PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 35, Fasc. 1 (2015), pp. 1–14

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 $\mathcal{O}(m^{p-1}n^{-p/2+1})$, 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 $\mathcal{O}(m^{p/2-1}n^{-p/2+1})$. 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 $B_n^2 = \mathrm{Var}[S_n]$. A very important issue in probability theory and statistics is whether the central limit theorem holds or not, i.e., whether we have

$$\lim_{n\to\infty} |P(S_n \leqslant xB_n) - \Phi(x)| = 0, \tag{1.1}$$

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

$$\lim_{n\to\infty} d(P_{S_n/B_n}, P_Z)\, r_n < \infty, \tag{1.2}$$

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

$$\sup_{x\in\mathbb{R}} \Delta_n(x) := \sup_{x\in\mathbb{R}} |P(S_n \leqslant xB_n) - \Phi(x)| \quad \text{as } n \to \infty. \tag{1.3}$$

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

$$\Delta_n(x) := |P(S_n \leqslant xB_n) - \Phi(x)|, \tag{1.4}$$

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

$$\sup_{x\in\mathbb{R}} \Delta_n(x) = \mathcal{O}(\lambda_n),$$

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 [30] showed that it is possible to obtain a rate of $\mathcal{O}(\sqrt{n})$ in (1.3) under certain mixing assumptions, given a bounded support of the underlying sequence $\{X_k\}_{k\in\mathbb{Z}}$ (see also Bolthausen [7] for martingale difference sequences). Given more general assumptions such as α-mixing (see [10] for definitions), Tikhomirov [31] obtained a rate of $\mathcal{O}\big((\log n)^2 n^{-1/2}\big)$, provided that the underlying third moments exist (see also Bentkus et al. [4]). For related results in general Hilbert spaces, we refer to Bentkus [3] and the references therein. We remark that in this case, even considering linear processes results in a nonlinear nature; see, for instance, El Machkouri [15]. In contrast to the previously mentioned results, Tikhomirov [31] also obtained the rate $\mathcal{O}\big((\log n)^2 n^{-1/2}\big)$ in the more difficult case of the non-uniform metric. Similarly, Hörmann [19] obtained rates of the form $\mathcal{O}\big((m_n)^2 n^{-1/2}\big)$ 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. [22], [29], [32]). In particular, it contains examples of time series that are known to be not α-mixing (cf. [1]). Interestingly, this concept is also applicable in more number theoretic settings, see Ibragimov [20], Hörmann [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

$$\sum_{k=1}^{n} \|X_k - X_{km_n}\|_p = o(a_n).$$

We will abbreviate this with $\{X_k\} \in \mathbf{W}(L^2, \{m_n\}, \{a_n\})$.

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 $Y_k^{(\leqslant m)}$), it is possible to establish a rate of $\mathcal{O}\big((m_n)^{1/2}n^{-1/2}\big)$ in (1.3) given $p = 3$ moments, which is an improvement by the factor $m_n^{-3/2}$. Note that in the case of many time series such as GARCH and ARMA models, this results in an ultimate bound of $\mathcal{O}\big((\log n)^{1/2}n^{-1/2}\big)$ 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 $\mathcal{F}_k = \sigma(\epsilon_j,\, j \leqslant k)$. In the sequel, we will consider the sequence of zero mean random variables $X_k = g(\epsilon_k, \epsilon_{k-1}, \ldots)$, $k \geqslant 1$, where g is a measurable function such that X_k are proper random variables. Note that this implies that $\{X_k\}_{k\geqslant 1}$ 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' $\xi_k^{(m,')}$, $\xi_k^{(m,*)}$ as

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

and

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

We put

$$\xi_k' = \xi_k^{(k,')} = (\epsilon_k, \epsilon_{k-1}, \ldots, \epsilon_0', \epsilon_{-1}, \ldots)$$

and

$$\xi_k^* = \xi_k^{(k,*)} = (\epsilon_k, \epsilon_{k-1}, \ldots, \epsilon_0, \epsilon_{-1}', \ldots).$$

By analogy, we put $X_k^{(m,')} = g(\xi_k^{(m,')})$ and $X_k^{(m,*)} = g(\xi_k^{(m,*)})$; in particular, we have $X_k' = X_k^{(k,')}$ and $X_k^* = X_k^{(k,*)}$.

As a dependence measure, one may now consider the quantities $\|X_k - X_k'\|_p$ or $\|X_k - X_k^*\|_p$, $p \geqslant 1$. Dependence conditions of this type are often quite general and easy to verify in many cases; see, for instance, [14], [34] and Examples 2.1 and 2.3 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 $p \in (2,3]$:*

(i) $\mathbb{E}[X_k] = 0$,

(ii) $\sum_{k=1}^{\infty} \|X_k - X_k'\|_p < \infty$,

(iii) $\sigma^2 = \sum_{k=-\infty}^{\infty} \mathbb{E}[X_k X_0] > 0$.

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*

$$\Delta_n(x) \leqslant C_0(1+|x|)^{-p}\left(\sigma^{-p}\left(\frac{m_n}{n}\right)^{p/2-1} + \big(-\log(e_{p,n,m_n})\big)^{(p+1)/2} e_{p,n,m_n}^{p/(p+1)}\right),$$

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

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., $\|X_k - X_k'\|_p = \mathcal{O}(k^{-\alpha})$.

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

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

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* $\|X_k - X_k'\|_p = \mathcal{O}(\rho^{-k})$, $0 < \rho < 1$. *Then there exists a finite*, *absolute constant* $C_0 > 0$ *such that*

$$\Delta_n(x) \leqslant C_0(1+|x|)^{-p}\sigma^{-p}\left(\frac{\log n}{n}\right)^{p/2-1}.$$

Note that if we have $p = 3$ moments, then we obtain a convergence rate of $\mathcal{O}(\sqrt{\log n}n^{-1/2})$ under a very general dependence condition, which improves upon the results in [31] and [19] (in both cases $\mathcal{O}((\log n)^2 n^{-1/2})$ was obtained). Let us also mention that, by imposing stronger rates of decay (e.g., $\|X_k - X_k'\|_p = \mathcal{O}(e^{-k^\gamma})$, $\gamma > 1$), 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* $n\sum_{k=m_n}^{\infty}\|X_k - X_k'\|_p^2 = o(1)$. *Then there exists a finite*, *absolute constant* $C_0 > 0$ *such that*

$$\sup_{x\in\mathbb{R}}|\Delta_n(x)| \leqslant C_0\sigma^{-p}\left(\frac{m_n}{n}\right)^{p/2-1} + 2e_{p,n,m_n}^{p/(p+1)},$$

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

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* $\|X_k - X_k'\|_p = \mathcal{O}(k^{-\alpha})$, $\alpha > 1$. *Then there exists a finite*, *absolute constant* $C_0 > 0$ *such that*

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

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*$(\mathfrak{p},\mathfrak{q})$ *sequences*). Let $\{X_k\}_{k\in\mathbb{Z}}$ be a GARCH$(\mathfrak{p},\mathfrak{q})$ 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

$$L_k^2 = \mu + \alpha_1 L_{k-1}^2 + \ldots + \alpha_{\mathfrak{p}} L_{k-\mathfrak{p}}^2 + \beta_1 X_{k-1}^2 + \ldots + \beta_{\mathfrak{q}} X_{k-\mathfrak{q}}^2$$

with $\mu, \alpha_1, \ldots, \alpha_{\mathfrak{p}}, \beta_1, \ldots, \beta_{\mathfrak{q}} \in \mathbb{R}$. A very important quantity in this context is

$$\gamma_C = \sum_{i=1}^{r} \|\alpha_i + \beta_i \epsilon_i^2\|_2 \quad \text{with } r = \max\{\mathfrak{p}, \mathfrak{q}\}, \tag{2.1}$$

where we replace possible undefined α_i, β_i with zero. If $\gamma_C < 1$, then $\{X_k\}_{k\in\mathbb{Z}}$ is stationary (cf. [9], [8]). In particular, it was shown in [5] that $\{X_k\}_{k\in\mathbb{Z}}$ may then be represented as

$$X_k = \sqrt{\mu}\epsilon_k \big(1 + \sum_{n=1}^{\infty} \sum_{1\leqslant l_1,\ldots,l_n\leqslant r} \prod_{i=1}^{n} (\alpha_{l_i} + \beta_{l_i} \epsilon_{j-l_1-\ldots-l_i}^2)\big)^{1/2}. \tag{2.2}$$

Using this representation and the fact that $|x-y|^p \leqslant |x^2 - y^2|^{p/2}$ for $x, y \geqslant 0$, $p \geqslant 1$, one can follow the proof of Theorem 4.2 in [2] to show that

$$\|X_k - X_k'\|_p = \mathcal{O}(\rho^{-k}), \quad \text{where } 0 < \rho < 1. \tag{2.3}$$

Hence an application of Corollary 2.2 yields a rate of $\mathcal{O}\big((\log n)^{p/2-1} n^{-p/2+1}\big)$ 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

$$L_\epsilon = \sup_{x\neq y} \frac{|f(x,\epsilon) - f(y,\epsilon)|}{|x-y|} \tag{2.4}$$

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

$$\|X_k - X_k'\|_p = \mathcal{O}(\rho^{-k}), \quad \text{where } 0 < \rho < 1, \tag{2.5}$$

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 $\mathcal{O}\big((\log n)^{p/2-1} n^{-p/2+1}\big)$ for $p \in (2, 3]$.

EXAMPLE 2.3 (*Sums of the form* $\sum f(\omega 2^k)$). 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 $\widehat{f}(x) = f(x - \lfloor x \rfloor)$, i.e., $\widehat{f}$ 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

$$L_k = f\big(\sum_{j=1}^{\infty} \zeta_{k+j} 2^{-j}\big),$$

where $\{\zeta_k\}_{k\in\mathbb{Z}}$ is a sequence of i.i.d. random variables, ζ_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

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

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

$$\sup_{x\in\mathbb{R}} \Delta_n(x) \leqslant C_0 \sigma^{-p} \left(\frac{\log n}{n}\right)^{p/2-1}. \tag{2.6}$$

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 $i \in \{n, n-1, \dots\}$, and let $\epsilon_k = \zeta_{T_n(k)}$. Then we may write

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

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 $p \in (2,3]$

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

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 $\mathcal{P}_k(U_i)$ as

$$\mathcal{P}_k(U_i) = \mathbb{E}(U_i \mid \mathcal{F}_k) - \mathbb{E}(U_i \mid \mathcal{F}_{k-1}), \quad k \geqslant 1,\, i \in \mathbb{Z}.$$

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]):

$$X_k = \sum_{i=-\infty}^{k} \mathcal{P}_i(X_k) \quad \text{for all } k \in \mathbb{Z}.$$

As already outlined in the introduction, another essential tool will be approximations with m_n-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 $\mathcal{F}_k = \sigma(\epsilon_j,\, j \leqslant k)$, and define, in addition, the σ-algebra

$$\mathcal{F}_{k-m}^{k} = \sigma(\epsilon_j,\, k-m \leqslant j \leqslant k) \tag{3.1}$$

and the random variables

$$Y_k^{(\leqslant m)} = \mathbb{E}[X_k|\mathcal{F}_{k-m}^{k}], \quad Y_k^{(>m)} = X_k - Y_k^{(\leqslant m)}. \tag{3.2}$$

Let $\eta_{j,m} = m^{-1/2}\sum_{i=jm+1}^{(j+1)m\wedge n} Y_i^{(\leqslant m)}$, where we point out that $\{\eta_{j,m}\}$ is a two-dependent sequence. Let $N = \lceil n/m \rceil + 1$ and put $S_N^{(\leqslant m)} = N^{-1/2}\sum_{j=1}^{N}\eta_{j,m}$. 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])

$$\mathcal{P}_{i+h}(X_{k+h}) \overset{d}{=} \mathcal{P}_i(X_k), \quad h \in \mathbb{Z}.$$

This implies, *in particular*, $\|\mathcal{P}_{i+h}(X_{k+h})\|_p = \|\mathcal{P}_i(X_k)\|_p$, $p \geqslant 1$.

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* 2.1 *be satisfied. Then*

$$\|\mathcal{P}_i(Y_k^{(\leqslant m)})\|_p \leqslant \|X_k - X_k^{(i,')}\|_p \quad \textit{and} \quad \|\eta_{j,m}\|_p = \mathcal{O}(1).$$

P r o o f. We have

$$\mathcal{P}_i(Y_k^{(\leqslant m)}) = \mathbb{E}[X_k^{(m,*)}|\mathcal{F}_i] - \mathbb{E}[X_k^{(m,*)}|\mathcal{F}_{i-1}] = \mathbb{E}[X_k - X_k^{(i,')}|\mathcal{F}_{k-m}^k].$$

Hence Jensen's inequality gives

$$\|\mathcal{P}_i(Y_k^{(\leqslant m)})\|_p \leqslant \|X_k - X_k^{(i,')}\|_p.$$

Using this inequality, applying Theorem 1 of [34] and the stationarity of $\{X_k\}_{k\in\mathbb{Z}}$, we obtain

$$\|\eta_{j,m}\|_p \leqslant C(p) \sum_{k=0}^{\infty} \|X_k - X_k'\|_p = \mathcal{O}(1). \quad \blacksquare \tag{3.3}$$

LEMMA 3.2. *Let the Assumption* 2.1 *be satisfied. Then*

$$\|Y_k^{(>m)}\|_p^2 \leqslant C \sum_{k=m}^{\infty} \|X_k - X_k'\|_p^2. \tag{3.4}$$

P r o o f. The lemma follows from combining Proposition 3 in [16] and Theorem 1 in [33]. $\blacksquare$

LEMMA 3.3. *Let the Assumption* 2.1 *be satisfied. Then* $\sigma^2 < \infty$, $n^{-1}B_n^2 \to \sigma^2$ *and* $\mathrm{Var}[\eta_{j,m}] \to \sigma^2$ *as* m *increases.*

P r o o f. 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. $\blacksquare$

LEMMA 3.4. *Let the Assumption* 2.1 *be satisfied. Then for all* $j \in \mathbb{N}$ *we have*

$$\lim_{n\to\infty} |\mathbb{E}[\eta_{j,m}\eta_{j+1,m}]| = 0.$$

P r o o f. Without loss of generality, we can assume that $j = 1$. We have

$$m\mathbb{E}[\eta_{1,m}\eta_{2,m}] = \sum_{k=1}^{m} \sum_{i=-\infty}^{k} \sum_{l=m+1}^{2m} \sum_{j=-\infty}^{l} \mathbb{E}[\mathcal{P}_i(Y_k^{(\leqslant m)})\mathcal{P}_j(Y_l^{(\leqslant m)})].$$

By orthogonality of the martingale difference sequences $\mathcal{P}_i(Y_k^{(\leqslant m)})$, $\mathcal{P}_j(Y_l^{(\leqslant m)})$ and an application of the Cauchy–Schwarz inequality, we get

$$m\mathbb{E}[\eta_{1,m}\eta_{2,m}] \leqslant \sum_{k=1}^{m} \sum_{i=-\infty}^{k} \sum_{l=m+1}^{2m} \|\mathcal{P}_i(Y_k^{(\leqslant m)})\|_2 \|\mathcal{P}_i(Y_l^{(\leqslant m)})\|_2.$$

Applying Lemma 3.1 and shifting the indices ($h = k - j$), we see that the above is bounded by

$$\sum_{k=1}^{m}\sum_{l=m+1}^{2m}\sum_{h=0}^{\infty}\|X_h - X_h'\|_2\|X_{l+h-k} - X_{l+h-k}'\|_2 \leqslant \sum_{k=1}^{m}\sum_{l=m+1}^{\infty}\sum_{h=0}^{\infty}\|X_h - X_h'\|_2\|X_{l-k} - X_{l-k}'\|_2.$$

Since $\sum_{k=0}^{\infty}\|X_k - X_k'\|_2 < \infty$, another shift in the indices implies that this is further bounded by

$$C\sum_{k=1}^{m}\sum_{l=0}^{\infty}\|X_{l-k+m} - X_{l-k+m}'\|_2 \leqslant C\sum_{l=0}^{\infty}(m \wedge l + 1)\|X_l - X_l'\|_2.$$

We thus obtain

$$\mathbb{E}[\eta_{1,m}\eta_{2,m}] \leqslant Cm^{-1}\sum_{l=0}^{\infty}(m \wedge l + 1)\|X_l - X_l'\|_2 = o(1) \quad \text{as } m \text{ increases},$$

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 $2 < p \leqslant 3$. Then*

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

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 $2 < p \leqslant 3$. Then there exists an absolute constant $c_0 > 0$ such that*

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

In the sequel, we require the notion of m-approximability

$$e_{p,n,m} = B_n^{-1}\sum_{k=1}^{n}\|X_k - X_{km}\|_p, \tag{3.5}$$

and the following preliminary estimate, which is Lemma 5.1 in [19].

LEMMA 3.7. *For every $\delta > 0$, every $m, n \geqslant 1$ and every $x \in \mathbb{R}$ the following estimate holds*:

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

where

$$A_0(x,\delta) = |\Phi(x) - \Phi(x+\delta)|, \quad A_1(m,n,\delta) = P(|S_n - S_{nm}| \geqslant \delta B_n),$$
$$A_2(m,n,x,\delta) = \big|P\big(S_{nm} \leqslant (x+\delta)B_n\big) - \Phi\big((x+\delta)B_n/B_{nm}\big)\big|,$$
$$A_3(m,n,x,\delta) = \big|\Phi\big((x+\delta)B_n/B_{nm}\big) - \Phi(x/B_n)\big|,$$
$$A_4(m,n,x,\delta) = A_2(m,n,x,-\delta) \quad \textit{and} \quad A_5(m,n,x,\delta) = A_3(m,n,x,-\delta).$$

We are now in a position to prove the main results. We will first deal with Theorem 2.1.

P r o o f o f T h e o r e m 2.1. 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

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

Recall the definition of $\{\eta_{j,m}\}_{j\in\mathbb{N}}$ and $S_N^{(\leqslant m)}$, and that $\{\eta_{j,m}\}_{j\in\mathbb{N}}$ constitutes a two-dependent sequence. Let $X_{km} = Y_k^{(\leqslant m)}$ be the approximating m-dependent sequence. Then $S_{nm} = \sqrt{n}S_N^{(\leqslant m)}$, and an application of Lemma 3.6 yields

$$A_2(m,n,x,\delta) \leqslant c_0(1+|x|)^{-p} \cdot 2^{p-1}\mathrm{Var}(S_N^{(\leqslant m)})^{-p/2}\sum_{j=1}^N \|\eta_{j,m}\|_p^p.$$

It follows from Lemma 3.1 that $\sum_{j=1}^N \|\eta_{j,m}\|_p^p = \mathcal{O}(N)$. Using the two-dependence of $\{\eta_{j,m}\}_{j\in\mathbb{N}}$, we also have $\mathrm{Var}(S_N^{(\leqslant m)}) \sim \sum_{j=1}^N \|\eta_{j,m}\|_2^2 + \sum_{j=1}^N \mathbb{E}[\eta_{j,m}\eta_{j+1,m}]$. Lemma 3.4 now gives $\mathbb{E}[\eta_{j,m}\eta_{j+1,m}] = o(1)$ as m increases, and hence Lemma 3.3 implies

$$\mathrm{Var}(S_N^{(\leqslant m)}) = N\big(\sigma^2 + o(1)\big). \tag{3.6}$$

Piecing all these facts together, we deduce that

$$\begin{aligned} A_2(m,n,x,\delta) &\leqslant (1+|x|)^{-p}\mathcal{O}\big((N\sigma^2)^{-p/2}N\big) \\ &= (1+|x|)^{-p}\mathcal{O}\bigg(\sigma^{-p}\Big(\frac{m}{n}\Big)^{p/2-1}\bigg). \end{aligned} \tag{3.7}$$

Now, since $B_n^2 = \mathcal{O}(\sigma^2 n)$ by Lemma 3.3, it follows from Lemma 3.2 and the assumptions that

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

and we conclude that $\{X_k\}_{k\in\mathbb{Z}} \in \mathbf{W}(L^2, \{m_n\}, \{B_n\})$. 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

$$\mathcal{O}\Big(\big(-\log(e_{p,n,m_n})\big)^{(p+1)/2} e_{p,n,m_n}^{p/(p+1)}\Big).$$

Hence, piecing everything together, we obtain the claim. ■

P r o o f o f T h e o r e m 2.2. The proof of Theorem 2.2 requires similar arguments. Setting

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

using the bound in (3.7) and the fact that $\{X_k\}_{k\in\mathbb{Z}} \in \mathbf{W}(L^2, \{m_n\}, \{B_n\})$, one may proceed as in the proof of Theorem 3.1 in [19]. ■

P r o o f o f C o r o l l a r y 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. ■

P r o o f o f C o r o l l a r y 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 $\sum_{k=m_n}^{\infty} \|X_k - X_k'\|_p = \mathcal{O}(n^{-(1+2/p)})$. This implies that $e_{p,n,m}^{p/(p+1)} = \mathcal{O}(n^{-1/2})$, hence the claim. ■

P r o o f o f C o r o l l a r y 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

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

[2] A. Aue, S. Hörmann, L. Horváth, and M. Reimherr, *Break detection in the covariance structure of multivariate time series models*, Ann. Statist. 37 (2009), pp. 4046–4087.

[3] V. Bentkus, *Asymptotic expansions for distributions of sums of independent random elements in a Hilbert space*, Litovsk. Mat. Sb. 4 (1984), pp. 29–48.

[4] V. Bentkus, F. Götze, and A. Tikhomirov, *Berry–Esseen bounds for statistics of weakly dependent samples*, Bernoulli 3 (1997), pp. 329–349.

[5] I. Berkes, S. Hörmann, and L. Horváth, *The functional central limit theorem for a family of GARCH observations with applications*, Statist. Probab. Lett. 78 (2008), pp. 2725–2730.

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

[7] E. Bolthausen, *Exact convergence rates in some martingale central limit theorems*, Ann. Probab. 10 (1982), pp. 672–688.

[8] P. Bougerol and N. Picard, *Stationarity of GARCH processes and of some nonnegative time series*, J. Econometrics 52 (1992), pp. 115–127.

[9] P. Bougerol and N. Picard, *Strict stationarity of generalized autoregressive processes*, Ann. Probab. 20 (1992), pp. 1714–1730.

[10] R. C. Bradley, *Introduction to Strong Mixing Conditions*, Vol. 1, Kendrick Press, Heber City, UT, 2007.

[11] L. Chen and Q. Shao, *Normal approximation under local dependence*, Ann. Probab. 32 (2004), pp. 1985–2028.

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

[13] P. Diaconis and D. Freedman, *Iterated random functions*, SIAM Rev. 41 (1999), pp. 45–76.

[14] P. Doukhan and O. Wintenberger, *An invariance principle for weakly dependent stationary general models*, Probab. Math. Statist. 27 (1) (2007), pp. 45–73.

[15] M. El Machkouri, *Berry–Esseen's central limit theorem for non-causal linear processes in Hilbert space*, Afr. Diaspora J. Math. 10 (2) (2010), pp. 81–86.

[16] M. El Machkouri, D. Volný, and W. B. Wu, *A central limit theorem for stationary random fields*, Stochastic Process. Appl. 123 (1) (2013), pp. 1–14.

[17] J. Gao, *Nonlinear Time Series: Semiparametric and Nonparametric Methods*, Monogr. Statist. Appl. Probab., Vol. 108, Chapman & Hall/CRC, Boca Raton, FL, 2007.

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

[19] S. Hörmann, *Berry–Esseen bounds for econometric time series*, ALEA Lat. Am. J. Probab. Math. Stat. 6 (2009), pp. 377–397.

[20] I. Ibragimov, *The central limit theorem for sums of functions of independent variables and for sums of the form* $\sum f(2^k t)$, Theory Probab. Appl. 12 (1967), pp. 596–607.

[21] J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, second edition, Grundlehren Math. Wiss., Vol. 288, Springer, Berlin 2003.

[22] M. Jirak, *Change-point analysis in increasing dimension*, J. Multivariate Anal. 111 (2012), pp. 136–159.

[23] M. Jirak, *On weak invariance principles for sums of dependent random functionals*, Statist. Probab. Lett. 83 (10) (2013), pp. 2291–2296.

[24] V. Ladokhin and D. Moskvin, *An estimate for the remainder term in the central limit theorem for sums of functions of independent variables and sums of the form* $\sum f(t2^k)$, Theory Probab. Appl. 16 (1971), pp. 116–125.

[25] D. L. McLeish, *A maximal inequality and dependent strong laws*, Ann. Probab. 3 (1975), pp. 829–839.

[26] W. Min and W. Wu, *On linear processes with dependent innovations*, Stochastic Process. Appl. 115 (2005), pp. 939–958.
[27] M. Peligrad and S. Utev, *A new maximal inequality and invariance principle for stationary sequences*, Ann. Probab. 33 (2005), pp. 798–815.
[28] M. Peligrad and S. Utev, *Central limit theorem for stationary linear processes*, Ann. Probab. 34 (2006), pp. 1608–1622.
[29] M. B. Priestley, *Non-linear and Non-stationary Time Series Analysis*, Academic Press, London 1988.
[30] E. Rio, *Sur le théorème de Berry–Esseen pour les suites faiblement dépendantes*, Probab. Theory Related Fields 104 (1996), pp. 255–282.
[31] A. Tikhomirov, *On the convergence rate in the central limit theorem for weakly dependent random variables*, Theory Probab. Appl. 25 (1981), pp. 790–809.
[32] H. Tong, *Non-linear Time Series: A Dynamical System Approach. With an Appendix by K. S. Chan*, Oxford Statist. Sci. Ser., Vol. 6, Oxford University Press, New York 1990.
[33] W. B. Wu, *Nonlinear system theory: Another look at dependence*, Proc. Natl. Acad. Sci. USA 102 (2005), pp. 14150–14154.
[34] W. B. Wu, *Strong invariance principles for dependent random variables*, Ann. Probab. 35 (2007), pp. 2294–2320.
[35] W. B. Wu and X. Shao, *Limit theorems for iterated random functions*, J. Appl. Probab. 41 (2004), pp. 425–436.

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