ON NON-UNIFORM BERRY–ESSEEN BOUNDS FOR TIME SERIES

by

MORITZ JIRAK (BERLIN)

Abstract. Given a stationary sequence \{X_k\}_{k \in \mathbb{Z}}, non-uniform bounds for the normal approximation in the Kolmogorov metric are established. The underlying weak dependence assumption includes many popular linear and nonlinear time series from the literature, such as ARMA or GARCH models. Depending on the number of moments \(p\), typical bounds in this context are of the size \(O(m^{-1}n^{-p/2+1})\), where we often find that \(m = m_n = \log n\). In our setup, we can essentially improve upon this rate by the factor \(m^{-p/2}\), yielding a bound of \(O(m^{p/2-1}n^{-p/2+1})\). Among other things, this allows us to recover a result from the literature, which is due to Ibragimov.

2000 AMS Mathematics Subject Classification: Primary: 60F05; Secondary: 91B84.

Key words and phrases: Berry–Esseen, weak dependence.

1. INTRODUCTION

Let \(\{X_k\}_{k \in \mathbb{Z}}\) be a zero mean process such that \(\mathbb{E}[X_k^2] < \infty\). Further, we introduce the partial sum \(S_n = \sum_{k=1}^n X_k\) and its variance \(B_n^2 = \text{Var}[S_n]\). A very important issue in probability theory and statistics is whether the central limit theorem holds or not, i.e., whether we have

\[
\lim_{n \to \infty} \left| P(S_n \leq xB_n) - \Phi(x) \right| = 0,
\]

where \(\Phi(x)\) denotes the standard normal distribution function. Going one step further, we can also ask ourselves about the possible rate of convergence in (1.1), more precisely, if for some explicit, increasing sequence \(r_n \to \infty\) we have

\[
\lim_{n \to \infty} d(P_{S_n/B_n}, P_Z) r_n < \infty,
\]

where \(d(\cdot, \cdot)\) is some probability metric, \(Z\) follows a standard normal distribution, and \(P_X\) denotes the probability measure induced by the random variable \(X\). This question has been addressed under numerous different setups with respect to the metric and underlying structure of the sequence \(\{X_k\}_{k \in \mathbb{Z}}\). Perhaps one of the most
popular metrics is the Kolmogorov (uniform) metric given as

\begin{equation}
\sup_{x \in \mathbb{R}} \Delta_n(x) := \sup_{x \in \mathbb{R}} |P(S_n \leq xB_n) - \Phi(x)| \quad \text{as } n \to \infty.
\end{equation}

(1.3)

In the case of a more difficult non-uniform analogue, we consider the error

\begin{equation}
\Delta_n(x) := |P(S_n \leq xB_n) - \Phi(x)|,
\end{equation}

(1.4)

and we are interested in bounds of the form \( \lambda_n(1 + |x|)^{-p} \), \( p \in (2, 3) \), with \( \lambda_n = o(1) \) as \( n \to \infty \). Note that we always have the relation

\[ \sup_{x \in \mathbb{R}} \Delta_n(x) = O(\lambda_n), \]

hence a bound for the non-uniform metric always gives a bound for the uniform metric. Particularly, the latter has been studied extensively in the literature under many different dependence assumptions. A common way to measure dependence is in terms of various mixing conditions. In the case of the uniform metric, Rio \cite{30} showed that it is possible to obtain a rate of \( O(\sqrt{n}) \) in (1.3) under certain mixing assumptions, given a bounded support of the underlying sequence \( \{X_k\}_{k \in \mathbb{Z}} \) (see also Bolthausen \cite{7} for martingale difference sequences). Given more general assumptions such as \( \alpha \)-mixing (see \cite{11} for definitions), Tikhomirov \cite{31} obtained a rate of \( O((\log n)^2n^{-1/2}) \), provided that the underlying third moments exist (see also Bentkus et al. \cite{4}). For related results in general Hilbert spaces, we refer to the references therein. We remark that in this case, even considering linear processes results in a nonlinear nature; see, for instance, El Machkouri \cite{15}. In contrast to the previously mentioned results, Tikhomirov \cite{31} also obtained the rate \( O((\log n)^2n^{-1/2}) \) in the more difficult case of the non-uniform metric. Similarly, Hörmann \cite{19} obtained rates of the form \( O((\log n)^2n^{-1/2}) \) under the notion of \( m_n \)-approximability in \( \| \cdot \|_p \) (see Definition 1.1 for details) both for the uniform and non-uniform metric. This weak dependence concept covers a wide range of very popular time series in the literature, such as ARCH, GARCH and many other nonlinear processes (cf. \cite{22}, \cite{29}, \cite{32}). In particular, it contains examples of time series that are known to be not \( \alpha \)-mixing (cf. \cite{1}). Interestingly, this concept is also applicable in more number theoretic settings, see Ibragimov \cite{29}, Hörmann \cite{19} and Example 2.3 below. To be more specific, let us introduce the notion of \( m_n \)-approximability in \( \| \cdot \|_p \):

**Definition 1.1.** Let \( p > 0 \), and put \( \| \cdot \|_p = \mathbb{E}[| \cdot |^p]^{1/p} \). Consider the sequence \( \{m_n\} \) of non-decreasing natural numbers. The process \( \{X_k\}_{k \in \mathbb{Z}} \) is called \( m_n \)-approximable in \( \| \cdot \|_p \) of size \( a_n \) if there exist \( m_n \)-dependent sequences \( \{X_{km}\}_{k \in \mathbb{Z}}, m = 1, 2, \ldots \) such that

\[ \sum_{k=1}^{n} \|X_k - X_{km}\|_p = o(a_n). \]

We will abbreviate this with \( \{X_k\} \in W(L^2, \{m_n\}, \{a_n\}) \).
A common method in the literature is to approximate $S_n$ with $S_{nm} = X_{1m} + \ldots + X_{nm}$, and then apply various blocking and truncation arguments to infer the result to relegate the problem to the i.i.d. case. This method has been used by Tikhomirov [31], Bentkus et al. [4] and many other. Hörmann [19] directly refers to the literature for this case, and concentrates on the error induced by the $m$-approximation. In contrast, our focus lies on controlling the error $\Delta_n(x)$ for $m$-dependent sequences. We will assume that the sequence $\{X_k\}_{k \in \mathbb{Z}}$ satisfies a weak dependence assumption that is related to the concept of $m_n$-approximability. By exploiting the weak dependence within our $m$-dependent approximating sequences $X_{km}$ (which will be denoted by $Y_{k}^{(m)}$), it is possible to establish a rate of $O((mn)^{1/2})$ in (1.3) given $p = 3$ moments, which is an improvement by the factor $m^{-3} = 2n^{-1}$. Note that in the case of many time series such as GARCH and ARMA models, this results in an ultimate bound of $O((\log n)^{1/2})$ in (1.3), see Corollary 2.3 and the following discussion.

The remainder of this paper is organised as follows. In Section 2, the main results together with some examples are presented. The proofs are given in Section 3.

2. DEPENDENCE CONDITION AND MAIN RESULTS

Let $\{\epsilon_k\}_{k \in \mathbb{Z}} \in \mathbb{R}^\mathbb{Z}$ be a sequence of zero mean i.i.d. random variables, and introduce the filtration $\mathcal{F}_k = \sigma(\epsilon_j, j \leq k)$. In the sequel, we will consider the sequence of zero mean random variables $X_k = g(\epsilon_k, \epsilon_{k-1}, \ldots)$, $k \geq 1$, where $g$ is a measurable function such that $X_k$ are proper random variables. Note that this implies that $\{X_k\}_{k \geq 1}$ is stationary and ergodic. For convenience, we will also write $X_k = g(\xi_k)$ with $\xi_k = (\epsilon_k, \epsilon_{k-1}, \ldots)$. The class of processes that fits into this framework is large, and contains a variety of linear and nonlinear processes including ARCH, GARCH and related processes; see, for instance, [22], [29], [32].

A very nice feature of the representation given above is that it allows us to give simple, yet very efficient and general dependence conditions. Following [33], let $\{\epsilon'_k\}_{k \in \mathbb{Z}}$ be an independent copy of $\{\epsilon_k\}_{k \in \mathbb{Z}}$ on the same probability space, and define the ‘filters’ $\xi_k^{(m',*)}$, $\xi_k^{(m,s)}$ as

$$\xi_k^{(m')} = (\epsilon_k, \epsilon_{k-1}, \ldots, \epsilon_k^{m'}, \epsilon_{k-m-1}, \ldots)$$

and

$$\xi_k^{(m,s)} = (\epsilon_k, \epsilon_{k-1}, \ldots, \epsilon_k^{m'}, \epsilon_{k-m-1}, \ldots).$$

We put

$$\xi'_k = \xi_k^{(k',*)} = (\epsilon_k, \epsilon_{k-1}, \ldots, \epsilon'_0, \epsilon_{-1}, \ldots)$$

and

$$\xi'^*_k = \xi_k^{(k,s)} = (\epsilon_k, \epsilon_{k-1}, \ldots, \epsilon_0, \epsilon'_{-1}, \ldots).$$
By analogy, we put \( X_k^{(m,')} = g(\xi_k^{(m,')}) \) and \( X_k^{(m, s)} = g(\zeta_k^{(m, s)}) \); in particular, we have \( X_k^+ = X_{k}^{(k, s)} \) and \( X_k^+ = X_{k}^{(k, s)} \).

As a dependence measure, one may now consider the quantities \( \|X_k - X_k'\|_p \) or \( \|X_k - X_k^+\|_p, p \geq 1 \). Dependence conditions of this type are often quite general and easy to verify in many cases; see, for instance, [13], [43] and Examples [44] and [45] below. The main results will be formulated in terms of the following assumptions.

**Assumption 2.1.** The sequence \( \{X_k\}_{k \in \mathbb{Z}} \) can be represented as \( \{g(\xi_k)\}_{k \in \mathbb{Z}} \) for some measurable function \( g(\cdot) \) and satisfies the following conditions for \( p \in (2, 3] \):

1. \( \mathbb{E}[X_k] = 0 \),
2. \( \sum_{k=1}^{\infty} \|X_k - X_k'\|_p < \infty \),
3. \( \sigma^2 = \sum_{k=-\infty}^{\infty} \mathbb{E}[X_kX_0] > 0 \).

We are now ready to formulate the main results.

**Theorem 2.1.** Let the Assumption 2.1 be satisfied and assume, in addition, that \( m_n \) is chosen such that \( n \sum_{k=m_n}^{\infty} \|X_k - X_k'\|_p^2 = o(1) \). Then there exists a finite, absolute constant \( C_0 > 0 \) such that

\[
\Delta_n(x) \leq C_0 (1 + |x|)^{-p} \left( \sigma^{-p} \left( \frac{m_n}{n} \right)^{p/2 - 1} + \left( - \log(c_{p, n, m_n}) \right)^{(p+1)/2} e_{p, n, m_n}^{(p+1)/2} \right),
\]

where \( e_{p, n, m_n}^{2} = O(\sigma^{-2n} \sum_{k=m_n}^{\infty} \|X_k - X_k'\|_p^2) \).

As can be seen from the above bound, this can be improved by balancing (optimising) the two error terms. By imposing additional conditions on the rate of decay of \( \|X_k - X_k'\|_p \), we get more compact expressions. We will first consider the case where we have an algebraic rate of decay, i.e., \( \|X_k - X_k'\|_p = O(k^{-\alpha}) \).

**Corollary 2.1.** Let the assumptions of Theorem 2.1 be satisfied. Assume, in addition, that \( \|X_k - X_k'\|_p = O(k^{-\alpha}), \alpha > 1 \). Then there exists a finite, absolute constant \( C_0 > 0 \) such that

\[
\sup_{x \in \mathbb{R}} \Delta_n(x) \leq C_0 (1 + |x|)^{-p} \sigma^{-p} n^{r(\alpha, p)} \left( 1 + \frac{1}{2(\alpha - 1)} (\log n)^{(p+1)/2} \right),
\]

where

\[
r(\alpha, p) = \frac{p(p - 2)(1 - \alpha)}{p(2\alpha + p - 2) - 2} < 0.
\]

Note that \( \lim_{\alpha \to \infty} r(\alpha, p) = -p/2 + 1 \), which is the optimal bound and corresponds to the case where \( \{X_k\}_{k \in \mathbb{Z}} \) constitutes an i.i.d. sequence. By imposing a stronger (exponential) rate of decay, we get the following result.
COROLLARY 2.2. Let the assumptions of Theorem 2.1 be satisfied. Assume, in addition, that \( \|X_k - X'_k\|_p = O(\rho^{-k}), \) \( 0 < \rho < 1 \). Then there exists a finite, absolute constant \( C_0 > 0 \) such that

\[
\sup_{x \in \mathbb{R}} |\Delta_n(x)| \leq C_0 (1 + |x|)^{-p} \sigma^{-p} \left( \frac{\log n}{n} \right)^{p/2-1}.
\]

Note that if we have \( p = 3 \) moments, then we obtain a convergence rate of \( O\left( \sqrt{\log n} n^{-1/2} \right) \) under a very general dependence condition, which improves upon the results in [31] and [19] (in both cases \( O\left( (\log n)^2 n^{-1/2} \right) \) was obtained).

Let us also mention that, by imposing stronger rates of decay (e.g., \( \|X_k - X'_k\|_p = O(e^{-k^\gamma}), \gamma > 1 \)), faster rates can be obtained.

As already mentioned, an estimate for the non-uniform metric always implies a bound for the uniform metric. However, more careful calculations give the following slightly improved result in case of the uniform metric.

THEOREM 2.2. Let the Assumption 2.1 be satisfied and assume, in addition, that \( m_n \) is chosen such that \( \sum_{k=m_n}^{\infty} \|X_k - X'_k\|_p^2 = o(1) \). Then there exists a finite, absolute constant \( C_0 > 0 \) such that

\[
\sup_{x \in \mathbb{R}} |\Delta_n(x)| \leq C_0 \sigma^{-p} \left( \frac{m_n}{n} \right)^{p/2-1} + 2e^{p/(p+1)}\cdot\epsilon_{p,n,m_n},
\]

where \( \epsilon_{p,n,m_n} = O(\sigma^{-2} \sum_{k=m_n}^{\infty} \|X_k - X'_k\|_p^2) \).

As before in Corollaries 2.1 and 2.2, one can derive more explicit results by imposing conditions on the decay rate of \( \|X_k - X'_k\|_p \).

COROLLARY 2.3. Let the assumptions of Theorem 2.2 be satisfied. Assume, in addition, that \( \|X_k - X'_k\|_p = O(k^{-\alpha}), \alpha > 1 \). Then there exists a finite, absolute constant \( C_0 > 0 \) such that

\[
\sup_{x \in \mathbb{R}} |\Delta_n(x)| \leq C_0 (1 + |x|)^{-p} \sigma^{-p} n^{r(\alpha,p)},
\]

where

\[
r(\alpha,p) = \frac{p(p-2)(1-\alpha)}{p(2\alpha + p - 2) - 2} < 0.
\]

In case of exponential decay, the same result as in Corollary 2.2 is obtained.

As already mentioned, our setup includes many popular time series that are used for modelling in many different fields. To highlight this fact, we will briefly discuss some prominent examples from the literature.

EXAMPLE 2.1 (GARCH(p, q) sequences). Let \( \{X_k\}_{k \in \mathbb{Z}} \) be a GARCH(p, q) sequence given by the relations

\[
X_k = \epsilon_k L_k,
\]
where \( \{\epsilon_k\}_{k \in \mathbb{Z}} \) is a zero mean i.i.d. sequence and
\[
L_k^2 = \mu + \alpha_1 L_{k-1}^2 + \ldots + \alpha_p L_{k-p}^2 + \beta_1 X_{k-1}^2 + \ldots + \beta_q X_{k-q}^2
\]
with \( \mu, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q \in \mathbb{R} \). A very important quantity in this context is
\[
\gamma_C = \sum_{i=1}^{r} \| \alpha_i + \beta_i \epsilon_i^2 \|_2
\]
with \( r = \max\{p, q\} \).

Using this representation and the fact that \( |x - y|^p \leq |x^2 - y^2|^{p/2} \) for \( x, y \geq 0, p \geq 1 \), one can follow the proof of Theorem 4.2 in [2] to show that
\[
\| X_k - X'_k \|_p = \mathcal{O}(\rho^{-k}), \quad \text{where } 0 < \rho < 1.
\]
Hence an application of Corollary 2.2 yields a rate of \( \mathcal{O}((\log n)^{p/2-1} n^{-p/2+1}) \) for \( p \in (2, 3) \).

**Example 2.2 (Iterated random functions).** Let \( \{X_k\}_{k \in \mathbb{Z}} \) be defined via the recursion
\[
X_k = f(X_{k-1}; \epsilon_k).
\]
Such a construction is often referred to as iterated random functions, see [13] for a general overview. Let
\[
L \epsilon = \sup_{x \neq y} \frac{|f(x, \epsilon) - f(y, \epsilon)|}{|x - y|}
\]
be the Lipschitz coefficient. If \( \mathbb{E}[L \epsilon] < 1 \) and \( \| f(x_0, \epsilon) \|_p < \infty \) for some \( x_0 \), then it follows that
\[
\| X_k - X'_k \|_p = \mathcal{O}(\rho^{-k}), \quad \text{where } 0 < \rho < 1,
\]
see, e.g., [35]. In particular, \( X_k \) can be presented as \( X_k = M(\epsilon_k, \epsilon_{k-1}, \ldots) \) for some measurable function \( M \). As a particular example, consider the function \( f(x, \epsilon) = 1/(1 + x^2) + \epsilon \). In this case, since
\[
\mathbb{E}[L \epsilon] = \sup_{x \neq y} \frac{|x + y|}{(1 + x^2)(1 + y^2)} < 1,
\]
the above conditions are clearly met. We may thus apply Corollary 2.2 to obtain a rate of \( \mathcal{O}((\log n)^{p/2-1} n^{-p/2+1}) \) for \( p \in (2, 3) \).
EXAMPLE 2.3 (Sums of the form $\sum f(2^k \omega)$). For the exposition of this particular example, we will borrow from the related discussion in [19]. Let $f$ be a function defined on the unit interval $[0, 1]$, such that

$$\int_0^1 f(\omega) d\omega = 0 \quad \text{and} \quad \int_0^1 |f(\omega)|^p d\omega < \infty, \quad p \in (2, 3].$$

For $x \in \mathbb{R}^+$, let $\hat{f}(x) = f(x - |x|)$, i.e., $\hat{f}$ is the one-periodic extension to the positive real line. Consider now the partial sum $S_n = \sum_{k=1}^n f(2^k \omega) = \sum_{k=1}^n L_k$.

Note that in this case we may write $L_k$ as

$$L_k = f\left(\sum_{j=1}^{\infty} \zeta_k + j \cdot 2^{-j}\right),$$

where $\{\zeta_k\}_{k \in \mathbb{Z}}$ is a sequence of i.i.d. random variables, $\zeta_k$ taking values zero and one with probability $1/2$. This representation originates from a binary expansion, see [19] and the next references. Quantity $S_n$ has been studied by many authors, see, for instance, [20], [25]. For our setup, the result in [20] is of particular interest.

Introduce the modulus of continuity $\omega_f(\delta)$ as

$$\omega_f(\delta) = \sup_{0 < s, t < 1, |s - t| < \delta} |f(t) - f(s)|, \quad \text{where} \ 0 < \delta < 1.$$ 

Provided that $\omega_f(h) \leq \text{const} \ h^\beta$, $\beta > 0$, Ibragimov [20] showed that

$$\sup_{x \in \mathbb{R}} \Delta_n(x) \leq C_0 \sigma^{-p} \left(\frac{\log n}{n}\right)^{p/2 - 1}.$$ 

A priori, the sequence $\{L_k\}_{k \in \mathbb{Z}}$ does not directly fit into our framework. However, this can be achieved by a simple time flip. Define the function $T_n(i) = n - i + 1$ for $i \in \{n, n - 1, \ldots\}$, and let $\epsilon_k = \zeta_{T_n(k)}$. Then we may write

$$L_{T_n(k)} = X_k = f\left(\sum_{j=1}^{\infty} \epsilon_k - j \cdot 2^{-j}\right).$$

Note that we have to perform this time flip for every $n \in \mathbb{N}$, which however has no impact on the applicability of our results. Using the same arguments as in Proposition 4.6 in [19], we find that for $p \in (2, 3)$

$$\|X_k - X_k\|_p = O(\omega_f(2^{-k})) = O(2^{-\beta k}),$$

hence the conditions of Corollary 2.2 are satisfied. Applying Corollary 2.2, we thus obtain a non-uniform version of (2.6), which in particular yields Ibragimov’s result in [20]. Note that in the case of $p = 3$ Ladokhin and Moskvin [24] have established a similar (slightly weaker) result.
3. PROOFS

Let \( \{U_k\}_{k \in \mathbb{Z}} \) be a stationary process adapted to the filtration \( \mathcal{F}_k \). Then we define the projection operator \( P_k(U_i) \) as

\[
P_k(U_i) = E(U_i | \mathcal{F}_k) - E(U_i | \mathcal{F}_{k-1}), \quad k \geq 1, \quad i \in \mathbb{Z}.
\]

Many of the following results are (implicitly) based on martingale approximations for partial sums. Various different approximating martingale sequences have been proposed in the literature, see, for instance, [18], [21], [27] and the references therein. In our setting, the following representation via martingale differences (projections) is useful (see [16], [34]):

\[
X_k = \sum_{i=-\infty}^{k} P_i(X_k) \quad \text{for all} \quad k \in \mathbb{Z}.
\]

As already outlined in the introduction, another essential tool will be approximations with \( m \)-dependent random variables. To this end, we introduce the following notation. Let \( \{\epsilon_k\}_{k \in \mathbb{Z}} \in \mathbb{R}^\mathbb{Z} \) be a sequence of zero mean i.i.d. random variables. Recall that \( \mathcal{F}_k = \sigma(\epsilon_j, j \leq k) \), and define, in addition, the \( \sigma \)-algebra

\[
\mathcal{F}_{k-m} = \sigma(\epsilon_j, k-m \leq j \leq k)
\]

and the random variables

\[
Y^{(\leq m)} = E[X_k | \mathcal{F}_{k-m}], \quad Y^{(> m)} = X_k - Y^{(\leq m)}.
\]

Let \( \eta_{j,m} = m^{-1/2} \sum_{i=jm+1}^{(j+1)m} Y^{(\leq m)}_i \), where we point out that \( \{\eta_{j,m}\} \) is a two-dependent sequence. Let \( N = [n/m] + 1 \) and put \( S_N^{(\leq m)} = N^{-1/2} \sum_{j=1}^{N} \eta_{j,m} \). In addition, in order to avoid any notational difficulties, we also put \( \eta_{j,m} = 0 \) for \( j > N \).

**Remark 3.1.** We will frequently make use of the following property. It follows that (cf. Lemma 3.84 in [21])

\[
P_{i+h}(X_{k+h}) \overset{d}{=} P_i(X_k), \quad h \in \mathbb{Z}.
\]

This implies, in particular, \( \|P_{i+h}(X_{k+h})\|_p = \|P_i(X_k)\|_p, \quad p \geq 1 \).

Throughout the proofs, \( C > 0 \) denotes an absolute constant that may vary from expression to expression. For convenience, we will also write \( m \) instead of \( m_n \), thereby dropping the index \( n \).

**Lemma 3.1.** Let the Assumption [21] be satisfied. Then

\[
\|P_i(Y^{(\leq m)}_k)\|_p \leq \|X_k - X_k^{(\epsilon')}\|_p \quad \text{and} \quad \|\eta_{j,m}\|_p = O(1).
\]
Proof. We have
\[ \mathcal{P}_i(Y_{k}^{(\leq m)}) = \mathbb{E}[X_{k}^{(m,s)} | \mathcal{F}_i] - \mathbb{E}[X_{k}^{(m,s)} | \mathcal{F}_{i-1}] = \mathbb{E}[X_k - X_{k}^{(i,')} | \mathcal{F}_{k-m}]. \]
Hence Jensen’s inequality gives
\[ \| \mathcal{P}_i(Y_{k}^{(\leq m)}) \|_p \leq \| X_k - X_{k}^{(i,')} \|_p. \]
Using this inequality, applying Theorem 1 of [34] and the stationarity of \( \{X_k\}_{k \in \mathbb{Z}} \), we obtain
\[ \| \mathcal{P}_i(Y_{k}^{(\leq m)}) \|_p \leq C(p) \sum_{k=0}^{\infty} \| X_k - X_{k}^{(i,')} \|_p = O(1). \]

**Lemma 3.2.** Let the Assumption [21] be satisfied. Then
\[ \| Y_{k}^{(> m)} \|_p^2 \leq C \sum_{k=m}^{\infty} \| X_k - X_{k}^{(i,')} \|_p^2. \]

Proof. The lemma follows from combining Proposition 3 in [16] and Theorem 1 in [33].

**Lemma 3.3.** Let the Assumption [21] be satisfied. Then \( \sigma^2 < \infty, n^{-1} B_n^2 \to \sigma^2 \) and \( \text{Var} \eta_{j,m} \to \sigma^2 \) as \( m \) increases.

Proof. The first and second claim follow by proceeding as in [23]. The third claim follows by employing additionally computations as in the proof of Lemma 3.4 given below.

**Lemma 3.4.** Let the Assumption [21] be satisfied. Then for all \( j \in \mathbb{N} \) we have
\[ \lim_{n \to -\infty} \| \mathbb{E}[\eta_{j,m} \eta_{j+1,m}] \| = 0. \]

Proof. Without loss of generality, we can assume that \( j = 1 \). We have
\[ m \mathbb{E}[\eta_{1,m} \eta_{2,m}] = \sum_{k=1}^{m} \sum_{i=-\infty}^{k} \sum_{l=m+1}^{2m} \sum_{j=-\infty}^{l} \mathbb{E}[\mathcal{P}_i(Y_{k}^{(\leq m)}) \mathcal{P}_j(Y_{l}^{(\leq m)})]. \]
By orthogonality of the martingale difference sequences \( \mathcal{P}_i(Y_{k}^{(\leq m)}), \mathcal{P}_j(Y_{l}^{(\leq m)}) \) and an application of the Cauchy–Schwarz inequality, we get
\[ m \mathbb{E}[\eta_{1,m} \eta_{2,m}] \leq \sum_{k=1}^{m} \sum_{i=-\infty}^{k} \sum_{l=m+1}^{2m} \| \mathcal{P}_i(Y_{k}^{(\leq m)}) \|_2 \| \mathcal{P}_j(Y_{l}^{(\leq m)}) \|_2. \]
Applying Lemma 3.1 and shifting the indices \((h = k - j)\), we see that the above is bounded by

\[
\sum_{k=1}^{m} \sum_{l=m+1}^{2m} \sum_{h=0}^{\infty} \|X_k - X'_h\| \cdot \|X_{l+h-k} - X'_{l+h-k}\| 
\leq \sum_{k=1}^{m} \sum_{l=m+1}^{\infty} \sum_{h=0}^{\infty} \|X_h - X'_h\| \cdot \|X_{l-k} - X'_{l-k}\|.
\]

Since \(\sum_{k=0}^{\infty} \|X_k - X'_k\| < \infty\), another shift in the indices implies that this is further bounded by

\[
C \sum_{k=1}^{m} \sum_{l=0}^{\infty} \|X_{l-k+m} - X'_{l-k+m}\| \leq C \sum_{l=0}^{\infty} (m \wedge l + 1) \|X_l - X'_l\|.
\]

We thus obtain

\[
\mathbb{E}[\eta_{1,m} \eta_{2,m}] \leq C m^{-1} \sum_{l=0}^{\infty} (m \wedge l + 1) \|X_l - X'_l\| = o(1) \quad \text{as} \quad m \text{ increases},
\]

which completes the proof.

The following two lemmas are special cases of Theorem 2.6 in [11].

**Lemma 3.5.** Let \(Z_1, Z_2, \ldots, Z_n\) be \(m\)-dependent random variables with zero mean and finite \(\|Z_i\|_p\) for \(2 < p \leq 3\). Then

\[
\sup_{x \in \mathbb{R}} |\Delta_n(x)| \leq 75(10m + 1)^{p-1} B_n^{-p} \sum_{i=1}^{n} \|Z_i\|_p^p.
\]

**Lemma 3.6.** Let \(Z_1, Z_2, \ldots, Z_n\) be \(m\)-dependent random variables with zero mean and finite \(\|Z_i\|_p\) for \(2 < p \leq 3\). Then there exists an absolute constant \(c_0 > 0\) such that

\[
|\Delta_n(x)| \leq c_0 (1 + |x|)^{-p} m^{p-1} B_n^{-p} \sum_{i=1}^{n} \|Z_i\|_p^p.
\]

In the sequel, we require the notion of \(m\)-approximability

\[
\epsilon_{p,n,m} = B_n^{-1} \sum_{k=1}^{n} \|X_k - X_{km}\|_p,
\]

and the following preliminary estimate, which is Lemma 5.1 in [13].
Lemma 3.7. For every $\delta > 0$, every $m, n \geq 1$ and every $x \in \mathbb{R}$ the following estimate holds:

$$\left| P(S_n \leq x B_n) - \Phi(x) \right| \leq A_0(x, \delta) + A_1(m, n, \delta) + \max \left\{ A_2(m, n, x, \delta) + A_3(m, n, x, \delta), A_4(m, n, x, \delta) + A_5(m, n, x, \delta) \right\},$$

where

$$A_0(x, \delta) = |\Phi(x) - \Phi(x + \delta)|, \quad A_1(m, n, \delta) = P(|S_n - S_{nm}| \geq \delta B_n),$$

$$A_2(m, n, x, \delta) = \left| P(S_{nm} \leq (x + \delta) B_n) - \Phi((x + \delta) B_n/B_{nm}) \right|,$$

$$A_3(m, n, x, \delta) = |\Phi((x + \delta) B_n/B_{nm}) - \Phi(x/B_n)|,$$

$$A_4(m, n, x, \delta) = A_2(m, n, x, -\delta) \quad \text{and} \quad A_5(m, n, x, \delta) = A_3(m, n, x, -\delta).$$

We are now in a position to prove the main results. We will first deal with Theorem 3.1.

Proof of Theorem 3.1. By Lemma 3.7 it suffices to bound the quantities $A_0(x, \delta), A_1(m, n, \delta), A_2(m, n, x, \delta)$ and $A_3(m, n, x, \delta)$, where we set

$$\delta = \delta_n(x) = e^{p/(p+1)}(1 + |x|).$$

Recall the definition of $\{\eta_{j,m}\}_{j \in \mathbb{N}}$ and $S_N^{(\leq m)}$, and that $\{\eta_{j,m}\}_{j \in \mathbb{N}}$ constitutes a two-dependent sequence. Let $X_{km} = Y_{k\leq m}$ be the approximating $m$-dependent sequence. Then

$$S_{nm} = \sqrt{n}S_N^{(\leq m)},$$

and an application of Lemma 3.6 yields

$$A_2(m, n, x, \delta) \leq c_0(1 + |x|)^{-p} \cdot 2^{p-1} \cdot \frac{1}{p} \sum_{j=1}^{N} \|\eta_{j,m}\|_p^p.$$
Now, since $B_n^2 = O(\sigma^2 n)$ by Lemma 3.3, it follows from Lemma 3.2 and the assumptions that

$$e_{p,n,m}^2 \leq C n \sum_{k=m}^{\infty} \|X_k - X'_k\|^2_p = o(1),$$

and we conclude that $\{X_k\}_{k \in \mathbb{Z}} \in W(L^2, \{m_n\}, \{B_n\})$. For the remaining parts $A_0(x, \delta), A_1(m, n, \delta)$ and $A_3(m, n, x, \delta)$, one may thus proceed exactly as in the proof of Theorem 3.2 in [19]. This implies that the total remaining error is of the magnitude

$$O\left((-\log(e_{p,n,m}))^{(p+1)/2} e^{p/(p+1)}\right).$$

Hence, piecing everything together, we obtain the claim.

**Proof of Theorem 2.2.** The proof of Theorem 2.2 requires similar arguments. Setting

$$\delta = \delta_n(x) = e_{p,n,m}^{p/(p+1)},$$

using the bound in (3.2) and the fact that $\{X_k\}_{k \in \mathbb{Z}} \in W(L^2, \{m_n\}, \{B_n\})$, one may proceed as in the proof of Theorem 3.1 in [19].

**Proof of Corollary 2.1.** Setting $m_n = n^q$, $q > 0$, it suffices to evaluate the bound in Theorem 2.1. Then, by a simple optimisation, we obtain

$$q = \frac{p^2 - 2}{2(\alpha - 1)p + p^2 + 2},$$

which implies the result.

**Proof of Corollary 2.2.** It again suffices to evaluate the bound in Theorem 2.2. Let $m_n = H \log n$. Then for sufficiently large $H > 0$ it follows that

$$\sum_{k=m_n}^{\infty} \|X_k - X'_k\|_p = O(n^{-1+2/p}).$$

This implies that $e_{p,n,m}^{p/(p+1)} = O(n^{-1/2})$, hence the claim.

**Proof of Corollary 2.3.** We may proceed as in the proof of Corollary 2.2 using Theorem 2.1.

**Acknowledgments.** I would like to thank the anonymous referee for a careful reading of the manuscript and for pointing out some relevant results from the literature that significantly shortened the proofs. Funding by the FG 1735 is gratefully acknowledged.
REFERENCES


Moritz Jirak
Humboldt-Universität zu Berlin
Institut für Mathematik
Unter den Linden 6
D-10099 Berlin, Germany
E-mail: m0ritz@yahoo.com

Received on 7.6.2013; revised version on 4.4.2014