Invariance principles for linear processes with application to isotonic regression

Jérôme Dedecker [a], Florence Merlevède [b] *and* Magda Peligrad [c] [1]

[a] Université Paris 6, LSTA, 175 rue du Chevaleret, 75013 Paris, FRANCE

[b] Université Paris Est-Marne la Vallée, LAMA and C.N.R.S UMR 8050, 5 Boulevard Descartes, 77 454 Marne La Vallée cedex 2, FRANCE

[c] Department of Mathematical Sciences, University of Cincinnati, PO Box 210025, Cincinnati, Oh 45221-0025, USA

Mathematical Subject Classification (2000): 60F17, 60G48, 60G10, 62G08.

Key words: linear processes, moment inequalities, generalizations of martingales, invariance principles, fractional Brownian motion, isotonic regression.

Abstract

In this paper we prove maximal inequalities and study the functional central limit theorem for the partial sums of linear processes generated by dependent innovations. Due to the general weights these processes can exhibit long range dependence and the limiting distribution is a fractional Brownian motion. The proofs are based on new approximations by a linear process with martingale difference innovations. The results are then applied to study an estimator of the isotonic regression when the error process is a (possibly long range dependent) time series.

1 Introduction and notations

Without loss of generality, we assume that all the strictly stationary sequences $(\xi_i)_{i\in\mathbf{Z}}$ considered in this paper are given by $\xi_i = \xi_0 \circ T^i$ where $T : \Omega \mapsto \Omega$ is a bijective bimeasurable transformation preserving the probability $\mathbf{P}$ on $(\Omega,\ \mathcal{A})$. We denote by $\mathcal{I}$ the σ-algebra of all T-invariant sets. For a subfield $\mathcal{F}_0$ satisfying $\mathcal{F}_0 \subseteq T^{-1}(\mathcal{F}_0)$, let $\mathcal{F}_i = T^{-i}(\mathcal{F}_0)$. Let $\mathcal{F}_{-\infty} = \bigcap_{n\geq 0}\mathcal{F}_{-n}$ and $\mathcal{F}_\infty = \bigvee_{k\in\mathbf{Z}}\mathcal{F}_k$. The sequence $(\mathcal{F}_i)_{i\in\mathbf{Z}}$ will be called a stationary filtration. We assume also that ξ_0 is *regular*, that is $\mathbf{E}(\xi_0|\mathcal{F}_{-\infty}) = 0$ and ξ_0 is $\mathcal{F}_\infty$-measurable. On $\mathbf{L}^2$, we define the projection operator P_j by

$$P_j(Y) = \mathbf{E}(Y|\mathcal{F}_j) - \mathbf{E}(Y|\mathcal{F}_{j-1}).$$

For any random variable Y, $\|Y\|_p$ denotes the norm in $\mathbf{L}^p$.

Recall that the linear process $X_k = \sum_{i\in\mathbf{Z}} a_i\xi_{k-i}$ is well defined in $\mathbf{L}^2$ for any $(a_i)_{i\in\mathbf{Z}}$ in ℓ^2 (i.e. $\sum_{i\in\mathbf{Z}} a_i^2 < \infty$) if and only if the stationary sequence

[1]Supported in part by a Charles Phelps Taft Memorial Fund grant and NSA grant, H98230-09-1-0005.

$(\xi_i)_{i\in\mathbf{Z}}$ has a bounded spectral density. Let $S_n = X_1 + \cdots + X_n$ and $c_{n,j} = a_{1-j} + \cdots + a_{n-j}$. In the case where ξ_0 is $\mathcal{F}_0$-measurable, Peligrad and Utev (2006-b) have proved that if the sequence $(\xi_i)_{i\in\mathbf{Z}}$ satisfies an appropriate weak dependence condition, then

$$\Big(\sum_{j\in\mathbf{Z}} c_{n,j}^2\Big)^{-1/2} S_n$$

converges in distribution to $\sqrt{\eta}N$ where η is a nonnegative $\mathcal{I}$ measurable random variable, and N is a standard normal random variable independent of η. Their result extends the classical result by Ibragimov (1962) from i.i.d ξ_i's, to the case of weakly dependent sequences. In particular, the result applies if

$$\sum_{i\in\mathbf{Z}} \|P_0(\xi_i)\|_2 < \infty\,. \tag{1}$$

Note that if this condition is satisfied, then the series $\sum_{k\in\mathbf{Z}} |\mathbf{E}(\xi_0\xi_k)|$ converges. Indeed since $\xi_k = \sum_{i\in\mathbf{Z}} P_i(\xi_k)$ and since $\mathbf{E}(P_i(\xi_0)P_j(\xi_k)) = 0$ if $i \neq j$, it follows that for any $k \in \mathbf{Z}$,

$$|\mathbf{E}(\xi_0\xi_k)| \leq \Big|\sum_{i\in\mathbf{Z}} \mathbf{E}(P_i(\xi_0)P_i(\xi_k))\Big| \leq \sum_{i\in\mathbf{Z}} \|P_0(\xi_i)\|_2 \|P_0(\xi_{k+i})\|_2\,,$$

so that $\sum_{k\in\mathbf{Z}} |\mathbf{E}(\xi_0\xi_k)| \leq \big(\sum_{i\in\mathbf{Z}} \|P_0(\xi_i)\|_2\big)^2$. In addition, under Condition (1), the nonnegative random variable η satisfies $\eta = \sum_{k\in\mathbf{Z}} \mathbf{E}(\xi_0\xi_k|\mathcal{I})$.

Condition (1) has been introduced by Hannan (1973), and by Heyde (1974) in a slightly weaker form, and is well adapted to the analysis of time series (see in particular the application to time series regression given in the paper by Hannan (1973)). As we shall see in our remark 3.3, Condition (1) is also satisfied if

$$\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \|\mathbf{E}(\xi_n|\mathcal{F}_0)\|_2 < \infty \text{ and } \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \|\xi_{-n} - \mathbf{E}(\xi_{-n}|\mathcal{F}_0)\|_2 < \infty\,, \tag{2}$$

which is weaker than the condition introduced by Gordin (1969). If ξ_0 is $\mathcal{F}_0$-measurable, the condition (2) leads to new interesting conditions for weakly dependent sequences, and can be successfully applied to functions of dynamical systems (see Section 3 in Peligrad and Utev (2006-b), and Section 6 in Dedecker, Merlevède and Volný (2007) for more details).

A natural question is now: what can we say about the weak convergence of the partial sum process

$$\Big\{\Big(\sum_{j\in\mathbf{Z}} c_{n,j}^2\Big)^{-1/2} S_{[nt]}, t \in [0,1]\Big\} \tag{3}$$

in the space $D([0,1])$ of cadlag functions equipped with the uniform topology? Since the paper by Davydov (1970) for i.i.d ξ_i's, we know that the question is not as simple as for the central limit question, and that the limiting process (when it exists) depends on the behavior of the normalizing sequence $v_n^2 = \sum_{j\in\mathbf{Z}} c_{n,j}^2$. More precisely if (1) holds, and if there exists $\beta \in]0,2]$ such that

$$\text{for any } t \in]0,1] \quad \lim_{n\to\infty} \frac{v_{[nt]}^2}{v_n^2} = t^\beta, \tag{4}$$

we show in Theorems 3.1 and 3.2 that the finite dimensional marginals of the process (3) converges in distribution to those of $\sqrt{\eta} W_H$ where W_H is a fractional Brownian motion independent of η, with Hurst index $H = \beta/2$. The question is now: under what conditions can we obtain the tightness in $D([0,1])$?

In Theorem 3.1 of Section 3.1, we show that if $\beta \in]1,2]$, then the condition (1) is sufficient for the weak convergence in $D([0,1])$. If $\beta \in]0,1]$, we point out in Theorem 3.1 that the convergence in $D([0,1])$ holds if (1) is replaced by the stronger condition

$$\sum_{i\in\mathbf{Z}} \|P_0(\xi_i)\|_q < \infty \quad \text{for } q > 2/\beta\,. \tag{5}$$

As a matter of fact, for $\beta = 1$, it is known from counter examples given in Wu and Woodroofe (2004) and also in Merlevède and Peligrad (2006) that if the sequence $(\xi_i)_{i\in\mathbf{Z}}$ is i.i.d. with $\mathbf{E}(\xi_0^2) < \infty$, then the weak invariance principle may not be true for the partial sums of the linear process, so that a reinforcement of (1) is necessary. The case $\beta = 1$, where $W_{1/2}$ is a standard Brownian motion, is of special interest, and is known as the weakly dependent case. In that case, we point out in Section 3.2 that if we make some additional assumptions on $(a_i)_{i\in\mathbf{Z}}$, then the condition (1) is sufficient for the weak invariance principle (Comments 3.1 and 3.2), or may be reinforced in a weaker way than (5) (Theorem 3.3).

Note that, with the notations above, the sum S_n may be written as

$$S_n = \sum_{i\in\mathbf{Z}} c_{n,i}\xi_i. \tag{6}$$

Consequently, to prove our main theorems, we give in Section 2 two preliminary results for linear statistics of type (6): first a moment inequality given in Proposition 2.1, and next a martingale approximation result given in Proposition 2.2, which enables to go back to the standard case where the ξ_i's are martingale differences. Both results are given in terms of Orlicz norms.

Our results provide, besides the invariance principles, estimates of the maximum of partial sums that make them appealing to studying statistics involving linear processes. In Section 4 we apply our results to the so-called isotonic regression problem

$$y_k = \phi\Big(\frac{k}{n}\Big) + X_k, \quad k = 1, 2, \ldots, n\,, \tag{7}$$

where ϕ is non-decreasing, and the error X_k is a linear process. We follow the general scheme given in Anevski and Hössjer (2006), who showed that in the context of dependent errors, the main tools to obtain the asymptotic distribution of the isotonic estimator $\hat{\phi}$ are the convergence in $D([0,1])$ of the partial sum process defined in (3), and a suitable maximal inequality for the rescaled stochastic term (see their condition (14)). Zhao and Woodroofe (2008) enlighten the fact that in addition of the weak invariance principle, it is in fact enough to prove a suitable maximal inequality directly on the partial sums of the error process. As in Anevski and Hössjer (2006), the rate of convergence of $\hat{\phi}$ is determined by the asymptotic behavior of the normalizing sequence $v_n^2 = \sum_{j\in\mathbf{Z}} c_{n,j}^2$, and the limiting distribution depends on the limiting process W_H.

2 Moments inequalities and Martingale approximation for Orlicz norms

For $\Psi : \mathbf{R}_+ \to \mathbf{R}_+$ a Young function (convex, increasing, $\Psi(0) = 0$ and $\lim_{x\to\infty} \Psi(x) = \infty$)), we denote by $\mathbf{L}_\Psi$ the Orlicz space defined as the space of all random variables X such that $\mathbf{E}\Psi(|X|/c) < \infty$ for some $c > 0$. It is a Banach space for the norm,

$$\|X\|_\Psi = \inf\{c > 0\,,\, \mathbf{E}\Psi(|X|/c) \le 1\}\,.$$

Note that when $\Psi(x) = x^q$, $1 \le q < \infty$, then $\mathbf{L}_\Psi = \mathbf{L}^q$.

Let us introduce also the following class of functions (see page 60 in de la Peña and Giné (1999)). For $\alpha > 0$, the class $\mathcal{A}_\alpha$ consists of functions $\Phi : \mathbf{R}_+ \to \mathbf{R}_+$, $\Phi(0) = 0$, Φ non-decreasing continuous and such that

$$\Phi(cx) \le c^\alpha \Phi(x) \text{ for all } c \ge 2,\ x \ge 0\,.$$

We denote also by $\mathcal{C}(\mathcal{A}_\alpha)$ the class of functions Ψ such that: Ψ is a Young function in $\mathcal{A}_\alpha$ and $x \mapsto \Psi(\sqrt{x})$ is a convex function.

Proposition 2.1 *Let $\{Y_k\}_{k\in\mathbf{Z}}$ be a sequence of random variables such that for all k, $\mathbf{E}(Y_k|\mathcal{F}_{-\infty}) = 0$ almost surely and Y_k is $\mathcal{F}_\infty$-measurable. Let Ψ be a function in $\mathcal{C}(\mathcal{A}_\alpha)$. Assume that*

$$||P_{k-j}(Y_k)||_\Psi \le p_j \quad \text{and} \quad D_\Psi \ := \sum_{j=-\infty}^{\infty} p_j < \infty\,.$$

For any positive integer m, let $\{c_{m,j}\}_{j\in\mathbf{Z}}$ be a sequence in ℓ^2. Define $S_m = \sum_{j\in\mathbf{Z}} c_{m,j} Y_j$. Then for all $m \ge 1$, there exists a positive constant C_α depending only on α such that

$$\|S_m\|_\Psi \le C_\alpha D_\Psi \Big(\sum_{j\in\mathbf{Z}} c_{m,j}^2\Big)^{1/2}\,. \tag{8}$$

Remark 2.1 *Under the notations of the above proposition, we get for the special function $\Psi(x) = x^q$ with $q \in [2, \infty[$, the following moment inequality. Assume that*

$$||P_{k-j}(Y_k)||_q \le p_j \quad \text{and} \quad D_q := \sum_{j=-\infty}^{\infty} p_j < \infty .$$

Then, for any $m \ge 1$,

$$||S_m||_q \le C_q \Big(\sum_{j \in \mathbf{Z}} c_{m,j}^2 \Big)^{1/2} D_q ,$$

where $C_q^q = 18 q^{3/2} / (q-1)^{1/2}$.

For all $j \in \mathbf{Z}$, let $d_j = \sum_{\ell \in \mathbf{Z}} P_j(\xi_\ell)$. Clearly $(d_j)_{j \in \mathbf{Z}}$ is a stationary sequence of martingale differences with respect to the filtration $(\mathcal{F}_j)_{j \in \mathbf{Z}}$.

Proposition 2.2 *For any positive integer n, let $\{c_{n,i}\}_{i \in \mathbf{Z}}$ be a sequence in ℓ^2. Let Ψ be a function in $\mathcal{C}(\mathcal{A}_\alpha)$. If $\sum_{j \in \mathbf{Z}} ||P_0(\xi_j)||_\Psi < \infty$ then we have the following martingale-difference approximation: for any positive integer m, there exists a positive constant C_α only depending on α such that*

$$\begin{aligned} \Big\| \sum_{i \in \mathbf{Z}} c_{n,i}(\xi_i - d_i) \Big\|_\Psi \;&\le\; 2C_\alpha \Big(\sum_{i \in \mathbf{Z}} c_{n,i}^2 \Big)^{1/2} \sum_{|k| \ge m} \|P_0(\xi_k)\|_\Psi \\ &\quad + 3C_\alpha m \Big(\sum_{j \in \mathbf{Z}} (c_{n,j} - c_{n,j-1})^2 \Big)^{1/2} \sum_{j \in \mathbf{Z}} ||P_0(\xi_j)||_\Psi . \end{aligned}$$

Corollary 2.1 *Let $(a_i)_{i \in \mathbf{Z}}$ be a sequence of real numbers in ℓ^2. Let Ψ be a function in $\mathcal{C}(\mathcal{A}_\alpha)$. Assume that $\xi_0 \in L_\Psi$ and $\sum_j ||P_0(\xi_j)||_\Psi < \infty$. Let $X_k = \sum_{j \in \mathbf{Z}} a_j \xi_{k-j}$ and $Y_k = \sum_{j \in \mathbf{Z}} a_j d_{k-j}$. Set $S_n = \sum_{k=1}^n X_k$ and $T_n = \sum_{k=1}^n Y_k$. Then for any positive m, there exist positive constants C_1 and C_2 such that*

$$\|S_n - T_n\|_\Psi \le C_1 v_n \sum_{|k| \ge m} \|P_0(\xi_k)\|_\Psi + C_2 m , \tag{9}$$

where $v_n^2 = \sum_{j \in \mathbf{Z}} c_{n,j}^2$, and $c_{n,j} = a_{1-j} + \cdots + a_{n-j}$.

Proof of Corollary 2.1 We apply Proposition 2.2 by noticing that $S_n - T_n = \sum_{j \in \mathbf{Z}} c_{n,j}(\xi_j - d_j)$ and that

$$\sum_{j \in \mathbf{Z}} (c_{n,j} - c_{n,j-1})^2 \le 4 \sum_{j \in \mathbf{Z}} a_j^2 .$$

$\diamond$

Using the Orlicz norms, we give the following maximal inequality which is a refinement of Inequality (6) in Proposition 1 of Wu (2007).

Lemma 2.1 *Let Ψ be a Young function. Let $p \geq 1$ and write $\Psi_p(x)$ for $\Psi(x^p)$. Let $(Y_i)_{1\leq i\leq 2^N}$ be a strictly stationary sequence of random variables such that $||Y_1||_{\Psi_p} < \infty$. Let $S_n = Y_1 + \cdots + Y_n$. Then*

$$\Big\| \max_{1\leq m\leq 2^N} |S_m| \Big\|_p \leq \sum_{L=0}^{N} \|S_{2^L}||_{\Psi_p} \Big(\Psi^{-1}(2^{N-L})\Big)^{1/p} .$$

Remark 2.2 *Clearly we can take $\Psi(x) = x$ in Lemma 2.1. Hence, in the stationary case, we recover the inequality (6) in Wu (2007).*

3 Invariance principle for linear processes

In this section we shall focus on the weak invariance principle for linear processes. Let $(a_i)_{i\in\mathbf{Z}}$ be a sequence of real numbers in ℓ^2. Let

$$X_k = \sum_{i\in\mathbf{Z}} a_i \xi_{k-i} \quad \text{and} \quad S_{[nt]} = \sum_{k=1}^{[nt]} X_k \,, \tag{10}$$

and

$$v_n^2 = \sum_{j\in\mathbf{Z}} c_{n,j}^2, \quad \text{where} \quad c_{n,j} = a_{1-j} + \cdots + a_{n-j} \quad . \tag{11}$$

The behavior of the process $\{S_{[nt]}, t \in [0,1]\}$, properly normalized, strongly depends on the behavior of the sequence $(a_i)_{i\in\mathbf{Z}}$.

In the next two sections we treat separately the case where the limit process is a mixture of Fractional Brownian motions and the case where it is a mixture of standard Brownian motions.

3.1 Convergence to a mixture of Fractional Brownian motions

Definition 3.1 *We say that a positive sequence $(v_n^2)_{n\geq 1}$ is regularly varying with exponent $\beta > 0$ if for any $t \in]0,1]$,*

$$\frac{v_{[nt]}^2}{v_n^2} \to t^\beta, \text{ as } n \to \infty \,. \tag{12}$$

We shall separate the case $\beta \in]1,2]$ from the case $\beta \in]0,1]$.

Theorem 3.1 *Let $(a_i)_{i\in\mathbf{Z}}$ in ℓ^2. Let $\beta \in]1,2]$ and assume that v_n^2 defined by (11) is regularly varying with exponent β. Let ξ_0 be a regular random variable such that $\|\xi_0\|_2 < \infty$, and let $\xi_i = \xi_0 \circ T^i$. Assume that condition (1) is satisfied. Then the process $\{v_n^{-1} S_{[nt]}, t \in [0,1]\}$ converges in $D([0,1])$ to $\sqrt{\eta} W_H$ where W_H is a standard fractional Brownian motion independent of η with Hurst index*

$H=\beta/2$, and $\eta=\sum_{k\in\mathbf{Z}}\mathbb{E}(\xi_0\xi_k|\mathcal{I})$ and there exists a positive constant C (not depending on n) such that

$$\mathbf{E}(\max_{1\le k\le n} S_k^2)\le C v_n^2\,. \tag{13}$$

Theorem 3.2 *Let $\beta\in]0,1]$ and assume that v_n^2 defined by (11) is regularly varying with exponent β. Let ξ_0 be a regular random variable such that $\|\xi_0\|_2<\infty$, and let $\xi_i=\xi_0\circ T^i$. Assume that condition (1) is satisfied. Then the finite dimensional distributions of $\{v_n^{-1}S_{[nt]},t\in[0,1]\}$ converges to the corresponding ones of $\sqrt{\eta}W_H$, where W_H is a standard fractional Brownian motion independent of η with Hurst index $H=\beta/2$, and $\eta=\sum_{k\in\mathbf{Z}}\mathbf{E}(\xi_0\xi_k|\mathcal{I})$. Assume in addition that for a $q>2/\beta$ we have $\|\xi_0\|_q<\infty$ and*

$$\sum_{j\in\mathbf{Z}}\|P_0(\xi_j)\|_q<\infty\,. \tag{14}$$

Then the process $\{v_n^{-1}S_{[nt]},t\in[0,1]\}$ converges in $D([0,1])$ to $\sqrt{\eta}W_H$ and (13) holds.

Remark 3.1 *According to Corollary 2 in Peligrad and Utev (2006-b), one has*

$$\lim_{n\to\infty}\frac{\mathrm{Var}(S_n)}{v_n^2}=\lim_{n\to\infty}\frac{\mathrm{Var}(\xi_1+\cdots+\xi_n)}{n}=v^2=\Big\|\sum_{j\in\mathbf{Z}}P_0(\xi_j)\Big\|_2^2\,.$$

Remark 3.2 *In the context of Theorem 3.1, condition (12) is necessary for the conclusion of this theorem (see Lamperti (1962)). This condition has been also imposed by Davydov (1970) to study the weak invariance principle of linear processes with i.i.d. innovations. To be more precise, Davydov proved that if (12) holds and if $\xi_0\in\mathbb{L}^q$ with $q\ge 4$ and $q>4(1/\beta-1)$, then $\{v_n^{-1}S_{[nt]},t\in[0,1]\}$ converges in $D([0,1])$ to $\sqrt{\mathbf{E}(\xi_0^2)}W_{\beta/2}$. Later, in the case $\beta>1$, Konstantopoulos and Sakhanenko (2004) sharpen Davydov's result showing that the weak invariance principle holds if the ξ_i's are iid and in $\mathbf{L}^2$.*

Example 1. For $0<d<1/2$, let us consider the linear process X_k defined by

$$X_k=(1-B)^{-d}\xi_k=\sum_{i\ge 0}a_i\xi_{k-i}\,, \tag{15}$$

where B is the lag operator, $a_0=1$ and $a_i=\frac{\Gamma(i+d)}{\Gamma(d)\Gamma(i+1)}$ for $i\ge 1$, and $(\xi_i)_{i\in\mathbf{Z}}$ is a strictly stationary sequence satisfying the condition of Theorem 3.1. In this case Theorem 3.1 applies with $\beta=2d+1$, since $a_k\sim(\Gamma(d))^{-1}k^{d-1}$.

Example 2. Now, if we consider the following choice of $(a_k)_{k\ge 0}$: $a_0=1$ and $a_i=(i+1)^{-\alpha}-i^{-\alpha}$ for $i\ge 1$ with $\alpha\in]0,1/2[$, then Theorem 3.2 applies. Indeed for this choice, $v_n^2\sim\kappa_\alpha n^{1-2\alpha}$, where κ_α is a positive constant depending on α.

Example 3. For the choice $a_i\sim i^{-\alpha}\ell(i)$ where ℓ is a slowly varying function at infinity and $1/2<\alpha<1$ then, $v_n^2\sim\kappa_\alpha n^{3-2\alpha}\ell^2(n)$ (see for instance Relations (12) in Wang *et al.* (2003)), where κ_α is a positive constant depending on α.

Example 4. Finally, if $a_i \sim i^{-1/2}(\log i)^{-\alpha}$ for some $\alpha > 1/2$, then $v_n^2 \sim n^2(\log n)^{1-2\alpha}/(2\alpha - 1)$ (see Relations (12) in Wang *et al.* (2003)). Hence (12) is satisfied with $\beta = 2$.

For the sake of applications, we now give a sufficient condition for (14) to hold.

Remark 3.3 *For any $q \in [2, \infty[$, the condition (14) is satisfied if we assume that*

$$\sum_{n=1}^{\infty} \frac{1}{n^{1/q}} \|\mathbf{E}(\xi_n|\mathcal{F}_0)\|_q < \infty \text{ and } \sum_{n=1}^{\infty} \frac{1}{n^{1/q}} \|\xi_{-n} - \mathbf{E}(\xi_{-n}|\mathcal{F}_0)\|_q < \infty . \qquad (16)$$

The fact that (16) implies (14) extends Corollary 2 in Peligrad and Utev (2006-b) and also Corollary 5 in Dedecker, Merlevède and Volný (2007) from the case $q = 2$ to more general situations.

For causal linear processes, Shao and Wu (2006) also showed that the weak invariance principle holds under the condition (14) as long as the coefficients of the linear processes satisfy a certain regularity condition. To be more precise, their condition on the coefficients of the linear processes lead either to $\beta > 1$ or to $\beta < 1$. For this last case, they specified the coefficients $(a_i)_{i\geq 0}$ as follows: for $1 < \alpha < 3/2$, $a_j = j^{-\alpha}\ell(j)$ for $j \geq 1$ (where $\ell(i)$ is a slowly varying function) and $\sum_{j=0}^{\infty} a_j = 0$ (see for instance their Lemma 4.1). For this choice, v_n^2 is regularly varying with coefficient $\beta = 3 - 2\alpha < 1$. Our Theorem 3.2 does not require conditions on the coefficients but only the fact that the variance is regularly varying which is a necessary condition.

3.2 Convergence to a mixture of Brownian motions

The case $\beta = 1$ deserves special attention. For this case the limit is a mixture of Brownian motions.

As an immediate consequence of Theorem 3.2 we formulate the following corollary for causal linear processes, under a recent condition introduced by Wu and Woodroofe (2004).

Corollary 3.1 *Let ξ_0 be a regular random variable such that $\|\xi_0\|_q < \infty$ for some $q > 2$, and let $\xi_i = \xi_0 \circ T^i$. Assume in addition that*

$$\sum_{j\in\mathbf{Z}} \|P_0(\xi_j)\|_q < \infty . \qquad (17)$$

Let $(a_i)_{i\in\mathbf{Z}}$ be a sequence of real numbers in ℓ^2 such that $a_i = 0$ for $i < 0$. Let $b_j = a_0 + \cdots + a_j$. Define $(X_k)_{k\geq 1}$ as above and assume that

$$\sum_{k=0}^{n-1} b_k^2 \to \infty, \text{ as } n \to \infty, \qquad (18)$$

and that

$$\sum_{j=0}^{\infty}(b_{n+j}-b_j)^2 = o\,\Big(\sum_{k=0}^{n-1} b_k^2\Big)\,. \tag{19}$$

Then $v_n^2 \sim nh(n)$, where $h(n)$ is a slowly varying function. Moreover, the process $\{v_n^{-1}S_{[nt]}, t\in[0,1]\}$ converges in $D([0,1])$ to $\sqrt{\eta}W$ where W is a standard Brownian motion independent of η, and $\eta=\sum_{k\in\mathbf{Z}}\mathbf{E}(\xi_0\xi_k|\mathcal{I})$. In addition (13) holds.

To prove this result, it suffices to apply Theorem 3.2, and to use the fact that under (18) and (19), $v_n^2 \sim nh(n)$ (see Wu and Woodroofe (2004)). Under the same conditions (18) and (19), Wu and Min (2005, Theorem 1) have also proved the weak invariance principle but under the stronger condition $\sum_{j\geq 0} j\|P_0(\xi_j)\|_q < \infty$ (in their paper the random variables ξ_j are adapted to the filtration $\mathcal{F}_j$).

Remark 3.4 *The above result fails if in (17) we take $q=2$. See Wu and Woodroofe (2004) and also Merlevède and Peligrad (2006, example 1 p. 657).*

Let us make some comments on the case where the condition (1) is sufficient for the weak convergence to the Brownian motion, with the normalization $\sqrt{n}$. The first case is already known, and the second case deserves a short proof.

Comment 3.1 *When $\sum_{i\in\mathbf{Z}}|a_i| < \infty$, (the short memory case) and condition (1) is satisfied one can use the result from Peligrad and Utev (2006-a) in the adapted case, showing that the invariance principle for the linear process is inherited from the innovations at no extra cost. For this case, the process $\{n^{-1/2}S_{[nt]}, t\in[0,1]\}$ converges in distribution in $D([0,1])$ to $\sqrt{\eta}W$, where W is a standard Brownian motion independent of η and $\eta = A^2\sum_{k\in\mathbf{Z}}\mathbf{E}(\xi_0\xi_k|\mathcal{I})$ with $A=\sum_{i\in\mathbf{Z}}a_i$. Moreover $\mathbf{E}(\max_{1\leq k\leq n}S_k^2)\leq Cn$. See Dedecker, Merlevède and Volný (2007), Corollaries 2 and 3 for the nonadapted case.*

Comment 3.2 *Let $(a_i)_{i\in Z}$ in ℓ^2 and assume that the series $\sum_{i\in\mathbf{Z}}a_i$ converges (meaning that the two series $\sum_{i\geq 0}a_i$ and $\sum_{i<0}a_i$ converge), and Heyde's (1975) condition (H) holds*

$$(H)\qquad \sum_{n=1}^{\infty}\Big(\sum_{k\geq n}a_k\Big)^2<\infty \quad \textit{and} \quad \sum_{n=1}^{\infty}\Big(\sum_{k\leq -n}a_k\Big)^2<\infty\,.$$

Assume also that condition (1) is satisfied. Then the same conclusion as in Comment 3.1 holds.

Example 5. Heyde's condition allows the following possibility: $\sum_{i\in\mathbf{Z}}|a_i|=\infty$ but $\sum_{i\in\mathbf{Z}}a_i$ converges. For instance, if for $n<0$, $a_n=0$, and for $n\geq 1$, $a_n=(-1)^n u_n$ for some sequence $(u_n)_{n\geq 1}$ of positive coefficients decreasing to zero, such that $\sum_{n\geq 1}u_n=\infty$, then Condition (H) is satisfied as soon as

$\sum_{n>0} u_n^2 < \infty$, which is a minimal condition. It is noteworthy to indicate that the Heyde's condition implies (19).

Now, if $\sum_{j\in\mathbf{Z}} |a_j| = \infty$ and (H) does not hold, condition (17) may still be weakened in some particular cases. The following result generalizes Corollary 4 in Dedecker, Merlevède and Volný (2007) to the case where the innovations of the linear process are not necessarily martingale differences sequences. Denote by

$$s_n^2 = n\Big(\sum_{i=-n}^{n} a_i\Big)^2 . \tag{20}$$

Theorem 3.3 *Let $(a_i)_{i\in\mathbf{Z}}$ be a sequence of real numbers in ℓ^2 but not in ℓ^1, and let s_n^2 be defined by (20). Define $(X_k)_{k\geq 1}$ as above and assume that*

$$\limsup_{n\to\infty} \frac{\sum_{i=-n}^{n} |a_i|}{\Big|\sum_{i=-n}^{n} a_i\Big|} < \infty \text{ and } \sum_{k=1}^{n} \sqrt{\sum_{|i|\geq k} a_i^2} = o(s_n)\,. \tag{21}$$

If one of the following two conditions holds

(a) $\sum_{j\in\mathbf{Z}} \|P_0(\xi_j)\|_{\Psi_{2,\alpha}} < \infty$, *where* $\Psi_{2,\alpha}(x) = x^2 \log^{\alpha}(1+x^2)$ *and* $\alpha > 2$.

or

(b) $\sum_{j\in\mathbf{Z}} \log(1+|j|)\|P_0(\xi_j)\|_2 < \infty$,

then $\{s_n^{-1} S_{[nt]}, t\in[0,1]\}$ converges weakly in $D([0,1])$ to $\sqrt{\eta}W$, where W is a standard Brownian motion independent of η and $\eta = \sum_{k\in\mathbf{Z}} \mathbf{E}(\xi_0\xi_k|\mathcal{I})$. In addition, there exists a positive constant C (not depending on n) such that

$$\mathbf{E}(\max_{1\leq k\leq n} S_k^2) \leq C s_n^2\,. \tag{22}$$

Remark 3.5 *For two positive sequences of numbers the notation $u_n \sim v_n$ means that $\lim_{n\to\infty} u_n/v_n = 1$. According to Remark 12 in Dedecker, Merlevède and Volný (2007), we have that*

$$s_n^2 \sim v_n^2 \sim nh(n)\,,$$

where $h(n)$ is a slowly varying function at infinity. In addition if we assume the first part of Condition (21) and $\sum_{j\in\mathbf{Z}} |a_j| = \infty$, we get that $s_n/\sqrt{n} \to \infty$, as $n\to\infty$.

Example 6. If we consider the following choice of $(a_k)_{k\in\mathbf{Z}}$: $a_0 = 1$ and $a_i = 1/|i|$ for $i \neq 0$, then Theorem 3.3 applies. Indeed for this choice, Condition (21) holds and $s_n \sim 2\sqrt{n}(\log n)$.

We give now a useful sufficient condition for the validity of condition (b) of Theorem 3.3.

Remark 3.6 *The condition (b) of Theorem 3.3 is satisfied if we assume that*

$$\sum_{n=1}^{\infty} \log n \frac{\|\mathbf{E}(\xi_n|\mathcal{F}_0)\|_2}{\sqrt{n}} < \infty \text{ and } \sum_{n=1}^{\infty} \log n \frac{\|\xi_{-n} - \mathbf{E}(\xi_{-n}|\mathcal{F}_0)\|_2}{\sqrt{n}} < \infty\,. \tag{23}$$

4 Application to isotonic regression

Let ϕ be a nondecreasing function on the unit interval and let

$$y_k = \phi\Big(\frac{k}{n}\Big) + X_k, \quad k = 1, 2, \ldots, n\,. \tag{24}$$

where (X_k) is a strictly stationary sequence of random variables such that $\mathbf{E}(X_k) = 0$ and $\mathbf{E}(X_k^2) < \infty$. The problem is then to estimate ϕ in a non parametric way. We denote by $S_n = \sum_{k=1}^n X_k$.

Taking advantage of the monotonicity of the regression function, isotonic estimates have been suggested. Let $\mu_k = \phi(k/n)$. It is well known that the least squares estimator

$$\hat{\mu} = \operatorname{argmin}\Big\{ \sum_{k=1}^n (y_k - \mu_k)^2, \mu_1 \le \cdots \le \mu_n \Big\},$$

is such that

$$\hat{\mu}_k = \max_{i \le k} \min_{j \ge k} \frac{y_i + \cdots + y_j}{j - i + 1}\,.$$

In addition, setting

$$Y_n(t) = \frac{1}{n}\Big(\sum_{k=1}^{[nt]} y_k\Big) \text{ and } \widetilde{Y}_n = \mathrm{GCM}(Y_n)\,,$$

where GCM designates the Greatest Convex Minorant, then

$$\hat{\mu}_k = \widetilde{Y}_n'\Big(\frac{k}{n}\Big)\,,$$

where the derivative in taken on the left (see Robertson, Wright and Dykstra (1988)). Let now $\hat{\phi}_n(.)$ be the left continuous step function on $[0, 1]$ such that $\hat{\phi}_n(k/n) = \hat{\mu}_k$ at the knots k/n for $k = 1, \ldots, n$.

When the error process (X_k) in the model (24) is short range dependent and satisfies suitable weak dependence conditions, Zhao and Woodroofe (2008) have obtained the asymptotic behavior of $\hat{\phi}_n(t)$. In their paper an application to global warming is given. Some other situations are considered in the paper by Anevski and Hössjer (2006): in their Theorem 3 (iii), they consider the case where (X_k) can exhibit long range dependence, and they assume that X_k is a function of a Gaussian process such that its Hermite polynomials expansion is of rank greater than one. When no shape assumption is imposed on the regression function, non parametric regression analysis when data can exhibit long range dependence has been also studied by other authors (see for instance the paper by Robinson (1997), or more recently the paper by Gao and Wang (2006) who introduced random designs in the nonparametric trend model). The motivation of studying such models is that, in order to avoid misrepresenting the mean function or the conditional mean function of long range dependent data, one should let the data “speak” for themselves in terms of specifying the true form of the mean function or the conditional mean function. Situations where the error process (X_k) in the model (24) is long range dependent often occurs when considering financial or climatology time series. For instance the annual series of winter means of the NAO index (North Atlantic Oscillation index) exhibits long range dependence (see Stephenson *et al.* (2000)) and also an increasing trend for the last decade (which can possibly be explained by global warming). Concerning financial time series, we refer to the paper by Pesee (2008) where daily exchange rate data are studied. For instance the daily changes of the US-Dollar against the Deutsch Mark is a financial series that exhibits long range dependence with a long period of monotonic trend. For other data examples of long-memory processes, we refer to the book by Beran (1994). In particular, concerning the monthly temperature for the northern hemisphere, Beran suggests (page 29 of his book) that the series could be long-range dependent (see the Figures 1.12a, 1.12b and 1.12c, page 31).

The aim of this section is then to derive the asymptotic behavior of $\hat{\phi}_n(t)$ when X_k is a linear process which can exhibit short or long memory. Recall that by the well-known Wold decomposition, a stationary process in $\mathbf{L}^2$ that is purely non deterministic and such that its one-step mean squared error is positive, can be represented by a linear process generated by orthogonal random variables.

As it is implicitly mentioned in Anevski and Hössjer (2006) and enlightened in Zhao and Woodroofe (2008), the two main tools to obtain the asymptotic behavior of $\hat{\phi}_n(t)$ are a weak invariance principle for the partial sums process $\{S_{[nt]}, t \in [0,1]\}$ properly normalized, and a suitable moment inequality for $\max_{1\le k\le n} S_k^2$.

Theorem 4.1 *Let $(a_i)_{i\in\mathbf{Z}}$ and $(\xi_i)_{i\in\mathbf{Z}}$ be as in Comments 3.1 or 3.2. Let us consider the model (24) with X_k defined by (10). For any $t \in (0,1)$ such that $\phi'(t) > 0$,*

$$n^{1/3}\kappa^{-1}(\hat{\phi}_n(t) - \phi(t)) \Rightarrow (\sqrt{\eta})^{2/3}\mathrm{argmin}\{B(s) + s^2, s \in \mathbf{R}\}\,,$$

where B denotes a standard two-sided Brownian motion independent of η, $\eta = \sum_{k\in\mathbf{Z}} \mathbf{E}(\xi_0\xi_k|\mathcal{I})$, and $\kappa = 2\Big(\frac{1}{2}A^2\phi'(t)\Big)^{1/3}$ with $A = \sum_{j\in\mathbf{Z}} a_j$.

Let $\beta \in]0,2]$, and let h be a slowly varying function at infinity. Let now

$$L(x) = \Big(\frac{1}{h(x^{2/(4-\beta)})}\Big)^{1/2}, \tag{25}$$

and notice that $L(x)$ is also a slowly varying function at infinity. Denote then by L^* the asymptotic conjugate of L, which means that L^* satisfies

$$\lim_{x\to\infty} L^*(x)L(xL^*(x)) = 1\,. \tag{26}$$

Define then

$$d_n = \frac{1}{n^{(2-\beta)/(4-\beta)}}\ell(n) \text{ where } \ell(n) = \big(L^*(n)\big)^{2/(4-\beta)}. \tag{27}$$

Theorem 4.2 *Let $(a_i)_{i\in\mathbf{Z}}$ and $(\xi_i)_{i\in\mathbf{Z}}$ be as in Theorem 3.3. For $\beta = 1$ and $h(n) = |\sum_{i=-n}^{n} a_i|^2$, let d_n be defined by (27). Let us consider the model (24) with X_k defined by (10). For any $t \in (0,1)$ such that $\phi'(t) > 0$,*

$$d_n^{-1}\kappa^{-1}(\hat{\phi}_n(t) - \phi(t)) \Rightarrow (\sqrt{\eta})^{2/3}\mathrm{argmin}\{B(s) + s^2, s \in \mathbf{R}\}\,,$$

where B denotes a standard two-sided Brownian motion independent of η, $\eta = \sum_{k\in\mathbf{Z}} \mathbf{E}(\xi_0\xi_k|\mathcal{I})$, and $\kappa = 2\Big(\frac{1}{2}\phi'(t)\Big)^{1/3}$.

Example 7. In case of the linear process defined in Example 6, Theorem 4.2 applies with $d_n = n^{-1/3}\big(4\ln(n)/3\big)^{2/3}$.

Theorem 4.3 *Let $(a_i)_{i\in\mathbf{Z}}$ and $(\xi_i)_{i\in\mathbf{Z}}$ be as in Theorem 3.1 or 3.2, for some $\beta \in]0,2[$. By assumption, v_n^2 defined by (11) is regularly varying with exponent β. For this β and for $h(n) = v_n^2 n^{-\beta}$, let d_n be defined by (27). Let us consider the model (24) with X_k defined by (10). Then for any $t \in (0,1)$ such that $\phi'(t) > 0$,*

$$d_n^{-1}\kappa_\beta^{-1}(\hat{\phi}_n(t) - \phi(t)) \Rightarrow (\sqrt{\eta})^{1/(2-H)}\mathrm{argmin}\{B_H(s) + s^2, s \in \mathbf{R}\}\,,$$

where B_H denotes a standard two-sided fractional Brownian motion independent of η, with Hurst index $H = \beta/2$, $\eta = \sum_{k\in\mathbf{Z}} \mathbf{E}(\xi_0\xi_k|\mathcal{I})$, and the constant κ_β is given by $\kappa_\beta = 2(\phi'(t)/2)^{(2-\beta)/(4-\beta)}$.

Example 8. In case of the linear process defined in Example 1, Theorem 4.3 applies with $\beta = 2d+1$ and $d_n = \tau_d n^{(1-2d)/(3-2d)}$ where τ_d is a positive constant depending only on d.

Proofs of Theorems 4.1, 4.2 and 4.3. For any $t \in (0,1)$ and any $s \in [-td_n^{-1}, d_n^{-1}(1-t)]$, let

$$Z_n(s) = d_n^{-2}\big(Y_n(t + d_n s) - Y_n(t) - \phi(t)d_n s\big)\,.$$

Then $d_n^{-1}(\hat{\phi}_n(t) - \phi(t)) = \widetilde{Z}_n'(0)$, the left hand derivative of the GCM of Z_n at $s = 0$. Hence the key for establishing the result is the study of the GCM of the process Z_n. This can be done by following the arguments given in the Section 3 of the paper by Anevski and Hössjer (2006), and also in the paper by Zhao and Woodroofe (2008). More precisely, a careful analysis of the proofs given in both papers shows that the following lemma is valid.

Lemma 4.1 *Assume that there exists a positive sequence $m_n \to \infty$ satisfying for any $t > 0$,*

$$m_{[nt]}/m_n \to t^H \text{ where } H \in]0,1[, \tag{28}$$

and such that

1. *The process $\{m_n^{-1}S_{[nt]}, t \in [0,1]\}$ converges in $D([0,1])$ to $\sqrt{\eta}W_H$, where η is a positive random variable and W_H is a standard fractional Brownian motion (with Hurst index H) independent of η,*

2. $\mathbf{E}(\max_{1\le k\le n} S_k^2) \le Cm_n^2$.

Then, for any positive sequence $d_n \to 0$ *such that* $nd_n \to \infty$ *and* $d_n^{-2}n^{-1}m_{[nd_n]} \to 1$, *and for any* $t \in (0,1)$ *such that* $\phi'(t) > 0$,

$$d_n^{-1}\kappa_H^{-1}(\hat{\phi}_n(t) - \phi(t)) \Rightarrow (\sqrt{\eta})^{1/(2-H)}\mathrm{argmin}\{B_H(s) + s^2, s \in \mathbf{R}\}\,.$$

where $B_H(.)$ *denotes a standard two-sided fractional Brownian motion independent of* η, *with Hurst index* $H \in]0,1[$, *and* $\kappa_H = 2(\phi'(t)/2)^{(1-H)/(2-H)}$.

Proof of Lemma 4.1 We proceed as in the proof of Theorem 3 in Anevski and Hössjer (2006). The main point is then to verify their assumptions A1-A7 in order to apply their Corollary 1. Since $nd_n \to \infty$, Assumption A2 follows from the arguments given in the proof of Theorem 3(i) in Anevski and Hössjer (2006). By the properties of our limiting process, $\sqrt{\eta}W_H$, the assumptions A5 and A7 are satisfied. Now if A1 holds then by Proposition 2 in Anevski and Hössjer (2006) and the properties of the fractional Brownian motion, the assumption A6 also holds. Note that their Proposition 2 allows to apply the continuous mapping theorem to the functional h from $D[-c,c]$ (the space of cadlag functions on $[-c,c]$) to $\mathbb{R}$ defined as the left hand derivative of GCM (x) at 0. To verify their assumptions A3 and A4, it suffices to apply their Proposition 1. According to the proofs of their Lemmas B1 and B2, the condition (14) of their Proposition 1 is satisfied as soon as their condition (87) and our condition (28) are. Now their condition (87) is clearly satisfied provided Item 2 of Lemma 4.1 holds.

It remains to prove Assumption A1 of Anevski and Hössjer (2006); namely that the process

$$\{n^{-1}d_n^{-2}S_{[nd_nt]}, t \in [0,1]\}$$

converges in $D[0,1]$ to $\sqrt{\eta}W_H$, where η is a positive random variable and W_H is a standard fractional Brownian motion (with Hurst index H) independent of η. This holds by Item 1 of Lemma 4.1 and the fact that $d_n^{-2}n^{-1}m_{[nd_n]} \to 1$. This completes the proof of Lemma 4.1. $\diamond$

We go back to the proofs of Theorems 4.1, 4.2 and 4.3. Note that the conditions of Items 1 and 2 are clearly satisfied by using either Comment 3.1 or 3.2 (with $m_n = \sqrt{n}$), either Theorem 3.3 (with $m_n = \sqrt{n}|\sum_{i=-n}^{n} a_i|$) or Theorem 3.1 or 3.2 (with $m_n = v_n$). In addition, in all these situations, we have that $m_n = (n^\beta h(n))^{1/2}$ and the selection of d_n leads to

$$\begin{aligned} d_n^{-2}n^{-1}m_{[nd_n]} &\sim d_n^{(\beta-4)/2}n^{(\beta-2)/2}\sqrt{h(nd_n)} \\ &\sim \big(L^*(n)\big)^{-1}\sqrt{h\Big((nL^*(n))^{2/(4-\beta)}\Big)} \\ &\sim \big(L^*(n)\big)^{-1}\big(L(nL^*(n))\big)^{-1}, \end{aligned}$$

which converges to 1 by (26). $\diamond$

5 Proofs

5.1 Proof of Proposition 2.1

Without restricting the generality we shall assume $D_\Psi = 1$ and $\sum_{j\in\mathbf{Z}} c_{m,j}^2 = 1$, since otherwise we can divide each coefficient $c_{m,j}$ by $(\sum_{j\in\mathbf{Z}} c_{m,j}^2)^{1/2}$ and each variable by D_Ψ. Start with the decomposition

$$Y_k = \sum_{j=-\infty}^{\infty} P_{k-j}(Y_k) = \sum_{j=-\infty}^{\infty} p_j P_{k-j}(Y_k)/p_j \,.$$

Then

$$S_m = \sum_{j=-\infty}^{\infty} p_j \sum_{k\in\mathbf{Z}} c_{m,k} P_{k-j}(Y_k)/p_j \,.$$

By using the facts that Ψ is convex and non-decreasing, and $p_j \geq 0$ with $\sum_{j\in\mathbf{Z}} p_j = D_\Psi = 1$, we obtain that

$$\mathbf{E}\Psi(|S_m|) \leq \sum_{j=-\infty}^{\infty} p_j \mathbf{E}\Psi(|\sum_{k\in\mathbf{Z}} c_{m,k} P_{k-j}(Y_k)/p_j|\,).$$

Consider the martingale difference $U_k = c_{m,k}P_{k-j}(Y_k)/p_j$, $k \in \mathbf{Z}$. By Burkholder's inequality (see Theorem 6.6.2. in de la Peña and Giné (1999)), we obtain that

$$\mathbf{E}\Psi(|\sum_{k\in\mathbf{Z}} c_{m,k} P_{k-j}(Y_k)/p_j|) \leq K_\alpha \mathbf{E}\Psi((\sum_{k\in\mathbf{Z}} c_{m,k}^2 P_{k-j}^2(Y_k)/p_j^2)^{1/2})\,,$$

where K_α is a constant depending only on α. Let $\Phi(x) = \Psi(\sqrt{x})$. Since Φ is convex and $\sum_{k\in\mathbf{Z}} c_{m,k}^2 = 1$, it follows that

$$\begin{aligned}
\mathbf{E}\Psi(|\sum_{k\in\mathbf{Z}} c_{m,k} P_{k-j}(Y_k)/p_j|) &\leq K_\alpha \mathbf{E}\Phi(\sum_{k\in\mathbf{Z}} c_{m,k}^2 P_{k-j}^2(Y_k)/p_j^2) \\
&\leq K_\alpha \sum_{k\in\mathbf{Z}} c_{m,k}^2 \mathbf{E}\Phi(P_{k-j}^2(Y_k)/p_j^2) \\
&\leq K_\alpha \sum_{k\in\mathbf{Z}} c_{m,k}^2 \mathbf{E}(\Psi(|P_{k-j}(Y_k)|/p_j))\,.
\end{aligned}$$

Therefore

$$\mathbf{E}\Psi(|S_m|) \leq K_\alpha \sum_{k\in\mathbf{Z}} c_{m,k}^2 \sum_{j=-\infty}^{\infty} p_j\, \mathbf{E}(\Psi(|P_{k-j}(Y_k)|/p_j))$$

Now notice that $||P_{k-j}(Y_k)||_\Psi \leq p_j$, hence using the fact that $\sum_{k\in\mathbf{Z}} c_{m,k}^2 = 1$ and $D_\Psi = \sum_{j=-\infty}^{\infty} p_j = 1$, we get that

$$\mathbf{E}\Psi(|S_m|) \leq K_\alpha\,,$$

and so the desired result. $\diamond$

5.2 Proof of Proposition 2.2

Fix a positive integer m and define

$$\theta_{0,m} = \sum_{k=0}^{2m-2} \sum_{i=k-m+1}^{m-1} P_i(\xi_k)\,, \text{and} \quad \theta_{j,m} = \theta_{0,m} \circ T^j\,.$$

Observe that, by stationarity,

$$\|\theta_{0,m}\|_\Psi = \| \sum_{k=0}^{2m-2} \sum_{i=k-m+1}^{m-1} P_i(\xi_k)\|_\Psi \le 2m \sum_{i\in\mathbf{Z}} \|P_0(\xi_i)\|_\Psi < \infty \ .$$

Simple computations lead to the decomposition

$$\sum_{i=-m+1}^{m-1} P_i(\xi_0) - \sum_{\ell=1}^{2m-1} P_m(\xi_\ell) = \theta_{0,m} - \theta_{1,m}\,,$$

implying that

$$\xi_0 - (\sum_k P_0(\xi_k)) \circ T^m = \theta_{0,m} - \theta_{1,m} + \sum_{|i|\ge m} P_i(\xi_0) - (\sum_{|k|\ge m} P_0(\xi_k)) \circ T^m \ .$$

With our notation ($d_0 = \sum_k P_0(\xi_k)$), we obtain

$$\xi_0 - d_0 = d_0 \circ T^m - d_0 + \theta_{0,m} - \theta_{1,m} + \sum_{|i|\ge m} P_i(\xi_0) - (\sum_{|k|\ge m} P_0(\xi_k)) \circ T^m. \quad (29)$$

By stationarity we obtain similar decompositions for each $\xi_j - d_j$. We shall treat the terms from the error of approximation $\sum_{i\in\mathbf{Z}} c_{n,i}(\xi_i - d_i)$ separately. First notice that

$$\begin{aligned} R_1 := \sum_{j=-\infty}^{\infty} c_{n,j}(\ d_j \circ T^m - d_j) &= \sum_{j=-\infty}^{\infty} (c_{n,j-m} - c_{n,j}) d_j \\ &= \sum_{k=0}^{m-1} \sum_{j=-\infty}^{\infty} (c_{n,j-k-1} - c_{n,j-k}) d_j\,. \end{aligned}$$

According to Proposition 2.1,

$$\|R_1\|_\Psi \le C_\alpha m \|d_0\|_\Psi (\sum_{j=-\infty}^{\infty} (c_{n,j} - c_{n,j-1})^2)^{1/2}\,.$$

To treat the second difference in the error, notice that

$$R_2 := \sum_{i=-\infty}^{\infty} c_{n,i}(\theta_{i,m} - \theta_{i+1,m}) = \sum_{i=-\infty}^{\infty} (c_{n,i} - c_{n,i-1})\theta_{i,m}\,.$$

By the definition of $\theta_{0,m}$ we have that

$$\sum_{j\in\mathbf{Z}} \|P_j(\theta_{0,m})\|_\Psi \le \sum_{k=0}^{2m-2} \sum_{i=k-m+1}^{m-1} \sum_{j\in\mathbf{Z}} \|P_j(P_i(\xi_k))\|_\Psi \,.$$

Now $P_j(P_i(f)) = 0$ for $j \neq i$. It follows that

$$\sum_{j\in\mathbf{Z}} \|P_j(\theta_{0,m})\|_\Psi \le \sum_{k=0}^{2m-2} \sum_{\ell=k-m+1}^{m-1} \|P_0(\xi_\ell)\|_\Psi \le (2m-1) \sum_{\ell=-m+1}^{m-1} \|P_0(\xi_\ell)\|_\Psi \,,$$

and by Proposition 2.1 we conclude that

$$||R_2||_\Psi \le 2C_\alpha m(\sum_{j=-\infty}^{\infty} (c_{n,j} - c_{n,j-1})^2)^{1/2} \sum_{\ell\in\mathbf{Z}} \|P_0(\xi_\ell)\|_\Psi \,.$$

For the term $R_3 := \sum_{i=-\infty}^{\infty} c_{n,i}(\sum_{|j|\ge m} P_j(\xi_0)) \circ T^i$ we apply Proposition 2.1 to get

$$||R_3||_\Psi \le C_\alpha(\sum_{i=-\infty}^{\infty} c_{n,i}^2)^{1/2} \sum_{|j|\ge m} \|P_j(\xi_0)\|_\Psi \,.$$

To deal with the last term $R_4 := \sum_{i=-\infty}^{\infty} c_{n,i}(\sum_{|k|\ge m} P_0(\xi_k)) \circ T^{m+i}$, we apply again Proposition 2.1, which gives

$$||R_4||_\Psi \le C_\alpha(\sum_{i=-\infty}^{\infty} c_{n,i}^2)^{1/2} \sum_{|k|\ge m} ||P_0(\xi_k)||_\Psi \,.$$

Combining all the bounds we obtain the desired approximation. $\diamond$

5.3 Proof of Lemma 2.1

For any $m \in [1, 2^N]$, write m in basis 2 as follows:

$$m = \sum_{i=0}^{N} b_i(m)2^i, \quad \text{with } b_i(m) = 0 \text{ or } b_i(m) = 1\,.$$

Set $m_L = \sum_{i=L}^{N} b_i(m)2^i$. So for any $p \ge 1$, we have

$$|S_m|^p \le \Big(\sum_{L=0}^{N} |S_{m_L} - S_{m_{L+1}}|\Big)^p \,.$$

Hence setting

$$\alpha_L = ||S_{2^L}||_{\Psi_p}\Big(\Psi^{-1}(2^{N-L})\Big)^{1/p} \text{ and } \lambda_L = \frac{\alpha_L}{\sum_{L=0}^{N} \alpha_L} \,,$$

we get by convexity

$$|S_m|^p \le \sum_{L=0}^{N} \lambda_L^{1-p} |S_{m_L} - S_{m_{L+1}}|^p \,.$$

Now $m_L \neq m_{L+1}$ only if $b_L(m) = 1$, and in that case $m_L = k_m 2^L$ with k_m odd. It follows that

$$\max_{1\le m\le 2^N} |S_m|^p \le \sum_{L=0}^{N} \lambda_L^{1-p} \max_{1\le k\le 2^{N-L}, k \text{ odd}} |S_{k2^L} - S_{(k-1)2^L}|^p \,.$$

Now, we apply Lemma 11.3 in Ledoux and Talagrand (1991) to the variables

$$Z_k = \frac{|S_{k2^L} - S_{(k-1)2^L}|^p}{A^p}, \text{ with } A = ||S_{2^L}||_{\Psi_p},$$

and to the Young function Ψ. Since

$$\mathbf{E}(\Psi(Z_k)) = \mathbf{E}\Psi_p\Big(\frac{|S_{2^L}|}{A}\Big) \le 1 \,,$$

and since Ψ^{-1} is concave, we get that for any measurable set B,

$$\mathbf{E}(Z_k \mathbf{1}_B) \le P(B)\Psi^{-1}\Big(\frac{1}{P(B)}\Big) \,,$$

so that the assumptions of Lemma 11.3 in Ledoux and Talagrand (1991) are satisfied. It follows that

$$\mathbf{E}\Big(\max_{1\le k\le 2^{N-L}, k \text{ odd}} |S_{k2^L} - S_{(k-1)2^L}|^p\Big) \le A^p \Psi^{-1}(2^{N-L}).$$

Finally, we conclude that

$$\mathbf{E}(\max_{1\le m\le 2^N} |S_m|^p) \le \Big(\sum_{L=0}^{N} \alpha_L\Big)^p \,,$$

which is the desired result. $\diamond$

5.4 Proof of Theorems 3.1 and 3.2

By the weak convergence theory of random functions, it suffices to establish the convergence of the finite dimensional distributions and the tightness of $\{v_n^{-1} S_{[nt]}, t \in [0,1]\}$. For the finite-dimensional distribution we shall use the following proposition which was basically established in Peligrad and Utev (1997, 2006-b).

Proposition 5.1 *Let $\{\xi_k\}_{k\in\mathbf{Z}}$ be a strictly stationary sequence of centered and regular random variables in $\mathbf{L}^2$ such that $\sum_j \|P_0(\xi_j)\|_2 < \infty$. For any positive integer n, let $\{b_{n,i}, -\infty \le i \le \infty\}$ be a triangular array of numbers satisfying*

$$\sum_i b_{n,i}^2 \to 1 \ \text{ and } \ \sum_j (b_{n,j} - b_{n,j-1})^2 \to 0 \ \text{ as } n \to \infty \ , \tag{30}$$

and

$$\sup_j |b_{n,j}| \to 0 \ \text{ as } \ n \to \infty \ . \tag{31}$$

Then $\{S_n = \sum_j b_{n,j}\xi_j\}$ converges in distribution to $\sqrt{\eta}N$ where N is a standard Gaussian random variable independent of η, and $\eta = \sum_{k\in\mathbf{Z}} \mathbf{E}(\xi_0\xi_k|\mathcal{I})$.

Proof of Proposition 5.1. We give here the proof for completeness. By using Proposition 2.2 it suffices to prove this proposition with $d_j = d_0 \circ T^j$ in place of ξ_j, where $d_0 = \sum_j P_0(\xi_j)$. Hence we just have to apply the central limit theorem for triangular arrays of martingales (see Theorem 3.6 in Hall and Heyde (1980)). The Lindeberg condition has been established by Peligrad and Utev (1997) provided that Condition (31) and the first part of Condition (30) are satisfied. Now in the proof of their proposition 4, Peligrad and Utev (2006-b) have established that (30) implies that

$$\sum_j b_{n,j}^2 d_j^2 \to \eta \quad \text{in probability as } n \to \infty,$$

which ends the proof of the proposition. $\diamond$

We return to the proof of Theorems 3.1 and 3.2. To prove the convergence of the finite dimensional distributions, we shall apply the Cramér-Wold device. For all integer $1 \le \ell \le m$, let $n_\ell = [nt_\ell]$ where $0 < t_1 < t_2 < \cdots < t_m \le 1$. For $\lambda_1, \dots, \lambda_m \in \mathbf{R}$, notice that

$$\frac{\sum_{\ell=1}^m \lambda_\ell S_{n_\ell}}{v_n} = \sum_{j\in\mathbf{Z}} \Big(\sum_{\ell=1}^m \frac{\lambda_\ell c_{n_\ell,j}}{v_n}\Big)\xi_j \ , \tag{32}$$

where $c_{n,j} = a_{1-j} + \cdots + a_{n-j}$ for all $j \in \mathbf{Z}$, and $v_n^2 = \sum_{j\in\mathbf{Z}} c_{n,j}^2$. Let

$$b_{n,j} = \frac{1}{\Lambda_{m,\beta}} \sum_{\ell=1}^m \frac{\lambda_\ell c_{n_\ell,j}}{v_n} \ , \tag{33}$$

where

$$\Lambda_{m,\beta}^2 = \frac{1}{2} \sum_{\ell,k=1}^m \lambda_\ell\lambda_k \big(t_\ell^\beta + t_k^\beta - |t_k - t_\ell|^\beta\big) \ .$$

We apply Proposition 5.1 to $b_{n,j}$ and the ξ_j's defined as $\Lambda_{m,\beta}\xi_j$. First, we have to calculate the limit over n of the following quantity

$$\sum_{j\in\mathbf{Z}} b_{n,j}^2 = \frac{1}{\Lambda_{m,\beta}^2} \frac{\sum_{j\in\mathbf{Z}} \sum_{\ell=1}^m \sum_{k=1}^m \lambda_\ell\lambda_k c_{n_\ell,j} c_{n_k,j}}{v_n^2} \ .$$

For any $1 \le \ell \le k \le m$, by using the fact that for any two real numbers A and B we have $A(A+B) = 1/2(A^2 + (A+B)^2 - B^2)$, we get that

$$\begin{aligned}\frac{1}{v_n^2}\sum_{j\in\mathbf{Z}} c_{n_\ell,j}c_{n_k,j} &= \frac{1}{2v_n^2}\sum_{j\in\mathbf{Z}}\left(c_{n_\ell,j}^2 + c_{n_k,j}^2 - \left(c_{n_\ell,j} - c_{n_k,j}\right)^2\right)\\ &= \frac{1}{2v_n^2}\sum_{j\in\mathbf{Z}}\left(c_{n_\ell,j}^2 + c_{n_k,j}^2 - c_{n_k-n_\ell,j}^2\right).\end{aligned}$$

By using now the condition (12), we derive that, for any $1 \le \ell \le k \le m$,

$$\frac{\sum_{j\in\mathbf{Z}} b_{n_\ell,j}b_{n_k,j}}{v_n^2} \to \frac{1}{2}\left(t_\ell^\beta + t_k^\beta - (t_k - t_\ell)^\beta\right). \tag{34}$$

It follows from (34) that

$$\lim_{n\to\infty}\sum_{j\in\mathbf{Z}} b_{n,j}^2 = 1\,. \tag{35}$$

As a consequence the first part of condition (30) holds. On the other hand, by using Lemma A.1 in Peligrad and Utev (2006-b), the second part of the condition (30) is satisfied. Now by the proof of Corollary 2.1 in Peligrad and Utev (1997) we get that

$$\frac{\max_j |c_{n,j}|}{v_n} \to 0\,,$$

which together with (12) implies (31). Applying now Proposition 5.1, we derive that

$$\frac{\sum_{\ell=1}^m \lambda_\ell S_{n_\ell}}{v_n} \text{ converges in distribution to } \Lambda_{m,\beta}\sqrt{\eta}N\,,$$

ending the proof of the convergence of the finite dimensional distribution.

We turn now to the proof of the tightness of $\{v_n^{-1}S_{[nt]}, t\in[0,1]\}$. By using Proposition 2.1, we get for $q \ge 2$ that

$$||S_k||_q \le C_q\Big(\sum_{j\in\mathbf{Z}} b_{k,j}^2\Big)^{1/2}\sum_{m\in\mathbf{Z}}\|P_0(\xi_m)\|_q = C_q v_k \sum_{m\in\mathbf{Z}}\|P_0(\xi_m)\|_q\,, \tag{36}$$

provided that $\sum_{m\in\mathbf{Z}}\|P_0(\xi_m)\|_q < \infty$. Therefore the conditions of Lemma 2.1 p. 290 in Taqqu (1975) are satisfied with $q > 2/\beta$, and the tightness follows.

Finally to prove (13), we use (36) together with Lemma 2.1 applied with $\psi(x) = x$ by taking into account that v_n^2 is regularly varying with exponent β. $\diamond$

5.5 Proof of Remarks 3.3 and 3.6

To prove Remark 3.3, we apply lemma 6.1 from the appendix with $b_i = 1$ and $u_i = \|P_{-i}(\xi_0)\|_q$. Hence we get

$$\sum_{n=1}^\infty \|P_{-n}(\xi_0)\|_q \le C_q \sum_{n=1}^\infty\Big(\frac{1}{n}\sum_{k=n}^\infty \|P_{-k}(\xi_0)\|_q^q\Big)^{1/q}.$$

Applying the Rosenthal's inequality given in Theorem 2.12 in Hall and Heyde (1980), we then derive that for any $q \in [2, \infty[$, there exists a constant c_q depending only on q such that

$$\sum_{k=n}^{\infty} \|P_{-k}(\xi_0)\|_q^q \leq c_q \| \sum_{k=n}^{\infty} P_{-k}(\xi_0)\|_q^q = c_q \|\mathbf{E}(\xi_n|\mathcal{F}_0)\|_q^q \, .$$

The same argument works with $P_{-i}(\xi_0)$ replaced by $P_i(\xi_0)$, and the result follows by applying Rosenthal's inequality and by noticing that $\|\xi_{-n} - \mathbf{E}(\xi_{-n}|\mathcal{F}_0)\|_q = \| \sum_{k=n}^{\infty} P_{k+1}(\xi_0)\|_q$.

To prove Remark 3.6, we apply Lemma 6.1 from the appendix with $b_n = \log(n)$ and $u_n = \|P_0(\xi_n)\|_2$. We then get that

$$\sum_{n=1}^{\infty} \log n \|P_0(\xi_n)\|_2 \leq C \sum_{n=1}^{\infty} \frac{\log n}{\sqrt{n}} \Big(\sum_{k=n}^{\infty} \|P_0(\xi_k)\|_2^2 \Big)^{1/2} .$$

Notice now that

$$\sum_{k=n}^{\infty} \|P_0(\xi_k)\|_2^2 = \|\mathbf{E}(\xi_n|\mathcal{F}_0)\|_2^2 \, ,$$

and then

$$\sum_{n=1}^{\infty} \log n \|P_0(\xi_n)\|_2 \leq C \sum_{n=1}^{\infty} \log n \frac{\|\mathbf{E}(\xi_n|\mathcal{F}_0)\|_2}{\sqrt{n}} < \infty \, .$$

The same argument works with $P_0(\xi_i)$ replaced by $P_0(\xi_{-i})$. $\diamond$

5.6 Proof of Theorem 3.3

For all $j \in \mathbf{Z}$, let $d_j = \sum_{\ell \in \mathbf{Z}} P_j(\xi_\ell)$. Note that, if either Condition (a) or Condition (b) is satisfied, $(d_j)_{j \in \mathbf{Z}}$ is a sequence of martingale differences in $\mathbf{L}^2$. We set

$$Y_k = \sum_{i \in \mathbf{Z}} a_i d_{k-i} \text{ and } T_n = \sum_{k=1}^{n} Y_k \, ,$$

and apply Corollary 4 in Dedecker, Merlevède and Volný (2007). By taking into account Remark 3.5, we derive that under (21),

$$\{s_n^{-1} T_{[nt]}, t \in [0,1]\} \text{ converges in distribution in } (D([0,1]), d) \text{ to } \sqrt{\mathbf{E}\big(d_0^2|\mathcal{I}\big)} W,$$

where W is a standard Brownian motion independent of $\mathcal{I}$. It follows that in order to prove that $\{s_n^{-1} S_{[nt]}, t \in [0,1]\}$ converges in distribution in $(D([0,1]), d)$ to $\sqrt{\mathbf{E}\big(d_0^2|\mathcal{I}\big)} W$ it is sufficient to show that

$$\frac{\| \max_{1 \leq k \leq n} |S_k - T_k| \|_2}{s_n} \to 0 \text{ , as } n \to \infty \, . \tag{37}$$

Now for any n, let N be such that $2^{N-1} < n \leq 2^N$. By using Remark 3.5 and the properties of the slowly varying function, we get that $s_n \sim s_{2^N}$. So, the proof (37), is reduced to showing that

$$\frac{\| \max_{1\leq k\leq 2^N} |S_k - T_k| \|_2}{s_{2^N}} \to 0 \text{ , as } N \to \infty \,. \tag{38}$$

We first prove that (38) holds under Condition (a). By using Corollary 2.1 together with Lemma 2.1, we get that for any positive integer m,

$$\begin{aligned} \| \max_{1\leq k\leq 2^N} |S_k - T_k| \|_2 \quad &\leq \quad C_1 \sum_{|k|\geq m} \|P_0(\xi_k)\|_{\Psi_{2,\alpha}} \sum_{L=0}^{N} v_{2^L} \big(g^{-1}(2^{N-L})\big)^{1/2} \\ &\quad + C_2 m \sum_{L=0}^{N} \big(g^{-1}(2^{N-L})\big)^{1/2} \,, \end{aligned}$$

where $g(x) = x \log^\alpha(1+x)$. Noticing that for $g^{-1}(x) \sim \frac{x}{\log^\alpha(1+x)}$ as x goes to infinity, by taking into account Remark 3.5 and the first part of Condition (21) we get that

$$\| \max_{1\leq k\leq 2^N} |S_k - T_k| \|_2 \leq C s_{2^N} \sum_{|k|\geq m} \|P_0(\xi_k)\|_{\Psi_{2,\alpha}} + C m \epsilon(N) s_{2^N} \,, \tag{39}$$

where $\epsilon(N) \to 0$ as $N \to \infty$. By using now (39) and letting first N tend to infinity and next m tend to infinity, we derive (38) under Condition (a).

We turn now to the proof of (38) under Condition (b). Taking $m = m_{2^L} = 2^{L/4}$ in Corollary 2.1 and using Lemma 2.1 with $p = 2$ and $\psi(x) = x$, we get that

$$\begin{aligned} \frac{\| \max_{1\leq k\leq 2^N} |S_k - T_k| \|_2}{s_{2^N}} &\leq C \frac{2^{N/2}}{s_{2^N}} \sum_{L=0}^{N} \frac{m_{2^L}}{2^{L/2}} \\ &\quad + C \frac{2^{N/2}}{s_{2^N}} \sum_{L=0}^{N} \frac{v_{2^L}}{2^{L/2}} \sum_{|k|\geq m_{2^L}} \|P_0(\xi_k)\|_2 \,. \end{aligned} \tag{40}$$

By Remark 3.5 we have that $\lim_{N\to\infty} \frac{s_{2^N}}{2^{N/2}} = \infty$ which together with the selection of m_{2^L} imply that the first term on the right hand of the above inequality tends to zero as $n \to \infty$. Now, to treat the last term, we first fix a positive integer p and we write that

$$\begin{aligned} \frac{2^{N/2}}{s_{2^N}} \sum_{L=0}^{N} \frac{v_{2^L}}{2^{L/2}} \sum_{|k|\geq m_{2^L}} \|P_0(\xi_k)\|_2 \quad &\leq \quad p \frac{2^{N/2}}{s_{2^N}} \max_{0\leq L<p} \frac{v_{2^L}}{2^{L/2}} \sum_{|k|\geq m_{2^L}} \|P_0(\xi_k)\|_2 \\ &\quad + \frac{2^{N/2}}{s_{2^N}} \sum_{L=p}^{N} \frac{v_{2^L}}{2^{L/2}} \sum_{|k|\geq m_{2^L}} \|P_0(\xi_k)\|_2 \,. \end{aligned}$$

Since $\lim_{N\to\infty} \frac{s_{2^N}}{2^{N/2}} = \infty$, the first term on the right-hand side of the above inequality tends to zero as $N \to \infty$. To treat the second one, we notice that if N and p are large enough,

$$\frac{2^{N/2}}{s_{2^N}} \sum_{L=p}^{N} \frac{v_{2^L}}{2^{L/2}} \sum_{|k|\ge m_{2^L}} \|P_0(\xi_k)\|_2 \le C \sum_{L=p}^{N} \frac{h(2^L)}{h(2^N)} \sum_{|k|\ge m_{2^L}} \|P_0(\xi_k)\|_2 \,,$$

where $h(n) = \left| \sum_{i=-n}^{n} a_i \right|$. By the first part of Condition (21),

$$\limsup_{N\to\infty} \max_{p\le L\le N} \frac{h(2^L)}{h(2^N)} < \infty \,.$$

Hence, for N and p large enough and taking into account the selection of m_{2^L}, we get that

$$\frac{2^{N/2}}{s_{2^N}} \sum_{L=p}^{N} \frac{v_{2^L}}{2^{L/2}} \sum_{|k|\ge m_{2^L}} \|P_0(\xi_k)\|_2 \le C \sum_{|k|\ge 2^{p/4}} \log k \|P_0(\xi_k)\|_2 \,,$$

which converges to zero as $p \to \infty$ by using Condition (b). Hence starting from (40) and taking into account the previous considerations, we get that (38) holds under Condition (b). The proof of (22) is straightforward, following the arguments used to derive (37). $\diamond$

5.7 Proof of Comment 3.2

The justification of this result is due to the following coboundary decomposition. Define

$$Z_0 = \sum_{\ell=1}^{\infty} \sum_{k=\ell}^{\infty} a_k \xi_{-\ell} - \sum_{\ell=0}^{\infty} \sum_{k=-\infty}^{-\ell-1} a_k \xi_\ell \,. \tag{41}$$

Since condition (1) implies that the sequence $(\xi_i)_{i\in\mathbf{Z}}$ has a bounded spectral density, the random variable Z_0 is well defined in $\mathbf{L}^2$ under under condition (H). Now

$$Z_0 - Z_0 \circ T = \sum_{\ell=1}^{\infty} a_\ell \xi_{-\ell} - \xi_0 \sum_{k=1}^{\infty} a_k - \xi_0 \sum_{k=1}^{\infty} a_{-k} + \sum_{\ell=1}^{\infty} a_{-\ell} \xi_\ell \,.$$

Whence,

$$A\xi_0 + Z_0 - Z_0 \circ T = a_0 \xi_0 + \sum_{j\in\mathbf{Z}\setminus\{0\}} a_j \xi_{-j} = X_0 \,.$$

We derive that for any $k \ge 1$,

$$S_k = A \sum_{i=1}^{k} \xi_i + Z_1 - Z_{k+1} \,, \tag{42}$$

where $Z_k = Z_0 \circ T_k$. Since under Condition (1), the partial sums process $\{n^{-1/2} \sum_{k=1}^{[nt]} \xi_k, t \in [0,1]\}$ converges in distribution in $D([0,1])$ to $\sqrt{\lambda} W$, with $\lambda = \sum_{j \in \mathbf{Z}} \mathbf{E}(\xi_0 \xi_j | \mathcal{I})$, we just have to show that

$$\limsup_{n \to \infty} \mathbf{P}\Big(\max_{1 \le k \le n} |Z_{k+1}| \ge \varepsilon \sqrt{n} \Big) = 0 \,,$$

which holds because $Z_0 \in \mathbf{L}^2$ (see the inequality (5.30) in Hall and Heyde (1980)). $\diamond$

6 Appendix

6.1 Fact about series

Lemma 6.1 *Let $q > 1$ and $\alpha = 2(q-1)/q$. Let $(b_j)_{j \in \mathbf{N}}$ be a sequence of non-negative numbers such that $n^\alpha b_n \le K_\alpha \sum_{k=1}^n k^{\alpha-1} b_k$, for some positive constant K_α depending only on α. Then for any sequence of non-negative numbers $(u_j)_{j \in \mathbf{N}}$ the following inequality holds*

$$\sum_{n=1}^\infty b_n u_n \le C_q \ \sum_{n=1}^\infty b_n \left(\frac{1}{n} \sum_{k=n}^\infty u_k^q \right)^{1/q}$$

where C_q is a constant depending only on q.

Proof. We write

$$\sum_{n=1}^\infty b_n u_n \ \le K_\alpha \sum_{n=1}^\infty n^{-\alpha} u_n \Big(\sum_{k=1}^n \ b_k k^{\alpha-1} \Big) \le K_\alpha \sum_{k=1}^\infty b_k k^{\alpha-1} \Big(\sum_{n \ge k} \ n^{-\alpha} u_n \Big) \,.$$

Then, Hölder's inequality gives

$$\sum_{n=1}^\infty b_n u_n \le C'_q \sum_{k=1}^\infty b_k k^{\alpha-1} \Big(\sum_{n \ge k} \ n^{-2} \ \Big)^{\alpha/2} \Big(\sum_{n \ge k} \ u_n^q \Big)^{1/q}$$

and the result follows. $\diamond$

Acknowledgment. The authors would like to thank Sergey Utev for valuable discussions on this topic. They are also grateful to the referees and the associate editor for helpful comments that improve the presentation of this paper.

References

[1] Anevski, D. and Hössjer, O. (2006). A general asymptotic scheme for inference under order restrictions. *Ann. Statist.* **34**, 1874–1930.

[2] Beran, J. (1994). *Statistics for long-memory processes.* Monographs on Statistics and Applied Probability, **61**. Chapman and Hall, New York.

[3] Davydov, Y.A. (1970). The invariance principle for stationary processes. *Theor. Probability Appl.*, **15**, 487-498.

[4] Dedecker, J., Merlevède, F. and Volný, D. (2007). On the weak invariance principle for non-adapted sequences under projective criteria. *J. Theoret. Probab.*, **20**, 971-1004.

[5] De La Peña, V. and Giné, E. (1999). *Decoupling: From dependence to independence: Randomly stopped processes, U-statistics and Processes, Martingales and Beyond*. Springer-Verlag, New York.

[6] Gao, J. and Wang, Q. (2006). Long-range dependent time series specification. *Preprint.* http://www.maths.usyd.edu.au/u/pubs/publist/preprints/2006/gao-26.pdf.

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

[8] Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and its Applications.* Academic Press, New York-London.

[9] Hannan, E. J. (1973). Central limit theorems for time series regression. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete.* **26**, 157–170.

[10] Heyde, C. C. (1974). On the central limit theorem for stationary processes. *Z. Wahrscheinlichkietstheorie und Verw. Gebiete.* **30**, 315-320.

[11] Ibragimov, I. A. (1962). Some limit theorems for stationary processes. *Theor. Probab. Appl.* **7**, 349-382.

[12] Konstantopoulos, T. and Sakhanenko, A. (2004). Convergence and convergence rate to fractional Brownian motion for weighted random sums. *Sib. lektron. Mat. Izv.* **1**, 47-63

[13] Lamperti, J. (1962). Semi-stable stochastic processes. *Trans. Amer. Math. Soc.* **104**, 62-78.

[14] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach spaces. Isoperimetry and processes.* Springer-Verlag, Berlin.

[15] Merlevède, F. and Peligrad, M. (2006). On the weak invariance principle for stationary sequences under projective criteria. *J. Theoret. Probab.*, **19**, 647-689.

[16] Peligrad, M. and Utev, S. (1997). Central limit theorem for linear processes. *Ann. Probab.* **25**, 443-456.

[17] Peligrad, M. and Utev, S. (2006-a). Invariance principle for stochastic processes with short memory. *IMS lecture notes-monograph series.* **51** 18-32.

[18] Peligrad, M. and Utev, S. (2006-b). Central limit theorem for stationary linear processes. *Ann. Probab.* **34**, 1608-1622.

[19] Pesee C. (2008). Long-range dependence of financial time series data. *WASET* **44**, 163-167.

[20] Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order restricted statistical inference.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Ltd., Chichester.

[21] Robinson, P. M. (1997). Large-sample inference for nonparametric regression with dependent errors. *Ann. Statist.* **25**, 2054-2083.

[22] Shao, X. and Wu, W.B. (2006). Invariance principles for fractionally integrated nonlinear processes. *IMS lecture notes-monograph series* **50**. 20-30.

[23] Stephenson, D. B., Pavan V. and Bojariu R. (2000). Is the North Atlantic Oscillation a random walk ? *International Journal of Climatology.* **20**. 1-18.

[24] Taqqu, M. S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **31**. 287-302.

[25] Wang, Q., Lin, X-Y. and Gulati, C. M. (2001). Asymptotics for moving average processes with dependent innovations. *Stat. and Probab. Letters.* **54**, 347-356.

[26] Wu, W. B. (2007). Strong invariance principles for dependent random variables. *Ann. Probab.* **35**, 2294-2320.

[27] Wu, W.B. and Min, W. (2005). On linear processes with dependent innovations. *Stochastic Process. Appl.* **115**, 939-958.

[28] Wu, W. B. and Woodroofe, M. (2004). Martingale approximations for sums of stationary processes. *Ann. Probab.* **32** 1674-1690.

[29] Zhao, O. and Woodroofe, M. (2008). Estimating a monotone trend. arXiv: 0812.3188.