ESAIM: PS 17 (2013) 120–134
DOI: 10.1051/ps/2011144

ESAIM: Probability and Statistics
www.esaim-ps.org

A CENTRAL LIMIT THEOREM FOR TRIANGULAR ARRAYS OF WEAKLY DEPENDENT RANDOM VARIABLES, WITH APPLICATIONS IN STATISTICS

Michael H. Neumann[1]

Abstract. We derive a central limit theorem for triangular arrays of possibly nonstationary random variables satisfying a condition of weak dependence in the sense of Doukhan and Louhichi [*Stoch. Proc. Appl.* **84** (1999) 313–342]. The proof uses a new variant of the Lindeberg method: the behavior of the partial sums is compared to that of partial sums of *dependent* Gaussian random variables. We also discuss a few applications in statistics which show that our central limit theorem is tailor-made for statistics of different type.

Mathematics Subject Classification. 60F05, 62F40, 62G07, 62M15.

Received May 18, 2010.

1. Introduction

For a long time mixing conditions have been the dominating way for imposing a restriction on the dependence between time series data. The most popular notions of this type are strong mixing (α-mixing) which was introduced by Rosenblatt [25], absolute regularity (β-mixing) introduced by Volkonski and Rozanov [27] and uniform mixing (ϕ-mixing) proposed by Ibragimov [15]. These concepts were extensively used by statisticians since many processes of interest fulfill such conditions under natural restrictions on the process parameters; see Doukhan [13] as one of the standard references for examples and available tools.

On the other hand, it turned out that some classes of processes statisticians like to work with are not mixing although a decline of the influence of past states takes place as time evolves. The simplest example of such a process is a stationary AR(1)-process, $X_t = \theta X_{t-1} + \varepsilon_t$, where the innovations are independent with $P(\varepsilon_t = 1) = P(\varepsilon_t = -1) = 1/2$ and $0 < |\theta| \le 1/2$. The stationary distribution of this process is supported on $[-1/(1-|\theta|), 1/(1-|\theta|)]$ and it follows from the above model equation that X_t has always the same sign as ε_t. Hence, we could perfectly recover $X_{t-1}, X_{t-2}, \ldots$ from X_t which clearly excludes any of the common mixing properties to hold. (Rosenblatt [26] mentions the fact that a process similar to $(X_t)_{t\in\mathbb{Z}}$ is purely deterministic going backwards in time. A rigorous proof that it is not strong mixing is given by Andrews [1]). On the other hand, it can be seen from the equation $X_t = \varepsilon_t + \theta\varepsilon_{t-1} + \ldots + \theta^{t-s-1}\varepsilon_{s+1} + \theta^{t-s}X_s$ that the impact of X_s on X_t declines as $t-s$ tends to infinity. Besides this simple example, there are several other processes of this type which are of considerable interest in statistics and for which mixing properties cannot be proved. To give a more

Keywords and phrases. Central limit theorem, Lindeberg method, weak dependence, bootstrap.

[1] Friedrich-Schiller-Universität Jena, Institut für Stochastik, Ernst-Abbe-Platz 2, 07743 Jena, Germany.
michael.neumann@uni-jena.de

Article published by EDP Sciences

© EDP Sciences, SMAI 2013

https://doi.org/10.1051/ps/2011144 Published online by Cambridge University Press

relevant example, for bootstrapping a linear autoregressive process of finite order, it is most natural to estimate first the distribution of the innovations by the empirical distribution of the (possibly re-centered) residuals and to generate then a bootstrap process iteratively by drawing independent bootstrap innovations from this distribution. It turns out that commonly used techniques for proving mixing of autoregressive processes fail; because of the discreteness of the bootstrap innovations it is in general impossible to construct a successful coupling of two versions of the process with different initial values.

Inspired by such problems, Doukhan and Louhichi [14] and Bickel and Bühlmann [3] proposed alternative notions called weak dependence and ν-dependence, respectively. Basically, they imposed a restriction on covariances of smooth functions of blocks of random variables separated by a certain time gap and called a process weakly dependent (or ν-dependent, respectively) if these covariances tend to zero as the time gap increases. It has been shown that many classes of processes for which mixing is problematic satisfy appropriate versions of these weaker conditions; see *e.g.* Dedecker *et al.* [12] for numerous examples.

The main result in this paper is a central limit theorem for triangular arrays of weakly dependent random variables. Under conditions of this type, central limit theorems were given in Corollary A in Doukhan and Louhichi [14] for sequences of stationary random variables and in Theorem 1 in Coulon-Prieur and Doukhan [6] for triangular arrays of asymptotically sparse random variables as they appear with nonparametric kernel density estimators. Using their notion of ν-mixing Bickel and Bühlmann [3] proved a CLT for linear processes of infinite order and their (smoothed) bootstrap counterparts. CLTs for triangular arrays of weakly dependent, row-wise stationary random variables were also given in Theorem 6.1 in Neumann and Paparoditis [21] and in Theorem 1 in Bardet *et al.* [2]. The latter one is formulated for situations where the dependence between the summands declines as $n \to \infty$. It can also be applied to dependent processes in conjunction with an additional blocking step. Under projective conditions, versions of a CLT were derived by Dedecker [9], Dedecker and Rio [11] and Dedecker and Merlevède [10]. In the present paper we generalize previous versions of a CLT under weak dependence. Having applications in statistics such as to bootstrap processes and to nonparametric curve estimators in mind, it is important to formulate the theorem for triangular arrays; see Sections 3.2 and 3.3 for a discussion of such applications. Moreover, motivated by a particular application sketched in Section 3.1, we also allow that the random variables in each row are nonstationary. The assumptions of our CLT are quite weak; beyond finiteness of second moments only Lindeberg's condition has to be fulfilled and only some minimal dependence conditions are imposed. Nevertheless, this result can be applied in situations offering different challenges such as nonstationarity, sparsity, and with a triangular scheme. The versatility of our CLT is indicated by the applications discussed in Section 3.

To prove this result, we adapt the approach introduced by Lindeberg [18] in the case of independent random variables. Actually, we use a variant of this method where the asymptotic behavior of the partial sums is explicitly compared to the behavior of partial sums of Gaussian random variables. This approach has also been chosen in Neumann and Paparoditis [21] for proving a CLT for triangular schemes but under stationarity. However, since in our more general context the increments of the variances of the partial sum process are not necessarily nonnegative we cannot take independent Gaussian random variables as an appropriate counterpart. Our new idea is to choose here *dependent*, jointly Gaussian random variables, having the same covariance structure as the original ones. This modification of the classical Lindeberg method seems to be quite natural in the case of dependent random variables and it turns out that it does not create any essential additional problems compared to situations where the partial sum process can be compared to a process with independent Gaussian summands.

This paper ist organised as follows. In Section 2 we state and discuss the main result. Section 3 is devoted to a discussion of some possible applications in statistics. The proof of the CLT is contained in Section 4.

https://doi.org/10.1051/ps/2011144 Published online by Cambridge University Press

2. Main result

Theorem 2.1. *Suppose that* $(X_{n,k})_{k=1,\ldots,n}$, $n \in \mathbb{N}$, *is a triangular scheme of random variables with* $EX_{n,k} = 0$ *and* $\sum_{k=1}^{n} EX_{n,k}^2 \leq v_0$, *for all* n, k *and some* $v_0 < \infty$. *We assume that*

$$\sigma_n^2 := \operatorname{var}(X_{n,1} + \ldots + X_{n,n}) \underset{n\to\infty}{\longrightarrow} \sigma^2 \in [0, \infty), \tag{2.1}$$

and that

$$\sum_{k=1}^{n} E\{X_{n,k}^2 \mathbb{1}(|X_{n,k}| > \epsilon)\} \underset{n\to\infty}{\longrightarrow} 0 \tag{2.2}$$

holds for all $\epsilon > 0$. *Furthermore, we assume that there exists a summable sequence* $(\theta_r)_{r\in\mathbb{N}}$ *such that, for all* $u \in \mathbb{N}$ *and all indices* $1 \leq s_1 < s_2 < \ldots < s_u < s_u + r = t_1 \leq t_2 \leq n$, *the following upper bounds for covariances hold true: for all measurable functions* $g : \mathbb{R}^u \longrightarrow \mathbb{R}$ *with* $\|g\|_\infty = \sup_{x\in\mathbb{R}^u} |g(x)| \leq 1$,

$$|\operatorname{cov}(g(X_{n,s_1}, \ldots, X_{n,s_u}) X_{n,s_u}, X_{n,t_1})| \leq \left(EX_{n,s_u}^2 + EX_{n,t_1}^2 + n^{-1}\right) \theta_r \tag{2.3}$$

and

$$|\operatorname{cov}(g(X_{n,s_1}, \ldots, X_{n,s_u}), X_{n,t_1} X_{n,t_2})| \leq \left(EX_{n,t_1}^2 + EX_{n,t_2}^2 + n^{-1}\right) \theta_r. \tag{2.4}$$

Then

$$X_{n,1} + \ldots + X_{n,n} \xrightarrow{d} \mathcal{N}(0, \sigma^2).$$

Remark 2.2. To prove this theorem, we use the approach introduced by Lindeberg [18] in the case of independent random variables. Later this method has been adapted by Billingsley [4] and Ibragimov [16] to the case of stationary and ergodic martingale difference sequences, by Rio [24] to the case of triangular arrays of strongly mixing and possibly nonstationary processes, by Dedecker [9] to stationary random fields, by Dedecker and Rio [11] for deriving a functional central limit theorem for stationary processes, and by Dedecker and Merlevède [10] and Neumann and Paparoditis [21] for triangular arrays of weakly dependent random variables. The core of Lindeberg's method is that, for any suffiently regular test funtion h, $E\{h(X_{n,1} + \ldots + X_{n,k}) - h(X_{n,1} + \ldots + X_{n,k-1})\}$ is approximated by $E\{h''(X_{n,1} + \ldots + X_{n,k-1})v_k/2\}$, where $v_k = \operatorname{var}(X_{n,1} + \ldots + X_{n,k}) - \operatorname{var}(X_{n,1} + \ldots + X_{n,k-1})$ is the increment of the variance of the partial sum process. Since an analogous approximation also holds for the partial sums process of related Gaussian random variables one can make the transition from the non-Gaussian to the Gaussian case. In most of these papers, this idea is put into practice by a comparison of $Eh(X_{n,1} + \ldots + X_{n,n})$ with $Eh(Z_{n,1} + \ldots + Z_{n,n})$, where $Z_{n,1}, \ldots, Z_{n,n}$ are independent Gaussian random variables with $EZ_{n,k} = 0$ and $EZ_{n,k}^2 = \operatorname{var}(X_{n,1} + \ldots + X_{n,k}) - \operatorname{var}(X_{n,1} + \ldots + X_{n,k-1})$. To simplify computations, sometimes blocks of random variables rather than single random variables are considered; see *e.g.* Dedecker and Rio [11] and Dedecker and Merlevède [10]. In the case considered here, however, we have to deviate from this common practice. Since $X_{n,1}, \ldots, X_{n,n}$ are not necessarily stationary we cannot guarantee that the increments of the variances, $\operatorname{var}(X_{n,1} + \ldots + X_{n,k}) - \operatorname{var}(X_{n,1} + \ldots + X_{n,k-1})$, are always nonnegative. Moreover, it is even not clear if blocks of fixed length would help here. Therefore, we compare the behavior of $X_{n,1} + \ldots + X_{n,n}$ with that of $Z_{n,1} + \ldots + Z_{n,n}$, where $Z_{n,1}, \ldots, Z_{n,n}$ are no longer independent. Rather, we choose them as centered and jointly Gaussian random variables with the same covariance structure as $X_{n,1}, \ldots, X_{n,n}$. Nevertheless, it turns out that the proof goes through without essential complications compared to the common situation with independent $Z_{n,k}$'s. Note that Rio [24] also proved a central limit theorem for triangular arrays of not necessarily stationary, mixing random variables. This author managed the problem with possibly negative increments of the variances of the partial sums by omitting an explicit use of Gaussian random variables; rather he derived an approximation to the characteristic function, on the Gaussian side.

https://doi.org/10.1051/ps/2011144 Published online by Cambridge University Press

3. Applications in statistics

The purpose of this section is twofold. First, we intend to demonstrate the versatility of our CLT by a few applications of different type. And second, we also want to show how this can be accomplished in an effective manner.

3.1. **Asymptotics of quadratic forms**

For the analysis of nonstationary time series, Dahlhaus [7] introduced the notion of locally stationary processes as a suitable framework for a rigorous asymptotic theory. Basically, one has to deal with a triangular scheme of random variables $(Y_{n,t})_{t=1,\dots,n}$, $n \in \mathbb{N}$, where the covariance structure of $\dots, Y_{n,[nv]-1}, Y_{n,[nv]}, Y_{n,[nv]+1}, \dots$ is asymptotically constant for any fixed $v \in [0,1]$. In this sense, the time points $1, \dots, n$ are renormalized to the unit interval. As an empirical version of the corresponding local spectral density, Neumann and von Sachs [23] proposed the so-called preperiodogram,

$$I_n(v,\omega) \;=\; \frac{1}{2\pi} \sum_{k\in\mathbb{Z}:\, 1\le [nv+1/2-k/2],[nv+1/2+k/2]\le n} Y_{n,[nv+1/2-k/2]} Y_{n,[nv+1/2+k/2]} \cos(\omega\, k).$$

The integrated and normalized preperiodogram is defined as

$$J_n(u,\nu) \;=\; \sqrt{n} \int_0^\nu \int_0^u \{I_n(v,\omega) \;-\; EI_n(v,\omega)\}\, \mathrm{d}v\, \mathrm{d}\omega.$$

The process $J_n = (J_n(u,\nu))_{u\in[0,1],\nu\in[0,\pi]}$ can be used in many contexts, ranging from testing the hypothesis of stationarity to estimation of model parameters; see Dahlhaus [8] for several examples. This author proposed the test statistics

$$T_n \;=\; \sup_{u\in[0,1],\nu\in[0,\pi]} \left| \sqrt{n} \int_0^\nu \left[\int_0^u I_n(v,\omega)\, \mathrm{d}v \;-\; u \int_0^1 I_n(v,\omega)\, \mathrm{d}v \right] \mathrm{d}\omega \right|.$$

Under the null hypothesis of stationarity, this test statistic can be approximated by $\sup_{u,\nu} |J_n(u,\nu) - uJ_n(1,\nu)|$. Neumann and Paparoditis [22] develop details for a bootstrap-based test for stationarity based on T_n. In the case of the null hypothesis, convergence of $J_n = (J_n(\nu,u))_{u\in[0,1],\nu\in[0,\pi]}$ to a certain Gaussian process $J = (J(\nu,u))_{u\in[0,1],\nu\in[0,\pi]}$ can be proved. In what follows we will sketch how Theorem 2.1 can be employed for proving asymptotic normality of its finite-dimensional distributions, which yields in conjunction with stochastic equicontinuity the convergence to a certain Gaussian process. The advantage of this somehow unusual approach is that finiteness of fourth moments suffices whereas traditionally used cumulant methods require finiteness of all moments.

Under the hypothesis of stationarity we don't have a triangular scheme and we denote the underlying process by $(Y_t)_{t\in\mathbb{Z}}$. Note that the preperiodogram can be rewritten as

$$I_n(v,\omega) \;=\; \underline{Y}_n' A_n \underline{Y}_n,$$

where $\underline{Y}_n = (Y_1, \dots, Y_n)'$ and

$$(A_n)_{i,j} \;=\; \begin{cases} (2\pi)^{-1} \cos(\omega(i-j)), & \text{if } (i+j-1)/(2n) \le v < (i+j+1)/(2n), \\ 0, & \text{otherwise.} \end{cases}$$

Accordingly, we have that

$$J_n(u,\nu) \;=\; \underline{Y}_n' B_n(u,\nu) \underline{Y}_n \;-\; E\underline{Y}_n' B_n(u,\nu) \underline{Y}_n,$$

https://doi.org/10.1051/ps/2011144 Published online by Cambridge University Press

where

$$(B_n(u,\nu))_{i,j} = \begin{cases} \frac{\sqrt{n}}{2\pi}\,\frac{\sin(\nu(i-j))}{i-j}\,\lambda^1\left([0,u]\cap[\frac{i+j-1}{2n},\frac{i+j+1}{2n})\right), & \text{for } i\neq j,\\ \frac{\sqrt{n}}{2\pi}\,\nu\,\lambda^1\left([0,u]\cap[\frac{2i-1/2}{n},\frac{2i+1/2}{n})\right), & \text{for } i=j \end{cases}$$

and λ^1 denotes the Lebesgue measure on $(\mathbb{R},\mathcal{B})$. The desired convergence of the finite-dimensional distributions of J_n to those of J, *i.e.*,

$$(J_n(u_1,\nu_1),\ldots,J_n(u_d,\nu_d))' \xrightarrow{d} J = (J_1,\ldots,J_d)', \tag{3.1}$$

for arbitrary $u_1,\ldots,u_d\in[0,1]$, $\nu_1,\ldots,\nu_d\in[0,\pi]$ and $d\in\mathbb{N}$, will follow by the Cramér–Wold device from

$$Z_n = \sum_{i=1}^{d} c_i J_n(u_i,\nu_i) \xrightarrow{d} Z \sim \mathcal{N}(0,v),$$

where $v=\sum_{j,k=1}^{d} c_j\,c_k\,\mathrm{cov}(J_k,J_l)$, for any $c_1,\ldots,c_d\in\mathbb{R}$. We assume that

(A1) (i) $\sum_{j=-\infty}^{\infty}|\mathrm{cov}(Y_0,Y_j)|<\infty$,
(ii) $EY_0^4<\infty$ and $\sum_{j,k,l}|\mathrm{cum}(Y_0,Y_j,Y_k,Y_l)|<\infty$.

Z_n can be rewritten as

$$Z_n = \underline{Y}_n' C_n \underline{Y}_n - E\underline{Y}_n' C_n \underline{Y}_n,$$

where C_n is an $(n\times n)$-matrix with

$$(C_n)_{i,j} = \sum_{k=1}^{d} c_k B_n(u_k,\nu_k)_{i,j} = O\left(\frac{1}{\sqrt{n}}\left(1\wedge\frac{1}{|i-j|}\right)\right).$$

Since the off-diagonal terms decay fast enough as $|i-j|$ grows we can approximate Z_n by

$$Z_n^{(K)} = \underline{Y}_n' C_n^{(K)} \underline{Y}_n - E\underline{Y}_n' C_n^{(K)} \underline{Y}_n,$$

where $(C_n^{(K)})_{i,j} = (C_n)_{i,j}\mathbb{1}(|i-j|\le K)$. Let $\Sigma_n=\mathrm{Cov}(\underline{Y}_n)$. It follows from (A1) that

$$\begin{aligned} &E\left(\{\underline{Y}_n' C_n \underline{Y}_n - E\underline{Y}_n' C_n \underline{Y}_n\} - \left\{\underline{Y}_n' C_n^{(K)} \underline{Y}_n - E\underline{Y}_n' C_n^{(K)} \underline{Y}_n\right\}\right)^2\\ &\quad = \mathrm{var}\left(\underline{Y}_n'(C_n - C_n^{(K)})\underline{Y}_n\right)\\ &\quad = \sum_{i,j,k,l=1}^{n} (C_n - C_n^{(K)})_{i,j}\,(C_n - C_n^{(K)})_{k,l}\,\mathrm{cum}(Y_i,Y_j,Y_k,Y_l)\\ &\qquad + 2\,\mathrm{tr}\left((C_n - C_n^{(K)})\Sigma_n(C_n - C_n^{(K)})\Sigma_n\right)\\ &\quad = O(1/K^2) + O(1/K). \end{aligned} \tag{3.2}$$

The latter equality follows from $\lambda_{\max}(\Sigma_n)\le\|\Sigma_n\|_\infty = \sup_i\sum_{j=1}^{n}|\mathrm{cov}(Y_i,Y_j)|$ and the estimate $\mathrm{tr}((C_n - C_n^{(K)})\Sigma_n(C_n - C_n^{(K)})\Sigma_n) \le (\lambda_{\max}(\Sigma_n))^2\,\mathrm{tr}((C_n - C_n^{(K)})^2)$; see *e.g.* Lütkepohl [20] (p. 44). According to Theorem 4.2 in Billingsley [5], it follows from (3.2) that it remains to prove that, for all fixed K,

$$Z_n^{(K)} \xrightarrow{d} \mathcal{N}(0,v^{(K)}), \tag{3.3}$$

https://doi.org/10.1051/ps/2011144 Published online by Cambridge University Press

with $v^{(K)} \to_{n\to\infty} v$. The matrix $C_n^{(K)}$, however, has almost a diagonal form which will allow us to derive (3.3) from Theorem 2.1. To this end, we rewrite $Z_n^{(K)}$ as

$$Z_n^{(K)} = \sum_{i=1}^{n} X_{n,i},$$

where $X_{n,i} = \sum_{j=\max\{i-K,1\}}^{\min\{i+K,n\}} (C_n)_{i,j}(Y_iY_j - E(Y_iY_j))$. Now we assume that

(A2) For all fixed $K \in \mathbb{N}$, there exists a summable sequence $(\theta_r)_{r\in\mathbb{N}}$ such that, for $s_1 < \ldots < s_u < t_1 \le t_2$, $\|g\|_\infty \le 1$ and $-K \le k_1, k_2, l_1, l_2 \le K$,
(i) $|\mathrm{cov}(g(Y_{s_1-K}, \ldots, Y_{s_u+K})Y_{s_u+k_2}Y_{s_u+k_2}, Y_{t_1+l_1}Y_{t_1+l_2})| \le \theta_{t_1-s_u}$,
(ii) $|\mathrm{cov}(g(Y_{s_1-K}, \ldots, Y_{s_u+K}), Y_{t_1+k_1}Y_{t_1+k_2}Y_{t_2+l_1}Y_{t_2+l_2})| \le \theta_{t_1-s_u}$.

It can be easily verified that the triangular scheme $(X_{n,i})_{i=1,\ldots,n}$ satisfies the assumptions of Theorem 2.1. Here we note that the $X_{n,i}$ are in general not stationary, even if the Y_i are. Therefore, it is essential to have a CLT which allows nonstationary random variables in each row. Moreover, it also follows from (3.2) that

$$\sup_n \left|\mathrm{var}(\underline{Y}_n' C_n^{(K)} \underline{Y}_n) - \mathrm{var}(\underline{Y}_n' C_n \underline{Y}_n)\right| \underset{K\to\infty}{\longrightarrow} 0.$$

This, together with

$$\mathrm{var}(\underline{Y}_n' C_n \underline{Y}_n) \underset{n\to\infty}{\longrightarrow} v,$$

yields (3.1) under the assumptions (A1) and (A2).

3.2. **Consistency of bootstrap methods**

Since our central limit theorem is formulated for triangular schemes it is tailor-made for applications to bootstrap processes. We demonstrate this for the simplest possible example, the sample mean.

Assume that we have a strictly stationary process $(X_t)_{t\in\mathbb{Z}}$ with $EX_0 = \mu$ and $EX_0^2 < \infty$. Additionally we assume that there exists a summable sequence $(\theta_r)_{r\in\mathbb{N}}$ such that, for all $u \in \mathbb{N}$ and all indices $1 \le s_1 < s_2 < \ldots < s_u < s_u + r = t_1 \le t_2 \le n$, the following upper bounds for covariances hold true: for all measurable functions $g : \mathbb{R}^u \longrightarrow \mathbb{R}$ with $\|g\|_\infty \le 1$,

$$|\mathrm{cov}\,(g(X_{s_1}, \ldots, X_{s_u})X_{s_u}, X_{t_1})| \le \left(EX_{s_u}^2 + EX_{t_1}^2 + 1\right)\theta_r \tag{3.4}$$

and

$$|\mathrm{cov}\,(g(X_{s_1}, \ldots, X_{s_u}), X_{t_1}X_{t_2})| \le \left(EX_{t_1}^2 + EX_{t_2}^2 + 1\right)\theta_r. \tag{3.5}$$

Then the triangular scheme $(X_{n,k})_{k=1,\ldots,n}$, $n \in \mathbb{N}$, given by $X_{n,k} = (X_k - \mu)/\sqrt{n}$ satisfies the conditions of Theorem 2.1. In particular, it follows from (3.4) and majorized convergence that

$$\begin{aligned}
\sigma_n^2 &:= \mathrm{var}(X_{n,1} + \ldots + X_{n,n}) \\
&= \mathrm{var}(X_0) + 2\sum_{k=1}^{n-1}(1 - k/n)\,\mathrm{cov}(X_0, X_k) \\
&\underset{n\to\infty}{\longrightarrow} \sigma^2 := \mathrm{var}(X_0) + 2\sum_{k=1}^{\infty}\mathrm{cov}(X_0, X_k) \in [0, \infty).
\end{aligned}$$

Now we obtain from Theorem 2.1, for $\bar{X}_n = n^{-1}\sum_{t=1}^n X_t$, that

$$\sqrt{n}(\bar{X}_n - \mu) = X_{n,1} + \ldots + X_{n,n} \overset{d}{\longrightarrow} Z \sim \mathcal{N}(0, \sigma^2). \tag{3.6}$$

https://doi.org/10.1051/ps/2011144 Published online by Cambridge University Press

Assume now that, conditioned on $X_1, \ldots, X_n$, a stationary bootstrap version $(X_t^*)_{t\in\mathbb{Z}}$ is given. We assume that

$$P_*^{X_0^*, X_k^*} \Longrightarrow P^{X_0, X_k} \qquad \text{in probability} \tag{3.7}$$

holds for all $k \in \mathbb{N}$ and that

$$E_*[(X_0^*)^2] \xrightarrow{P} E[X_0^2]. \tag{3.8}$$

(Starred symbols such as E_* and P_* refer as usual to the conditional distribution, given $X_1, \ldots, X_n$). (3.7) and (3.8) imply that

$$\mathrm{cov}_*(X_0^*, X_k^*) \xrightarrow{P} \mathrm{cov}(X_0, X_k) \qquad \forall k = 0, 1, \ldots \tag{3.9}$$

Assume further that there exist sequences $(\theta_{n,r})_{r\in\mathbb{N}}$, $n \in \mathbb{N}$, which may depend on the original sample $X_1, \ldots, X_n$, such that, for all $u \in \mathbb{N}$ and all indices $1 \le s_1 < s_2 < \ldots < s_u < s_u + r = t_1 \le t_2 \le n$, the following upper bounds for covariances hold true: for all measurable functions $g : \mathbb{R}^u \longrightarrow \mathbb{R}$ with $\|g\|_\infty \le 1$,

$$\left|\mathrm{cov}_*\left(g(X_{s_1}^*, \ldots, X_{s_u}^*)X_{s_u}^*, X_{t_1}^*\right)\right| \le \left(E(X_{s_u}^*)^2 + E(X_{t_1}^*)^2 + 1\right)\theta_{n,r} \tag{3.10}$$

and

$$\left|\mathrm{cov}_*\left(g(X_{s_1}^*, \ldots, X_{s_u}^*), X_{t_1}^* X_{t_2}^*\right)\right| \le \left(E(X_{t_1}^*)^2 + E(X_{t_2}^*)^2 + 1\right)\theta_{n,r}, \tag{3.11}$$

and that

$$P\left(\theta_{n,r} \le \bar{\theta}_r \quad \forall r \in \mathbb{N}\right) \underset{n\to\infty}{\longrightarrow} 1, \tag{3.12}$$

for some summable sequence $(\bar{\theta}_r)_{r\in\mathbb{N}}$. Actually, (3.10) to (3.12) can often be verified for model-based bootstrap schemes if for the underlying model (3.4) and (3.5) are fulfilled; see Neumann and Paparoditis [21] for a few examples. We define $X_{n,k}^* = (X_k^* - \mu_n)/\sqrt{n}$, where $\mu_n = E_* X_0^*$. It follows, again from (3.7) and (3.8), that

$$\begin{aligned}
&\sum_{k=1}^n E_*\left\{(X_{n,k}^*)^2 \mathbb{1}(|X_{n,k}^*| > \epsilon)\right\} \\
&\quad = E_*\left\{(X_0^* - \mu_n)^2 \mathbb{1}(|X_0^* - \mu_n| > \epsilon\sqrt{n})\right\} \xrightarrow{P} 0.
\end{aligned}$$

Hence, the triangular scheme $(X_{n,k}^*)_{k\in\mathbb{N}}$, $n \in \mathbb{N}$, satisfies the conditions of Theorem 2.1 "in probability". The latter statement means that there exist measurable sets $\Omega_n \subseteq \mathbb{R}^n$ with $P((X_1, \ldots, X_n)' \in \Omega_n) \longrightarrow_{n\to\infty} 1$ and, for any sequence $(\omega_n)_{n\in\mathbb{N}}$ with $\omega_n \in \Omega_n$, we have that, conditioned on $(X_1, \ldots, X_n)' = \omega_n$, the triangular scheme $(X_{n,k}^*)_{k\in\mathbb{N}}$ satisfies the conditions of our CLT. This, however, implies that

$$\sqrt{n}(\bar{X}_n^* - \mu_n) = X_{n,1}^* + \ldots + X_{n,n}^* \xrightarrow{d} Z, \qquad \text{in probability.} \tag{3.13}$$

(3.6) and (3.13) yield that

$$\sup_x \left|P\left(\sqrt{n}(\bar{X}_n - \mu) \le x\right) - P_*\left(\sqrt{n}(\bar{X}_n^* - \mu_n) \le x\right)\right| \xrightarrow{P} 0,$$

which means that the bootstrap produces an asymptotically correct confidence interval for μ.

3.3. **Nonparametric density estimation under uniform mixing**

As a further test case for the suitability of our CLT for triangular schemes we consider asymptotics for a nonparametric density estimator under a uniform (ϕ-) mixing condition. Here we have the typical aspect of sparsity, that is, in the case of a kernel function with compact support, there is a vanishing share of nonzero summands. It turns out that a result by Dedecker and Merlevède [10] can also be proved by our Theorem 2.1.

https://doi.org/10.1051/ps/2011144 Published online by Cambridge University Press

Assume that observations $X_1, \ldots, X_n$ from a strictly stationary process $(X_t)_{t\in\mathbb{Z}}$ with values in $\mathbb{R}^d$ are available. We assume that X_0 has a density which we denote by f. The commonly used kernel density estimator $\widehat{f}_n$ of f is given by

$$\widehat{f}_n(x) \;=\; \frac{1}{nh_n^d}\sum_{k=1}^{n} K\left(\frac{x - X_k}{h_n}\right),$$

where $K : \mathbb{R}^d \to \mathbb{R}$ is a measurable "kernel" function and $(h_n)_{n\in\mathbb{N}}$ is a sequence of (nonrandom) bandwidths. Favorable asymptotic properties of $\widehat{f}_n(x)$ are usually guaranteed if f is continuous at x, K possesses certain regularity properties and, as a minimal condition for consistency, if $h_n \longrightarrow_{n\to\infty} 0$ and $nh_n^d \longrightarrow_{n\to\infty} \infty$. Here we will make the following assumptions:

(B1) (i) f is bounded and continuous at x.
(ii) For all $k \in \mathbb{N}$, the joint density of X_0 and X_k, f_{X_0,X_k}, is bounded.
(iii) $h_n \longrightarrow_{n\to\infty} 0$ and $nh_n^d \longrightarrow_{n\to\infty} \infty$.

(B2) (i) $(X_t)_{t\in\mathbb{Z}}$ is uniform mixing with coefficients ϕ_r satisfying $\sum_{r=1}^{\infty}\sqrt{\phi_r} < \infty$.
(ii) $\int_{\mathbb{R}^d} |K(u)|\,\lambda^d(\mathrm{d}u) < \infty$ and $\int_{\mathbb{R}^d} K(u)^2\,\lambda^d(\mathrm{d}u) < \infty$.

(B3) (i) $(X_t)_{t\in\mathbb{Z}}$ is uniform mixing with coefficients ϕ_r satisfying $\sum_{r=1}^{\infty}\phi_r < \infty$.
(ii) K is bounded and $\int_{\mathbb{R}^d} |K(u)|\,\lambda^d(\mathrm{d}u) < \infty$.

(B1) and (B2) are similar to the conditions of a functional CLT in Billingsley [5] (p. 174), while Dedecker and Merlevède [10] proved asymptotic normality of a kernel density estimator under conditions similar to (B1) and (B3). We will show how Theorem 2.1 can be employed for proving asymptotic normality of $\widehat{f}_n(x)$.

Lemma 3.1. *If (B1) and either (B2) or (B3) are fulfilled, then*

$$\sqrt{nh_n^d}\left(\widehat{f}_n(x) \;-\; E\widehat{f}_n(x)\right) \;\xrightarrow{d}\; \mathcal{N}(0, \sigma_0^2),$$

where $\sigma_0^2 \;=\; f(x)\;\int_{\mathbb{R}^d} K^2(u)\,\lambda^d(\mathrm{d}u)$.

Proof. First of all, it is easy to see that the Lindeberg condition (2.2) is fulfilled by the triangular scheme in both cases. To verify the other conditions of our central limit theorem we need appropriate covariance inequalities for ϕ-mixing processes. For $s_1 < s_2 < \ldots < s_u < t_1 < t_2 < \ldots < t_v$, it follows from Lemma 20.1 in Billingsley [5] that

$$|\mathrm{cov}\,(g(X_{s_1}, \ldots, X_{s_u}), h(X_{t_1}, \ldots, X_{t_v}))| \le 2\,\sqrt{\phi_{t_1 - s_u}}\,\sqrt{E\{g^2(X_{s_1}, \ldots, X_{s_u})\}}\,\sqrt{E\{h^2(X_{t_1}, \ldots, X_{t_v})\}}. \quad (3.14)$$

Furthermore, equation (20.28) in Billingsley [5] yields that

$$|\mathrm{cov}\,(g(X_{s_1}, \ldots, X_{s_u}), h(X_{t_1}, \ldots, X_{t_v}))| \le 2\,\phi_{t_1 - s_u}\;E\,\{|g\,(X_{s_1}, \ldots, X_{s_u})|\}\;\|h\|_\infty. \quad (3.15)$$

In order to get condition (2.4) from our CLT satisfied, we will consider the process $(X_t)_{t\in\mathbb{Z}}$ in *time-reversed* order. We define, for $k = 1, \ldots, n$ and $n \in \mathbb{N}$,

$$X_{n,k} \;=\; (nh_n^d)^{-1/2}\,\{K((x - X_{n-k+1})/h_n) \;-\; EK((x - X_{n-k+1})/h_n)\}\,.$$

Now let $1 \le s_1 < s_2 < \ldots < s_u < t_1 \le t_2 \le n$ and $g : \mathbb{R}^{du} \longrightarrow \mathbb{R}$ be measurable with $\|g\|_\infty \le 1$. We consider first the case where (B1) and (B2) are fulfilled. It follows from (3.14) that

$$|\mathrm{cov}\,(g(X_{n,s_1}, \ldots, X_{n,s_u})X_{n,s_u}, X_{n,t_1})| \;\le\; 2\,\sqrt{\phi_{t_1 - s_u}}\,\sqrt{EX_{n,s_u}^2}\,\sqrt{EX_{n,t_1}^2}, \quad (3.16)$$

https://doi.org/10.1051/ps/2011144 Published online by Cambridge University Press

and from (3.15) that

$$|\mathrm{cov}\,(g(X_{n,s_1},\ldots,X_{n,s_u}),X_{n,t_1}X_{n,t_2})| \;\le\; 2\;\phi_{t_1-s_u}\;EX_{n,t_1}^2. \tag{3.17}$$

Therefore, the weak dependence conditions (2.3) and (2.4) are fulfilled for $\theta_r = 2\sqrt{\phi_r}$. It remains to indentify the limiting variance of $X_{n,1}+\ldots+X_{n,n}$. Since $EX_{n,k}^2 \le n^{-1}E\{h_n^{-d}K^2((x-X_{n-k+1})/h_n)\} \le Cn^{-1}$, for some $C<\infty$, we obtain from (3.16) that

$$|\mathrm{cov}\,(X_{n,k},X_{n,k+l})| \;\le\; 2\;\sqrt{\phi_l}\;C\;n^{-1} \qquad \forall l\in\mathbb{N}. \tag{3.18}$$

On the other hand, it follows from the boundedness of the joint densities that there exist finite constants $C_1,C_2,\ldots$ such that

$$|\mathrm{cov}\,(X_{n,k},X_{n,k+l})| \;\le\; h_n^d\;C_l\;n^{-1}$$

holds for all $l\in\mathbb{N}$. Both inequalities together yield, by majorized convergence, that

$$\sum_{k=1}^{n}\sum_{l\neq k}|\mathrm{cov}\,(X_{n,k},X_{n,l})| \underset{n\to\infty}{\longrightarrow} 0. \tag{3.19}$$

Therefore, we have, as in the case of independent random variables, that

$$\mathrm{var}(X_{n,1}+\ldots+X_{n,n}) \;=\; n\;\mathrm{var}\,(X_{n,1}) \;+\; o(1) \underset{n\to\infty}{\longrightarrow} \sigma_0^2.$$

Hence, the assertion follows from Theorem 2.1.

Now we consider the other case where (B1) and (B3) are fulfilled. We only have to modify the two covariance estimates (3.16) and (3.18). Instead of these two inequalities we use the following alternative estimates:

$$\begin{aligned}
&|\mathrm{cov}\,(g(X_{n,s_1},\ldots,X_{n,s_u})X_{n,s_u},X_{n,t_1})| \\
&\quad\le\quad 2\;\phi_{t_1-s_u}\;\mathrm{ess\;sup}|X_{n,s_u}|\;E|X_{n,t_1}| \\
&\quad\le\quad 2\;\phi_{t_1-s_u}\;\|K\|_\infty\;\|f\|_\infty\int_{\mathbb{R}^d}|K(u)|\,\lambda^d(\mathrm{d}u)\;n^{-1}
\end{aligned} \tag{3.20}$$

and

$$|\mathrm{cov}\,(X_{n,k},X_{n,k+l})| \;\le\; 2\;\phi_l\;\|K\|_\infty\;\|f\|_\infty\int_{\mathbb{R}^d}|K(u)|\,\lambda^d(\mathrm{d}u)\;n^{-1} \qquad \forall l\in\mathbb{N}. \tag{3.21}$$

This time the weak dependence conditions (2.3) and (2.4) are fulfilled for $\theta_r = 2\phi_r\max\{\|K\|_\infty^2\|f\|_\infty(2M)^d,1\}$, and we obtain again the assertion. □

Remark 3.2. The example of kernel density estimators is included as a further test case for the suitability of our CLT for triangular schemes. Here we are faced with the aspect of sparsity, where only a vanishing share of the summands is of non-negligible size. It turns out that again moment conditions of second order suffice for asymptotic normality. It is not clear if the summability condition imposed for the uniform mixing coefficients is the best possible. For stationary ρ-mixing sequences, Ibragimov [17] proved a CLT under the condition $\sum_{n=1}^{\infty}\rho(2^n)<\infty$. Since $\rho(n)\le 2\sqrt{\phi_n}$ the condition $\sum_{n=1}^{\infty}\sqrt{\phi_{2^n}}<\infty$ suffices for the CLT to hold in the case of a stationary uniform mixing sequence.

4. Proof of the main result

Proof of Theorem 2.1. Let $h:\mathbb{R}\longrightarrow\mathbb{R}$ be an arbitrary bounded and three times continuously differentiable function with $\max\{\|h'\|_\infty,\|h''\|_\infty,\|h^{(3)}\|_\infty\}\le 1$. Furthermore, let $Z_{n,1},\ldots,Z_{n,n}$ be centered and jointly Gaussian random variables which are independent of $X_{n,1},\ldots,X_{n,n}$ and have the same covariance structure as $X_{n,1},\ldots,X_{n,n}$, that is, $\mathrm{cov}(Z_{n,j},Z_{n,k})=\mathrm{cov}(X_{n,j},X_{n,k})$, for $1\le j,k\le n$.

https://doi.org/10.1051/ps/2011144 Published online by Cambridge University Press

Since $Z_{n,1}+\ldots+Z_{n,n} \sim N(0,\sigma_n^2) \Longrightarrow N(0,\sigma^2)$ it follows from Theorem 7.1 in Billingsley [5] that it suffices to show that

$$Eh\left(X_{n,1}+\ldots+X_{n,n}\right) - Eh\left(Z_{n,1}+\ldots+Z_{n,n}\right) \underset{n\to\infty}{\longrightarrow} 0. \tag{4.1}$$

We define $S_{n,k}=\sum_{j=1}^{k-1} X_{n,j}$ and $T_{n,k}=\sum_{j=k+1}^{n} Z_{n,j}$. Then,

$$Eh\left(X_{n,1}+\ldots+X_{n,n}\right) - Eh\left(Z_{n,1}+\ldots+Z_{n,n}\right) = \sum_{k=1}^{n}\Delta_{n,k},$$

where

$$\Delta_{n,k} = E\left\{h(S_{n,k}+X_{n,k}+T_{n,k}) - h(S_{n,k}+Z_{n,k}+T_{n,k})\right\}.$$

We set $v_{n,k}=\frac{1}{2}EX_{n,k}^2+\sum_{j=1}^{k-1}EX_{n,k}X_{n,j}$ and $\bar{v}_{n,k}=\frac{1}{2}EZ_{n,k}^2+\sum_{j=k+1}^{n}EZ_{n,k}Z_{n,j}$. Note that $2v_{n,k}=\mathrm{var}(S_{n,k}+X_{n,k})-\mathrm{var}(S_{n,k})$ are the commonly used increments of the variances of the partial sum process while $2\bar{v}_{n,k}=\mathrm{var}(T_{n,k}+Z_{n,k})-\mathrm{var}(T_{n,k})$ are the increments of the variances of the reversed partial sum process. Now we decompose further $\Delta_{n,k}=\Delta_{n,k}^{(1)}-\Delta_{n,k}^{(2)}+\Delta_{n,k}^{(3)}$, where

$$\begin{aligned}
\Delta_{n,k}^{(1)} &= Eh(S_{n,k}+X_{n,k}+T_{n,k}) - Eh(S_{n,k}+T_{n,k}) - v_{n,k}Eh''(S_{n,k}+T_{n,k}),\\
\Delta_{n,k}^{(2)} &= Eh(S_{n,k}+Z_{n,k}+T_{n,k}) - Eh(S_{n,k}+T_{n,k}) - \bar{v}_{n,k}Eh''(S_{n,k}+T_{n,k}),\\
\Delta_{n,k}^{(3)} &= (v_{n,k}-\bar{v}_{n,k})Eh''(S_{n,k}+T_{n,k}).
\end{aligned}$$

We will show that

$$\sum_{k=1}^{n}\Delta_{n,k}^{(i)} \underset{n\to\infty}{\longrightarrow} 0, \qquad \text{for } i=1,2,3. \tag{4.2}$$

(i) *Upper bound for* $|\sum_{k=1}^{n}\Delta_{n,k}^{(1)}|$

Let $\epsilon>0$ be arbitrary. We will actually show that

$$\left|\sum_{k=1}^{n}\Delta_{n,k}^{(1)}\right| \le \epsilon \qquad \text{for all } n\ge n(\epsilon) \tag{4.3}$$

and $n(\epsilon)$ sufficiently large.

It follows from a Taylor series expansion that

$$\Delta_{n,k}^{(1)} = EX_{n,k}h'(S_{n,k}+T_{n,k}) + E\left\{\frac{X_{n,k}^2}{2}h''(S_{n,k}+\tau_{n,k}X_{n,k}+T_{n,k})\right\} - v_{n,k}Eh''(S_{n,k}+T_{n,k}),$$

for some appropriate random $\tau_{n,k}\in(0,1)$. (It is not clear that $\tau_{n,k}$ is measurable, however, $X_{n,k}^2h''(S_{n,k}+\tau_{n,k}X_{n,k}+T_{n,k})$ as a sum of measurable quantities is. Therefore, the above expectation is correctly defined). Since $S_{n,1}=0$ and, therefore, $EX_{n,k}h'(S_{n,1}+T_{n,k})=0$ we obtain that

$$\begin{aligned}
EX_{n,k}h'(S_{n,k}+T_{n,k}) &= \sum_{j=1}^{k-1}EX_{n,k}\left\{h'(S_{n,j+1}+T_{n,k}) - h'(S_{n,j}+T_{n,k})\right\}\\
&= \sum_{j=1}^{k-1}EX_{n,k}X_{n,j}h''(S_{n,j}+\mu_{n,k,j}X_{n,j}+T_{n,k}),
\end{aligned}$$

https://doi.org/10.1051/ps/2011144 Published online by Cambridge University Press

again for some random $\mu_{n,k,j} \in (0,1)$. This allows us to decompose $\Delta_{n,k}^{(1)}$ as follows:

$$\begin{aligned}\Delta_{n,k}^{(1)} &= \sum_{j=1}^{k-d-1} E\left[X_{n,k}X_{n,j}\left\{h''(S_{n,j}+\mu_{n,k,j}X_{n,j}+T_{n,k}) - Eh''(S_{n,k}+T_{n,k})\right\}\right] \\ &\quad + \sum_{j=k-d}^{k-1} E\left[X_{n,k}X_{n,j}\left\{h''(S_{n,j}+\mu_{n,k,j}X_{n,j}+T_{n,k}) - Eh''(S_{n,k}+T_{n,k})\right\}\right] \\ &\quad + \frac{1}{2}E\left[X_{n,k}^2\left\{h''(S_{n,k}+\tau_{n,k}X_{n,k}+T_{n,k}) - Eh''(S_{n,k}+T_{n,k})\right\}\right] \\ &= \Delta_{n,k}^{(1,1)} + \Delta_{n,k}^{(1,2)} + \Delta_{n,k}^{(1,3)}, \end{aligned} \tag{4.4}$$

say.

By choosing d sufficiently large we obtain from (2.3) that

$$\left|\sum_{k=1}^{n}\Delta_{n,k}^{(1,1)}\right| \le (2v_0+1)\sum_{j=d+1}^{n}\theta_j \le \frac{\epsilon}{3} \qquad \text{for all } n. \tag{4.5}$$

The term $\Delta_{n,k}^{(1,2)}$ will be split up as

$$\begin{aligned}\Delta_{n,k}^{(1,2)} &= \sum_{j=k-d}^{k-1} EX_{n,k}X_{n,j}\left\{h''(S_{n,j}+\mu_{n,k,j}X_{n,j}+T_{n,k}) - h''(S_{n,j-d'}+T_{n,k})\right\} \\ &\quad + \sum_{j=k-d}^{k-1} EX_{n,k}X_{n,j}\left\{h''(S_{n,j-d'}+T_{n,k}) - Eh''(S_{n,j-d'}+T_{n,k})\right\} \\ &\quad + \sum_{j=k-d}^{k-1} EX_{n,k}X_{n,j}\left\{Eh''(S_{n,j-d'}+T_{n,k}) - Eh''(S_{n,k}+T_{n,k})\right\} \\ &= \Delta_{n,k}^{(1,2,1)} + \Delta_{n,k}^{(1,2,2)} + \Delta_{n,k}^{(1,2,3)}, \end{aligned} \tag{4.6}$$

say. (The proper choice of d' will become clear from what follows).

Using the decomposition $X_{n,k} = X_{n,k}\mathbb{1}(|X_{n,k}| > \epsilon') + X_{n,k}\mathbb{1}(|X_{n,k}| \le \epsilon')$, for some $\epsilon' > 0$, we obtain by the Cauchy-Schwarz inequality and the Lindeberg condition (2.2) that

$$\begin{aligned}\left|\sum_{k=1}^{n}\Delta_{n,k}^{(1,2,1)}\right| &\le 2\,d\sqrt{\sum_{k=1}^{n}EX_{n,k}^2\mathbb{1}(|X_{n,k}|>\epsilon')}\sqrt{\sum_{j=1}^{n}EX_{n,j}^2} + \epsilon'\sqrt{\sum_{k=1}^{n}\sum_{j=k-d}^{k-1}EX_{n,j}^2} \\ &\quad \times\sqrt{\sum_{k=1}^{n}\sum_{j=k-d}^{k-1}E\{h''(S_{n,j}+\mu_{n,k,j}X_{n,j}+T_{n,k}) - h''(S_{n,j-d'}+T_{n,k})\}^2} \\ &= o(1) + O(\epsilon'). \end{aligned} \tag{4.7}$$

Using condition (2.4) we obtain that

$$\left|\sum_{k=1}^{n}\Delta_{n,k}^{(1,2,2)}\right| \le d\,\theta_{d'}\,(2v_0+1). \tag{4.8}$$

https://doi.org/10.1051/ps/2011144 Published online by Cambridge University Press

Furthermore, it follows from the Lindeberg condition (2.2) that

$$\max_{1\le k\le n}\{\mathrm{var}(X_{n,k})\} \underset{n\to\infty}{\longrightarrow} 0, \tag{4.9}$$

which implies that

$$\left|\sum_{k=1}^{n}\Delta_{n,k}^{(1,2,3)}\right| \underset{n\to\infty}{\longrightarrow} 0.$$

This implies, in conjunction with (4.7) and (4.8) that

$$\left|\sum_{k=1}^{n}\Delta_{n,k}^{(1,2)}\right| \le \frac{\epsilon}{3} \qquad \text{for all } n\ge n^{(1)}(\epsilon), \tag{4.10}$$

provided that d' and $n^{(1)}(\epsilon)$ are chosen sufficiently large. Finally, we obtain in complete analogy to the computations above that

$$\left|\sum_{k=1}^{n}\Delta_{n,k}^{(1,3)}\right| \le \frac{\epsilon}{3} \qquad \text{for all } n\ge n^{(2)}(\epsilon), \tag{4.11}$$

which completes, in conjunction with (4.5) and (4.10), the proof of (4.3).

(ii) *Upper bound for* $|\sum_{k=1}^{n}\Delta_{n,k}^{(2)}|$

Similarly to (4.4), we decompose $\Delta_{n,k}^{(2)}$ as

$$\begin{aligned}\Delta_{n,k}^{(2)} &= E\left[Z_{n,k}h'(S_{n,k}+T_{n,k+d})\right]\\ &\quad+\sum_{j=k+1}^{k+d}E\left[Z_{n,k}Z_{n,j}\left\{h''(S_{n,k}+\bar\mu_{n,k,j}Z_{n,j}+T_{n,j})-Eh''(S_{n,k}+T_{n,k})\right\}\right]\\ &\quad+\frac{1}{2}E\left[Z_{n,k}^2\left\{h''(S_{n,k}+\bar\tau_{n,k}Z_{n,k}+T_{n,k})-Eh''(S_{n,k}+T_{n,k})\right\}\right]\\ &= \Delta_{n,k}^{(2,1)}+\Delta_{n,k}^{(2,2)}+\Delta_{n,k}^{(2,3)},\end{aligned}$$

say, where $\bar\mu_{n,k,j}$ and $\bar\tau_{n,k}$ are appropriate random quantities with values in $(0,1)$. It follows from (2.3) that

$$|\mathrm{cov}(Z_{n,j},Z_{n,k})| = |\mathrm{cov}(X_{n,j},X_{n,k})| \le \left(EX_{n,j}^2+EX_{n,k}^2+n^{-1}\right)\theta_{|j-k|}. \tag{4.12}$$

For our following estimates we make use of Lemma 1 from Liu [19] which states that, for $X:=(X_1,\ldots,X_k)'\sim\mathcal{N}(\mu,\Sigma)$ and any function $h:\mathbb{R}^k\longrightarrow\mathbb{R}$ such that $\partial h/\partial x_i$ exists almost everywhere and $E|\frac{\partial}{\partial x_i}h(X)|<\infty$, $i=1,\ldots,k$, the following identity holds true:

$$\mathrm{cov}(X,h(X)) = \Sigma\left(E\left[\frac{\partial}{\partial x_1}h(X)\right],\ldots,E\left[\frac{\partial}{\partial x_k}h(X)\right]\right)'. \tag{4.13}$$

Since $S_{n,k}$ is by construction independent of $Z_{n,1},\ldots,Z_{n,n}$ we obtain from (4.13) that

$$\begin{aligned}\Delta_{n,k}^{(2,1)} &= \mathrm{cov}\left(Z_{n,k},h'(S_{n,k}+Z_{n,k+d+1}+\ldots+Z_{n,n})\right)\\ &= \sum_{j=k+d+1}^{n}\mathrm{cov}(Z_{n,k},Z_{n,j})\,Eh''(S_{n,k}+T_{n,k+d}),\end{aligned}$$

https://doi.org/10.1051/ps/2011144 Published online by Cambridge University Press

which implies by (4.12) that

$$\left|\sum_{k=1}^{n}\Delta_{n,k}^{(2,1)}\right| \leq (2v_0+1)\sum_{j=d+1}^{\infty}\theta_j. \tag{4.14}$$

Similarly to (4.6), we decompose $\Delta_{n,k}^{(2,2)}$ as

$$\begin{aligned}\Delta_{n,k}^{(2,2)} &= \sum_{j=k+1}^{k+d} E\left[Z_{n,k}Z_{n,j}\left\{h''(S_{n,k}+\bar{\mu}_{n,k,j}Z_{n,j}+T_{n,j})-h''(S_{n,k}+T_{n,j+d'})\right\}\right]\\ &\quad+\sum_{j=k+1}^{k+d} E\left[Z_{n,k}Z_{n,j}\left\{h''(S_{n,k}+T_{n,j+d'})-Eh''(S_{n,k}+T_{n,j+d'})\right\}\right]\\ &\quad+\sum_{j=k+1}^{k+d} E\left[Z_{n,k}Z_{n,j}\left\{Eh''(S_{n,k}+T_{n,j+d'})-Eh''(S_{n,k}+T_{n,k})\right\}\right]\\ &= \Delta_{n,k}^{(2,2,1)}+\Delta_{n,k}^{(2,2,2)}+\Delta_{n,k}^{(2,2,3)},\end{aligned} \tag{4.15}$$

say. It follows from (4.9) that $\sum_{k=1}^{n}E|Z_{k,n}|^3 \longrightarrow_{n\to\infty} 0$, which implies that

$$\left|\sum_{k=1}^{n}\Delta_{n,k}^{(2,2,1)}\right| \underset{n\to\infty}{\longrightarrow} 0. \tag{4.16}$$

To estimate $\Delta_{n,k}^{(2,2,2)}$, we use the following identities:

$$\begin{aligned}&E\left[Z_{n,k}Z_{n,j}\left\{h''(S_{n,k}+T_{n,j+d'})-Eh''(S_{n,k}+T_{n,j+d'})\right\}\right]\\ &= \operatorname{cov}\left(Z_{n,k}Z_{n,j},h''(S_{n,k}+T_{n,j+d'})\right)\\ &= \operatorname{cov}\left(Z_{n,k},Z_{n,j}h''(S_{n,k}+T_{n,j+d'})\right)-E[Z_{n,k}Z_{n,j}]\,Eh''(S_{n,k}+T_{n,j+d'})\\ &= \sum_{l\geq d'+1}\operatorname{cov}(Z_{n,k},Z_{n,j+l})\,E[Z_{n,j}h^{(3)}(S_{n,k}+T_{n,j+d'})].\end{aligned}$$

The latter equation follows again from (4.13). This implies that

$$\left|\sum_{k=1}^{n}\Delta_{n,k}^{(2,2,2)}\right| = O\left(\sum_{j=1}^{d}\sum_{l=d'+1}^{\infty}\theta_{j+l}\right). \tag{4.17}$$

Finally, $\sum_{k=1}^{n}\Delta_{n,k}^{(2,2,3)} = O(\sum_{k=1}^{n}E|Z_{n,k}|^3) = o(1)$ is obvious. This yields, in conjunction with (4.16) and (4.17), that

$$\sum_{k=1}^{n}\Delta_{n,k}^{(2,2)} \underset{n\to\infty}{\longrightarrow} 0. \tag{4.18}$$

The term $\sum_{k=1}^{n}\Delta_{n,k}^{(2,3)}$ can be estimated analogously, which gives, together with (4.14) and (4.18), that

$$\sum_{k=1}^{n}\Delta_{n,k}^{(2)} \underset{n\to\infty}{\longrightarrow} 0. \tag{4.19}$$

https://doi.org/10.1051/ps/2011144 Published online by Cambridge University Press

(iii) *Upper bound for* $|\sum_{k=1}^n \Delta_{n,k}^{(3)}|$

We split up:

$$\begin{aligned}\sum_{k=1}^n \Delta_{n,k}^{(3)} &= \sum_{k=1}^n \left(\sum_{j=k-d}^{k-1} EX_{n,k}X_{n,j} - \sum_{j=k+1}^{k+d} EX_{n,k}X_{n,j} \right) Eh''(S_{n,k} + T_{n,k}) \\ &\quad + \sum_{k=1}^n \left(\sum_{j=1}^{k-d-1} EX_{n,k}X_{n,j} - \sum_{j=k+d+1}^{n} EX_{n,k}X_{n,j} \right) Eh''(S_{n,k} + T_{n,k}) \\ &= \sum_{k=1}^n \sum_{j=1}^d E\left[X_{n,k}X_{n,k-j} \left\{ Eh''(S_{n,k} + T_{n,k}) - Eh''(S_{n,k-j} + T_{n,k-j}) \right\} \right] \\ &\quad + O\left(\sum_{r=d+1}^{\infty} \theta_r \right). \end{aligned} \tag{4.20}$$

Since $\|h^{(3)}\|_\infty \le 1$ we obtain from (4.9) that

$$\max_{1\le k\le n,\, 1\le j\le d} \left| Eh''(S_{n,k} + T_{n,k}) - Eh''(S_{n,k-j} + T_{n,k-j}) \right| \underset{n\to\infty}{\longrightarrow} 0.$$

Therefore, the first term on the right-hand side of (4.20) converges to 0 as $n \to \infty$. The second one can be made arbitrarily small if d is chosen large enough. This completes the proof of the theorem. □

Acknowledgements. This research was supported by the German Research Foundation DFG (project: NE 606/2-1). I thank two anonymous referees for their helpful comments.

References

[1] D.W.K. Andrews, Non-strong mixing autoregressive processes, *J. Appl. Probab.* **21** (1984) 930–934.
[2] J.M. Bardet, P. Doukhan, G. Lang and N. Ragache, Dependent Lindeberg central limit theorem and some applications. *ESAIM: PS* **12** (2008) 154–172.
[3] P.J. Bickel and P. Bühlmann, A new mixing notion and functional central limit theorems for a sieve bootstrap in time series. *Bernoulli* **5** (1999) 413–446.
[4] P. Billingsley, The Lindeberg-Lévy theorem for martingales. *Proc. Amer. Math. Soc.* **12** (1961) 788–792.
[5] P. Billingsley, *Convergence of Probability Measures*. Wiley, New York (1968).
[6] C. Coulon-Prieur and P. Doukhan, A triangular central limit theorem under a new weak dependence condition. *Stat. Probab. Lett.* **47** (2000) 61–68.
[7] R. Dahlhaus, Fitting time series models to nonstationary processes. *Ann. Stat.* **25** (1997) 1–37.
[8] R. Dahlhaus, Local inference for locally stationary time series based on the empirical spectral measure. *J. Econ.* **151** (2009) 101–112.
[9] J. Dedecker, A central limit theorem for stationary random fields. *Probab. Theory Relat. Fields* **110** (1998) 397–426.
[10] J. Dedecker and F. Merlevède, Necessary and sufficient conditions for the conditional central limit theorem. *Ann. Probab.* **30** (2002) 1044–1081.
[11] J. Dedecker and E. Rio, On the functional central limit theorem for stationary processes. *Ann. Inst. Henri Poincaré Série B* **36** (2000) 1–34.
[12] J. Dedecker, P. Doukhan, G. Lang, J.R. León, S. Louhichi and C. Prieur, *Weak Dependence: With Examples and Applications*. Springer-Verlag. *Lect. Notes Stat.* **190** (2007).
[13] P. Doukhan, *Mixing: Properties and Examples*. Springer-Verlag. *Lect. Notes Stat.* **85** (1994).
[14] P. Doukhan and S. Louhichi, A new weak dependence condition and application to moment inequalities. *Stoch. Proc. Appl.* **84** (1999) 313–342.
[15] I.A. Ibragimov, Some limit theorems for stationary processes. *Teor. Veroyatn. Primen.* **7** (1962) 361–392 (in Russian). [English translation: *Theory Probab. Appl.* **7** (1962) 349–382].

https://doi.org/10.1051/ps/2011144 Published online by Cambridge University Press

[16] I.A. Ibragimov, A central limit theorem for a class of dependent random variables. *Teor. Veroyatnost. i Primenen.* **8** (1963) 89–94 (in Russian). [English translation: *Theor. Probab. Appl.* **8** (1963) 83–89].
[17] I.A. Ibragimov, A note on the central limit theorem for dependent random variables. *Teor. Veroyatnost. i Primenen.* **20** (1975) 134–140 (in Russian). [English translation: *Theor. Probab. Appl.* **20** (1975) 135–141].
[18] J.W. Lindeberg, Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung, *Math. Zeitschr.* **15** (1922) 211–225.
[19] J.S. Liu, Siegel's formula *via* Stein's identities. *Statist. Probab. Lett.* **21** (1994) 247–251.
[20] H. Lütkepohl, *Handbook of Matrices*. Wiley, Chichester (1996).
[21] M.H. Neumann and E. Paparoditis, Goodness-of-fit tests for Markovian time series models: Central limit theory and bootstrap approximations. *Bernoulli* **14** (2008) 14–46.
[22] M.H. Neumann and E. Paparoditis, *A test for stationarity*. Manuscript (2011).
[23] M.H. Neumann and R. von Sachs, Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra. *Ann. Statist.* **25** (1997) 38–76.
[24] E. Rio, About the Lindeberg method for strongly mixing sequences. *ESAIM: PS* **1** (1995) 35–61.
[25] M. Rosenblatt, A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. USA* **42** (1956) 43–47.
[26] M. Rosenblatt, Linear processes and bispectra. *J. Appl. Probab.* **17** (1980) 265–270.
[27] V.A. Volkonski and Y.A. Rozanov, Some limit theorems for random functions, Part I. *Teor. Veroyatn. Primen.* **4** (1959) 186–207 (in Russian). [English translation: *Theory Probab. Appl.* **4** (1959) 178–197].

https://doi.org/10.1051/ps/2011144 Published online by Cambridge University Press