adapted case refers to the special case where X_t is $\mathcal{M}_t$-measurable. The natural filtration is written as $\mathcal{M}_t = \sigma(Y_i, i \leqslant t)$ for independent and identically distributed inputs $(Y_t)_{t\in\mathbb{Z}}$; thus X_t can be written as a function of the past inputs:

$$X_t = H(Y_t, Y_{t-1}, \ldots). \tag{1.2}$$

Then only the first series in (1.1) needs to be considered. Using the Lindeberg technique, Dedecker and Rio [6] relax (1.1). Bernstein's blocks method allowed also Peligrad and Utev [22] to improve on (1.1). Such projective conditions are related to dependence coefficients; Dedecker and Doukhan obtain sharp results for the causal θ-dependence in [5] and Merlevède et al. address the mixing cases in a nice survey paper [20].

Martingale difference approximation is not always easy, for instance in the particular case where a natural filtration does not exist. The most striking example is given by associated sequences $(X_t)_{t\in\mathbb{Z}}$. Let us recall this notion. A series is said to be *associated* if $\operatorname{Cov}(f_1, f_2) \geqslant 0$ for any two coordinatewise nondecreasing functions f_1 and f_2 of $(X_{t_1}, \ldots, X_{t_m})$ with $\operatorname{Var}(f_1) + \operatorname{Var}(f_2) < \infty$. However, Newman and Wright [21] obtain a weak invariance principle under the existence of second order moments and

$$\sigma^2 = \sum_{t\in\mathbb{Z}} \operatorname{Cov}(X_0, X_t) < \infty. \tag{1.3}$$

Theorems 2.1 and 2.2 propose invariance principles under general assumptions: they apply to the non-causal Bernoulli shifts with weakly dependent inputs $(Y_t)_{t\in\mathbb{Z}}$,

$$X_t = H(Y_{t-j}, j \in \mathbb{Z}). \tag{1.4}$$

Heredity of weak dependence through such non-linear functionals follows from a new λ-weak dependence property; a function of a λ-weak dependence process is λ-weakly dependent, see Section 3.2. Analogous models with dependent inputs are already considered by [3]. If $X_t = \sum_{j\in\mathbb{Z}} \alpha_j Y_{t-j}$, Peligrad and Utev [23] prove that the Donsker invariance principle holds for X as soon as it holds for the innovation process Y. The non-linearity of H considered here is an important feature which has not been frequently discussed in the past. The condition of moments with order greater than 4 on the observations needed in [10] is reduced to a one of the moments with order greater than 2 and the results rely on specific decays of the dependence coefficients. We do not reach the second order moment condition of [15] (or projective conditions) and [21]. We *conjecture* that some time series satisfy weak dependence conditions with fast enough decay rates in order to ensure a Donsker type theorem but they satisfy neither condition (1.1) nor other projective criterion (see [20]) nor association nor Gaussianity. More general models (1.4) are considered here while causal models (1.2) fit to the adapted case and to projective conditions. However, proving this conjecture is really difficult since condition (1.1) has to be checked for each σ-algebra $\mathcal{M}_0$.

The paper is organized as follows. In Section 2 we introduce various weak dependent coefficients in order to state our main results. Section 3 is devoted to examples of weak dependent models for which we discuss our results. We shall focus on examples of λ-weakly dependent sequences. Proofs are given in the last section; we first derive conditions ensuring the convergence of the series σ^2. A bound of the Δ-moment of a sum (with $2 < \Delta < m$) is proved in Section 4.2; this bound is of an independent interest since for example it directly yields the strong laws of large numbers. The standard Lindeberg method with Bernstein's blocks is developed in Section 4.3 and yields our versions of the Donsker theorem. Convergence rates of the CLT are obtained in Section 4.4.

2. DEFINITIONS AND MAIN RESULTS

2.1. Weak dependence assumptions

DEFINITION 2.1 (Doukhan and Louhichi [10]). The process $(X_t)_{t\in \mathbf{Z}}$ is said to be (ε, ψ)*-weakly dependent* if there exist a sequence $\varepsilon(r)\downarrow 0$ (as $r\uparrow\infty$) and a function $\psi: \mathbf{N}^2\times(\mathbf{R}^+)^2\to\mathbf{R}^+$ such that

$$|\mathrm{Cov}(f(X_{s_1}, \ldots, X_{s_u}), g(X_{t_1}, \ldots, X_{t_v}))| \leqslant \psi(u, v, \mathrm{Lip} f, \mathrm{Lip} g)\,\varepsilon(r)$$

for any $r \geqslant 0$ and any $(u+v)$-tuples such that $s_1 \leqslant \ldots \leqslant s_u \leqslant s_u+r \leqslant t_1 \leqslant \ldots \leqslant t_v$, where the real-valued functions f and g are defined on $\mathbf{R}^u$ and $\mathbf{R}^v$, respectively, satisfy $\|f\|_\infty \leqslant 1$, $\|g\|_\infty \leqslant 1$ and are such that $\mathrm{Lip} f+\mathrm{Lip} g < \infty$, where

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

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 write $\varepsilon(r)$ as $\kappa(r)$;

• κ' (causal) weak dependence for which $\psi(u, v, a, b) = vab$, in this case we simply write $\varepsilon(r)$ as $\kappa'(r)$; this is the causal counterpart of κ coefficients which is recalled only for completeness;

• η-weak dependence, $\psi(u, v, a, b) = ua+vb$, in this case we write $\varepsilon(r) = \eta(r)$ for short;

• θ-weak dependence is a causal dependence which refers to the function $\psi(u, v, a, b) = vb$, in this case we simply put $\varepsilon(r) = \theta(r)$ (see [5]); this is the causal counterpart of η coefficients which is recalled only for completeness;

• λ-weak dependence $\psi(u, v, a, b) = uvab+ua+vb$, in this case we write $\varepsilon(r) = \lambda(r)$.

Remark 2.1. Besides the fact that it includes η- and κ-weak dependences, this new notion of λ-weak dependence will be proved to be convenient, for example, for the Bernoulli shifts with associated inputs (see Lemma 2.1 below).

Remark 2.2. If functions f and g are complex-valued, the previous inequalities remain true if we substitute $\varepsilon(r)/2$ for $\varepsilon(r)$. A useful case of such complex-valued functions is $f(x_1, \ldots, x_u) = \exp\big(it(x_1+\ldots+x_u)\big)$ for each $t \in \boldsymbol{R}$, $u \in \boldsymbol{N}^*$ and $(x_1, \ldots, x_u) \in \boldsymbol{R}^u$ (see Section 4.3). This indeed corresponds to the characteristic function adapted to derive the convergence in distribution.

2.2. Main results. Let $(X_t)_{t\in\boldsymbol{Z}}$ be a real-valued stationary sequence of mean zero satisfying

$$E|X_0|^m < \infty \quad \text{for a real number } m > 2. \tag{2.1}$$

Let us assume that

$$\sigma^2 = \sum_{k\in\boldsymbol{Z}} \operatorname{Cov}(X_0, X_k) = \sum_{k\in\boldsymbol{Z}} EX_0X_k > 0. \tag{2.2}$$

Denote by W the standard Brownian motion and by W_n the partial sums process:

$$W_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} X_i \quad \text{for } t \in [0, 1],\ n \geqslant 1. \tag{2.3}$$

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

THEOREM 2.1 (κ-dependence). *Assume that the* 0*-mean* κ*-weakly dependent stationary process* $(X_t)_{t\in\boldsymbol{Z}}$ *satisfies the condition* (2.1) *and* $\kappa(r) = \mathcal{O}(r^{-\kappa})$ *(as* $r \uparrow \infty$*) for* $\kappa > 2+1/(m-2)$. *Then* σ^2 *defined by* (2.2) *is finite and*

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

Remark 2.3. Under the more restrictive κ' condition, Bulinski and Sashkin [4] obtain invariance principles with the sharper assumption $\kappa' > 1+1/(m-2)$. Our loss is explained by the fact that κ'-weakly dependent sequences satisfy $\kappa'(r) \geqslant \sum_{s\geqslant r} \kappa_s$. This simple bound follows from the definitions.

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

THEOREM 2.2 (λ-dependence). *Assume that the* 0-*mean* λ-*weakly dependent stationary inputs satisfy the condition* (2.1) *and* $\lambda(r) = \mathcal{O}(r^{-\lambda})$ (*as* $r \uparrow \infty$) *for* $\lambda > 4+2/(m-2)$. *Then* σ^2 *defined by* (2.2) *is finite and*

$$W_n(t) \xrightarrow{D} \sigma W(t) \text{ as } n \to \infty \quad \text{in the Skorokhod space } D([0, 1]).$$

Remark 2.4. In comparison with the result obtained by Dedecker and Doukhan [5], our outcomes are not as good under θ-weak dependence. We work under more restrictive moment conditions than these authors. The same remark applies for all projective measures of dependence; here we refer to results in [15], [21], [6] and [22].

Remark 2.5. However, the examples of Section 3.2 stress the fact that such results are not systematically better than those of Theorem 2.2; for such general examples, we even conjecture that theorems of [15], [21], [6] or [22] do not apply.

Remark 2.6. The technique of the proofs is based on the Lindeberg method. In fact, we prove that $|E(\phi(S_n/\sqrt{n})-\phi(\sigma N))| = o(n^{-c})$ (ϕ denotes here the characteristic function) for $0 < c < c^*$, where c^* depends only on the parameters m and κ or λ, respectively. If m and κ (or λ) both tend to infinity, we notice that $c^* \to \frac{1}{4}$. As κ or λ tends to infinity and $m < 3$, c^* always remains smaller than $(m-2)/(2m-2)$ (see Proposition 4.2 in Section 4.4 for more details).

Remark 2.7. Using a smoothing lemma also yields an analogous bound for the uniform distance:

$$\sup_{x\in\mathbf{R}} \left| P\left(\frac{1}{\sqrt{n}} S_n \leqslant x\right) - P(\sigma N \leqslant x) \right| = o(n^{-c}) \quad \text{for some } c < c'.$$

A first and easy way to control c' is to let $c' = c^*/4$ but the corresponding rate is a really bad one (see e.g. [11]). The Esseen inequality holds with the optimal exponent 1/2 in the independent and identically distributed case (see [24]) and Rio [26] reaches the exponent 1/3 in the case of strongly mixing sequences. In Proposition 4.2 we achieve $c' > c^*/4$. Analogous results have been settled in [8] for weakly dependent random fields. Previous results in [15], [21], [6] or [22] do not derive such convergence rates for the Kolmogorov distance.

Let us denote by $\mathbf{R}^{(\mathbf{Z})} = \bigcup_{I>0} \{z \in \mathbf{R}^{\mathbf{Z}} \mid z_i = 0, |i| > I\}$ the set of finite sequences of real numbers. We consider functions $H\colon \mathbf{R}^{(\mathbf{Z})} \to \mathbf{R}$ such that if $x, y \in \mathbf{R}^{(\mathbf{Z})}$ coincide for all indices but one, say let $s \in \mathbf{Z}$, then

$$|H(x)-H(y)| \leqslant b_s(\|z\|^l \vee 1)|x_s-y_s|, \tag{2.4}$$

where $z \in \mathbf{R}^{(\mathbf{Z})}$ is defined by $z_s = 0$ and $z_i = x_i = y_i$ for $i \neq s$. Here $\|x\| = \sup_{i \in \mathbf{Z}} |x_i|$. In Section 3.2, we prove the existence of the sequence

$$X_n = \lim_{I \to \infty} H\big((Y_{n-j} \mathbb{1}_{\{j \leqslant I\}})_{j \in \mathbf{Z}}\big),$$

where $(Y_t)_{t \in \mathbf{Z}}$ is a weakly dependent real-valued input process. We denote this process by $X_n = H(Y_{n-j}, j \in \mathbf{Z})$ for simplicity and we derive its λ-weak invariance properties. Various asymptotic results, among which our weak invariance principle, follow Theorem 2.2.

COROLLARY 2.1. *Let $(Y_t)_{t \in \mathbf{Z}}$ be a stationary λ-weakly dependent process (with dependence coefficients $\lambda_Y(r)$) and $H: \mathbf{R}^{(\mathbf{Z})} \to \mathbf{R}$ satisfying the condition* (2.4) *for some $l \geqslant 0$. Let us assume that there exist real numbers m, m' with $E|Y_0|^{m'} < \infty$ such that $m > 2$ and $m' \geqslant (l+1)m$.*

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

Geometric case: *$b_r \leqslant Ce^{-b|r|}$ and $\lambda_Y(r) \leqslant De^{-ar}$ for $a, b, C, D > 0$.*

Riemannian case: *$b_r \leqslant C(1+|r|)^{-b}$ and $\lambda_Y(r) \leqslant Dr^{-a}$ for $a, C, D > 0$ with the conditions*

$$a > \frac{1+b}{b-1}\left(4+\frac{2}{m-2}\right) \qquad \textit{for } l = 0,\ b > 1;$$

$$\text{(2.5)}\qquad a > \frac{b(m'-1+l)}{(b-2)(m'-1-l)}\left(4+\frac{2}{m-2}\right) \qquad \textit{for } l > 0,\ b > 2.$$

Remark 2.8. The previous conditions are also tractable in the mixed cases. We explicitly state them for $l > 0$:

$$b_r \leqslant Ce^{-b|r|} \quad \text{and} \quad \lambda_Y(r) \leqslant Dr^{-a}$$

if moreover

$$a > \frac{m'-1+l}{m'-1-l}\left(4+\frac{2}{m-2}\right) \quad \text{and} \quad b, C, D > 0,$$

or

$$b_r \leqslant C|r|^{-b} \quad \text{and} \quad \lambda_Y(r) \leqslant De^{-ar}$$

if $a, C, D > 0$ and $b > (6m-10)/(m-2)$.

3. EXAMPLES

Theorem 2.1 is useful to derive the weak invariance principle in various cases. This section is aimed at a detailed treatment of the Bernoulli shifts with dependent inputs. The important class of Lipschitz functions of dependent inputs is presented in a separate section. The importance of our results is highlighted by the models of the first subsection. More general non-linear models are considered in the second subsection. Some of those examples illustrate the conjecture we made in the Introduction but we were not able to prove it formally.

3.1. Lipschitz processes with dependent inputs. Consider Lipschitz functions $H: \mathbf{R}^{(\mathbf{Z})} \to \mathbf{R}$, i.e. the functions for which the condition (2.4) is satisfied for $l = 0$. A simple example of this situation is the two-sided linear sequence

$$X_t = \sum_{i \in \mathbf{Z}} \alpha_i Y_{t-i} \tag{3.1}$$

with dependent inputs $(Y_t)_{t \in \mathbf{Z}}$. As quoted in [17] for the case of linear processes with dependent input, there exists a very general solution; essentially, any Donsker type theorem for the stationary inputs implies the central limit theorem for any linear process driven by such inputs. More precisely, Theorem 5 of [23] states that this process even satisfies the Donsker invariance principle if $\sum_j |\alpha_j| < \infty$.

A simple example of the Lipschitz non-linear functional of dependent inputs is

$$X_t = \Big|\sum_{i \in \mathbf{Z}} \alpha_i Y_{t-i}\Big| - E\Big|\sum_{i \in \mathbf{Z}} \alpha_i Y_{-i}\Big|. \tag{3.2}$$

In this case the inequality (2.4) holds with $l = 0$ and $b_r \leqslant |\alpha_r|$.

Another example of this situation is the following stationary process:

$$X_t = Y_t\big(a + \sum_{j \neq 0} a_j Y_{t-j}\big) - EY_t\big(a + \sum_{j \neq 0} a_j Y_{t-j}\big),$$

where the inputs $(Y_t)_{t \in \mathbf{Z}}$ are bounded. In this case, the inequality (2.4) also holds with $l = 0$ and $b_s \leqslant 2\|Y_0\|_\infty |a_s|$.

To apply our result, we compute the weak dependence coefficients of such models.

LEMMA 3.1. *Let $(Y_t)_{t \in \mathbf{Z}}$ be a strictly stationary process with a finite moment of order $m > 2$ and $H: \mathbf{R}^{(\mathbf{Z})} \to \mathbf{R}$ satisfying the condition* (2.4) *for $l = 0$ and some nonnegative sequence $(b_s)_{s \in \mathbf{Z}}$ such that $L = \sum_j b_j < \infty$. Then:*

- *The process $X_n = H(Y_{n-j}, j \in \mathbf{Z}) := \lim_{I \to \infty} H(Y_{n-j} \mathbf{1}_{\{j \leqslant I\}}, j \in \mathbf{Z})$ is a strictly stationary process with finite moments of order m.*

- *If the input process $(Y_t)_{t \in \mathbf{Z}}$ is λ-weakly dependent (the weak dependence coefficients are denoted by $\lambda_Y(r)$), then $(X_t)_{t \in \mathbf{Z}}$ is λ-weakly dependent and*

$$\lambda(k) = \inf_{2r \leqslant k} \big[2 \sum_{|i| \geqslant r} b_i \|Y_0\|_1 + (2r+1)^2 L^2 \lambda_Y(k-2r)\big].$$

- *If the input process $(Y_t)_{t \in \mathbf{Z}}$ is η-weakly dependent (the weak dependence coefficients are denoted by $\eta_Y(r)$), then $(X_t)_{t \in \mathbf{Z}}$ is η-weakly dependent and*

$$\eta(k) = \inf_{2r \leqslant k} \big[2 \sum_{|i| \geqslant r} b_i \|Y_0\|_1 + (2r+1) L \eta_Y(k-2r)\big].$$

Remark 3.1. Let $(Y_t)_{t\in\mathbf{Z}}$ be a strictly stationary process with a finite moment of order $m > 2$. If $L = \sum_j |\alpha_j| < \infty$, the process $X_n = \sum_{j\in\mathbf{Z}} \alpha_j Y_{n-j}$ is a strictly stationary process with finite moments of order m which satisfies the assumptions of Lemma 3.1 with $b_j = |\alpha_j|$. Even if the weak invariance principle is already given in [23], our result is of an independent interest, for example for functional estimation purposes. For non-linear Lipschitz functionals it yields new central limit theorems.

The result of Theorems 2.1 and 2.2 holds systematically in geometric cases. Then Riemannian decays are assumed, i.e. there exist $\alpha, C > 0$ such that

$$b_r \leqslant C r^{-\alpha}.$$

The conditions from [15] are compared below with the conditions of Theorems 2.1 and 2.2 for specific classes of inputs $(Y_t)_{t\in\mathbf{Z}}$.

3.1.1. *LARCH*(∞) *inputs.* A vast literature is devoted to the study of conditionally heteroskedastic models. A simple equation in terms of a vector-valued process allows a unified treatment of those models, see [12]. Let $(\xi_t)_{t\in\mathbf{Z}}$ be an independent and identically distributed centered real-valued sequence and $a, a_j, j\in N^*$, be real numbers. LARCH(∞) models are solutions of the recurrence equation

$$Y_t = \xi_t\Big(a + \sum_{j=1}^{\infty} a_j Y_{t-j}\Big). \tag{3.3}$$

We provide below sufficient conditions for the following chaotic expansion:

$$Y_t = \xi_t\Big(a + \sum_{k=1}^{\infty} \sum_{j_1,\ldots,j_k \geqslant 1} a_{j_1}\xi_{t-j_1} a_{j_2} \ldots a_{j_k} \xi_{t-j_1-\ldots-j_k} a\Big). \tag{3.4}$$

Assume that $\Lambda = \|\xi_0\|_m \sum_{j\geqslant 1} |a_j| < 1$. Then one (essentially unique) stationary solution of equation (3.3) in $\mathbf{L}^m$ is given by (3.4). This solution is θ-weakly dependent with $\theta_Y(r) \leqslant K r^{1-a} \log^{a-1} r$ for some constant $K > 0$. This implies the same bound on their coefficients $(\lambda_Y(r))_{r\geqslant 0}$. The condition (2.5) gives the weak invariance principle for $(X_t)_{t\in\mathbf{Z}}$ under the conditions $E|\xi_0|^m < +\infty$ for $m > 2$, $\alpha > 1$, and

$$a > \frac{1+\alpha}{\alpha-1}\Big(4 + \frac{2}{m-2}\Big) + 1.$$

The model (3.1) is also Heyde's martingale difference approximation (1.1) as soon as

$$\sum_{k\geqslant 1} \sqrt{\sum_{i\geqslant k} \alpha_i^2} < +\infty.$$

Necessary conditions for the weak invariance principle are satisfied if $\alpha > \frac{3}{2}$, $|a_j| \leqslant C j^{-a}$ for some $a > 1$, $E\xi_0^2 < +\infty$, and $\|\xi_0\|_2 \sum_{j\geqslant 1} |a_j| < 1$. These conditions are not optimal since in this case the process is adapted to the filtration $\mathcal{M}_t = \sigma(\xi_i, i \leqslant t)$. Peligrad and Utev [22] extend the Donsker theorem to the cases where $\alpha > \frac{1}{2}$. Thus, our conditions are not optimal when compared to those of [23] in the linear case as in equation (3.1). However, for the non-linear Lipschitz functional, the result seems to be new.

3.1.2. *Non-causal LARCH(∞) inputs.* The provious approach extends for the case of non-causal LARCH(∞) inputs

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

Doukhan et al. [12] prove the same results of existence as for the previous causal case (just replace summation over $j > 0$ by summation over $j \neq 0$) and the dependence becomes of η-type with

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

where $A(x) = \sum_{|j| \geqslant x} |a_j|$, $A = \|\xi_0\|_\infty \sum_{j \geqslant 1} |a_j| < 1$. By the condition (2.5) the weak invariance principle holds for $(X_t)_{t \in \mathbf{Z}}$ if $\|\xi_0\|_\infty < \infty$, $\alpha > 1$ and

$$a > \frac{1+\alpha}{\alpha-1}\left(4 + \frac{2}{m-2}\right) + 1.$$

Notice that a very restrictive new assumption is that inputs need to be uniformly bounded in this non-causal case. This result is new, the conjecture is that (1.1) does not hold.

3.1.3. *Non-causal and non-linear inputs.* The weak dependence properties of non-causal and non-linear inputs Y_t are recalled, see [10] for more details. Let $H: (\mathbf{R}^d)^{\mathbf{Z}} \to \mathbf{R}$ be a measurable function. If the sequence $(\xi_n)_{n \in \mathbf{Z}}$ is independent and identically distributed on $\mathbf{R}^d$, the Bernoulli shift with input process $(\xi_n)_{n \in \mathbf{Z}}$ is defined as

$$Y_n = H\big((\xi_{n-i})_{i \in \mathbf{Z}}\big), \qquad n \in \mathbf{Z}.$$

Such Bernoulli's shifts are *η-weakly dependent* (see [10]) with $\eta(r) \leqslant 2\delta_{[r/2]}$ if

$$E|H(\xi_j, j \in \mathbf{Z}) - H(\xi_j \mathbf{1}_{|j| \leqslant r}, j \in \mathbf{Z})| \leqslant \delta_r. \tag{3.5}$$

Then condition (2.5) leads to the invariance principle for $(X_t)_{t \in \mathbf{Z}}$ if $E|Y_0|^m < \infty$ for $m > 2$, $\alpha > 1$ and $\delta_r \leqslant K r^{-\delta}$ for

$$\delta > \frac{1+\alpha}{\alpha-1}\left(4 + \frac{2}{m-2}\right).$$

Conditions (1.1) of [15] do not give clear restrictions on coefficients for these models. We do not know other weak invariance principle in that general context.

3.1.4. *Associated inputs.* A process is *associated* if $\mathrm{Cov}(f(Y^{(n)}), g(Y^{(n)})) \geqslant 0$ for any coordinatewise non-decreasing functions $f, g: \mathbf{R}^n \to \mathbf{R}$ such that the previous covariance makes sense with $Y^{(n)} = (Y_1, \dots, Y_n)$. The κ-weak dependence condition is known to hold for associated or Gaussian sequences. In both cases we have the relation

$$\kappa(r) = \sup_{j \geqslant r} |\mathrm{Cov}(Y_0, Y_j)|.$$

Notice that the absolute values are needed only in the second case since for associated processes these covariances are nonnegative. Independent sequences are associated as well and Pitt [25] proves that a Gaussian process with nonnegative covariances is also associated. Finally, we recall that non-decreasing functions of associated sequences remain associated. Associated models are classically built this way of independent and identically distributed sequences (see [18]).

Suppose that the inputs $(Y_t)_{t \in \mathbf{Z}}$ are such that $\kappa(r) \leqslant Cr^{-a}$ (for some $a, C > 0$). For the associated cases and model (3.1), the invariance principle of [21] follows from the remark in [19] as soon as $EY^2 < +\infty$, $a > 1$ and $\alpha > 1$. These conditions are optimal, they correspond to $\sum_j \mathrm{Cov}(X_0, X_j) < \infty$. Such strong conditions are due to the fact that zero correlation implies independence for associated processes. Our conditions for the invariance principle are much stronger: $E|Y|^m < +\infty$ with $m > 2$, $\alpha > 1$ and

$$a > \frac{1+\alpha}{\alpha - 1}\left(4 + \frac{2}{m \dot{-} 2}\right).$$

In the special case of κ-weak dependent inputs that are not associated, the optimal weak invariance principle of [21] does not apply, see e.g. [10].

3.2. The Bernoulli shifts with dependent inputs. Let $H: \mathbf{R}^{(\mathbf{Z})} \to \mathbf{R}$ be a measurable (not necessarily Lipschitz) function and $X_n = H(Y_{n-i}, i \in \mathbf{Z})$. Such models are proved to exhibit either λ- or η-weak dependence properties. Because the Bernoulli shifts of κ-weak dependent inputs are neither κ- nor η-weakly dependent, the κ case is here included in the λ one.

Consider the non-Lipschitz function H defined by

$$H(x) = \sum_{k=0}^{K} \sum_{j_1, \dots, j_k} a^{(k)}_{j_1, \dots, j_k} x_{j_1} \dots x_{j_k}.$$

In this case, Lemma 3.1 does not apply. To derive weak dependence properties of such processes, we assume that H satisfies the condition (2.4) with $l \neq 0$, which remains a stronger assumption than that for the case of independent inputs (see (3.5)). Relaxing the Lipschitz assumption on H is possible if we assume the existence of higher moments for the inputs. The following lemma gives both the existence and the weak dependence properties of such models.

LEMMA 3.2. *Let $(Y_i)_{i\in Z}$ be a stationary process and assume that $H\colon R^{(Z)} \to R$ satisfies the condition* (2.4) *for some $l > 0$ and some sequence $b_j \geqslant 0$ such that $\sum_j |j|\, b_j < \infty$. Let us assume that there exists a pair of real numbers (m, m') with $E|Y_0|^{m'} < \infty$ such that $m > 2$ and $m' \geqslant (l+1)m$. Then*:

- *The process $X_n = H(Y_{n-i}, i\in Z)$ is well defined in L^m i.e., it is a strictly stationary process.*
- *If the input process $(Y_i)_{i\in Z}$ is λ-weakly dependent (the weak dependence coefficients are denoted by $\lambda_Y(r)$), then X_n is λ-weakly dependent and there exists a constant $c > 0$ such that*

$$\lambda(k) = c \inf_{r\leqslant[k/2]} \Big[\sum_{|j|\geqslant r} |j|\, b_j + (2r+1)^2 \lambda_Y(k-2r)^{(m'-1-l)/(m'-1+l)}\Big].$$

- *If the input process $(Y_i)_{i\in Z}$ is η-weakly dependent (the weak dependence coefficients are denoted by $\eta_Y(r)$), then X_n is η-weakly dependent and there exists a constant $c > 0$ such that*

$$\eta(k) = c \inf_{r\leqslant[k/2]} \Big[\sum_{|j|\geqslant r} |j|\, b_j + (2r+1)^{1+l/(m'-1)} \eta_Y(k-2r)^{(m'-2)/(m'-1)}\Big].$$

Such models were already mentioned in the mixing case by Billingsley [2] and Borovkova et al. [3]. The proofs are given in Section 4.6.

3.2.1. *Volterra models with dependent inputs.* Consider the function H defined by

$$H(x) = \sum_{k=0}^{K} \sum_{j_1,\ldots,j_k} a^{(k)}_{j_1,\ldots,j_k} x_{j_1} \ldots x_{j_k}.$$

Then if x, y are as in (2.4), we have

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

From the triangular inequality we thus derive that Lemma 3.2 may be written with $l = K-1$,

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

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

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

For example,

$$|a^{(k)}_{j_1,\dots,j_k}| \leqslant C(j_1 \vee \dots \vee j_k)^{-\alpha} \quad \text{or} \quad |a^{(k)}_{j_1,\dots,j_k}| \leqslant C\exp\big(-\alpha(j_1 \vee \dots \vee j_k)\big)$$

yield $b_s \leqslant C' s^{d-1-\alpha}$ or $b_s \leqslant C' e^{-\alpha s}$, respectively, for some constant $C' > 0$.

3.2.2. *Markov stationary inputs.* Markov stationary sequences satisfy a recurrence equation

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

where $(\xi_n)_{n\in\mathbf{Z}}$ is a sequence of independent and identically distributed random variables. In this case $Y_n = (Z_n, \dots, Z_{n-d+1})$ is a Markov chain $Y_n = M(Y_{n-1}, \xi_n)$ with

$$M(x_1, \dots, x_d, \xi) = \big(F(x_1, \dots, x_d, \xi), x_1, \dots, x_{d-1}\big). \tag{3.6}$$

Theorem 1.IV.24 of [13] proves that equation (3.6) has a stationary solution $(Z_n)_{n\in\mathbf{Z}}$ in $\mathbf{L}^m$ for $m \geqslant 1$ if $\|F(0, \xi)\|_m < \infty$ and there exist a norm $\|\cdot\|$ on $\mathbf{R}^d$ and a real number $a \in [0, 1[$ such that $\|F(x, \xi) - F(y, \xi)\|_m \leqslant a\|x-y\|$. In this setting, θ-dependence holds with $\theta_Z(r) = \mathcal{O}(a^{r/d})$ (as $r \uparrow \infty$). We shall not give more details about the significative examples provided in [9]. Indeed, we already mentioned that our results are suboptimal in such causal cases; such dependent sequences may however also be used as inputs for the Bernoulli shifts.

3.2.3. *Explicit dependence rates.* We now specify the decay rates from Lemma 3.2. For standard decays of the previous sequences, it is easy to get the following explicit bounds. Here b, c, C, D, λ, $\eta > 0$ are constants which may differ from one case to the other.

- Assume that $b_j \leqslant C(|j|+1)^{-b}$. If $\lambda_Y(j) \leqslant Dj^{-\lambda}$ or $\eta_Y(j) \leqslant Dj^{-\eta}$, then by a simple calculation we optimize both terms in order to prove that

$$\lambda(k) \leqslant ck^{A_1}, \quad \text{where } A_1 = -\lambda\left(1-\frac{2}{b}\right)\frac{m'-1-l}{m'-1+l},$$

or

$$\eta(k) \leqslant ck^{A_2}, \quad \text{where } A_2 = -\eta\,\frac{(b-2)(m'-2)}{(b-1)(m'-1)-l},$$

respectively.

Note that in the case where $m' = \infty$ 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.

- Assume that $b_j \leqslant Ce^{-|j|b}$. If $\lambda_Y(j) \leqslant De^{-j\lambda}$ or $\eta_Y(j) \leqslant De^{-j\eta}$, we obtain

$$\lambda(k) \leqslant ck^2 \exp\left(-\lambda k\frac{b(m'-1-l)}{b(m'-1+l)+2\eta(m'-1-l)}\right)$$

or

$$\eta(k) \leqslant ck^{(m'-1-l)/(m'-1)} \exp\left(-\eta k \frac{b(m'-2)}{b(m'-1)+2\eta(m'-2)}\right),$$

respectively.

The geometric decay of both $(b_j)_{j\in Z}$ and the weak dependence coefficients of the inputs ensure the geometric decay of the weak dependence coefficients of the Bernoulli shift.

• If we assume that the coefficients $(b_j)_{j\in Z}$ associated with the Bernoulli shift have a geometric decay, say $b_j \leqslant Ce^{-|j|b}$, and that $\lambda_Y(j) \leqslant Dj^{-\lambda}$ or $\eta_Y(j) \leqslant Dj^{-\eta}$, we obtain the bounds

$$\lambda(k) \leqslant ck^{-\lambda[(m'-1-l)/(m'-1+l)]}\log^2 k \quad \text{or} \quad \eta(k) \leqslant ck^{-\eta[(m'-2)/(m'-1)]}\log^{1+l/(m'-1)} k,$$

respectively.

If $m' = \infty$, tightness is reduced by a factor $\log^2 k$ with respect to the dependence coefficients of the input dependent series $(Y_t)_{t\in Z}$.

• If we assume that the coefficients $(b_j)_{j\in Z}$ associated with the Bernoulli shift have a Riemannian decay, say $b_j \leqslant C(|j|+1)^{-b}$, and that $\lambda_Y(j) \leqslant De^{-j\lambda}$ or $\eta_Y(j) \leqslant De^{-j\eta}$, we find $\lambda(k) = ck^{2-b}$ or $\eta(k) \leqslant ck^{2-b}$, respectively.

All models or functions of models we present here are λ-weakly dependent. We treat some basic examples in detail when a discussion with other results is possible. We believe that for some models λ-weak invariance properties follow by easy computations, and then statistical results like our weak invariance principle.

4. PROOFS OF THE MAIN RESULTS

Our proof for central limit theorems is based on a truncation method. For a truncation level $T \geqslant 1$ we shall write $\bar{X}_k = f_T(X_k) - Ef_T(X_k)$ with $f_T(X) = X \vee (-T) \wedge T$. From now on, we shall use the convenient notation $a_n \lesssim b_n$ for two real sequences $(a_n)_{n\in N}$ and $(b_n)_{n\in N}$ when there exists some constant $C > 0$ such that $|a_n| \leqslant Cb_n$ for each integer n. We also remark that $\bar{X}_k$ has moments of all orders because it is bounded. In the sequel, we put $\mu = E|X_0|^m$. For any $a \leqslant m$, we control the moment $E|f_T(X_0) - X_0|^a$ with Markov inequality

$$E|f_T(X_0) - X_0|^a \leqslant E|X_0|^a \mathbb{1}_{\{|X_0| \geqslant T\}} \leqslant \mu T^{a-m}.$$

Thus using the Jensen inequality yields

$$\|\bar{X}_0 - X_0\|_a \leqslant 2\mu^{1/a} T^{1-m/a}. \tag{4.1}$$

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

In this section we prove that the central limit theorems corresponding to the convergence $W_n(1) \to W(1)$ in both Theorems 2.1 and 2.2 hold and we shall provide convergence rates corresponding to these central limit theorems. The weak invariance principle is obtained in a standard way from such central limit theorems and tightness, which follows from Lemma 3.2, by using the classical Kolmogorov–Centsov tightness criterion (see [2]). In the last subsection, we prove Lemma 4.2 that states the properties of our (new) Bernoulli's shifts with dependent inputs.

4.1. Variances

LEMMA 4.1 (Variances). *If one of the following conditions holds:*

$$\sum_{k=0}^{\infty} \kappa(k) < \infty, \tag{4.2}$$

$$\sum_{k=0}^{\infty} \lambda(k)^{(m-2)/(m-1)} < \infty, \tag{4.3}$$

then the series σ^2 is convergent.

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

$$|\operatorname{Cov}(\bar{X}_0, \bar{X}_k)| \leqslant \kappa(k) \quad \text{or} \quad |\operatorname{Cov}(\bar{X}_0, \bar{X}_k)| \leqslant (4T+1)\lambda(k), \tag{4.4}$$

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

$$|\operatorname{Cov}(X_0, X_k)| \leqslant \kappa(k). \tag{4.5}$$

In the following, we shall only consider λ dependence. We develop

$$\operatorname{Cov}(X_0, X_k) = \operatorname{Cov}(\bar{X}_0, \bar{X}_k) + \operatorname{Cov}(X_0 - \bar{X}_0, X_k) + \operatorname{Cov}(\bar{X}_0, X_k - \bar{X}_k).$$

We use a truncation T (to be determined) and the bounds given in (4.1) and (4.4); then the Hölder inequality with the exponents $1/a + 1/m = 1$ yields

$$\begin{aligned} |\operatorname{Cov}(X_0, X_k)| &\leqslant (4T+1)\lambda(k) + 2\|X_0\|_m \|\bar{X}_0 - X_0\|_a \\ &\leqslant (4T+1)\lambda(k) + 4\mu^{1/a+1/m} T^{1-m/a} \\ &\leqslant (4T+1)\lambda(k) + 4\mu T^{2-m}. \end{aligned}$$

Choosing $T^{m-1} = \mu/\lambda(k)$ we obtain

$$|\operatorname{Cov}(X_0, X_k)| \leqslant 9\mu^{1/(m-1)}\lambda(k)^{(m-2)/(m-1)}. \ \blacksquare \tag{4.6}$$

4.2. A Δ-order moment bound

LEMMA 4.2. *Let $(X_t)_{t\in\mathbb{Z}}$ be a stationary and centered process. Let us assume that $E|X_0|^m < \infty$ and that this process is either κ-weakly dependent with $\kappa(r) = \mathcal{O}(r^{-\kappa})$ or λ-weakly dependent with $\lambda(r) = \mathcal{O}(r^{-\lambda})$. If $\kappa > 2 + 1/(m-2)$ or*

$\lambda > 4+2/(m-2)$, *then for all* $\Delta > 2$ *small enough there exists a constant* $C > 0$ *such that*

$$\|S_n\|_\Delta \leqslant C\sqrt{n}.$$

Remark 4.1. $\Delta \in]2, 2+A\wedge B\wedge 1[$, where A and B are constants smaller than $m-2$ and depend on m and on κ or λ, respectively. Equations (4.10) and (4.11) below precise the previously involved constants A and B.

Remark 4.2. The constant satisfies

$$C > \left(\frac{5}{2^{(\Delta-2)/2}-1}\right)^{1/\Delta} \sum_{k\in\mathbf{Z}} |\mathrm{Cov}(X_0, X_k)|.$$

Under the conditions of Lemma 4.2, using Lemma 4.1 we obtain

$$c \equiv \sum_{k\in\mathbf{Z}} |\mathrm{Cov}(X_0, X_k)| < \infty.$$

Remark 4.3. The result is sketched from Bulinski and Shashkin [4]. However, their dependence condition is of causal nature while our is not. It explains a loss with respect to the exponents λ and κ. In their κ'-weak dependence setting the best possible value of the exponent is 1 while it is 2 for our non-causal dependence.

Proof of Lemma 4.2. For convenience, let us put in the sequel $\Delta = 2+\delta$ and $m = 2+\zeta$. Like in [16] or [4], we proceed by induction on k for $n \leqslant 2^k$ to prove that

$$\| 1+|S_n| \|_\Delta \leqslant C\sqrt{n}. \tag{4.7}$$

We assume that the condition (4.7) is satisfied for all $n \leqslant 2^{K-1}$. Setting $N = 2^K$ we have to find a bound for $\|1+|S_N|\|_\Delta$. We can divide the sum S_N into three blocks: the first two blocks have the same size $n \leqslant 2^{K-1}$ and are denoted by Q and R; the third block V, located between Q and R, has cardinality $q < n$. We then have

$$\|1+|S_N|\|_\Delta \leqslant \|1+|Q|+|R|\|_\Delta + \|V\|_\Delta.$$

By the recurrence assumption, the term $\|V\|_\Delta$ is directly bounded by $\|1+|V|\|_\Delta \leqslant C\sqrt{q}$. Writing $q = \xi N^b$ with $b < 1$ and $0 < \xi < 1$, we see that this term is of order strictly smaller than $\sqrt{N}$. For $\|1+|Q|+|R|\|_\Delta$, we have

$$\begin{aligned} E(1+|Q|+|R|)^\Delta &\leqslant E(1+|Q|+|R|)^2(1+|Q|+|R|)^\delta \\ &\leqslant E\big(1+2|Q|+2|R|+(|Q|+|R|)^2\big)(1+|Q|+|R|)^\delta. \end{aligned}$$

We expand the right-hand side of this expression. Then the following terms (i), (ii), (iii) appear:

(i) $$E(1+|Q|+|R|)^\delta \leqslant 1+|Q|_2^\delta+|R|_2^\delta \leqslant 1+2c^\delta(\sqrt{n})^\delta.$$

(ii)
$$E|Q|(1+|Q|+|R|)^{\delta} \leqslant E|Q|((1+|R|)^{\delta}+|Q|^{\delta}) \leqslant E|Q|(1+|R|)^{\delta}+E|Q|^{1+\delta}.$$

The term $E|Q|^{1+\delta}$ is bounded by $\|Q\|_2^{1+\delta}$, and then by $c^{1+\delta}(\sqrt{n})^{1+\delta}$. Using the Hölder inequality, we see that the term $E|Q|(1+|R|)^{\delta}$ is bounded by $\|Q\|_{1+\delta/2}\|1+|R|\|_{\Delta}^{\delta}$. It is at least of order $cC^{\delta}(\sqrt{n})^{1+\delta}$, analogous to the latter one, where we exchange the roles of Q and R.

(iii)
$$E(|Q|+|R|)^2(1+|Q|+|R|)^{\delta}.$$

For this term, we use an inequality from [4] and we obtain

$$E(|Q|+|R|)^2(1+|Q|+|R|)^{\delta} \leqslant E|Q|^{\Delta}+E|R|^{\Delta}+5\big(EQ^2(1+|R|)^{\delta}+ER^2(1+|Q|)^{\delta}\big).$$

Now, by (4.7), we get $E|Q|^{\Delta} \leqslant C^{\Delta}(\sqrt{n})^{\Delta}$. The second term is its analogue with R substituted for Q. The third term has to be handled with a particular care as follows.

We use the weak dependence notion to control $EQ^2(1+|R|)^{\delta}$ and $ER^2(1+|Q|)^{\delta}$. Denote by $\bar{X}$ the variable $X \vee T \wedge (-T)$ for a real $T>0$ to be determined later. By extension, $\bar{Q}$ and $\bar{R}$ denote the truncated sums of the variables X_i. We have

$$E|Q|^2(1+|R|)^{\delta} \leqslant EQ^2\big||R|-|\bar{R}|\big|^{\delta}+E(Q^2-\bar{Q}^2)(1+|\bar{R}|)^{\delta}+E\bar{Q}^2(1+|\bar{R}|)^{\delta}.$$

We begin with a control of $EQ^2\big||R|-|\bar{R}|\big|^{\delta}$. Using the Hölder inequality with $2/m+1/m'=1$ yields

$$EQ^2\big||R|-|\bar{R}|\big|^{\delta} \leqslant \|Q\|_m^2\,\big\|\,\big||R|-|\bar{R}|\big|^{\delta}\,\big\|_{m'}.$$

$\|Q\|_{\Delta}$ is bounded by using (4.7) and

$$\big||R|-|\bar{R}|\big|^{\delta m'} \leqslant |R|^{\delta m'}\mathbb{1}_{\{|R|>T\}} \leqslant |R|^{\delta m'}\mathbb{1}_{|R|>T}.$$

We then bound $\mathbb{1}_{|R|>T} \leqslant (|R|/T)^{\alpha}$ with $\alpha=m-\delta m'$, and hence

$$E\big||R|-|\bar{R}|\big|^{\delta m'} \leqslant E|R|^m T^{\delta m'-m}.$$

By convexity and stationarity, we have $E|R|^m \leqslant n^m E|X_0|^m$, so that

$$EQ^2(|R|-|\bar{R}|)^{\delta} \preceq n^{2+m/m'}T^{\delta-m/m'}.$$

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

$$EQ^2(|R|-|\bar{R}|)^{\delta} \preceq n^m T^{\Delta-m}.$$

We get the same bound for the second term:

$$E(Q^2-\bar{Q}^2)(1+|\bar{R}|)^{\delta} \preceq n^m T^{\Delta-m}.$$

For the third one, we introduce a covariance term

$$E\bar{Q}^2(1+|\bar{R}|)^{\delta} \leqslant \operatorname{Cov}(\bar{Q}^2, (1+|\bar{R}|)^{\delta}) + E\bar{Q}^2 E(1+|\bar{R}|)^{\delta}.$$

The latter is bounded with $|Q|_2^2 |R|_2^{\delta} \leqslant c^{\Delta}(\sqrt{n})^{\Delta}$. The covariance is verified as follows by using the 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 the truncation $T^{m-\delta} = n^{m-3}/\lambda(q)$. At this point, the tree terms of the decomposition are of the same order:

$$E|Q|^2(1+|R|)^{\delta} \lesssim \begin{cases} \left(n^{3m-2\Delta}\kappa(q)^{m-\Delta}\right)^{1/(m-\delta-1)} & \text{under } \kappa\text{-dependence},\\ \left(n^{5m-3\Delta}\lambda(q)^{m-\Delta}\right)^{1/(m-\delta)} & \text{under } \lambda\text{-dependence}.\end{cases}$$

Let $q = N^b$. We note that $n \leqslant N/2$ and this term is of order:
$N^{(3m-2\Delta+b\kappa(\Delta-m))/(m-\delta-1)}$ under κ-weak dependence,
$N^{(5m-3\Delta+b\lambda(\Delta-m))/(m-\delta)}$ under λ-weak dependence.

Those terms are thus negligible with respect to $N^{\Delta/2}$ if

$$\kappa > \frac{3m-2\Delta-(\Delta/2)(m-\delta-1)}{b(m-\Delta)} \quad \text{under } \kappa\text{-dependence}, \tag{4.8}$$

$$\lambda > \frac{5m-3\Delta-(\Delta/2)(m-\delta)}{b(m-\Delta)} \quad \text{under } \lambda\text{-dependence}. \tag{4.9}$$

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

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

Using the relation between C and c, we conclude that the inequality (4.7) is also true at step N if the constant C satisfies the condition

$$(2^{-\delta/2}+\xi^{\Delta})C^{\Delta}+5\cdot 2^{-\delta/2}c^{\Delta}+a_1 N^{-a_2} \leqslant C^{\Delta}.$$

Choose

$$C > \left(\frac{5c^{\Delta}+a_1 2^{\delta/2}}{2^{\delta/2}-1}\right)^{1/\Delta} \quad \text{with } c = \sum_{k\in\mathbf{Z}} |\operatorname{Cov}(X_0, X_k)|.$$

Then the previous relation holds for some $0 < \xi < 1$. Finally, we use (4.8) and (4.9) to find a condition on δ.

In the case of κ-weak dependence, we rewrite inequality (4.8) as

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

which leads to the following condition on δ:

$$\delta < \frac{\sqrt{(2\kappa-3-\zeta)^2+4(\kappa\zeta-2\zeta-1)}+\zeta+3-2\kappa}{2} \wedge 1 = A. \tag{4.10}$$

We do the same in the case of the λ-weak dependence and we obtain

$$\delta < \frac{\sqrt{(2\lambda-6-\zeta)^2+4(\lambda\zeta-4\zeta-2)}+\zeta+6-2\lambda}{2} \wedge 1 = B. \blacksquare \tag{4.11}$$

Remark 4.4. The bounds A and B are always smaller than ζ.

4.3. Proofs of Theorems 2.1 and 2.2. Let $S = (1/\sqrt{n})S_n$ 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 $k = k(n) = n/[p(n)+q(n)]$,

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

where $B_j =](p+q)(j-1), (p+q)(j-1)+p] \cap N$ 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 $q = q(n)$ from each other. We denote by 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, let

$$\sigma_p^2 = \mathrm{Var}(U_1)/p = \sum_{|i|<p} (1-|i|/p) EX_0 X_i$$

and

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

where the Gaussian variables V_j are mutually independent and also independent of the sequence $(X_n)_{n\in Z}$. We also consider a sequence $U_1^*, \ldots, U_k^*$ of mutually independent random variables with the same distribution as U_1 and we let $Z^* = (U_1^* + \ldots + U_k^*)/\sqrt{n}$. In the whole section, we fix $t \in R$ and we define $f: R \to C$ by $f(x) = \exp(itx)$. Then

$$Ef(S)-f(\sigma N) = Ef(S)-f(Z)+Ef(Z)-f(Z^*)+Ef(Z^*)-f(Y)+Ef(Y)-f(\sigma N).$$

The Lindeberg method is devoted to prove that this expression converges to 0 as $n \to \infty$. The first and last terms in this equality 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 analogous 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, 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 separate subsections. In the last one, we collect these calculations to obtain the central limit theorem.

4.3.1. *Auxiliary terms.* Using Taylor expansions up to the second order, we obtain

$$|Ef(S)-f(Z)| \leqslant \|f'\|_\infty E|S-Z|$$

and

$$|Ef(Y)-f(\sigma N)| \leqslant \frac{\|f''\|_\infty^2}{2} |EY^2-\sigma^2|.$$

We note that $Z-S = (V_1+\ \ldots\ +V_k)/\sqrt{n}$ is a sum of X_i's for which the number of terms is less than or equal to $(k+1)q+p$. Then inequalities (4.6) and (4.5), under conditions (4.3) or (4.2), respectively, entail

$$(E|Z-S|)^2 \leqslant E|Z-S|^2 \lesssim ((k+1)q+p)/n.$$

Now $Y \sim \sqrt{kp/n}\,\sigma_p N$ and, consequently,

$$|EY^2-\sigma^2| \leqslant |kp/n-1|\,\sigma_p^2+|\sigma_p^2-\sigma^2|.$$

Since $|kp/n-1|^2 \lesssim ((k+1)q+p)/n$, it remains to bound the quantity

$$|\sigma_p^2-\sigma^2| \leqslant \sum_{|i|<p} \frac{|i|}{p} |EX_0X_i| + \sum_{|i|>p} |EX_0X_i|.$$

Let $a_i = |EX_0X_i|$. Under the condition (4.3) or (4.2) (respectively), the series $\sum_{i=0}^{\infty} a_i$ converge. Thus $s_j = \sum_{i=j}^{\infty} a_i \to 0$ as $j \to \infty$ and

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

The Cesàro lemma entails that the term $|\sigma_p^2-\sigma^2|$ converges to 0. Hence $|Ef(S)-f(Z)|+|Ef(Y)-f(\sigma N)|$ tends to 0 as $n \uparrow \infty$. To determine the convergence rate, we assume that $a_i = \mathcal{O}(i^{-\alpha})$ for some $\alpha > 1$; then

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

Observing that $a_i = EX_0 X_i = \mathrm{Cov}(X_0, X_i)$, we then use inequalities (4.5) and (4.6) and we find $\alpha = \kappa$ or $\alpha = \lambda(m-2)/(m-1)$ depending on the weak-dependence setting. With $p = n^a$, $q = n^b$ for two constants a and b and by the relation $\|f^{(j)}\|_\infty \leqslant |t|^j$, those bounds become, up to a constant,

$$|t|\,(n^{(b-a)/2}+n^{(a-1)/2})+t^2\,(n^{b-a}+n^{a(1-\kappa)})$$

for κ-weak dependence, and

$$|t|\,(n^{(b-a)/2}+n^{(a-1)/2})+t^2\,(n^{b-a}+n^{a(1-\lambda(m-2)/(m-1))})$$

for λ-weak dependence.

4.3.2. *Main terms.* It remains to verify the second and third terms of the sum. They are bounded as usual as follows:

$$|Ef(Z)-f(Z^*)| \leqslant \sum_{j=1}^{k} |E\Delta_j|, \qquad |Ef(Z^*)-f(Y)| \leqslant \sum_{j=1}^{k} |E\Delta_j'|,$$

where

$$\Delta_j = f(W_j+x_j)-f(W_j+x_j^*) \qquad \text{for } j = 1, \ldots, k$$

with

$$x_j = \frac{1}{\sqrt{n}}U_j, \quad x_j^* = \frac{1}{\sqrt{n}}U_j^*, \quad W_j = w_j+\sum_{i>j} x_i^*, \quad w_j = \sum_{i<j} x_i$$

and

$$\Delta_j' = f(W_j'+x_j^*)-f(W_j'+x_j') \qquad \text{for } j = 1, \ldots, k$$

with

$$x_j' = \frac{1}{\sqrt{n}}U_j', \qquad W_j' = \sum_{i<j} x_i^* + \sum_{i>j} x_i'.$$

Using the special form of f and the independence properties of the variables U_i^* and U_i', we can write

$$E\Delta_j = \big(Ef(w_j)\,f(x_j)-Ef(w_j)\,Ef(x_j^*)\big)\,Ef\big(\sum_{i>j} x_i^*\big),$$

$$E\Delta_j' = \big(Ef(x_j^*)-Ef(x_j')\big)\,Ef(W_j').$$

We then verify the two terms $Ef(\sum_{i>j} x_i^*)$ and $Ef(W_j')$ by the fact that $\|f\|_\infty \leqslant 1$ and we use the coupling to introduce a covariance term:

$$|E\Delta_j| \leqslant \big|\mathrm{Cov}\big(f(\sum_{i<j} x_i),\, f(x_j)\big)\big|, \qquad |E\Delta_j'| \leqslant |Ef(x_j^*)-Ef(x_j')|.$$

- For Δ_j, we use the weak dependence. To do this, write

$$|E\Delta_j| = |\mathrm{Cov}\,[F(X_m,\, m\in B_i,\, i<j),\, G(X_m,\, m\in B_j)]|,$$

with $F(z_1, \ldots, z_{kp}) = f(\sum_{i<j} u_i/\sqrt{n})$, where $u_i = \sum_{l\in B_i} z_l$. We shall verify that $\|F\|_\infty \leqslant 1$ and we control $\operatorname{Lip} F$:

$$\left|f\left(\frac{1}{\sqrt{n}}\sum_{i<j}\sum_{l\in B_i} z_l\right)-f\left(\frac{1}{\sqrt{n}}\sum_{i<j}\sum_{l\in B_i} z_l'\right)\right|$$
$$\leqslant \left|1-\exp\left(it\left(\frac{1}{\sqrt{n}}\sum_{i<j}\sum_{l\in B_i}(z_l-z_l')\right)\right)\right| \leqslant \frac{|t|}{\sqrt{n}}\sum_{l=1}^{kp}|z_l-z_l'|.$$

Similarly, for $G(z_1, \ldots, z_p) = f(\sum_{i=1}^{p} z_i/\sqrt{n})$, we have $\|G\|_\infty = 1$ and $\operatorname{Lip} G \preceq |t|/\sqrt{n}$. We then distinguish the two cases of weak dependence, observing that the gap between the left and right terms in the covariance is at least q.

– Under the κ-weak dependence condition:

$$|E\Delta_j| \preceq kp\cdot p\cdot\frac{|t|}{\sqrt{n}}\cdot\frac{|t|}{\sqrt{n}}\cdot\kappa(q).$$

– Under the λ-weak dependence condition:

$$|E\Delta_j| \preceq \left(kp\cdot p\cdot\frac{|t|}{\sqrt{n}}\cdot\frac{|t|}{\sqrt{n}}+kp\cdot\frac{|t|}{\sqrt{n}}+p\cdot\frac{|t|}{\sqrt{n}}\right)\cdot\lambda(q).$$

Note that these bounds do not depend on j:

$$|Ef(Z)-f(Z^*)| \preceq \begin{cases} kp\cdot t^2\cdot\kappa(q) & \text{under } \kappa,\\ kp\cdot(t^2+|t|\sqrt{k/p})\cdot\lambda(q) & \text{under } \lambda.\end{cases}$$

Since $p = n^a$, $q = n^b$, $\kappa(r) = \mathcal{O}(r^{-\kappa})$ or $\lambda(r) = \mathcal{O}(r^{-\lambda})$, these convergence rates become $n^{1-\kappa b}$ or $n^{1+(1/2-a)_+-\lambda b}$, respectively, in the κ or λ dependence context.

- For Δ_j', Taylor expansions up to order 2 or 3, respectively, give

$$|f(x_j^*)-f(x_j')| \leqslant |x_j^*-x_j'|\,\|f'\|_\infty+\tfrac{1}{2}(x_j^*-x_j')^2\,\|f''\|_\infty+r_j,$$
$$r_j \leqslant \tfrac{1}{2}\|f''\|_\infty(x_j^*-x_j')^2 \quad\text{ or }\quad r_j \leqslant \tfrac{1}{6}\|f'''\|_\infty|x_j^*-x_j'|^3.$$

For an arbitrary $\delta\in[0, 1]$, we have

$$\begin{aligned} Er_j &\preceq E\left(t^2(|x_j^*|^2+|x_j'|^2)\wedge|t|^3(|x_j^*|^3+|x_j'|^3)\right)\\ &\preceq E(t^2|x_j^*|^2\wedge|t|^3|x_j^*|^3)+E(t^2|x_j'|^2\wedge|t|^3|x_j'|^3)\\ &\preceq |t|^{2+\delta}(E|x_j^*|^{2+\delta}+E|x_j'|^{2+\delta}).\end{aligned}$$

By the stationarity of the sequence $(X_i)_{i\in\mathbf{Z}}$, we obtain

$$|E\Delta_j'| \preceq |t|^{2+\delta}n^{-1-\delta/2}(E|S_p|^{2+\delta}\vee p^{1+\delta/2}).$$

Lemma 4.2 allows us to find a bound for $E|S_p|^{2+\delta}$. If $\kappa > 2+1/\zeta$ or $\lambda > 4+2/\zeta$, where $\kappa(r) = \mathcal{O}(r^{-\kappa})$ or $\lambda(r) = \mathcal{O}(r^{-\lambda})$, then there exist $\delta\in]0, \zeta\wedge 1[$ and

$C > 0$ such that

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

We then obtain

$$|Ef(Z^*) - f(Y)| \lesssim |t|^{2+\delta} k(p/n)^{1+\delta/2}.$$

Since $p = n^a$, this bound is of order $n^{(a-1)\delta/2}$ in both κ- and λ-weak dependence settings.

We now collect the previous bounds to conclude that a multidimensional CLT holds under assumptions of both Theorems 2.1 and 2.2. Tightness follows from the Kolmogorov–Chentsov criterion (see [2]) and Lemma 4.2; thus both Theorems 2.1 and 2.2 follow from repeated application of the previous CLT. ■

4.4. Rates of convergence. Rates of convergence are now presented in two propositions of independent interest. We evaluate explicit bounds for both the difference of characteristic functions and the Berry–Esseen inequalities.

PROPOSITION 4.1. *Let $(X_t)_{t\in\mathbf{Z}}$ be a weakly dependent stationary process satisfying* (2.1) *with $m = 2+\zeta$, then the difference between the characteristic functions is bounded by*

$$\left|E\left(\exp(itS_n/\sqrt{n}) - \exp(it\sigma N)\right)\right| = o(n^{-c})$$

for some $c < c^*$ and all $t \in \mathbf{R}$, where c^* depends on the weak dependent coefficients as follows:

- *under κ-weak dependence, if $\kappa(r) = \mathcal{O}(r^{-\kappa})$ for $\kappa > 2+1/\zeta$, then*

$$c^* = \frac{(\kappa-1)A}{A+2\kappa(1+A)},$$

where

$$A = \frac{\sqrt{(2\kappa-3-\zeta)^2+4(\kappa\zeta-2\zeta-1)}+\zeta+3-2\kappa}{2} \wedge 1;$$

- *under λ-weak dependence, if $\lambda(r) = \mathcal{O}(r^{-\lambda})$ for $\lambda > 4+2/\zeta$, then*

$$c^* = \frac{(\lambda+1)B}{2+B+2\lambda(1+B)},$$

where

$$B = \frac{\sqrt{(2\lambda-6-\zeta)^2+4(\lambda\zeta-4\zeta-2)}+\zeta+6-2\lambda}{2} \wedge 1.$$

We use the following Esseen inequality in Proposition 4.2 below.

THEOREM 4.1 ([24], Theorem 5.1, p. 142). *Let X and Y be two random variables and assume that Y is Gaussian. Let F and G be their distribution* *functions with corresponding characteristic functions f and g. Then, for every $T > 0$, we have for suitable constants b and c:*

$$\sup_{x\in\mathbb{R}} |F(x)-G(x)| \leqslant b \int_{-T}^{T} \left|\frac{f(t)-g(t)}{t}\right| dt + \frac{c}{T}. \tag{4.12}$$

PROPOSITION 4.2 (A rate in the Berry–Esseen bounds). *Let $(X_t)_{t\in\mathbb{Z}}$ be a real stationary process satisfying the assumptions of Proposition 4.1. Then*

$$\sup_x |F_n(x)-\Phi(x)| = o(n^{-c})$$

with $c < c'$, where $c' = c^*/(3+A)$ or $c' = c^*/(3+B)$ in κ- or λ-weak dependence contexts, respectively (A, B and c^* are defined in Proposition 4.1).

Proof of Proposition 4.1. In the previous section, the different terms have already been bounded as follows:

• In the κ-weak dependence case, the exponents of n in the bounds obtained in Section 4.3 are

– for the auxiliary terms: $(b-a)/2$, $(a-1)/2$ and $a(1-\kappa)$,

– for the main terms: $1-\kappa b$ and $(a-1)\delta/2$.

Since $\delta < 1$ and $b < a < 1$, we have $(a-1)\delta/2 > (a-1)/2$ and $1-\kappa b > a(1-\kappa)$. The only rate of the auxiliary term it remains to consider is $(b-a)/2$ and we obtain

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

We conclude with standard calculations and with the help of the inequality $\delta < A$ (see (4.10)).

• In the λ-weak dependence case we have

– for the auxiliary terms: $(b-a)/2$, $(a-1)/2$ and $a(1-\lambda)$,

– for the main terms: $1+(1/2-a)_+-\lambda b$ and $(a-1)\delta/2$.

Only three rates give the asymptotic: $(a-1)\delta/2$, $1+(\frac{1}{2}-a)_+-\lambda b$ and $(b-a)/2$. In the previous case, the optimal choice of a^* was smaller than $\frac{1}{2}$. Then we have to consider here the rate $\frac{3}{2}-a-\lambda b$ and not $1-\lambda b$. Thus

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

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

Finally, we obtain a rate of n^{-c^*} using the inequality (4.11). ■

Proof of Proposition 4.2. Let us choose a^* and b^* as in the proof of Proposition 4.1. Now we need to make precise the impact of t on the different term of the bound of the $\mathbb{L}^1$ distance between the characteristic functions of S and σN. Up to a constant independent of t, the Kolmogorov distance is bounded by $(|t|+t^2+|t|^{2+C})n^{-c*}$. Here $C = A$ or B in the two contexts of dependence. Using Theorem 4.1 for a well-chosen value of T, we obtain the assertion of Proposition 4.2. ∎

4.5. Proof of Lemma 3.1. The case of Lipschitz functions of dependent inputs is divided into two sections devoted, respectively, to the definition of such models and to their weak dependence properties.

4.5.1. *Existence.* Let $Y^{(s)} = (Y_{-i}\mathbf{1}_{|i|<s})_{i\in\mathbf{Z}}$, $Y_+^{(s)} = (Y_{-i}\mathbf{1}_{-s<i\leqslant s})_{i\in\mathbf{Z}}$ for $s\in\mathbf{Z}$ and $H(Y^{(\infty)}) = \lim_{s\to\infty} H(Y^{(s)})$. In order to prove the existence of the Bernoulli shift with dependent inputs, we show that X_0 is the sum of a normally convergent series in $\mathbf{L}^m$; formally,

$$X_0 = H(Y^{(\infty)}) = H(0)+\big(H(Y^{(1)})-H(0)\big)$$

$$+\sum_{s=1}^{\infty} H(Y^{(s+1)})-H(Y_+^{(s)})+\big(H(Y_+^{(s)})-H(Y^{(s)})\big).$$

From (2.4) we obtain

$$\|H(Y^{(1)})-H(0)\|_m \leqslant b_0\|Y_0\|_m,$$

$$\|H(Y^{(s+1)})-H(Y_+^{(s)})\|_m \leqslant b_{-s}\|Y_{-s}\|_m,$$

$$\|H(Y_+^{(s)})-H(Y^{(s)})\|_m \leqslant b_s\|Y_s\|_m.$$

By $(Y_t)_{t\in\mathbf{Z}}$'s stationarity we get

$$\|X_0\|_m \leqslant \|H(Y^{(1)})-H(0)\|_m+\sum_{s=1}^{\infty}\|H(Y^{(s+1)})-H(Y_+^{(s)})\|_m \tag{4.13}$$

$$+\|H(Y_+^{(s)})-H(Y^{(s)})\|_m \leqslant \sum_{i\in\mathbf{Z}} b_i\|Y_0\|_m.$$

Analogously, the process $X_t = H(Y_{t-i},\, i\in\mathbf{Z})$ is well defined as the sum of a normally convergent series in $\mathbf{L}^m$. The stationarity of $(X_t)_{t\in\mathbf{Z}}$ follows from that of the input process $(Y_t)_{t\in\mathbf{Z}}$.

4.5.2. *Weak dependence properties.* Let $X_n^{(r)} = H(Y^{(r)})$ and

$$X_{\mathbf{s}} = (X_{s_1},\ldots,X_{s_u}),\qquad X_{\mathbf{t}} = (X_{t_1},\ldots,X_{t_v})$$

for any $k\geqslant 0$ and any $(u+v)$-tuple such that $s_1<\ldots<s_u\leqslant s_u+k\leqslant t_1<\ldots<t_v$. Then we have, for all functions f, g satisfying $\|f\|_\infty, \|g\|_\infty\leqslant 1$ and $\operatorname{Lip} f+\operatorname{Lip} g<\infty$,

$$\begin{aligned}|\operatorname{Cov}(f(X_{\mathbf{s}}), g(X_{\mathbf{t}}))| &\leqslant |\operatorname{Cov}(f(X_{\mathbf{s}})-f(X_{\mathbf{s}}^{(r)}), g(X_{\mathbf{t}}))|\\ &\quad+|\operatorname{Cov}(f(X_{\mathbf{s}}^{(r)}), g(X_{\mathbf{t}})-g(X_{\mathbf{t}}^{(r)}))|\\ &\quad+|\operatorname{Cov}(f(X_{\mathbf{s}}^{(r)}), g(X_{\mathbf{t}}^{(r)}))|\\ &:= T_1+T_2+T_3.\end{aligned} \tag{4.14}$$

Using the fact that $\|g\|_\infty \leqslant 1$, we bound the first term T_1 on the right-hand side of the inequality (4.14):

$$2\mathrm{Lip}\, f \cdot E\Big|\sum_{i=1}^{u} (X_{s_i}-X_{s_i}^{(r)})\Big| \leqslant 2u\,\mathrm{Lip}\, f \max_{1\leqslant i\leqslant u} E\,|X_{s_i}-X_{s_i}^{(r)}|.$$

Applying the inequality (4.13) in the case where $m = 1$, we obtain

$$E\,|X_{s_i}-X_{s_i}^{(r)}| \leqslant \sum_{i\geqslant r} b_i\,\|Y_0\|_1.$$

The second term T_2 is bounded in a similar way. The last term T_3 can be written as

$$\big|\mathrm{Cov}\big(F^{(r)}(Y_{s_i+j},\ 1\leqslant i\leqslant u,\ |j|\leqslant r),\ G^{(r)}(Y_{t_i+j},\ 1\leqslant i\leqslant v,\ |j|\leqslant r)\big)\big|,$$

where $F^{(r)}\colon \boldsymbol{R}^{u(2r+1)}\to \boldsymbol{R}$ and $G^{(r)}\colon \boldsymbol{R}^{v(2r+1)}\to \boldsymbol{R}$. Under the assumption $r\leqslant [k/2]$, we use the ε- or λ-weak dependence of Y (where $\varepsilon=\eta$) in order to bound this covariance term by

$$\psi\big(\mathrm{Lip}\, F^{(r)},\ \mathrm{Lip}\, G^{(r)},\ u(2r+1),\ v(2r+1)\big)\,\varepsilon_{k-2r}$$

with $\psi(u, v, a, b) = ua+vb$ or $\psi(u, v, a, b) = uvab+ua+vb$, respectively. We compute

$$\mathrm{Lip}\, F^{(r)} = \sup \frac{\big|f\big(H(x_{s_i+l},\ 1\leqslant i\leqslant u,\ |l|\leqslant r)\big)-f\big(H(y_{s_i+l},\ 1\leqslant i\leqslant u,\ |l|\leqslant r)\big)\big|}{\sum_{i=1}^{u}\sum_{-r\leqslant l\leqslant r}|x_{s_i+l}-y_{s_i+l}|},$$

where the supremum extends to $x\neq y$, where $x, y\in \boldsymbol{R}^{u(2r+1)}$. Notice now that if x, y are sequences with $x_i = y_i = 0$ for $|i|\geqslant r$, then the repeated application of the condition (2.4) yields

$$|H(x)-H(y)| \leqslant \sum_{|i|\leqslant r} b_i\,|x_i-y_i| \leqslant L\sum_{|i|\leqslant r}|x_i-y_i|, \tag{4.15}$$

where $L = \sum_{i\in \boldsymbol{Z}} b_i$. Repeating the inequality (4.15), we obtain

$$|F^{(r)}(x)-F^{(r)}(y)| = (\mathrm{Lip}\, f)\,L\sum_{i=1}^{u}\ \sum_{-r\leqslant l\leqslant r}|x_{s_i+l}-y_{s_i+l}|$$

and we get $\mathrm{Lip}\, F^{(r)}\leqslant (\mathrm{Lip}\, f)\,L$. Similarly, $\mathrm{Lip}\, G^{(r)}\leqslant (\mathrm{Lip}\, g)\,L$.

Under η-weak dependent inputs, we bound the covariance

$$\big|\mathrm{Cov}\big(f(X_s),\ g(X_t)\big)\big| \leqslant (u\,\mathrm{Lip}\, f+v\,\mathrm{Lip}\, g)\Big[2\sum_{|i|\geqslant r} b_i\,\|Y_0\|_1+(2r+1)\,L\eta_Y(k-2r)\Big].$$

Under λ-weak dependent inputs we have

$$\big|\mathrm{Cov}\big(f(X_s),\ g(X_t)\big)\big| \leqslant (u\,\mathrm{Lip}\, f+v\,\mathrm{Lip}\, g+uv\,\mathrm{Lip}\, f\ \mathrm{Lip}\, g)\times\Big[\Big\{2\sum_{|i|\geqslant r} b_i\,\|Y_0\|_1+(2r+1)\,L\lambda_Y(k-2r)\Big\}\vee(2r+1)^2L^2\lambda_Y(k-2r)\Big]. \ \blacksquare$$

4.6. Proof of Lemma 3.2

4.6.1. *Existence.* We decompose X_0 as above in the case $l = 0$. Here, we bound each of the terms by

$$|H(Y^{(1)}) - H(0)| \leqslant b_0 |Y_0|,$$

$$|H(Y^{(s+1)}) - H(Y_+^{(s)})| \leqslant b_{-s}(\|Y_+^{(s)}\|_\infty^l \vee 1)|Y_{-s}|,$$

$$|H(Y_+^{(s)}) - H(Y^{(s)})| \leqslant b_s(\|Y^{(s)}\|_\infty^l \vee 1)|Y_s|.$$

Using the Hölder inequality yields

$$E|H(Y^{(1)}) - H(0)| + \sum_{s=1}^{\infty} E|H(Y^{(s+1)}) - H(Y_+^{(s)})| + E|H(Y_+^{(s)}) - H(Y^{(s)})|$$
$$\leqslant \sum_{i \in \mathbf{Z}} 2|i| b_i (\|Y_0\|_1 + \|Y_0\|_{l+1}^{l+1}).$$

Hence the assumptions $l+1 \leqslant m'$ and $\sum_{i\in\mathbf{Z}} |i| b_i < \infty$ together imply that the variable $H(Y)$ is well defined in L^1. In the same manner, the process $X_n = H(Y_{n-i}, i \in \mathbf{Z})$ is well defined. The proof extends in L^m if $m \geqslant 1$ is such that $(l+1)m \leqslant m'$.

4.6.2. *Weak dependence properties.* Here, we exhibit some Lipschitz functions, and then we truncate inputs. We write $\bar{Y} = Y \vee (-T) \wedge T$ for a truncation T set below. Let us put $X_n^{(r)} = H(Y^{(r)})$ and $\bar{X}_n^{(r)} = H(\bar{Y}^{(r)})$. Furthermore, for any $k \geqslant 0$ and any $(u+v)$-tuple such that $s_1 < \ldots < s_u \leqslant s_u + k \leqslant t_1 < \ldots < t_v$, we set

$$X_s = (X_{s_1}, \ldots, X_{s_u}), \quad X_t = (X_{t_1}, \ldots, X_{t_v})$$

and

$$\bar{X}_s^{(r)} = (\bar{X}_{s_1}^{(r)}, \ldots, \bar{X}_{s_u}^{(r)}), \quad \bar{X}_t^{(r)} = (\bar{X}_{t_1}^{(r)}, \ldots, \bar{X}_{t_v}^{(r)}).$$

Then we have, for all f, g satisfying $\|f\|_\infty, \|g\|_\infty \leqslant 1$ and $\mathrm{Lip}\, f + \mathrm{Lip}\, g < \infty$,

$$\begin{aligned} (4.16) \qquad |\mathrm{Cov}(f(X_s), g(X_t))| &\leqslant |\mathrm{Cov}(f(X_s) - f(\bar{X}_s^{(r)}), g(X_t))| \\ &\quad + |\mathrm{Cov}(f(\bar{X}_s^{(r)}), g(X_t) - g(\bar{X}_t^{(r)}))| \\ &\quad + |\mathrm{Cov}(f(\bar{X}_s^{(r)}), g(\bar{X}_t^{(r)}))| \\ &:= U_1 + U_2 + U_3. \end{aligned}$$

Using the fact that $\|g\|_\infty \leqslant 1$, the first term U_1 on the right-hand side of the inequality (4.16) is bounded by

$$2u \,\mathrm{Lip}\, f\big(\max_{1\leqslant i\leqslant u} E|X_{s_i} - X_{s_i}^{(r)}| + \max_{1\leqslant i\leqslant u} E|X_{s_i}^{(r)} - \bar{X}_{s_i}^{(r)}|\big).$$

By the same arguments as those used in the proof of the existence of $H(Y^{(\infty)})$, we infer that the term U_1 is bounded by

$$\sum_{i \geqslant s} 2|i|\, b_i (\|Y_0\|_1 + \|Y_0\|_{l+1}^{l+1}).$$

Notice now that if x, y are sequences with $x_i = y_i = 0$ for $|i| \geqslant r$, then an infinitely repeated application of the inequality (2.4) yields

$$|H(x) - H(y)| \leqslant L(\|x\|_\infty^l \vee \|y\|_\infty^l \vee 1)\,\|x - y\|, \tag{4.17}$$

where $L = \sum_{i \in \mathbf{Z}} b_i < \infty$ because $\sum_{i \in \mathbf{Z}} |i|\, b_i < \infty$. The second term U_2 is bounded by using (4.17):

$$\begin{aligned} E|X_{s_i}^{(r)} - \bar{X}_{s_i}^{(r)}| &= E|H(Y^{(r)}) - H(\bar{Y}^{(r)})| \\ &\leqslant LE\Big(\big(\max_{-r \leqslant i \leqslant r} |Y_i|\big)^l \sum_{-r \leqslant j \leqslant r} |Y_j|\, \mathbf{1}_{Y_j \geqslant T}\Big) \\ &\leqslant L(2r+1)^2 E\big(\max_{-r \leqslant i, j \leqslant r} |Y_i|^l |Y_j|\, \mathbf{1}_{|Y_j| \geqslant T}\big) \\ &\leqslant L(2r+1)^2 \|Y_0\|_{m'}^{m'} T^{l+1-m'}. \end{aligned}$$

The last term U_3 can be written as

$$\big|\mathrm{Cov}\big(\bar{F}^{(r)}(Y_{s_i+j},\ 1 \leqslant i \leqslant u,\ |j| \leqslant r),\ \bar{G}^{(r)}(Y_{t_i+j},\ 1 \leqslant i \leqslant v,\ |j| \leqslant r)\big)\big|,$$

where $\bar{F}^{(r)}\colon \mathbf{R}^{u(2r+1)} \to \mathbf{R}$ and $\bar{G}^{(r)}\colon \mathbf{R}^{u(2r+1)} \to \mathbf{R}$. Under the assumption $r \leqslant [k/2]$, we use the ε- or λ-weak dependence of Y (where $\varepsilon = \eta$) in order to bound this covariance term by

$$\psi(\mathrm{Lip}\, \bar{F}^{(r)},\ \mathrm{Lip}\, \bar{G}^{(r)},\ u(2r+1),\ v(2r+1))\, \varepsilon_{k-2r}$$

with $\psi(u, v, a, b) = uvab$ or $\psi(u, v, a, b) = uvab + ua + vb$, respectively. We evaluate

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

where the supremum extends to $(x_1, \ldots, x_u) \neq (y_1, \ldots, y_u)$, where $x_i, y_i \in \mathbf{R}^{2r+1}$. Using (4.17) we get

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

We thus obtain $\mathrm{Lip}\, F^{(r)} \leqslant \mathrm{Lip}\, f \cdot L \cdot T^l$. Similarly, $\mathrm{Lip}\, G^{(r)} \leqslant \mathrm{Lip}\, g \cdot L \cdot T^l$.

Under η-weak dependent inputs, we bound the covariance

$$|\mathrm{Cov}(f(X_s), g(X_t))| \leqslant (u \operatorname{Lip} f + v \operatorname{Lip} g)\Big\{4 \sum_{|i| \geqslant r} |i|\, b_i (\|Y_0\|_1 + \|Y_0\|_{l+1}^{l+1})$$

$$+ (2r+1) L\big((2r+1) 2 \|Y_0\|_{m'}^{m'} T^{l+1-m'} + T^l \eta_Y(k-2r)\big)\Big\}.$$

We then fix the truncation

$$T^{m'-1} = \frac{2(2r+1)\|Y_0\|_{m'}^{m'}}{\eta_Y(k-2r)}$$

to obtain the assertion of Lemma 3.2 in the η-weak dependent case.

Under λ-weak dependent inputs we have

$$|\mathrm{Cov}(f(X_s), g(X_t))|$$

$$\leqslant (u \operatorname{Lip} f + v \operatorname{Lip} g + uv \operatorname{Lip} f \operatorname{Lip} g)\Big(\Big\{4 \sum_{|i| \geqslant r} |i|\, b_i (\|Y_0\|_1 + \|Y_0\|_{l+1}^{l+1})$$

$$+ (2r+1) L\big(2(2r+1) T^{l+1-m'} \|Y_0\|_{m'}^{m'} + T^l \lambda_Y(k-2r)\big)\Big\}$$

$$\vee \big\{(2r+1)^2 L^2 T^{2l} \lambda_Y(k-2r)\big\}\Big).$$

We then set a truncation such that

$$T^{l+m'-1} = \frac{2\|Y_0\|_{m'}^{m'}}{L\lambda_Y(k-2r)}$$

to obtain the assertion of Lemma 3.2 in the η-weak dependent case. ∎

Acknowledgements. We wish to thank anonymous referees for their comments; the first one pointed out the reference [15]; we then exhibited examples for which this method based on martingale arguments does not seem to apply. The second referee helped us to improve on the readability of the manuscript. We are deeply grateful to Alain Latour who made a critical review of the drafts and with whom we have worked on the final version of this paper.

REFERENCES

[1] D. W. K. Andrews, *Nonstrong mixing autoregressive processes*, J. Appl. Probab. 21 (4) (1984), pp. 930–934.

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

[3] S. Borovkova, R. Burton, and H. Dehling, *Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation*, Trans. Amer. Math. Soc. 353 (11) (2001), pp. 4261–4318.

[4] A. Bulinski and A. Shashkin, *Strong invariance principle for dependent random fields*, preprint, 2005.

[5] J. Dedecker and P. Doukhan, *A new covariance inequality and applications*, Stochastic Process. Appl. 106 (1) (2003), pp. 63–80.

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

[7] P. Doukhan, *Mixing*, Lecture Notes in Statist. 85, Springer, New York 1994.

[8] P. Doukhan and G. Lang, *Rates in the empirical central limit theorem for stationary weakly dependent random fields*, Stat. Inference Stoch. Process. 5 (2) (2002), pp. 199–228.
[9] P. Doukhan, G. Lang, D. Surgailis, and M.-C. Viano, *Functional limit theorem for the empirical process of a class of Bernoulli shifts with long memory*, J. Theoret. Probab. 18 (1) (2005), pp. 161–186.
[10] P. Doukhan and S. Louhichi, *A new weak dependence condition and applications to moment inequalities*, Stochastic Process. Appl. 84 (2) (1999), pp. 313–342.
[11] P. Doukhan and F. Portal, *Principe d'invariance faible pour la fonction de répartition empirique dans un cadre multidimensionnel et mélangeant*, Probab. Math. Statist. 8 (1987), pp. 117–132.
[12] P. Doukhan, G. Teyssière, and P. Winant, *A LARCH(∞) vector valued process*, in: *Dependence in Probability and Statistics*, P. Bertail, P. Doukhan, and P. Soulier (Eds.), Lecture Notes in Statist. 187, Springer, New York 2006.
[13] M. Duflo, *Algorithmes stochastiques*, Math. Appl. (Berlin) 23 (1986).
[14] M. I. Gordin, *The central limit theorem for stationary processes*, Dokl. Akad. Nauk SSSR 188 (1969), pp. 739–741.
[15] C. C. Heyde, *On the central limit theorem and iterated logarithm law for stationary processes*, Bull. Austral. Math. Soc. 12 (1975), pp. 1–8.
[16] I. A. Ibragimov, *A note on the central limit theorem for dependent random variables*, Theory Probab. Appl. 20 (1975), pp. 135–141.
[17] G. Lang and P. Soulier, *Convergence de mesures spectrales aléatoires et applications à des principes d'invariance*, Stat. Inference Stoch. Process. 3 (1–2) (2000), pp. 41–51.
[18] S. Louhichi, *Rates of convergence in the CLT for some weakly dependent random variables*, Teor. Veroyatnost. i Primenen. 46 (2) (2001), pp. 345–364.
[19] S. Louhichi and P. Soulier, *Marcinkiewicz–Zygmund strong laws for infinite variance time series*, Stat. Inference Stoch. Process. 1 (2) (2000), pp. 31–40.
[20] F. Merlevède, M. Peligrad, and S. Utev, *Recent advances in invariance principles for stationary sequences*, Probability Surveys 3 (2006), pp. 1–36.
[21] C. M. Newman and A. L. Wright, *An invariance principle for certain dependent sequences*, Ann. Probab. 9 (4) (1981), pp. 671–675.
[22] M. Peligrad and S. Utev, *Central limit theorem for stationary linear processes*, Ann. Probab. (2006).
[23] M. Peligrad and S. Utev, *Invariance principle for stochastic processes with short memory*, IMS Lecture Notes Monogr. Ser. (2006).
[24] V. V. Petrov, *Limit theorems of probability theory*, Oxford Stud. Probab. 4 (1995).
[25] L. D. Pitt, *Positively correlated normal variables are associated*, Ann. Probab. 10 (2) (1982), pp. 496–499.
[26] E. Rio, *Théorie asymptotique des processus aléatoires faiblement dépendants*, Math. Appl. (Berlin) 31 (2000).
[27] D. Volný, *Approximating martingales and the central limit theorem for strictly stationary process*, Stochastic Process. Appl. 44 (1993), pp. 41–74.

Laboratoire de Statistique, CREST, France
Timbre J340
3, avenue Pierre Larousse
92240 Malakoff, France
E-mail: doukhan@ensae.fr

SAMOS, Statistique Appliquée et
Modélisation Stochastique
Université Paris 1
Centre Pierre Mendès–France
90 rue de Tolbiac
F-75634 Paris Cedex 13, France
E-mail: olivier.wintenberger@univ-paris1.fr

Received on 20.6.2006