Strong approximation for a class of stationary processes

Weidong Liu, Zhengyan Lin*

Department of Mathematics, Zhejiang University, Hangzhou 310027, China

Received 27 May 2007; received in revised form 7 November 2007; accepted 22 January 2008

Available online 12 February 2008

Abstract

Strong approximation for sums of a class of stationary processes with optimal bound is established. The main tools are $m$-dependent approximation and block techniques. Some previous results are improved. © 2008 Elsevier B.V. All rights reserved.

MSC: 60F05; 60F17

Keywords: Causal process; Strong invariance principle; $m$-dependent approximation

1. Introduction

Let $\{\varepsilon_n; n \in \mathbb{Z}\}$ be i.i.d. random elements and $g$ denote a measurable function such that

$$X_n = g(\ldots, \varepsilon_{n-1}, \varepsilon_n)$$

is a well-defined $R^d$ valued ($d \geq 1$) random vector. $\{X_n\}$ is a causal process. A large class of processes, including a variety of nonlinear time series models, can be represented in this way. In this paper we study strong invariance principles for $S_n = \sum_{i=1}^{n}(X_i - \mathbb{E}X_i)$.

We denote by $|\cdot|$ the $d$-dimensional Euclidean norm in $R^d$. As an application, Wu [30] obtained strong invariance principles for $S_n$ with the rate $O_{a.s.}(n^{1/p}(\log n)^{1/2})$ $(2 < p < 4)$ under $d = 1$, $\mathbb{E}|X_0|^p < \infty$ and some other conditions. The basic tool used in his paper is a new version of martingale approximation. It is well known that the martingale approximation method is quite effective. Studies on the asymptotic behavior for $S_n$ could usually be reduced to that of the approximating martingale. However, there are essential difficulties in getting the optimal

* Corresponding author.
E-mail addresses: liuweidong99@gmail.com (W. Liu), zlin@zju.edu.cn (Z. Lin).

0304-4149/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.spa.2008.01.012
bound \( o_{d.s.}(n^{1/p}) \) for strong invariance principles based on the martingale embedding method. Moreover, Monrad and Philipp [18] proved that it is impossible to embed a general \( R^d \) valued \( (d \geq 2) \) martingale in an \( R^d \) valued Gaussian process.

In order to reach the rate \( o_{d.s.}(n^{1/p}) \) of invariance principles for \( R^d \) valued stationary processes, some new method should be developed. In the present paper, we approximate \( S_n \) by the partial sums of \( m \)-dependent random variables (see Lemma A.1). With the help of blocking arguments, the optimal bound is reached under some easily verifiable conditions. Applications to linear and nonlinear processes are discussed.

The strong invariance principles are quite useful and have received considerable attention in probability theory. They play an important role in statistical inference. Strassen [24,25] initiated the study for i.i.d. random variables and stationary and ergodic martingale differences. Some optimal results for i.i.d. random variables were obtained by Komlós et al. [15,16]. For dependent random variables, see [2,3,9,26,23,17,22,7,6].

The plan of the paper is as follows. Our main results are presented in Section 2. Applications to linear and nonlinear processes are discussed in Section 3. The proofs are stated in Sections 4 and 5. The basic tools (Lemmas A.1 and A.2) are proven in the Appendix. Throughout, we let \( C, C_\cdot \) denote positive constants which may be different in every place. For a random vector \( Z \) write \( Z \in \mathcal{L}^p, \ p > 0 \), if \( \|Z\|_p = (\mathbb{E}|Z|^p)^{1/p} < \infty \). For two real sequences \( \{a_n\} \) and \( \{b_n\} \), write \( a_n = O(b_n) \) if there exists a constant \( C \) such that \( |a_n| \leq C|b_n| \) holds for large \( n \), \( a_n = o(b_n) \) if \( \lim_{n \to \infty} a_n/b_n = 0 \) and \( a_n \asymp b_n \) if \( C_1 b_n \leq a_n \leq C_2 b_n \).

2. Main results

We introduce the following notation. Let the shift process \( \xi_k = (\ldots , \varepsilon_{k-1}, \varepsilon_k) \). Following [29], (1.1) can be viewed as a physical system with \( \xi_i \) being the input, \( X_i \) being the output and \( g \) being a filter or transform. The dependence then is interpreted as the degree of dependence of output on input. Let \( \{\varepsilon_n; n \in Z\} \) be an i.i.d. copy of \( \{\varepsilon_n; n \in Z\} \) and \( \xi'_i = (\xi_{-1}, \varepsilon_0', \varepsilon_1, \ldots, \varepsilon_i) \). Define \( X^*_n = g(\xi'_n), \ n \in Z \). Assume \( X_0 \in \mathcal{L}^p, \ p > 0 \). Let

\[
\theta_{n,p} = \|X_n - X^*_n\|_p.
\]

Wu [29] called \( \theta_{n,p} \) the physical dependence measure. Throughout the paper we assume

\[
\Theta_{n,p} = \sum_{i=n}^{\infty} \theta_{i,p} < \infty, \quad n = 0, 1, 2, \ldots.
\]

Let \( p > 1 \). Following Theorem 1 in [30], there are stationary and ergodic \( \mathcal{L}^p \) martingale differences \( \{D_n\} \) with respect to \( \sigma(\xi_n) \) and the corresponding martingale \( M_k = \sum_{i=1}^{k} D_i \) satisfying

\[
\|S_n - M_n\|_{p'} \leq C_{p,d} \sum_{j=1}^{n} \Theta_{j,p'}.
\]

where \( p' = \min(2, p) \). If \( p \geq 2 \), then we have \( \Gamma_n := \text{Cov}(S_n)/n \to \text{Cov}(D_0) =: \Gamma \) as \( n \to \infty \). (Wu [30] proved (2.3) for the one-dimensional case \( d = 1 \) but the proof is valid for all \( d \).) Throughout, we assume that

\[
\text{when } d \geq 2, \quad \Gamma \text{ is positive definite}.
\]
Next, we introduce a technical condition. Set
\[ U_j(\delta) = \sum_{i=1}^{2^j} |X_i|, \quad j \geq 1, \delta > 0. \]

Define \( \tilde{\theta}_{n,p} \) as the physical dependence measure for \( \{|X_n|\} \). Let \( \tilde{\Theta}_{n,p} = \sum_{i=n}^{\infty} \tilde{\theta}_{i,p} \). Obviously, \( \tilde{\theta}_{n,p} \leq \theta_{n,p} \) and \( \tilde{\Theta}_{n,p} \leq \Theta_{n,p} \). Now, let
\[ \chi_p(n) = \begin{cases} \sqrt{n \log_2 n} & \text{if } p = 2 \\ n^{1/p} & \text{if } 2 < p < 4. \end{cases} \]

The following technical condition is needed.

**Condition A.** Let \( 2 \leq p < 4 \). Suppose that there exists a constant \( C \) satisfying \( 0 < C \leq 1/p \) such that for every \( 0 < \delta < C \) and every \( \varepsilon > 0 \)
\begin{equation}
\sum_{j=1}^{\infty} 2^{j(1-\delta)} P\left( U_j(\delta) \geq \varepsilon \chi_p(2^j) \right) < \infty.
\end{equation}

It should be indicated that (2.5) is an easily verifiable condition. Since \( C \) may be arbitrarily small, Condition A holds if \( E|X_0|^{p+\tau} < \infty \) for some \( \tau > 0 \). For many linear and nonlinear time series, such as the functional of linear processes, GARCH processes, generalized random coefficient autoregressive models, nonlinear AR models (including the threshold autoregressive model, the exponential autoregressive model), bilinear models etc., (2.5) is true under some appropriate conditions on \( \varepsilon_0 \). We will treat these time series separately in Section 3.

We are ready to state our main results now.

**Theorem 2.1.** Let \( 2 \leq p < 4 \) and (2.4), Condition A hold. Suppose that \( E|X_0| = 0, E|X_0|^p < \infty \) and
\begin{equation}
\begin{cases}
\sum_{n=1}^{\infty} \frac{(\log n)^2}{n \log_2 n} \theta_{n,2}^2 < \infty & \text{if } p = 2 \\
\theta_{n,p} = O(n^{-(p-2)/(2(4-p))}) & \text{if } 2 < p < 4
\end{cases}
\end{equation}
for some \( \tau > 0 \). Then on a richer probability space, there exists an \( R^d \) valued Brownian motion \( B(t) \) with covariance matrix \( \Gamma \) such that
\begin{equation}
|S_n - B(n)| = o_{a.s.}(\chi_p(n)).
\end{equation}

**Remark 2.1.** Brownian motion \( B(t) \) with covariance matrix \( \Gamma \) is a Gaussian process \( B(t) \) with values in \( R^d \), independent increments, \( B(0) = 0 \) such that \( B(t) - B(s) \) has normal distribution with mean 0 and covariance matrix \( (t - s)\Gamma, 0 \leq s < t \).

**Remark 2.2.** Let \( 2 < p < 4 \) and \( d = 1 \). We compare our result to Wu’s Theorem 3 [30] in the case of causal functions of an i.i.d. sequence. Theorem 3(ii) and Corollary 4 in his paper proved that if \( E|X_0|^p < \infty \) and
\begin{equation}
\sum_{i=1}^{\infty} i \theta_{i,p} < \infty,
\end{equation}
then $|S_n - B(n)| = O_{a.s.}(n^{1/p} (\log n)^{1/2+1/p} (\log_2 n)^{2/p})$. It is obvious that (2.7) is better than his conclusion. Moreover, (2.8) implies $\Theta_{n,p} = o(n^{-1})$, which is stronger than (2.6) when $2 < p < 10/3$. Of course, we should point out that Wu’s paper deals with stationary processes more general than ours.

**Remark 2.3.** Let $\xi_n' = (\ldots, \varepsilon'_{n-1}, \varepsilon_n')$ and $\xi_n^* = (\xi_{0}', \ldots, \xi_n')$. Define $g_1(\xi_n) = E[g(\xi_{n+1}) | \xi_n]$ and

$$
\widetilde{\alpha}_k = \|g_1(\xi_k) - g_1(\xi_k^*)\|_p, \quad \alpha_k^* = \|g_1(\xi_k) - g_1(\xi_k^*)\|_p, \\
\beta_k^* = \|g(\xi_k) - g(\xi_k^*)\|_p. 
$$

Let $2 < p < 4$ and $d = 1$. An application of Theorem 3(i) and Proposition 3(iii) of [30] shows that if $E|X_0|^p < \infty$, $\Theta_{n,p} = O(n^{1/p-1/2}(\log n)^{1/2})$ and

$$
\sum_{k=1}^{\infty} \left\{ \beta_k^* + \sum_{i=k}^{\infty} \min(\alpha_i^*, \widetilde{\alpha}_{i-k}) \right\} < \infty, \quad (2.9)
$$

then $|S_n - B(n)| = O_{a.s.}(n^{1/p} (\log n)^{1/2})$. Clearly (2.7) gives a better bound. To compare (2.9) with (2.6), we consider the functionals of the linear process example in Section 3.1. Let $h$ be Lipschitz continuous on $R$, $\{X_n\}$ be defined in Section 3.1 and $a_n = n^{-\alpha}$ with $\alpha > 1$. If no other conditions are imposed on $h$, the best possible bounds for $\alpha_n^*$ and $\widetilde{\alpha}_n$ that can be derived by the general methods are

$$
\widetilde{\alpha}_n \leq C a_n, \quad \alpha_n^* \leq C \left( \sum_{i=n}^{\infty} a_i^2 \right)^{1/2} \sim C n^{1/2-\alpha}.
$$

(Take $h(x) = x$ to see the above bounds.) With elementary manipulations,

$$
\sum_{i=k}^{\infty} \min(i^{1/2-\alpha}, (i-k)^{-\alpha}) = \sum_{i=k}^{k+[k^{1-(2\alpha)-1}]} \min(i^{1/2-\alpha}, (i-k)^{-\alpha}) + \sum_{i=k+[k^{1-(2\alpha)-1}]+1}^{\infty} \min(i^{1/2-\alpha}, (i-k)^{-\alpha}) \\
\leq C k^{3/2-\alpha-1/(2\alpha)}.
$$

In order to ensure (2.9), we should let $\alpha > (5 + \sqrt{17})/4 > 2.28$ at least. On the other hand, (2.6) requires $\alpha > (6 - p)/(2(4 - p))$. It can be shown that $(6 - p)/(2(4 - p)) < 2.28$ when $p < 3.438$. Moreover, Condition A will be proved under $E|\varepsilon_0|^p < \infty$ in Section 3.1.

Now we give a theorem which does not need Condition A.

**Theorem 2.2.** Let $2 < p < 4$ and (2.4). Suppose that $E X_0 = 0$, $E|X_0|^p < \infty$ and

$$
\Theta_{n,p} = O(n^{-\eta}), \quad \eta > 0. \quad (2.10)
$$

Set $\tau = \max(1 - 2\eta/(1 + 4\eta), 2/p)$. Then on a richer probability space, there exists an $R^d$ valued Brownian motion $B(t)$ with covariance matrix $\Gamma$ such that

$$
|S_n - B(n)| = o_{a.s.}(n^{\tau/2+\delta}) \quad \text{for any } \delta > 0. \quad (2.11)
$$
Remark 2.4. Let \( d = 1 \). An application of Theorem 4 and Proposition 3(iii) of [30] shows that if \( \mathbb{E}|X_0|^p < \infty \) \((2 < p \leq 4)\), \( \Theta_{n,p} = O(n^{-\eta}) \) and
\[
\beta_n^* + \sum_{i=n}^\infty \min(\alpha_i^*, \tilde{\alpha}_{i-n}) = O(n^{-\eta}), \quad \eta > 0, \tag{2.12}
\]
then \( |S_n - B(n)| = o_{d.s.}(n^{\gamma/2}(\log n)^{3/2}) \), where \( \gamma = \max(1 - \eta, 2/p) \). It is easy to see that when
\[
0 < \eta < \min \left( \frac{1}{4}, \frac{p - 2}{p}, \frac{p - 2}{2(4 - p)} \right),
\]
we have \( \tau > \gamma \). Observing another condition (2.12), we continue the example of Remark 2.3. The largest value for \( \eta \) in (2.12) that we can get is \( \alpha + 1/(2\alpha) - 3/2 \). So \( \gamma = \max(5/2 - \alpha - (2\alpha)^{-1}, 2/p) \). Noting that \( \tau = \max((2\alpha - 1)/(4\alpha - 3), 2/p) \), we can show that when \( 1 < \alpha < 3/2, \)
\[
5/2 - \alpha - (2\alpha)^{-1} > (2\alpha - 1)/(4\alpha - 3), \quad \text{i.e. } \tau \leq \gamma.
\]
For example, if we let \( p = 3 \) and \( \alpha = 7/5 \), then \( \tau = 9/13 < 26/35 = \gamma \). It should be noticed that \( \eta > 1/4 \) in this case. On the other hand, when \( \alpha > 3/2 \), \( \tau \) may be larger than \( \gamma \).

3. Applications

In this section, we suppose that \( \{\varepsilon_n\} \) are i.i.d. real valued random variables. Some applications of Theorem 2.1 will be given. Condition A and (2.6) will be checked.

3.1. The functionals of the linear process

Let \( Y_n = \sum_{i=0}^\infty a_i \varepsilon_{n-i} \) with \( \{a_i\} \) satisfying \( \sum_{i=0}^\infty |a_i| < \infty \). We consider the following functionals of linear processes:
\[
X_n = h(Y_n) - \mathbb{E}h(Y_n)
\]
for some measurable function \( h \). Assume that, for some \( r \geq 1 \),
\[
|h(x) - h(y)| \leq C(|x|^{r-1} + |y|^{r-1})|x - y| \quad \text{for any } x, y \in \mathbb{R}. \tag{3.1}
\]

Corollary 3.1. Let \( 2 \leq p < 4 \). If \( \mathbb{E}|\varepsilon_0|^p < \infty \) and
\[
\{a_n\} \text{ satisfies } \begin{cases}
\sum_{n=1}^\infty \frac{(\log n)^2}{n \log^2 n} \left( \sum_{i=n}^\infty |a_i| \right)^2 < \infty \quad \text{if } p = 2 \\
\sum_{i=n}^\infty |a_i| = O(n^{-(p-2)/(2(4-p)) - \tau}) \quad \text{if } 2 < p < 4
\end{cases} \tag{3.2}
\]
for some \( \tau > 0 \), then (2.7) holds.

Remark 3.1. Let \( p = 2 \), \( h \) be Lipschitz continuous on \( \mathbb{R} \) (namely, \( r = 1 \)) and \( \mathbb{E}\varepsilon_0^2 < \infty \). By Corollary 3.1, we have
\[
\limsup_{n \to \infty} \frac{\pm S_n}{\sqrt{n \log^2 n}} = r^{1/2} \quad \text{a.s.} \tag{3.3}
\]
if
\[
\sum_{n=1}^{\infty} \frac{(\log n)^2}{n \log^2 n} \left( \sum_{i=n}^{\infty} |a_i| \right)^2 < \infty.
\] (3.4)

To obtain (3.3), Theorem 2(ii) of [30] requires

\[
\sum_{n=1}^{\infty} \frac{1}{n} \left( \sum_{i=n}^{\infty} |a_i| \right)^{2/3} < \infty.
\] (3.5)

In the following, we show that (3.4) is weaker than (3.5). From (3.5), \( \sum_{n=1}^{\infty} (\sum_{i=2^n}^{\infty} |a_i|)^{2/3} < \infty \). So \( n (\sum_{i=2^n}^{\infty} |a_i|)^{2/3} = o(1) \), which implies \( \sum_{i=n}^{\infty} |a_i| = o((\log n)^{-3/2}) \). Hence we have

\[
\sum_{n=1}^{\infty} \frac{(\log n)^2}{n} \left( \sum_{i=n}^{\infty} |a_i| \right)^2 \leq \sum_{n=1}^{\infty} \frac{1}{n} \left( \sum_{i=n}^{\infty} |a_i| \right)^{2/3} \left( \sum_{i=n}^{\infty} |a_i| \right)^{2/3} = \sum_{n=1}^{\infty} \frac{1}{n} \left( \sum_{i=n}^{\infty} |a_i| \right)^{2/3} < \infty,
\]

which, of course, yields (3.4).

**Remark 3.2.** Let \( 2 < p < 4 \), \( h \) be Lipschitz continuous on \( \mathbb{R} \) and \( \mathbb{E}|\epsilon_0|^p < \infty \). Wu [30] (Section 3.1 of his paper) showed that under

\[
\sum_{i=1}^{\infty} i |a_i| < \infty,
\] (3.6)

it holds that \( |S_n - B(n)| = O_{a.s.}(\tau_p(n)) \) with \( \tau_p(n) = n^{1/p} (\log n)^{1/2+1/p} (\log_2 n)^{2/p} \).

**Corollary 3.1** improves the rate \( O_{a.s.}(\tau_p(n)) \) to the optimal rate \( o_{a.s.}(n^{1/p}) \). Moreover, it is easy to see that (3.6) implies \( \sum_{i=n}^{\infty} |a_i| = o(n^{-1}) \). So our condition (2.6) is weaker than (3.6) when \( 2 < p < 10/3 \).

**Remark 3.3.** When \( h(x) = x \), Proposition 2 in [30] gives sharp conditions on the linear coefficients to get the rate \( o_{a.s.}(n^{1/p}) \) for any \( p > 2 \). This result cannot be obtained with Corollary 3.1.

### 3.2. Sample autocovariance function of linear process

Let \( Y_n = \sum_{i=0}^{\infty} a_i \epsilon_{n-i} \) with \( \{|a_i|\} \) satisfying \( \sum_{i=0}^{\infty} |a_i| < \infty \). Define \( X_n = Y_n Y_{n-m} - \mathbb{E}Y_n Y_{n-m} \) with fixed \( m \geq 0 \) and \( S_n = \sum_{k=1}^{n} X_k \).

**Corollary 3.2.** Let \( 2 \leq p < 4 \). Suppose that \( \mathbb{E}|\epsilon_0|^{2p} < \infty \) and (3.2) holds. Then we have (2.7).

### 3.3. Augmented GARCH(1, 1) process

Consider a GARCH(1, 1) process \( \{Y_t\} \) satisfying the equations

\[
Y_t = h_t \epsilon_t,
\] (3.7)
and

\[ \Lambda(h_k^2) = c(\varepsilon_{k-1})\Lambda(h_{k-1}^2) + d(\varepsilon_{k-1}), \]

where \( \Lambda(x) \), \( c(x) \) and \( d(x) \) are real valued functions, \( h_t \) are nonnegative random variables and \( \Lambda^{-1}(x) \) exists. This augmented GARCH(1, 1) process was introduced by Duan [8]. For \( \mathbb{E} \log^+ |d(\varepsilon_0)| < \infty, \mathbb{E} \log^+ |c(\varepsilon_0)| < \infty \) and \( \mathbb{E} \log |c(\varepsilon_0)| < 0 \), Aue et al. [1] showed that

\[ \Lambda(h_k^2) = \sum_{i=1}^{\infty} d(\varepsilon_{k-i}) \prod_{j=1}^{i-1} c(\varepsilon_{k-j}) \]

(3.9)
is the only stationary solution to (3.8). They also obtained strong invariance principles for \( \sum_{i=1}^{n} (Y_i^2 - \mathbb{E} Y_i^2) \) with the rate \( o_{a.s.}(n^{5/12+\varepsilon}) \) for any \( \varepsilon > 0 \), under the following smoothness conditions:

\[ \Lambda(h_0^2) \geq \omega \quad \text{with some } \omega > 0, \]

(3.10)
and there are \( C, \gamma \) such that

\[ \left| \frac{1}{\Lambda' (\Lambda^{-1}(x))} \right| \leq Cx^\gamma \quad \text{for all } x \geq \omega. \]

(3.11)

Denote the stationary solution to (3.7) and (3.8) by \( \{Y_n\} \) and let \( X_n = |Y_n|^r - \mathbb{E} |Y_n|^r \) for \( r \geq 1 \). Under the conditions of Corollary 3.3, Theorem 2.2 of [1] ensures that \( \mathbb{E} |X_0|^p < \infty \).

**Corollary 3.3.** Let \( 2 \leq p < 4 \) and (3.10) and (3.11) hold. Suppose that \( \mathbb{E} |\varepsilon_0|^p r < \infty, \mathbb{E} |c(\varepsilon_0)|^{pr(1+\gamma)}/2 < 1 \) and \( \mathbb{E} |d(\varepsilon_0)|^{pr(1+\gamma)}/2 < \infty \). Then (2.7) holds.

**Remark 3.4.** Corollary 3.3 improves Theorem 2.4(i) of [1], in which it was proved that

\[ \sum_{i=1}^{n} (Y_i^2 - \mathbb{E} Y_i^2) - B(n) = o_{a.s.}(n^{5/12+\varepsilon}) \quad \text{for any } \varepsilon > 0 \]

(3.12)
when \( \mathbb{E} |\varepsilon_0|^{8+\delta} < \infty \) for some \( \delta > 0 \), \( \mathbb{E} |c(\varepsilon_0)|^\nu < 1 \) and \( \mathbb{E} |d(\varepsilon_0)|^\nu < \infty \) for some \( \nu > 4(1+\gamma) \). Our result shows that the rate \( o_{a.s.}(n^{5/12+\varepsilon}) \) can be improved to \( o_{a.s.}(n^{1/p}) \) with \( 2 < p < 4 \) under \( \mathbb{E} |\varepsilon_0|^{2p} < \infty, \mathbb{E} |c(\varepsilon_0)|^{p(1+\gamma)}/2 < 1 \) and \( \mathbb{E} |d(\varepsilon_0)|^{p(1+\gamma)}/2 < \infty \). (Theorem 2.4(i) of [1] did not state the condition \( \mathbb{E} |c(\varepsilon_0)|^\nu < 1 \). This condition is needed in fact, observing Lemmas 5.1–5.3 in their paper.)

**Remark 3.5.** Strong approximation for the sample autocovariance function of \( \{Y_n\} \) can be established similarly.

**Remark 3.6.** Aue et al. [1] illustrated the usefulness of (3.12) with some applications from change-point analysis.

### 3.4. Generalized random coefficient autoregressive model

Let

\[ X_n = A_n X_{n-1} + B_n, \quad n \in \mathbb{Z}, \]

(3.13)
where \( \{(A_n, B_n)\} \) are i.i.d. random variables with values in \( M(d) \times R^d \), \( M(d) \) denotes the set of \( d \times d \) real matrices. \( A_n \) is not assumed to be independent of \( B_n \) in this model. For this reason, (3.13) is called the generalized random coefficient autoregressive model [20]. Let

\[
X_n = B_n + \sum_{k=1}^{\infty} A_n A_{n-1} \cdots A_{n-k+1} B_{n-k}.
\]

(3.14)

Define the norm on \( M(d) \) by \( |M| = \sup_{|x| \leq 1} |Mx|, x \in R^d \) for \( M \in M(d) \). If \( E \log |A_0| < 0 \) and \( E \log^+ |B_0| < \infty \), then (3.13) admits a unique strictly stationary solution (3.14); see [4,5]. Define \( S_n = \sum_{i=1}^{n} (X_i - EX_i) \).

**Corollary 3.4.** Let \( 2 \leq p < 4 \) and suppose that \( E|A_0|^p < 1 \) and \( E|B_0|^p < \infty \). Then (2.7) holds.

### 3.5. Bilinear models

In this paragraph, we use the same introduction of bilinear models as Fan and Yao [12], pp 184–185. Let

\[
X_t = \sum_{j=1}^{a} b_j X_{t-j} + \varepsilon_t + \sum_{k=1}^{b} a_k \varepsilon_{t-k} + \sum_{j=0}^{P} \sum_{k=1}^{Q} c_{jk} X_{t-j-k} \varepsilon_{t-k}, \quad t \in Z.
\]

(3.15)

Let \( d = \max\{a, P + b, P + Q\}, m = d - \max\{b, Q\} \), and \( b_{a+j} = a_{b+j} = c_{P+i, Q+j} = 0 \) for all \( i, j \geq 1 \). It has been established by Pham [19,21] that \( X_t \) defined by (3.15) has the state-space representation

\[
X_t = h^\top Z_{t-1} + \varepsilon_t,
\]

and

\[
Z_t = (A + B \varepsilon_t) Z_{t-1} + c \varepsilon_t + d \varepsilon_t^2,
\]

where the state-space variable \( Z_t \) is a \( d \times 1 \) vector with \( X_{t-m+i} \) as its \( i \)th component for \( i = 1, \ldots, m \) and

\[
\sum_{k=j}^{m} b_k X_{t+j-k} + \sum_{k=j}^{n-m} a_k + \sum_{l=0}^{P} c_{lk} X_{t+j-k-l} \varepsilon_{t+j-k}
\]

as its \((m+j)\)th element for \( j = 1, \ldots, d - m \), \( h \) is a \( d \times 1 \) vector with the \((m+1)\)th element 1 and all others 0, \( c \) is a \( d \times 1 \) vector with the first \( m - 1 \) elements 0 followed by 1, \( b_1, a_1, \ldots, b_{d-m}, a_{d-m} \), \( d \) is a \( d \times 1 \) vector with the first \( m \) elements 0 followed by \( c_{01}, \ldots, c_{0,d-m} \), \( B \) is a \( d \times d \) matrix with

\[
\begin{bmatrix}
c_{m1} & \cdots & c_{01} \\
\vdots & \ddots & \vdots \\
c_{m,d-m} & \cdots & c_{0,d-m}
\end{bmatrix}
\]

as the \((d-m) \times (m+1)\) submatrix at the bottom left corner and all of the other elements 0, and \( A \) is a \( d \times d \) matrix with 1 as its \((i, i+1)\) element for \( i = 1, \ldots, d - 1 \), \( b_j \) as its \((m+j, m+1)\)
element for \( j = 1, \ldots, d - m \), and \( b_{d-1-k} \) as its \((d, k)\)th element for \( k = 1, \ldots, m + 1 \) and 0 as all of the other elements.”

Let \( A_t = A + B \epsilon_t, \ c_t = c \epsilon_t + d \epsilon_t^2 \). Then

\[
Z_t = A_t Z_{t-1} + c_t, \tag{3.16}
\]

is the generalized random coefficient autoregressive model. For \( \mathbb{E} \log |A_0| < 0 \) and \( \mathbb{E} \log^+ |c_0| < \infty \), Brandt [4] and Bougerol and Picard [5] proved that (3.16) has the unique strictly stationary solution

\[
Z_t = c_t + \sum_{k=1}^{\infty} A_t A_{t-1} \cdots A_{t-k+1} c_{t-k}.
\]

Now let \( X_t \) be the strictly stationary solution to (3.15) and \( S_n = \sum_{i=1}^{n}(X_t - \mathbb{E}X_t) \).

**Corollary 3.5.** Let \( 2 \leq p < 4 \) and suppose that \( \mathbb{E}|\epsilon_0|^{2p} < \infty, \mathbb{E}|A_0|^p < 1 \) and \( \mathbb{E}|c_0|^p < \infty \). Then (2.7) holds.

3.6. **Nonlinear AR model**

Define the nonlinear autoregressive model by

\[
X_n = f(X_{n-1}) + \epsilon_n, \quad n \in \mathbb{Z}, \tag{3.17}
\]

where \( |f(x) - f(y)| \leq \rho |x - y|, \ 0 < \rho < 1 \). Special cases of (3.17) include the TAR model [28] \( X_n = a \max(X_{n-1}, 0) + b \min(X_{n-1}, 0) + \epsilon_n \) with \( \max(|a|, |b|) < 1 \) and the exponential autoregressive model [13] \( X_n = (a + b \exp(-cX_n^2))X_{n-1} + \epsilon_n \) with \( |a| + |b| < 1 \) and \( c > 0 \). If \( \mathbb{E}|\epsilon_0|^p < \infty \), then \( X_n \) can be represented as \( g(\ldots, \epsilon_{n-1}, \epsilon_n) \) and satisfies \( \theta_{n,p} \leq C \rho^n \) for some \( 0 < \rho < 1 \); see [31]. Let \( S_n = \sum_{i=1}^{n}(X_i - \mathbb{E}X_i) \).

**Corollary 3.6.** Let \( 2 \leq p < 4 \) and \( \mathbb{E}|\epsilon_0|^p < \infty \). Then (2.7) holds.

3.7. **Linear processes with dependent innovations**

Note that the time series in Sections 3.3–3.6 satisfy the following geometric-moment contraction condition: there exist \( C > 0 \) and \( r \in (0, 1) \) such that for all \( n \in \mathbb{N} \),

\[
\|g(\xi_n) - g(\xi_n^*)\|_p \leq C r^n, \tag{3.18}
\]

where \( \xi_n^* = (\ldots, \epsilon_{-1}', \epsilon_0', \epsilon_1', \ldots, \epsilon_n) \). In this subsection, we let

\[
X_n = \sum_{i=0}^{\infty} a_i \eta_{n-i}, \tag{3.19}
\]

where \( \eta_n = g(\xi_n) \) is a real valued random variable satisfying (3.18). We will obtain the strong approximation for \( S_n = \sum_{i=1}^{n}(X_i - \mathbb{E}X_i) \) with slightly slower rate than \( \chi_p(n) \) when \( 2 < p < 4 \). Define \( \phi_p(n) = n^{1/p}/(\log n)^\alpha \) for \( 2 < p < 4 \).

**Corollary 3.7.** Let \( 2 < p < 4 \). Assume that

\[
\sum_{i=n}^{\infty} |a_i| = O(n^{1/p-1/2}/(\log n)^\alpha) \tag{3.20}
\]
improves the rate in Corollary 5(ii) of \[(3.20)\]

For any \(2 < p < 4\) and \(\kappa_p(n) = n^{1/p}(\log n)^{1/2}\) if \(2 < p < 4\) and \(\kappa_p(n) = n^{1/p}(\log n)^{1/2}(\log_2 n)^{1/4}\) if \(p = 4\). On the other hand, (3.20) is slightly stronger than the assumption there of \(\sum_{i=1}^{\infty} |a_i| = O(n^{1/p-1/2}/\log n)\) and that result allows \(p = 4\).

Remark 3.8. If \(\{\eta_n\}_{n \in \mathbb{Z}}\) are the time series considered in Sections 3.3–3.6, then the optimal bound \(o_{a.s.}(n^{1/p})\) can be derived.

4. Proof of the theorems

Let \(I_i\) denote the interval \([2^i, 2^{i+1})\), \(i \geq 0\). For \(0 < a < 1\), \(0 < b < a(1+a)/2\), let \(p_i = \lceil 2^{ai} \rceil\), \(q_i = \lceil 2^{bi} \rceil\), \(k_i = \lceil 2^{i/(p_i + q_i)} \rceil\) and the blocks

\[
\begin{align*}
I_i(j) &= [2^i + (j - 1)(p_i + q_i), 2^i + j p_i + (j - 1)q_i - 1]; \\
J_i(j) &= [2^i + j p_i + (j - 1)q_i, 2^i + j (p_i + q_i) - 1]; \\
1 \leq j \leq k_i, \\
I_i(k_i + 1) &= [2^i + k_i(p_i + q_i), 2^{i+1}).
\end{align*}
\]

Let \(H_p(n)\) be a sequence of real numbers satisfying the following conditions:

1. For any \(t > 1\),

\[
\frac{H_p(n)}{H_p(n^t)} = O(n^{(1-t)/p}).
\]

2. For any \(\nu > 0\), \(n^{1/p-\nu}/H_p(n)\) is eventually non-increasing. Moreover, \(H_p(2n)/H_p(n) = O(1)\).

3. When \(2 < p < 4\), \(H_p(n)/n^{1/2-\gamma}\) is eventually non-increasing for some \(\gamma > 0\).

4. When \(p = 2\), \(\sum_{n=1}^{\infty} \exp \left(-C2^{-n}H_2^2(2^n)\right) < \infty\) for some \(C > 0\).

5. For some \(\epsilon > 0\), \(n^\epsilon/H_2^2(n) = O(n^{-\epsilon})\) and \(n^{1-a+b}/H_2^2(n) = O(n^{-\epsilon})\).

6. We have

\[
T_p := \sum_{n=1}^{\infty} 2^{pn/2} \Theta_{[2^{ln}],\nu}(n^2 I\{p = 2\} + 1) < \infty.
\]

7. Let \(2 \leq p < 4\). Suppose that there exists a constant \(C\) satisfying \(0 < C \leq 1/p\) such that for every \(0 < \delta < C\) and every \(\epsilon > 0\)

\[
\sum_{j=1}^{\infty} 2^{j(1-\delta)} P \left( U_j(\delta) \geq \epsilon H_p(2^j) \right) < \infty.
\]

After choosing some appropriate values of \(a\) and \(b\), \(H_p(n) = n^\alpha h_n\) would satisfy (1)–(5), where \(1/p \leq \alpha < 1/2\) and \(h_n\) is slowly varying and increasing. This holds true also for \(H_2(n) = \sqrt{n \log_2 n}\).

The following lemma is available for using and checking (7).
Lemma 4.1. Let \( 2 \leq p < 4 \) and \( \mathbb{E}|X_0|^p < \infty \). Under (1) and (6), (4.3) is equivalent to
\[
\sum_{j=1}^{\infty} 2^{j(1-\delta)} \mathbb{P} \left( U_j'(\delta) \geq \varepsilon H_p(2^j) \right) < \infty
\]
for every \( \varepsilon > 0 \), where
\[
U_j'(\delta) = \sum_{i=1}^{2^{\delta j}} |X_i|', \quad |X_i|' = \mathbb{E}[|X_i||\varepsilon_{i-2^{\delta j}}, \ldots, \varepsilon_i], \quad 1 \leq i \leq [2^{\delta j}] .
\]

Proof. By (1), Lemma A.1 and the arguments in (4.11) below, we can get the lemma immediately. \( \square \)

In order to prove Theorems 2.1 and 2.2, we only need to prove a general theorem based on (1)–(7).

Theorem 4.1. Let \( 2 \leq p < 4 \), (2.4) and (1)–(7) hold. Suppose that \( \mathbb{E}X_0 = 0 \) and \( \mathbb{E}|X_0|^p < \infty \). Then on a richer probability space, there exists an \( R^d \) valued Brownian motion \( B(t) \) with covariance matrix \( \Gamma \) such that
\[
|S_n - B(n)| = o_{a.s.}(H_p(n)). \tag{4.4}
\]

The proof of Theorem 4.1 is technical and consists of a series of lemmas. First of all, using Lemma A.1, we approximate \( S_n \) by the partial sums of \( m \)-dependent random vectors. Then the partial sums are decomposed into two parts, which we call big blocks and small blocks. Small blocks are negligible in view of Lemmas 4.2 and 4.3. With the help of Gaussian approximation results obtained by Einmahl [10,11], we use Lemmas 4.4–4.6 to conclude that big blocks can be approximated by an \( R^d \) valued Brownian motion.

Proof of Theorem 4.1. Let \( \bar{X}_j = \mathbb{E}[X_j|\varepsilon_{j-q_i-1}, \ldots, \varepsilon_j] \) for \( 2^i \leq j < 2^{i+1} \). By Lemma A.1 in the Appendix and (6), we can conclude that for every \( \varepsilon > 0 \),
\[
\sum_{i=1}^{\infty} \mathbb{P} \left( \max_{2^i \leq j < 2^{i+1}} \sum_{k=2^i}^{j} (X_k - \bar{X}_k) \geq \varepsilon H_p(2^i) \right) \leq C_{p,d,\varepsilon} \sum_{i=0}^{\infty} \frac{2^{i p/2} \Theta_{q_i} \rho_i}{(H_p(2^i))^p} (i^2 I\{p = 2\} + 1) < \infty .
\]

By the Borel–Cantelli lemma, we have
\[
\max_{2^i \leq j < 2^{i+1}} \left| \sum_{k=2^i}^{j} (X_k - \bar{X}_k) \right| = o_{a.s.}(H_p(2^i)). \tag{4.5}
\]

A simple proof using (2) implies that there exists \( c > 1 \) such that \( H_p(2^{n+1})/H_p(2^n) > c \) when \( n \) is large. Hence we can infer from standard arguments that (4.5) implies
\[
\max_{1 \leq j \leq 2^i} \left| \sum_{k=1}^{j} (X_k - \bar{X}_k) \right| = o_{a.s.}(H_p(2^i)). \tag{4.6}
\]

Put
\[
\xi_i(j) = \sum_{k \in I_i(j)} \bar{X}_k, \quad \eta_i(j) = \sum_{k \in J_i(j)} \bar{X}_k, \quad 1 \leq j \leq k_i ,
\]
\[ \eta_i(k_i + 1) = \sum_{k \in I_i(k_i + 1)} X_k. \]

For every \( n > 0 \), there exist integers \( m_n \geq 0 \) and \( 1 \leq t_n \leq k_{m_n} + 1 \) such that \( 2^{m_n} \leq n < 2^{m_n + 1} \) and \( n \in \mathbb{I}_{m_n}(t_n) \cup \mathbb{I}_{m_n}(t_n) \) (define \( \mathbb{I}_{m_n}(k_{m_n} + 1) = \emptyset \)). Let \( N_{m_n} = 2^{m_n} + \text{card} \{ \bigcup_{j=1}^{t_n-1} \{ I_{m_n}(j) \cup I_{m_n}(j) \} \} \). From (4.6) and the blocks mentioned above, \( S_n \) may be decomposed into

\[
S_n = \left\{ \sum_{i=1}^{m_n-1} \sum_{j=1}^{t_n-1} \xi_i(j) + \sum_{j=1}^{t_n} \xi_{m_n}(j) \right\} + \left\{ \sum_{i=1}^{m_n-1} \sum_{j=1}^{t_n-1} \eta_i(j) + \sum_{j=1}^{t_n} \eta_{m_n}(j) \right\}
+ \sum_{i=N_{m_n}+1}^{n} X_i + o_{a.s.}(H_p(n)) =: S_{1,n} + S_{2,n} + S_{3,n} + o_{a.s.}(H_p(n)).
\]

We will show that the sums of small blocks \( S_{2,n} \) and \( S_{3,n} \) are negligible, while the sums \( S_{1,n} \) can be approximated by a Brownian motion.

**Lemma 4.2.** Under the conditions of **Theorem 4.1**, it holds that \( |S_{3,n}| = o_{a.s.}(H_p(n)). \)

**Proof.** We set \( a_{n,j} = 2^n + (j - 1)(p_n + q_n) \) for notational brevity. It is enough to prove

\[
\max_{1 \leq j \leq k_n+1} \max_{a_{n,j} \leq k \leq a_{n,j+1}} \left| \sum_{i=n_j}^{k_n+1} X_i \right| = o_{a.s.}(H_p(2^n)). \tag{4.7}
\]

By the stationarity and the Borel–Cantelli lemma, (4.7) follows immediately if we can show that for every \( \varepsilon > 0 \),

\[
\mathbb{Q} := \sum_{n=1}^{\infty} 2^{(1-a)n} \mathbb{P} \left( \max_{1 \leq j \leq 2p_n} \left| \sum_{k=1}^{j} X_{k,n}' \right| \geq \varepsilon H_p(2^n) \right) < \infty, \tag{4.8}
\]

where \( X_{k,n}' = \mathbb{E}[X_k|\varepsilon_{k-q_n-1}, \ldots, \varepsilon_k] \) for \( 1 \leq k \leq 2p_n \).

Let \( b < \Delta < 1 \) and \( r \) be an integer satisfying \( \Delta^r b < C \) (recall \( C \) in condition (7)). Define \( q_n(i) = [2\Delta^r bn] \) for \( 0 \leq i \leq r \) and

\[ X_{k,n}'(i) = \mathbb{E}[X_k|\varepsilon_{k-q_n-1(i)}, \ldots, \varepsilon_k], 1 \leq k \leq 2p_n, 0 \leq i \leq r. \]

It is easily seen that \( X_{k,n}'(i) - X_{k,n}'(i+1) = \sum_{i=0}^{r-1} (X_{k,n}'(i) - X_{k,n}'(i+1)). \) Hence

\[
\mathbb{Q} \leq \sum_{n=1}^{\infty} 2^{(1-a)n} \mathbb{P} \left( \max_{1 \leq j \leq 2p_n} \left| \sum_{k=1}^{j} X_{k,n}'(r) \right| \geq 2^{-1} \varepsilon H_p(2^n) \right) \]
\[ + \sum_{i=0}^{r-1} \sum_{n=1}^{\infty} 2^{(1-a)n} \mathbb{P} \left( \max_{1 \leq j \leq 2p_n} \left| \sum_{k=1}^{j} (X_{k,n}'(i) - X_{k,n}'(i+1)) \right| \geq C_r \varepsilon H_p(2^n) \right) \]
\[ =: \mathbb{Q}_1 + \mathbb{Q}_2. \]

First we use the blocking method to deal with \( \mathbb{Q}_2 \). For every \( 0 \leq i \leq r \), let us split the interval \([1, 2p_n]\) into \( 2d_n(i) := 2p_n/q_n(i) \) (without loss of generality, we assume that \( d_n(i) \) is an integer) small intervals having the length \( q_n(i) \). Denote these intervals by \( J_{1,n}(i),\]
\[ K_{1,n}(i), \ldots, J_{d_n(i),n}(i), K_{d_n(i),n}(i) \) and write \( \xi_j(i) = \sum_{m \in J_{j,n}(i)} (X_{m,n}'(i) - X_{m,n}'(i+1)), \]
Lemma A.2 is finite by (5) and letting and observing that Lemma A.1 (4.10) where we have used first term in By (4.10), we use \( \varsigma \) to denote \( \varsigma_j(i) \). Then \( \{\varsigma_j, 1 \leq j \leq d_n\} \) and \( \{\varsigma_j, 1 \leq j \leq d_n\} \) are two sets of independent random vectors. Note that

\[
\begin{align*}
\max_{1 \leq j \leq 2p_n} \left| \sum_{k=1}^{j} (X'_{k,n}(i) - X'_{k,n}(i+1)) \right| & \leq \max_{1 \leq j \leq d_n} \left| \sum_{k=1}^{j} \varsigma_k \right| + \max_{0 \leq j \leq d_n} \left| \sum_{k=1}^{j} \varsigma_k \right| \\
+ \max_{1 \leq j \leq 2d_n} \max_{0 \leq j \leq (j+1)q_n(i)} \left| \sum_{m=jq_n(i)+1}^{k} (X'_{m,n}(i) - X'_{m,n}(i+1)) \right|
\end{align*}
\]

\[
= L_1 + L_2 + L_3.
\]

(4.9)

By taking \( x = d^{-1}eH_p(2^n) \) and \( y = x/(12q) \) with \( q > 0 \) in the Fuk–Nagaev-type inequality (Lemma A.3), we can obtain that, for every \( q > 0 \) and every \( \varepsilon > 0 \),

\[
\sum_{n=1}^{\infty} 2^{(1-a)n} P(L_1 \geq \varepsilon H_p(2^n)) 
\]

\[
\leq C_{\varepsilon,q,d} \sum_{n=1}^{\infty} 2^{(1-a)n} \left( \frac{d_n E|\varsigma_1|^2}{(H_p(2^n))^2} \right)^q + C_{\varepsilon,q,d} \sum_{n=1}^{\infty} 2^{(1-a)n} d_n \frac{E|\varsigma_1|^p}{(H_p(2^n))^p}.
\]

(4.10)

By Lemma A.2, \( E|\varsigma_1|^2 = O(q_n(i)) \). Recall that \( d_n(i) \sim 2^{(a-\Delta b)n} \). It is readily seen that the first term in (4.10) is finite by (5) and letting \( q \) be large enough. We claim that the second term is also finite. Actually, by virtue of Lemma A.1,

\[
E|\varsigma_1|^p \leq C_{p,d}(q_n(i))^{p/2} \Theta_{[q_{n-1}(i)+1],p}^p \leq C_{p,d} 2^{\Delta b p n/2} \Theta_{[2^{\Delta b (n-1)+1}],p}^p.
\]

Hence the claim is proved by taking this estimate back into (4.10) and observing that

\[
\begin{align*}
\sum_{n=1}^{\infty} 2^{(1-\Delta b n + \Delta b p n/2)} \Theta_{[2^{\Delta b n/2}],p}^p (H_p(2^n))^{-p} \\
\leq C \sum_{n=1}^{\infty} \sum_{k=1}^{n+1} 2^{k + \Delta b (p-2)k/2} \Theta_{[2^{\Delta b k/2}],p}^p (H_p(2^k))^{-p} \\
\leq C \sum_{n=1}^{\infty} 2^{n/2} \Theta_{[2^n],p}^p (H_p(2^n))^{-p} \\
\leq C \sum_{n=1}^{\infty} 2^{pn/2} \Theta_{[2^n],p}^p (H_p(2^n))^{-p}
\end{align*}
\]

(4.11)

where we have used \( b < \Delta \), (1) and (6) in the third and fourth inequalities respectively. Similarly,

\[
\sum_{n=1}^{\infty} 2^{(1-a)n} P(L_2 + L_3 \geq \varepsilon H_p(2^n)) < \infty.
\]

Combining the inequalities above, we obtain that \( Q_2 < \infty \).
It remains for us to show that $\mathbb{Q}_1 < \infty$. Write $\alpha_{j,n} = \sum_{m \in J_{j,n}(r)} X_{m,n}'(r)$ and $\beta_{j,n} = \sum_{m \in K_{j,n}(r)} X_{m,n}'(r)$ for $1 \leq j \leq d_n(r)$. Then

$$
\mathbb{Q}_1 \leq 2 \sum_{n=1}^{\infty} 2^{(1-a)n} \mathbb{P} \left( \max_{1 \leq j \leq d_n(r)} \left| \sum_{k=1}^{j} \alpha_{k,n} \right| \geq 8^{-1} \varepsilon H_p(2^n) \right) \\
+ \sum_{n=1}^{\infty} 2^{(1-a)n} d_n(r) \mathbb{P} \left( \max_{1 \leq k \leq q_n(r)} \left| \sum_{i=1}^{k} X_{i,n}'(r) \right| \geq 4^{-1} \varepsilon H_p(2^n) \right) \\
=: \mathbb{Q}_{11} + \mathbb{Q}_{12}.
$$

By a proof analogous to that of (4.10), we can get $\mathbb{Q}_{11} < \infty$ if

$$
\sum_{n=1}^{\infty} 2^{(1-a)n} d_n(r) \mathbb{P} \left( |\alpha_{1,n}| \geq \varepsilon H_p(2^n) \right) < \infty \tag{4.12}
$$

for every $\varepsilon > 0$. Actually, this follows from (7) and Lemma 4.1 immediately since $\Delta' b < C$. Similarly, $\mathbb{Q}_{12} < \infty$, which entails $\mathbb{Q}_1 < \infty$, and hence completes the proof of the lemma. □

**Lemma 4.3.** Under the conditions of Theorem 4.1, we have $|S_{2,n}| = o_{a.s.}(H_p(n))$.

**Proof.** The proof of Lemma 4.3 is contained in the proof of Lemma 4.2. In fact, it can be concluded from (4.7) that

$$
\sum_{i=1}^{m_n-1} |\eta_i(k_i + 1)| = \sum_{i=1}^{m_n-1} o(H_p(2^i)) = o(H_p(n)) \quad \text{a.s.}
$$

Therefore it suffices to show that

$$
\max_{1 \leq m \leq n} \left\{ \sum_{i=1}^{m-1} \sum_{j=1}^{k_i} \eta_i(j) + \max_{1 \leq t \leq k_m} \sum_{j=1}^{t} \eta_m(j) \right\} = o_{a.s.}(H_p(2^n))
$$

which will be derived if

$$
\max_{1 \leq i \leq k_n} \left| \sum_{j=1}^{i} \eta_n(j) \right| = o_{a.s.}(H_p(2^n)). \tag{4.13}
$$

We note that $\{\eta_i(j), i \geq 1, 1 \leq j \leq k_i\}$ are independent random vectors. Then by the Borel–Cantelli lemma, Lemma A.3, (4.8) and observing that $k_n \mathbb{E} |\eta_n(1)|^2 = O(2^n(1-\alpha+b))$, (4.13) holds. □

In view of (7), we see that for any $0 < \delta < C$, there exists a sequence of positive numbers $\{\varepsilon_n(\delta)\}$ satisfying $\varepsilon_n(\delta) \to 0$ sufficiently slowly (for example, $\varepsilon_n(\delta) \log_3 n \to \infty$) such that for every $\varepsilon > 0$,

$$
\sum_{j=1}^{\infty} 2^{j(1-b)} \mathbb{P} \left( U_j(\delta) + U_j'(\delta) \geq \varepsilon \varepsilon_j(\delta) H_p(2^j) \right) < \infty
$$

and $\varepsilon_n(\delta) H_2(2^n)/2^n \to \infty$. (We remark that (4) implies $H_2(n)/n \to \infty$.) Now take $\theta_i = \varepsilon_i(\Delta' b)$ with $\Delta' b < C$ and let

$$
\xi_i'(j) = \xi_i(j) I\{|\xi_i(j)| \leq \theta_i H_p(2^i)\}, \quad \xi_i''(j) = \xi_i'(j) - \mathbb{E} \xi_i'(j), \quad 1 \leq j \leq k_i, i \geq 1.
$$
Lemma 4.4. Under the conditions of Theorem 4.1, we have
\[ \kappa_n \mathbb{E}[\xi_n(1)| I \{ |\xi_n(1)| \geq \theta_n H_p(2^n) \}] = o(H_p(2^n)). \] (4.14)

Proof. If \( p = 2 \), we have from Lemma A.2 that
\[ \kappa_n \mathbb{E}[\xi_n(1)| I \{ |\xi_n(1)| \geq \theta_n H_p(2^n) \}] \leq \theta_n^{-1} k_n (H_p(2^n))^{-1} \mathbb{E}[\xi_n(1)]^2 I \{ |\xi_n(1)| \geq \theta_n H_p(2^n) \} = o(H_p(2^n)), \]
and hence (4.14) holds. Suppose now that \( 2 < p < 4 \). It is easy to see that
\[ \mathbb{E}[\xi_n(1)| I \{ |\xi_n(1)| \geq \theta_n H_p(2^n) \}] \leq C \sum_{j=n}^{\infty} H_p(2^j) \mathbb{P}\left( \left| \sum_{k=1}^{n_j} X'_{k,n} \right| \geq \theta_n H_p(2^j) \right). \] (4.15)

In order to estimate the series in (4.15), we should note that the arguments from (4.8) to (4.12) imply
\[ \mathbb{P}\left( \max_{1 \leq i \leq n_j} \left| \sum_{k=1}^{i} X'_{k,n} \right| \geq \theta_n H_p(2^j) \right) \leq O \left( \left( \frac{2^{an}}{\theta_n^2 H_p^2(2^j)} \right)^q \right) + O(1) \sum_{i=0}^{r-1} d_n(i) \frac{2^{\Delta^j bpn/2} \Theta^p}{\theta_n^p H_p(2^j)^p} + O(1) d_n(r) \mathbb{P}\left( \max_{1 \leq i \leq n(r)} \left| \sum_{i=1}^{k} X'_{i,n}(r) \right| \geq C_{q,d} \theta_n H_p(2^j) \right), \] (4.16)
where \( q \) can be arbitrarily large, and
\[ \mathbb{P}\left( \max_{1 \leq i \leq n(r)} \left| \sum_{i=1}^{k} X'_{i,n}(r) \right| \geq C_{q,d} \theta_n H_p(2^j) \right) \leq \mathbb{P}\left( U_n(\Delta^j b) \geq C_{q,d} \theta_n H_p(2^j) \right) + C_{d,e,p} \frac{2^{\Delta^b d n/2} \Theta^p}{\theta_n^p H_p(2^j)^p}. \] (4.17)

By virtue of (3) and the fact \( U_n(\Delta^j b) \leq U_j(\Delta^j b) \) when \( j \geq n \),
\[ (H_p(2^n))^{-1} k_n \sum_{j=n}^{\infty} H_p(2^j) d_n(r) \mathbb{P}\left( U_n(\Delta^j b) \geq C_{q,d} \theta_n H_p(2^j) \right) \leq (H_p(2^n))^{-1} 2^{(1-\Delta^j b)n} \sum_{j=n}^{\infty} H_p(2^j) \mathbb{P}\left( U_j(\Delta^j b) \geq C_{q,d} \theta_j H_p(2^j) \right) \]
\[ = O(1) \sum_{j=n}^{\infty} 2^{(1-\Delta^j b)j} \mathbb{P}\left( U_j(\Delta^j b) \geq C_{q,d} \theta_j H_p(2^j) \right) \]
\[ = o(1). \]
Taking this estimate, (4.16) and (4.17) back into (4.15) and by tedious but simple calculations, we can get (4.14) immediately. \( \square \)

By virtue of (4.16) and (4.17), it is readily seen that
\[
\sum_{n=1}^{\infty} 2^{(1-a)n} P\left( \max_{1 \leq |i| \leq p_n} \left| \sum_{k=1}^{i} X_{k,n}^i \right| \geq \theta_n H_p(2^n) \right) < \infty. \quad (4.18)
\]

Write \( \Sigma_{n,j} := \text{Cov}(\xi_n^{(j)}), n \geq 1, 1 \leq j \leq k_n \). The stationarity of \( \{X_n\} \) yields \( \Sigma_{n,j} = \Sigma_{n,1} \) for \( 1 \leq j \leq k_n \).

**Lemma 4.5.** Suppose that \( \Gamma = I \) and the conditions of Theorem 4.1 hold. Then for every \( \varepsilon > 0 \),
\[
\mathbb{Q}_3 := \sum_{n=1}^{\infty} \exp\left( -\frac{\varepsilon H_p^2(2^n)}{k_n |\sqrt{p_n I - \Sigma_{n,1}^{1/2}|^2}} \right) < \infty,
\]
where the norm of a matrix \( A \) is defined by \( |A| = \sup_{|x| \neq 0} |Ax|/|x| \).

**Proof.** Set \( T_n(j) = \sum_{i \in E_n(j)} X_i = (T_{n,1}(j), \ldots, T_{n,d}(j)), 1 \leq j \leq k_n \).

Routine calculations imply that
\[
|\Sigma_{n,1} - \text{Cov}(\xi_n(1))| \leq C E|\xi_n(1)|^2 I \{ |\xi_n(1)| \geq \theta_n H_p(2^n) \}. \quad (4.19)
\]
By Lemma A.1 and (2.3), we have
\[
\mathbb{E}|T_n(1) - \xi_n(1)|^2 = O(2^{an} \Theta_{q_n,2}^2), \quad (4.20)
\]
\[
|\text{Cov}(T_n(1)) - \text{Cov}(\xi_n(1))|^2 = O(2^{2an} \Theta_{q_n,2}^2), \quad (4.21)
\]
and
\[
|\text{Cov}(T_n(1)) - p_n I|^2 = O(1)2^{an} \sum_{i=1}^{p_n} \Theta_{i,2}^2. \quad (4.22)
\]
Let us deal with the case of \( p = 2 \) first. Set \( S_n = (S_{n,1}, \ldots, S_{n,d}) \) and \( D_0 = (D_{0,1}, \ldots, D_{0,d}) \). Then by (2.3), \( S_{n,i}/\sqrt{n} \rightarrow N(0, \sigma_i^2) \) in distribution and \( \mathbb{E}S_{n,i}^2/n \rightarrow \sigma_i^2 \) for \( 1 \leq i \leq d \), where \( \sigma_i^2 = ED_{0,i} \). This yields that \( \{S_{n,i}^2/n\} \) is uniformly integrable. Hence
\[
\mathbb{E}|T_n(1)|^2 I \{ |T_n(1)| \geq \theta_n H_p(2^n) \} = o(2^{an}). \quad (4.23)
\]
Since for any semi-positive definite matrix \( A, |(I - A)^2| \leq |I - A^2|^2 \), we conclude from (4.19) to (4.23) that
\[
k_n|\sqrt{p_n I - \Sigma_{n,1}^{1/2}}| \leq k_n p_n^{-1} |p_n I - \Sigma_{n,1}|^2 = o(2^n). \quad (4.24)
\]
This together with (4) proves the lemma.

Next we will prove the lemma for when \( 2 < p < 4 \). Let \( r \) in (4.16) and (4.17) be large enough that \( \Delta^r b \leq \gamma \), where \( \gamma \) is defined in (3). As for (4.14), it follows easily from (4.16) and (4.17) and simple calculations that
\[
H_p^{-2}(2^n) k_n 2^{-na} \left( \mathbb{E}|\xi_n(1)|^2 I \{ |\xi_n(1)| \geq \theta_n H_p(2^n) \} \right)^2 \leq C H_p^{-2}(2^n) 2^{n-2an} \left( \sum_{j=n}^{\infty} \frac{2^{aqn}}{\theta_n^2 H_p^{2q-2}(2^j)} \right)^2
\]
Lemma 4.5

and the inequality \( e^{-x} \leq C x^{-q} \) for any \( q > 0 \) and \( x > 0 \), it is readily seen that

\[
Q_3 \leq C \sum_{n=1}^{\infty} \left( \frac{2^{(1-a)n} \sum_{i=1}^{p_n} \Theta_{i,2}^q}{H_p^2(2^n)} \right) + C \sum_{n=1}^{\infty} \frac{2^{qn} \Theta_{[2bn],p}^{2q}}{H_p^2(2^n)}
+ C \sum_{n=1}^{\infty} (I_1(n) + I_2(n) + I_3(n))^q < \infty.
\]

This completes the proof of Lemma 4.5. \( \square \)
Lemma 4.6. Under the conditions of Theorem 4.1, we have for some \( q > 2 \)

\[
\mathbb{Q}_4 := \sum_{n=1}^{\infty} k_n \frac{\mathbb{E}[|\xi_n''(1)|^q]}{(H_p(2^n)^q) \gamma} < \infty. \tag{4.25}
\]

Remark 4.1. From the proof we see that if \( p = 2 \) then \( q \) can be less than 4.

Proof. We have

\[
\mathbb{Q}_4 \leq C \sum_{n=1}^{\infty} 2^{(1-a)n} \sum_{k=1}^{\infty} \frac{\mathbb{P}(H_p(2^{k-1}) < |\xi_n(1)| \leq H_p(2^k))}{(H_p(2^n)^q) \gamma}
\]

\[
\leq C \sum_{n=1}^{\infty} 2^{(1-a)n} \sum_{k=1}^{\infty} \frac{\mathbb{P}(|\xi_n(1)| \geq H_p(2^{k-1}))}{(H_p(2^n)^q) \gamma} + C,
\tag{4.26}
\]

where \( 0 < \mu < 1 \) satisfying \( a\mu > b \) is taken as follows: Since \( b < a(1+a)/2 \) and \( (1+a)/2 > a \), we can let \( \delta > 0 \) be small enough that \( (1+a)/2 - \delta > a \vee (b/a) \). When \( p = 2 \), we take \( \mu = (1+a)/2 - \delta \) and \( q = \frac{4(1-a)}{1+a+\delta} \). So by (1),

\[
\sum_{n=1}^{\infty} 2^{(1-a)n} \sum_{k=1}^{\infty} \frac{\mathbb{P}(H_p(2^{k}) \leq |\xi_n(1)| \leq H_p(2^{k+1}))}{(H_p(2^n)^q) \gamma} \leq C \sum_{n=1}^{\infty} 2^{(1-a)n} 2^{(\mu-1)nq/2} < \infty.
\]

When \( p > 2 \), we let \( \max(\alpha/(\alpha + \epsilon), b/a) < \mu < 1 \) and \( q \) be large enough to ensure (4.26). (Recall \( \epsilon \) defined in (5).)

Now, for \( \mu n \leq k \leq n \), let us split the interval \( [1, p_n] \) into blocks \( J_j, K_j, 1 \leq j \leq d_n(k) \), with equal length \( p_k \). Here \( 2d_n(k) = p_n/p_k \) (we assume that \( d_n(k) \) is an integer for brevity, even if \( k = n \)). Clearly, \( d_n(k) \) is proportional to \( 2^{a(n-k)} \). Set

\[
\xi_{j,k} = \sum_{m \in J_j} X_{m,n}', \quad \eta_{j,k} = \sum_{m \in K_j} X_{m,n}', \quad 1 \leq j \leq d_n(k).
\]

Since \( a\mu > b, k \geq \mu n \), we have \( p_k \geq q_n \) and hence \( \{\xi_{j,k}, 1 \leq j \leq d_n(k)\} \) are independent random vectors. This is true also for \( \{\eta_{j,k}\} \). Let \( x = H_p(2^{k-1}), y = x/(12t) \) in Lemma A.3. By virtue of Lemmas A.3 and A.1, we have

\[
P(\xi_{j,k} \geq H_p(2^{k-1})) \leq O \left( \left( \frac{2^{an}}{(H_p(2^k))^2} \right)^{t} + d 2^{a(n-k)+1} \mathbb{P}(\xi_{1,k} \geq C_{t,d} H_p(2^{k-1})) \right)
\]

\[
\leq O \left( \left( \frac{2^{an}}{(H_p(2^k))^2} \right)^{t} + d 2^{a(n-k)+1} \mathbb{P} \left( \left\{ \sum_{i=1}^{p_k} X_{i,k}' \geq 2^{-1} C_{t,d} H_p(2^{k-1}) \right\} \right) \right)
\]

\[
+ C_{t,d,p} (H_p(2^{k-1}))^{-p} 2^{a(n-k)+apk/2} \Theta_{2^{k+1},p}^{p_k},
\]

where \( t \) is large enough. From (4) and the choice of \( \mu \) referred to above, we know that (i) when \( p = 2, H_p^2(2^n) \geq C 2^n \) and \( \mu > a \); (ii) if \( p > 2 \), then by (5) and \( \mu > a/(a + \epsilon) \) it holds that
\[ 2^{an} = O(1)2^{(a-\mu a-\mu e)n} H_p^2(2^{\mu n}). \] Therefore, we have for \( p \geq 2 \)

\[
\sum_{n=1}^{\infty} 2^{(1-a)n} \frac{\sum_{k=\mu n}^{n} H_p^q(2^k) \left( \frac{2^{an}}{(H_p(2^k))^q} \right)^t}{(H_p(2^n))^q} < \infty
\]

when \( t \) is large. By (2) and routine calculations, we can get

\[
\sum_{n=1}^{\infty} 2^{(1-a)n} \frac{\sum_{k=\mu n}^{n} (H_p(2^k))^{q-p} 2^{a(n-k)+apk/2} \Theta_{[2^{bk}, p]}^P}{H_p^q(2^n)} \leq C \sum_{k=1}^{\infty} 2^{a(p-2)k/2} (H_p(2^k))^{q-p} \Theta_{[2^{bk}, p]}^P \sum_{n=k}^{\infty} \frac{2^{n-(1/p-v)qn} 2^{(1/p-v)qn}}{H_p^q(2^n)} \leq C \sum_{k=1}^{\infty} 2^{pk/2} \Theta_{[2^{bk}, p]}^P < \infty.
\]

Taking these inequalities back into (4.26) yields

\[
\mathcal{Q}_4 \leq C \sum_{n=1}^{\infty} 2^{(1-a)n} \frac{\sum_{k=\mu n}^{n} H_p^q(2^k) 2^{a(n-k)} P \left( \left| \sum_{i=1}^{pk} X'_{i,k} \right| \geq C \sum_{i=1}^{pk} X'_{i,k} \right)}{(H_p(2^n))^q} + C
\]
\[
\leq C \sum_{k=1}^{\infty} 2^{(1-a)k} P \left( \left| \sum_{i=1}^{pk} X'_{i,k} \right| \geq C \sum_{i=1}^{pk} X'_{i,k} \right) + C.
\]

This together with (4.18) implies that \( \mathcal{Q}_4 < \infty \). \( \square \)

**Lemma 4.7.** Suppose that the conditions of Theorem 4.1 hold. Then on a richer probability space, there exists an \( R^d \) valued Brownian motion \( B(t) \) with covariance matrix \( \Gamma \) such that

\[
|S_{1,n} - B(n)| = o_{a.s.}(H_p(n)). \tag{4.27}
\]

**Proof.** Suppose that \( \Gamma > 0 \) when \( d = 1 \). Therefore, without loss of generality, we can assume that \( \Gamma = I. \)

By (4.14) and (4.18), in order to prove the lemma, we only need to infer that

\[
\left| \sum_{i=1}^{m_n-1} \sum_{j=1}^{k_i} \xi''_{i,j}(j) + \sum_{j=1}^{\ell_n-1} \xi''_{m_n}(j) - B(n) \right| = o_{a.s.}(H_p(n)). \tag{4.28}
\]

In the following, we introduce some notation. Write \( \xi''_{i,j}(j) = \Sigma_{i,j}^{-1/2} \xi''_{i,j}(j). \) (From (4.24), we see that \( p_n^{-1} \Sigma_{n,1} \to I. \) So we can assume that \( \Sigma_{n,1} \) is positive definite for all \( n. \) ) Clearly, \( \xi''_{i,j}(j) = \Sigma_{i,j}^{-1/2} \xi''_{i,j}(j). \) Put

\[
\widetilde{z}_n(t) = \sum_{i=1}^{n-1} \sum_{j=1}^{k_i} (\sqrt{p_i} I - \Sigma_{i,j}^{1/2}) \xi''_{i,j}(j) + \sum_{j=1}^{\ell_n-1} (\sqrt{p_n} I - \Sigma_{n,j}^{1/2}) \xi''_{m_n}(j),
\]

\[
\Sigma_n(t) = \sum_{i=1}^{n-1} \sum_{j=1}^{k_i} (\sqrt{p_i} I - \Sigma_{i,j}^{1/2}) \xi''_{i,j}(j) + \sum_{j=1}^{\ell_n-1} (\sqrt{p_n} I - \Sigma_{n,j}^{1/2}) \xi''_{m_n}(j).
\]
\[
\sum_{n=1}^{\infty} P \left( \max_{1 \leq t \leq k_n} \left| \tilde{S}_n(t) - \tilde{S}_n-1(k_{n-1}) \right| \geq \varepsilon H_p(2^n) \right) < \infty. \quad (4.29)
\]

By the Borel–Cantelli lemma, in order to prove (4.28), it suffices to show that
\[
|S'_{m_n}(t_n - 1) - B(n)| = o_{a.s.}(H_p(n)). \quad (4.30)
\]

The case \( p = 2 \). Let \( b_n = \sum_{i=1}^{n} k_i \). Define \( \{a_n\} \) in Lemma A.4 by \( a_k = H_q(2^n) \) if \( b_n-1 < k \leq b_n \). We also define \( \{Z_n\} \) in Lemma A.4 by
\[
Z_k = \sqrt{p_n} \xi''(k - b_{n-1}), \quad b_{n-1} < k \leq b_n.
\]

Observe that (4.25) implies that the conditions in Lemma A.4 hold. So in a richer probability space, we can construct independent centered normal random vectors \( \{\eta_n, n \geq 1\} \) with \( \text{Cov}(\eta_n) = \text{Cov}(Z_n) \) such that
\[
\max_{1 \leq k \leq b_n} \left| \sum_{i=1}^{k} (Z_i - \eta_i) \right| = o_{a.s.}(H_p(2^n)). \quad (4.31)
\]

Without loss of generality, we can write \( \eta_k \) as \( \sqrt{p_n} Y_n(k - b_{n-1}) \) for \( b_{n-1} < k \leq b_n \), where \( \{Y_n(j), n \geq 1, 1 \leq j \leq k_n\} \) are i.i.d. centered normal random vectors with \( \text{Cov}(Y_j(j)) = I \).

The case \( 2 < p < 4 \). Since (4.25) may not hold for \( 2 < q < 4 \), we should use Lemma A.5 instead of Lemma A.4. Note that \( \text{Cov}(\xi''(j)) = I \) and \( \xi''(1), \ldots, \xi''(k_n) \) are i.i.d. centered random vectors. Also, \( p_n^{-1/2} \Sigma_{n=1}^{1/2} \rightarrow I \). So \( \xi''(1) \leq C p^{-1/2} \theta_n H_p(2^n) \). Then it can be shown that when \( n \geq N_0 \), for some \( N_0 > 0 \),
\[
\alpha_n E[|\xi''(1)|^3] \leq 1
\]
with \( \alpha_n = \theta_n^{-1/2} p_n^{-1/2} (H_p(2^n))^{-1} \log H_p(2^n) \). Applying Lemma A.5, on a richer probability space, we can construct independent normal random vectors \( Y_n(1), \ldots, Y_n(k_n) \) with \( \text{E}Y_n(j) = 0, \text{Cov}(Y_n(j)) = I, 1 \leq j \leq k_n \), such that we have for \( x \geq 0 \),
\[
P \left( \max_{1 \leq k \leq k_n} \left| \sum_{j=1}^{k} (\xi''(j) - Y_n(j)) \right| \geq x \right) \leq c_12^{(1-a)n} \left[ \exp(-c_12 \alpha_n x) + \exp \left( -c_12 \left( \frac{x}{\gamma} \right)^{1/2} \right) \right] \quad (4.32)
\]
with \( \gamma = E[|\xi''(1)|^3] \). From Lemma A.2, \( \gamma \leq C[(\theta_n p_n^{-1/2} H_p(2^n))^{3-p} \vee 1] \). Now take \( x = \theta_n^{1/4} H_p(2^n)/p_n^{1/2} \). It easily follows from (4.32) and (5) that
\[
\sum_{n=1}^{\infty} P \left( \max_{1 \leq k \leq k_n} \left| \sum_{j=1}^{k} (\xi''(j) - Y_n(j)) \right| \geq \theta_n^{1/4} H_p(2^n)/p_n^{1/2} \right) < \infty.
\]
Applying the Borel–Cantelli lemma, we see that almost surely
\[
\max_{1 \leq m \leq n, 1 \leq t \leq k_m} \left| \sum_{i=1}^{m-1} \sum_{j=1}^{k_i} \sqrt{p_i} (\xi_i''(j)(\omega) - Y_i(j)(\omega)) + \sum_{j=1}^{t} \sqrt{p_m} (\xi_m''(j)(\omega) - Y_m(j)(\omega)) \right|
\]
\[
\leq K(\omega) + \sum_{i=1}^{n} q_i^{1/2} H_\rho(2^i)
\]
\[
\leq K(\omega) + o(H_\rho(2^n)),
\]
where \( K(\omega) \) is a finite constant.

Set
\[
S_n''(t) = \sum_{i=1}^{m-1} \sum_{j=1}^{k_i} \sqrt{p_i} Y_i(j) + \sum_{j=1}^{t} \sqrt{p_m} Y_m(j), \quad 1 \leq t \leq k_n.
\]

Without loss of generality, we assume that there exists an \( R^d \) valued Brownian motion \( B(t) \) with covariance matrix \( I \) such that \( S_n''(t_n - 1) = B(\sigma_n^2) \). Moreover, it can be shown that \( |\sigma_n^2 - n| \leq C(n^{-a+b} + na) \). Then by (5) and the tail probabilities of normal distribution, we can get
\[
|B(\sigma_n^2) - B(n)| = o_a.s.(H_\rho(n)).
\]

The proof of (4.27) is complete by (4.31), (4.33) and (4.34).

For \( d = 1 \) and \( \ell = 0 \), we see that \( D_k = 0 \) a.s., \( k \geq 0 \). By (2.3), Lemma A.1 and the proof of Lemma 4.5, we can get for every \( \varepsilon > 0 \),
\[
\sum_{n=1}^{\infty} \exp \left( - \frac{\varepsilon H_\rho^2(2^n)}{k_n \sum_{n=1}^{\infty} \xi''(j)} \right) < \infty.
\]

So by Lemma 3 in [10] and Lemma 4.6 again,
\[
\max_{1 \leq m \leq n, 1 \leq t \leq k_m} \left| \sum_{i=1}^{m-1} \sum_{j=1}^{k_i} \xi_i''(j) + \sum_{j=1}^{t} \xi_m''(j) \right| = o_a.s.(H_\rho(n))
\]
and therefore \( |S_{1,n}| = o_a.s.(H_\rho(n)) \). The proof of the lemma is terminated. □

**Proof of Theorem 2.1.** Take \( H_\rho(n) = \chi_{p}(n) \). Let \( a = 2/p - \varrho \) and \( b = (4 - p)/p - 2 \varrho \), where \( \varrho > 0 \) is sufficiently small. We can see that \( 0 < b < a(a + 1)/2 \), and (1)–(5), (7) are satisfied. When \( 2 < p < 4 \), (6) is easily proved. For the other case of \( p = 2 \),
\[
T_2 \leq C \sum_{i=1}^{\infty} \frac{i^{2} \Theta_{[i]}^{2}}{2^{i} \log i} \leq C \sum_{n=1}^{\infty} \frac{(\log n)^2 \Theta_{[i]}^{2}}{n \log_2 n}
\]
\[
\leq C \sum_{n=1}^{\infty} \sum_{i=\lceil n^{1/b} \rceil + 1} \frac{(\log i)^2 \Theta_{[i]}^{2}}{i \log_2 i} \leq C \sum_{n=1}^{\infty} \frac{(\log n)^2 \Theta_{[i]}^{2}}{n \log_2 n}
\]
\[
< \infty.
\]
This completes the proof of Theorem 2.1. □

**Proof of Theorem 2.2.** Let \( H_\rho(n) = n^{r+\delta}, a = \tau, b = 2\tau - 1 \). Noting that \( 2/p < \tau < 1 \), we have \( b < a(a + 1)/2 \), (1)–(7) are easily proved and the proof is omitted. □
5. Proof of corollaries

Proof of Corollary 3.1. By condition (3.1) on \( h(\cdot) \) and the Hölder inequality, we see that \( \theta_{n,p} = O(|a_i|) \), and hence (2.6) holds. It remains to check Condition A. Let \( C < 1/(pr) \) and write \( \delta_j = [2^{\delta_j}] \) with \( \delta < C \). By (3.1),

\[
|h(Y_i)| \leq C \left| \sum_{m=0}^{\delta_j} a_m \varepsilon_{i-m} \right|^r + C \left| \sum_{m=\delta_j+1}^{\infty} a_m \varepsilon_{i-m} \right|^r + C.
\]

This together with the fact \( \mathbb{E}|\sum_{m=\delta_j+1}^{\infty} a_m \varepsilon_{j-m}|^r = O(1) \) implies that

\[
U_j'(\delta) \leq C \left( \sum_{i=1}^{\delta_j} \sum_{m=0}^{\delta_j} |a_m||\varepsilon_{i-m}| \right)^r + C \delta_j.
\] (5.1)

Let \( \bar{\varepsilon}_i = |\varepsilon_i| - \mathbb{E}|\varepsilon_i|, i \in \mathbb{Z} \), and write

\[
\sum_{j=1}^{n} \sum_{m=0}^{n} |a_m| \bar{\varepsilon}_{j-m} = \sum_{t=-n}^{n-1} a_{n-t} \bar{\varepsilon}_t,
\]

where \( a_{n,t} = \sum_{j=-t}^{n} |a_{j+t}| \) for \( -n \leq t \leq -1 \), and \( a_{n,t} = \sum_{j=1}^{n-t} |a_{j+t}| \) for \( 0 \leq t \leq n-1 \).

Without loss of generality, we assume \( \sum_{i=0}^{\infty} |a_i| \leq 1 \) so that \( a_{n,t} \leq 1 \) for \( -n \leq t \leq n-1 \). Since \( \delta_j/(\chi_p(2^j))^{1/r} = o(1) \), we have from (5.1) that for every \( \varepsilon > 0 \),

\[
P \left( U_j'(\delta) \geq \varepsilon \chi_p(2^j) \right) \leq P \left( \sum_{i=1}^{\delta_j} \sum_{m=0}^{\delta_j} |a_m||\varepsilon_{i-m}| \geq C(\chi_p(2^j))^{1/r} \right)
\]

\[
\leq P \left( \sum_{i=1}^{\delta_j} \sum_{m=0}^{\delta_j} |a_m||\varepsilon_{i-m}| \geq C(\chi_p(2^j))^{1/r} \right)
\]

\[
= P \left( \sum_{m=\delta_j}^{\delta_j-1} |a_{\delta_j,m}||\varepsilon_{-m}| \geq C(\chi_p(2^j))^{1/r} \right)
\]

\[
\leq O \left( (\delta_j(\chi_p(2^j))^{-2/r})^t \right) + 2 \sum_{m=\delta_j}^{\delta_j-1} \mathbb{P} \left( |a_{\delta_j,m}||\varepsilon_{-m}| \geq (12t)^{-1}C(\chi_p(2^j))^{1/r} \right)
\]

\[
\leq O \left( (\delta_j(\chi_p(2^j))^{-2/r})^t \right) + 4\delta_j \mathbb{P} \left( |\bar{\varepsilon}_0|^r \geq (12t)^{-r}C(\chi_p(2^j)) \right),
\]

where we have used Lemma A.3 (by taking \( x = C(\chi_p(2^j))^{1/r} \) and \( y = x/(12t) \)) and \( t \) is large enough. This together with Lemma 4.1 proves Condition A, and hence completes the proof of Corollary 3.1. \( \square \)

Proof of Corollary 3.2. It follows easily from the conditions of Corollary 3.2 that \( \Theta_{n,p} = O(\sum_{i=n-m}^{\infty} |a_i|) \). So we only need to verify Condition A. By the inequality \( 2|x y| \leq x^2 + y^2 \),

\[
2U_j'(\delta) \leq \sum_{i=1}^{\delta_j} (X_i)^2 + \sum_{i=1}^{\delta_j} (X_{i-m})^2 + C \delta_j.
\]
Then Condition A follows from Lemma 4.1 and the proof of Corollary 3.1. □

**Proof of Corollary 3.3.** We may write $Y_n = g(\ldots, \varepsilon_{n-1}, \varepsilon_n)$ for some measurable function $g$. Set $Y_n^* = g(\ldots, \varepsilon_1^{'}, \ldots, \varepsilon_{n-1}^{'}, \varepsilon_n^{'})$, $X_n^* = |Y_n^*|^r - E|Y_n^*|^r$. By the Hölder inequality,

$$
E|X_n - X_n^*|^p \leq rE|Y_n - Y_n^*|^p (|Y_n^*|^{r-1} + |Y_n^*|^{r-1})^p \\
\leq C_{p,r}E|Y_n - Y_n^*|^p |Y_n|^{p(r-1)} \\
\leq C_{p,r}r (E|Y_n - Y_n^*|^{rp})^{1/r} (E|Y_n|^{pr})^{(r-1)/r}.
$$

In view of (3.11), we have

$$
|h_n^2 - h_n^{'2}| = \left| \int_{A(h_n^{'2})}^{A(h_n^2)} [A^{-1}(x)]'dx \right| \\
\leq C|A(h_n^2) - A(h_n^{'2})|((A(h_n^{'2}))^{1/y0} + |A(h_n^2)|^{1/y0}).
$$

Hence, by virtue of conditions of Corollary 3.1 and the inequality $|\sqrt{x} - \sqrt{y}| \leq \sqrt{|x - y|}$ for $x, y \geq 0$, it is readily seen that

$$
E|Y_n - Y_n^*|^{rp} \leq C E|h_n - h_n^{'2}|^{rp} \leq C E|h_n^2 - h_n^{'2}|^{rp/2} \\
\leq C (E|A(h_n^2) - A(h_n^{'2})|^{rp/2}) / (|A(h_n^{'2})|^{(1/y0)rp/2}) \\
\leq C (E|A(h_n^2) - A(h_n^{'2})|^{rp(1+y\vee 0)/2})^{1/(1+y\vee 0)} \\
\times (E|A(h_n^{'2})|^{rp(1+y\vee 0)/2})^{1/(1+y\vee 0)} \\
\leq C \left( E \left| \sum_{i=n}^{\infty} d(\varepsilon_{n-i}) \prod_{j=i}^{j-1} c(\varepsilon_{n-j}) \right|^{rp(1+y\vee 0)/2} \right)^{1/(1+y\vee 0)} \\
= C \left( E \left| \prod_{j=1}^{n-1} c(\varepsilon_{n-j}) \right|^{rp(1+y\vee 0)/2} \right)^{1/(1+y\vee 0)} \\
\leq C \rho^n
$$

for some $0 < \rho < 1$. This proves (2.6).

Now we check Condition A. Let $C < 1/(p\alpha r)$. By (3.11), $A^{-1}(x) \leq Cx^{1+y\vee 0} + C$ for $x \geq \omega$, which yields that

$$
|Y_n| \leq |h_n\varepsilon_n| \leq C|\varepsilon_n|(A(h_n^2))^{(1+y\vee 0)/2} + C|\varepsilon_n|. \quad (5.2)
$$

Set $\alpha = (1 + y \vee 0)/2$ and $\beta = \alpha^{-1}$. Suppose now that $\alpha r \geq 1$. Recall that $\delta_j = [2^{\delta_j}]$ with $\delta < C$. Since $X_n = |Y_n|^r - E|Y_n|^r$, (5.2) implies

$$
U_j(\delta) \leq C \left( \sum_{i=1}^{\delta_j} |\varepsilon_i|^\beta A(h_i^2) \right)^{\alpha r} + C \sum_{i=1}^{\delta_j} |\varepsilon_i|^r + C\delta_j.
$$

Since $E|\varepsilon_0|^{pr} < \infty$ and $\delta_j/\chi_p(2^j) = o(1)$, we only need to show that for every $\varepsilon > 0$,
Lemma A.3 easily. We now give the proof of (5.4). Let \( \rho' \) satisfy \( 0 < (E|c(\varepsilon_0)|^{par})^{1/(par)} < \rho' < 1 \) and set

\[
\eta_{i,m} = |\varepsilon_i|^\beta d(\varepsilon_{i-m}) \prod_{\ell=1}^{m-1} c(\varepsilon_{i-\ell}), \quad 1 \leq i \leq \delta_j, m \geq 1.
\]

Note that, for every fixed \( m \), \( \{\eta_{i,m}, i \geq 1\} \) are \( m \)-dependent random variables. Simple calculations yield that

\[
\mathbb{W} \leq \sum_{j=1}^{\infty} 2^j 2^{(1-\delta)} \sum_{m=1}^{[\delta_j^{1/2}]} \sum_{i=1}^{\delta_j} \mathbb{P}\left( \left| \sum_{i=1}^{\delta_j} \eta_{i,m} \right| \geq \varepsilon \rho'm (\chi_p(2^j))^{1/(ar)} \right)
\]

\[
\quad + \sum_{j=1}^{\infty} 2^j 2^{(1-\delta)} \sum_{m=[\delta_j^{1/2}]+1}^{\infty} \mathbb{P}\left( \left| \sum_{i=1}^{\delta_j} \eta_{i,m} \right| \geq \varepsilon \rho'm (\chi_p(2^j))^{1/(ar)} \right)
\]

\[
=: \mathbb{W}_1 + \mathbb{W}_2,
\]

where \( \varepsilon \) in every line may be different. By the Markov inequality and \( E|c(\varepsilon_0)| < \rho' \), we have for some \( 0 < \rho'' < 1 \),

\[
\mathbb{W}_2 \leq C \sum_{j=1}^{\infty} 2^j 2^{(1-\delta)} \delta_j (\chi_p(2^j))^{1/(ar)} \sum_{m=[\delta_j^{1/2}]+1}^{\infty} \rho''m < \infty.
\]

In order to deal with \( \mathbb{W}_1 \), we split the interval \([1, \delta_j]\) into blocks \( J_1, K_1, \ldots, J_M, K_M \) with equal length \( m \). (The blocks \( J_M, K_M \) can be incomplete, but we still assume that they have the same length \( m \) for the sake of brevity.) Clearly, \( M \) is proportional to \( \delta_j/m =: \delta'_j \). Introduce

\[
T_{i,m}^{(1)} = \sum_{j \in J_i} \eta_{j,m}, \quad T_{i,m}^{(2)} = \sum_{j \in K_i} \eta_{j,m}.
\]

Since \( E|c(\varepsilon_0)| < \rho' \) and \( \delta_j/(\chi_p(2^j))^{1/(ar)} = o(1) \), we have, uniformly for \( m \geq 1 \),

\[
\mathbb{P}\left( \left| \sum_{i=1}^{\delta'_j} T_{i,m}^{(1)} \right| \geq \varepsilon \rho'm (\chi_p(2^j))^{1/(ar)} \right) = o(1).
\]

So by the Hoffmann-Jørgensen inequality (cf. [14]),
and Condition A is satisfied in the case of $ar$. This, together with a similar proof of $\left|\chi_p(2^j)\right|^{1/(ar)}$, yields that

$$\sum_{j=1}^{\infty} \sum_{m=1}^{\delta_j} \rho^m (\chi_p(2^j))^{1/(ar)} \leq \sum_{j=1}^{\infty} \sum_{m=1}^{\delta_j} \rho^m (\chi_p(2^j))^{1/(ar)}$$

Therefore we prove (5.4) and Condition A is satisfied in the case of $ar \geq 1$.

When $ar < 1$, we have

$$U_j(\delta) \leq C \sum_{i=1}^{\delta_j} |\epsilon_i|^{r} (A(h^2_\epsilon))^{ar} + \sum_{i=1}^{\delta_j} |\epsilon_i|^{r} + C \delta_j$$

$$\leq C \sum_{i=1}^{\delta_j} |\epsilon_i|^{r} \sum_{m=1}^{\infty} |d(\epsilon_{i-1})|^{ar} \prod_{j=1}^{m-1} |c(\epsilon_{i-j})|^{ar} + C \sum_{i=1}^{\delta_j} |\epsilon_i|^{r} + C \delta_j.$$
Proof of Corollary 3.4. Let $\varepsilon_n = (A_n, B_n)$. We may write $X_n = g(\ldots, \varepsilon_{n-1}, \varepsilon_n)$ for some function $g$. Set

$$G_n = \sum_{k=1}^{\infty} A_n A_{n-1} \cdots A_{n-k+1} B_{n-k} =: \sum_{k=1}^{\infty} G_{n,k}.$$ 

Using arguments similar to those in the proof of $\mathbb{W} < \infty$ (comparing $G_{n,k}$ with $\eta_{i,m}$ there), we see that Condition A holds and

$$\sum_{j=1}^{\infty} 2^j \mathbb{P}\left(|G_n| \geq 2^j/p\right) < \infty.$$ 

This ensures $\mathbb{E}[|G_n|^p] < \infty$. The proof of (2.6) can be given as follows:

$$\mathbb{E}[X_n - X^*_n]^p \leq C_p \mathbb{E}\left|\prod_{j=2}^{n} A_j\right|^p |G_1|^p \leq C \rho^n$$

for some $0 < \rho < 1$. □

Proof of Corollary 3.5. The proof is similar to those of Corollaries 3.3 and 3.4, and hence is omitted. □

Proof of Corollary 3.6. We only need to check Condition A. Since $|f(x) - f(y)| \leq \rho|x - y|$, we have

$$|X_n| \leq \rho |X_{n-1}| + |\varepsilon_n| + |f(0)| \leq \cdots \leq \sum_{i=0}^{\infty} \rho^i |\varepsilon_{n-i}| + (1 - \rho)^{-1} |f(0)| \text{ a.s.}$$

The rest of the proof is similar to that of Corollary 3.1. □

Proof of Corollary 3.7. We first show that under condition (3.18),

$$\sum_{i=1}^{n} (\eta_i - \mathbb{E}[\eta_i]) - \mathbb{B}(\sigma^2 n) = o_{a.s.}(\phi_p(n)), \quad (5.5)$$

where $\mathbb{B}(t)$ is a standard Brownian motion. Take $H_p(n) = \phi_p(n)$, $a = 2/p - \varrho$ and $b = (4 - p)/p - 2\varrho$. Thus, (1)–(5) are satisfied. To prove (6), we should note that (3.18) implies that the physical dependence measure of $\eta_n$ satisfies $\Theta_{n,p} = O(r_1^n)$ for some $0 < r_1 < 1$. Hence (6) holds immediately. It remains to check (7). Recall that $\delta_j = [2^j]$ with $\delta < \mathcal{C}$. Set $|\eta_i'| = \mathbb{E}[|\eta_i||\varepsilon_i - [Q], \ldots, \varepsilon_i]$ for $1 \leq i \leq \delta_j$, where $Q$ is large enough. By Lemma A.1,

$$\sum_{j=1}^{\infty} 2^{j(1-\delta)} \mathbb{P}\left(\sum_{i=1}^{\delta_j} |\eta_i'| \geq 2\varepsilon \phi_p(2^j)\right) \leq (2\varepsilon^{-1})^p \sum_{j=1}^{\infty} 2^{j(1-\delta)} \frac{\mathbb{E}\left(\sum_{i=1}^{\delta_j} |\eta_i| - |\eta_i'|\right)^p}{\phi_p^p(2^j)}$$

$$= O(1) + \sum_{j=1}^{\infty} 2^{j(1-\delta)} \mathbb{P}\left(\sum_{i=1}^{\delta_j} |\eta_i'| \geq 2^{-1} \varepsilon \phi_p(2^j)\right). \quad (5.6)$$
Hence, (5.5) will follow if we prove that the series in (5.6) is convergent. Let

\[ u_k = \sum_{i=(k-1)[Qj]+1}^{(k(Qj))\wedge \delta_j} (|\eta_i’| - E|\eta_i’|), \quad 1 \leq k \leq \lfloor Qj/\lfloor Qj \rfloor \rfloor + 1. \]

We can see that \( u_1, u_2, \ldots \) are 1-dependent random variables. By Lemma A.3, for \( t \) large enough,

\[
\sum_{j=1}^{\infty} 2^{j(1-\delta)} P \left( \sum_{i=1}^{\delta_j} |\eta_i’| \geq 2^{-1} \varepsilon \phi_p(2^j) \right) \]

\[ = O(1) \sum_{j=1}^{\infty} 2^{j(1-\delta)} P \left( \sum_{k=1}^{\lfloor (Qj)/(Qj) \rfloor + 1} u_k \geq \frac{4^{-1} \varepsilon \phi_p(2^j)}{E u_k^2} \right) \]

\[ \leq O(1) \sum_{j=1}^{\infty} 2^{j(1-\delta)} \left( \sum_{k=1}^{\lfloor (Qj)/(Qj) \rfloor + 1} \frac{E u_k^2}{\phi_p^2(2^j)} \right)^t \]

\[ + 2 \sum_{j=1}^{\infty} 2^{j(1-\delta)} \sum_{k=1}^{\lfloor (Qj)/(Qj) \rfloor + 1} P \left( |u_k| \geq (96t)^{-1} \varepsilon \phi_p(2^j) \right) \]

\[ = O(1) + 2 \sum_{j=1}^{\infty} 2^{j(1-\delta)} \sum_{k=1}^{\lfloor (Qj)/(Qj) \rfloor + 1} P \left( |u_k| \geq (96t)^{-1} \varepsilon \phi_p(2^j) \right), \quad (5.7) \]

where the last equation follows from the fact \( \sum_{k=1}^{\lfloor (Qj)/(Qj) \rfloor + 1} E u_k^2 = O(\delta_j) \). We now estimate \( P \left( |u_k| \geq (96t)^{-1} \varepsilon \phi_p(2^j) \right) \). For the sake of convenience, we assume \( k = 1 \). The proof for the case \( k \geq 2 \) is similar. Let \( |\eta_i’| = E[|\eta_i’| | \varepsilon_i - \{Q \log j \}, \ldots, \varepsilon_i], 1 \leq i \leq \lfloor Qj \rfloor \), and set

\[ v_k = \sum_{i=(k-1)[Q \log j]+1}^{(k[Q \log j]) \wedge \lfloor Qj \rfloor} (|\eta_i’| - E|\eta_i’|), \quad 1 \leq k \leq \lfloor Qj/\lfloor Q \log j \rfloor \rfloor + 1 =: J. \]

By Lemmas A.1 and A.3, we have for every \( \varepsilon > 0 \) and \( t \) large enough,

\[
P \left( |u_1| \geq \varepsilon \phi_p(2^j) \right) = O \left( (Qj)^{p/2} \phi_p^{-p}(2^j) r_1^{p(Q \log j)} \right) \]

\[ + O(1) P \left( \left| \sum_{k=1}^{J} v_k \right| \geq 2^{-1} \varepsilon \phi_p(2^j) \right) \]

\[ \leq O \left( (Qj)^{p/2} \phi_p^{-p}(2^j) r_1^{Q \log j} \right) + O \left( (Qj \phi_p^{-2}(2^j)^{it}) \right) \]

\[ + 2 \sum_{k=1}^{J} P \left( |v_k| \geq (48t)^{-1} \varepsilon \phi_p(2^j) \right). \quad (5.8) \]

Using Lemma A.1 again and the fact \( \sum_{i=(k-1)[Q \log j]+1}^{(k[Q \log j]) \wedge \lfloor Qj \rfloor} E|\eta_i| = o(\phi_p(2^j)) \), we have
\[ P\left(|v_k| \geq (48t)^{-1} \varepsilon \phi_p(2^j)\right) \leq O\left((\log j)^{p/2} \phi_p^{-p} (2^j) \rho \right) \]

\[ + P\left(\sum_{i=1}^{[\log j]} |\eta_i| \geq (96t)^{-1} \varepsilon \phi_p(2^j)\right). \]  

Combining (5.7)–(5.9) and elementary manipulations,

\[
\sum_{j=1}^{\infty} 2^{j(1-\delta)} P \left( \sum_{i=1}^{\delta_j} |\eta_i|' \geq 2^{-1} \varepsilon \phi_p(2^j) \right)
\]

\[ = O(1) + O(1) \sum_{j=1}^{\infty} 2^{j} (\log j)^{-1} P \left( \sum_{i=1}^{[\log j]} |\eta_i| \geq C_{i,\varepsilon} \phi_p(2^j) \right)
\]

\[ = O(1) + O(1) \sum_{j=1}^{\infty} 2^{j} P \left( |\eta_1| \geq C_{Q, i, \varepsilon} \phi_p(2^j) / \log j \right)
\]

\[ = O(1) + O(1) E|\eta_1|^p. \]

This together with (5.6) implies Condition A. Hence (5.5) holds.

Set \( A_n = \sum_{i=n}^{\infty} a_i \). Then we have

\[ S_n = A_0 \sum_{i=1}^{n} \eta_i - \sum_{i=1}^{\infty} (A_i - A_{i-n} 1_{i>n}) \eta_{n+1-i} =: A_0 \sum_{i=1}^{n} \eta_i - \tilde{R}_n. \]

This decomposition was obtained by Wu [30] for when \( \{\eta_i\} \) are i.i.d. random variables. In fact, one can easily check that it holds for any \( \{\eta_i\} \) if \( E|\eta_i| = E|\eta_0| < \infty \). Observe that \( \max_{1 \leq k \leq n} |\tilde{R}_k| = \max_{1 \leq k \leq n} |\sum_{i=1}^{k} (A_0 \eta_i - X_i)| \) and \( \{A_0 \eta_i - X_i\}_{i \in Z} \) is a stationary process. By Proposition 1 in [30], we only need to show that

\[
\sum_{j=0}^{\infty} \left( 2^{-j} E|\tilde{R}_{2j}|^p \right)^{1/(p+1)} < \infty.
\]  

(5.10)

Let

\[ D_{k,n} = \sum_{i=k[M \log n]+1}^{(k+1)[M \log n]} (A_i - A_{i-n} 1_{i>n}) \eta_{n+1-i} \]

and

\[ D'_{k,n} = \sum_{i=k[M \log n]+1}^{(k+1)[M \log n]} (A_i - A_{i-n} 1_{i>n}) \eta'_{n+1-i}, \]

where \( M \) is large enough and \( \eta'_i = E[\eta_i | \varepsilon_i - [M \log n], \ldots, \varepsilon_i] \) for \( i \in Z \). We also let \( \tilde{R}_n' = \sum_{k=0}^{\infty} D'_{k,n} \). Then

\[
(E|\tilde{R}_n - \tilde{R}_n'|^p)^{1/p} \leq \sum_{i=1}^{\infty} |A_i - A_{i-n} 1_{i>n}|(E|\eta_{n+1-i} - \eta'_{n+1-i}|^p)^{1/p}
\]

\[ = O(1) \rho^{M \log n} \sum_{i=1}^{\infty} |A_i - A_{i-n} 1_{i>n}| = O \left( n \rho^{M \log n} \right) = O(1). \]
Moreover, \( \tilde{R}_n = \sum_{k=0}^{\infty} D'_{2k,n} + \sum_{k=0}^{\infty} D'_{2k+1,n} \). By noting that \( \{D'_{2k,n}\}_{k \geq 0} \) are independent random variables, we have

\[
E \left| \sum_{k=0}^{\infty} D'_{2k,n} \right|^p \leq C_p \left( \sum_{k=0}^{\infty} E(\sum_{k=0}^{\infty} D'_{2k,n})^2 \right)^{p/2} + C_p \sum_{k=0}^{\infty} E|D'_{2k,n}|^p
\]

\[
\leq C_p M (\log n)^{p/2} \left( \sum_{i=1}^{\infty} (A_i - A_{i-n} 1_{i>n})^2 \right)^{p/2}
\]

\[
+ C_p (\log n)^{-p-1} \sum_{i=1}^{\infty} |A_i - A_{i-n} 1_{i>n}|^p
\]

\[
\leq C_p M (\log n)^{p/2} \left( \sum_{i=1}^{\infty} A_i^2 \right)^{p/2} + C_p M (\log n)^{-p-1} \sum_{i=1}^{n} |A_i|^p
\]

\[
= O(n(\log n)^{-ap+p/2}).
\]

Similarly, \( E \left| \sum_{k=0}^{\infty} D'_{2k+1,n} \right|^p = O(n(\log n)^{-ap+p/2}) \). So \( E|\tilde{R}_n|^p = O(n(\log n)^{-ap+p/2}) \) and (5.10) holds. \( \square \)

Acknowledgements

This research was supported by the National Natural Science Foundation of China (10571159, 10671176) and the Specialized Research Fund for the Doctor Program of Higher Education (20060335032). The authors would like to thank the referee for many valuable comments.

Appendix

Let \( S'_n = \sum_{i=1}^{n} X'_i \), where \( X'_i = E(X_i | F_m(i)) \), \( F_m(i) = \sigma(\varepsilon_{i-m}, \ldots, \varepsilon_i) \), \( m \geq 0 \). Define \( R_n = S_n - S'_n \), \( R_n^* = \max_{1 \leq i \leq n} |R_i| \) and the projection \( P_m X_i = E(X_i | F_m(i)) - E(X_i | F_{m-1}(i)) \).

The following lemma gives the \( m \)-dependent approximation for the sum \( S_n \). It is of independent interest and may have wider applicability.

Lemma A.1. (i) Suppose that \( X_1 \in L^q \) for some \( q > 1 \). Let \( q' = \min(2, q) \). Then we have

\[
\| R_n \|_{q'}^q \leq C_{q,d,n} q' \Theta_{m,q}',
\]

where \( C_{q,d,n} \) is a constant only depending on \( q \) and \( d \).

(ii) If \( q > 2 \), then

\[
\| R_n^* \|_{q'}^2 \leq C_{q,d,n} 2^3 \theta_{m,q}^2.
\]

(iii) If \( 1 < q \leq 2 \), then

\[
\| R_n^* \|_{q'}^{q'} \leq C_{q,d,n} (\log n)^q \theta_{m,q}^2.
\]

Proof. By Proposition 1 in [30], (ii) and (iii) can be obtained from (i). (Proposition 1 in [30] is stated there for one-dimensional case \( d = 1 \) but the proof is true for all \( d \).) So we only need to
prove (i) here. Since \( X_i = \lim_{j \to \infty} E(X_i|\mathcal{F}_{i+j}(i)) \), we see that

\[
R_n = \sum_{i=1}^{n} \sum_{j=m-i+1}^{\infty} \mathcal{P}_{i+j} X_i = \sum_{j=m-n+1}^{\infty} \sum_{i=(m-j+1)\land 1}^{n} \mathcal{P}_{i+j} X_i =: \sum_{j=m-n+1}^{\infty} R_{n,j}.
\]

For every fixed \( n \) and \( m \), \( \{R_{n,j}, j \geq m - n + 1\} \) is a sequence of martingale differences with respect to \( \sigma(\varepsilon_{-j}, \varepsilon_{-j+1}, \ldots) \). If \( q \geq 2 \), we have by Burkholder’s inequality,

\[
\|R_n\|_q^2 \leq C_{q,d} \left( E \left[ \sum_{j=m-n+1}^{\infty} |R_{nj}|^2 \right] \right)^{2/q} \leq C_{q,d} \sum_{j=m-n+1}^{\infty} \|R_{nj}\|^2. \tag{A.4}
\]

Note that

\[
\|R_{nj}\|_q \leq \sum_{i=(m-j+1)\land 1}^{n} \|\mathcal{P}_{i+j} X_i\|_q \leq \sum_{i=(m-j+1)\land 1}^{n} \theta_{i+j,q}. \tag{A.5}
\]

By (A.4) and (A.5), it is readily seen that

\[
\|R_n\|_q^2 \leq C_{q,d} \sum_{j=m-n+1}^{\infty} \left( \sum_{i=(m-j+1)\land 1}^{n} \theta_{i+j,q} \right)^2 = C_{q,d} \sum_{j=m-n+1}^{m} \left( \sum_{i=m-j+1}^{n} \theta_{i+j,q} \right)^2 + C_{q,d} \sum_{j=m+1}^{\infty} \left( \sum_{i=1}^{n} \theta_{i+j,q} \right)^2 \leq C_{q,d} n \theta_{m,q}^2 + C_{q,d} \sum_{j=m+1}^{\infty} \left( \sum_{i=1}^{n} \theta_{i+j,q} \right) \theta_{m,q} \leq C_{q,d} \left( n \theta_{m,q}^2 + \sum_{i=1}^{n} \theta_{i+m,q} \theta_{m,q} \right),
\]

and hence (A.1) is proved. If \( 1 < q < 2 \), then

\[
E|R_n|^q \leq C_{q,d} E \left[ \sum_{j=m-n+1}^{\infty} |R_{nj}|^2 \right]^{q/2} \leq C_{q,d} \sum_{j=m-n+1}^{\infty} E|R_{nj}|^q. \]

By (A.5),

\[
E|R_n|^q \leq C_{q,d} \sum_{j=m-n+1}^{\infty} \left( \sum_{i=(m-j+1)\land 1}^{n} \theta_{i+j,q} \right)^q = C_{q,d} \sum_{j=m-n+1}^{m} \left( \sum_{i=m-j+1}^{n} \theta_{i+j,q} \right)^q + C_{q,d} \sum_{j=m+1}^{\infty} \left( \sum_{i=1}^{n} \theta_{i+j,q} \right)^q \leq C_{q,d} n \theta_{m,q}^q + C_{q,d} \sum_{j=m+1}^{\infty} \left( \sum_{i=1}^{n} \theta_{i+j,q} \right) \theta_{m,q}^{q-1} \leq C_{q,d} n \theta_{m,q}^q.
\]

So (A.1) holds. The proof is complete. \( \square \)
The following lemma follows from Lemma A.1 immediately.

Lemma A.2. Suppose that $X_1 \in \mathcal{L}^q$ for some $q \geq 2$, $\mathbb{E}X_1 = 0$ and $\Theta_{0,q} < \infty$. Set $S_n^q = \max_{1 \leq k \leq n} |S_k|$ and $S_n^{q'} = \max_{1 \leq k \leq n} |\Gamma_k|$. (i) For $q > 2$, we have $\mathbb{E}(S_n^q)^q \leq C n^{q/2}$ and $\mathbb{E}(S_n^{q'})^{q'} \leq Cn^{q'/2}$. (ii) For $q \geq 2$, $\mathbb{E}|S_n|^q \leq C n^{q/2}$ and $\mathbb{E}|S_n^{q'}|^q \leq C n^{q'/2}$, where $C$ is a finite constant which depends on $d, q, \Theta_{0,q}$ and $\|X_1\|_q$, but does not depend on $m$.

The following Fuk–Nagaev inequality can be found in [27]. (Shao proved it for the case $d = 1$. Extending to all $d$ is immediate.)

Lemma A.3. Let $Y_1, \ldots, Y_n$ be $\mathbb{R}^d$ valued independent centered random vectors. Then for $x > 0$ and $y > 0$,

$$
\mathbb{P}\left( \max_{1 \leq i \leq n} \left| \sum_{j=1}^i Y_j \right| \geq d x \right) \leq 2d \mathbb{P}( \max_{1 \leq k \leq n} |Y_k| > y ) + 4d \exp\left(-\frac{x^2}{8B_n}\right)
$$

$$
+ 4d \left( \frac{B_n}{4xy} \right)^{x/(12y)}
$$

where $B_n = \sum_{j=1}^n \mathbb{E}|Y_j|^2$.

The following lemma comes from [10].

Lemma A.4. Let $\{Z_k\}$ be a sequence of independent $\mathbb{R}^d$ valued random vectors with zero means and $\text{Cov}(Z_k) = \sigma_k^2 \Gamma$. Assume that the following holds true for some $q \in (2, 4)$:

$$
\sum_{n=1}^{\infty} \frac{\mathbb{E}|Z_n|^q}{a_n^q} < \infty, \quad 0 < a_k \uparrow \infty.
$$

Then on a richer probability space we can construct a sequence of independent normal random vectors $\{\eta_k\}$ with $\mathbb{E}\eta_k = 0$ and $\text{Cov}(\eta_k) = \sigma_k^2 \Gamma, k \in \mathbb{N}$, such that the partial sums $S_n = \sum_{k=1}^n Z_k, T_n = \sum_{k=1}^n \eta_k$ fulfill

$$
|S_n - T_n| = o(a_n) \quad a.s.
$$

The last lemma comes from [11], Theorem 12.

Lemma A.5. Let $X, X_1, \ldots, X_n$ be i.i.d. mean zero $\mathbb{R}^d$ valued random vectors with $\text{Cov}(X) = I$. Suppose that there exists an $\alpha \in (0, 1/2)$ such that

$$
\alpha \mathbb{E}|X|^3 \exp(\alpha |X|) \leq 1. \quad (A.6)
$$

Then on a richer probability space, we can construct independent normal random vectors $Y_1, \ldots, Y_n$ with $\mathbb{E}Y_k = 0, \text{Cov}(Y_k) = I, 1 \leq k \leq n$, such that we have for $x \geq 0$,

$$
\mathbb{P}\left( \max_{1 \leq k \leq n} \left| \sum_{i=1}^k (X_i - Y_i) \right| \geq x \right)
$$

$$
\leq c_{11} n \left[ \exp(-c_{12} \alpha x) + \exp\left(-c_{12} \left( \frac{x}{\gamma} \right)^{1/2}\right) \right],
$$

where $\gamma = \mathbb{E}|X|^3$, and $c_{11}, c_{12}$ are positive constants depending only on $d$.
References