AN INVARIANCE PRINCIPLE FOR TRIANGULAR ARRAYS
OF DEPENDENT VARIABLES WITH APPLICATION TO
AUTOCOVARIANCE ESTIMATION

BY
HUI CHEN and JOSEPH P. ROMANO

TECHNICAL REPORT NO. 506
DECEMBER 1996

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION
GRANT DMS95-04379

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
An Invariance Principle for Triangular Arrays of Dependent Variables with Application to Autocovariance Estimation

By

Hui Chen and Joseph P. Romano

Technical Report No. 506
December 1996

Prepared Under the Auspices Of
National Science Foundation
Grant DMS95-04379

Department of Statistics
Stanford University
Stanford, California
An Invariance Principle for Triangular Arrays of Dependent Variables with Application to Autocovariance Estimation

Hui Chen and Joseph P. Romano
Department of Statistics
Stanford University
Stanford, CA 94305
December 30, 1996

Abstract

The invariance principle for triangular array of dependent variables is studied. We use the concept of mixingale, proposed by McLeish [11]. Uniform bounds are imposed on the growth of conditional expectations with respect to distant predecessors. The theorem is applied to invariance principles for autocovariance estimates based on triangular arrays of time series data for weak mixing sequences and linear processes. Such results are required for bootstrap applications, which will be developed in subsequent work.

1 Introduction

A large literature can be found on central limit theorems and invariance principles for dependent variables. Among the dependency types explored are mixing sequences, martingale differences, and linear processes, often with a stationarity condition added. In McLeish's 1975 paper [11], the concept of mixingale, a hybrid of the notions of martingales and mixing sequences, is proposed. This ingenious idea provides a unified treatment of the aforementioned three kinds of dependencies, and it works in some nonstationary situation as well. Our goal is to extend McLeish's results to a double array of dependent variables, and develop some invariance principles and central limit theorems for arrays of mixing and linear processes, and sample autocovariance functions of them.

The motivation of this work lies in our attempts to prove consistency of bootstrap and other resampling schemes in time series analysis. Bootstrap procedures produce pseudo-data that form a portion of an array of random variables. Therefore the consistency results hinge on a central limit theorem for triangular arrays of one kind or another.

One method for proving central limit theorems for mixing processes, such as the way Ibragimov proved his two famous theorems for uniform mixing and strong mixing processes, is to divide the sequence of dependent random variables into alternative blocks, called the "small blocks" and "big blocks". However, we found this technique does not carry over to the triangular array case easily when the mixing conditions are generalized in a simple way. Other attempts rely on maximal inequalities which usually depend on stringent mixing assumptions (Doukhan [7]). McLeish's mixingale approach, on the other hand, makes use of general probabilistic theories for general stochastic processes, which enjoys relative ease of generalization.

Our invariance principles for general mixingale arrays and for arrays of mixing variables are extensions of McLeish [11]. The emphasis is on linear processes, and on central limit theorems for sample autocovariance functions, including the autocovariance estimates based on the tapered observations. The results can be readily applied to spectral analysis of time series.

We will treat an array of random variables \( \{X_{n,i} : n, i = 1, 2, \ldots \} \). Each row \( \{X_{n,i}, i = 1, 2, \ldots \} \) is a mixingale in itself, i.e. the size of the conditional expectation of each \( X_{n,i} \) with
respect to its distant predecessors on the same row is restricted. We shall assume these restrictions are uniform for all rows. In the application of these results for our purpose, each row is an approximation of a target process, and they should come closer and closer to the target process in some sense as \( n \) increases. Hence the uniform restriction should not be a rigid one. Moreover, the conditions for these invariance principles are minimal. For \( \phi \)-mixing and \( \alpha \)-mixing arrays, only a summability condition on the mixing coefficients is required, and no stationarity is assumed, although the growth of the variance of partial sums is restrained. This actually improves Theorem 18.5.2 and 18.5.3 of Ibragimov & Linnik [10]. For linear processes, our condition is comparable to the summability of the coefficients, and is weaker than conditions for the whole process to be strong mixing given in Withers [13].

2 Invariance Principles for Triangular Arrays of Dependent Variables

Central limit theorems and invariance principles for dependent variables are available in the literature for mainly two kinds of dependence: martingale difference sequences and mixing variables. See Hall and Heyde (1980) [9] and Philipp (1986) [12] for a survey of results. Also see Ibragimov and Linnik (1971). While central limit theorems for martingale arrays are readily available for triangular arrays, there is a lack of similar results for mixing arrays in general. Central limit theorems for mixing sequences are usually proved by dividing the sequence into alternating blocks. The sums of odd-numbered blocks are nearly independent, and the sums of even-numbered blocks are negligible in probability.

McLeish (1975) [11] proposed the concept of mixingale and gave a theorem for which martingales and mixing processes both become special cases. In addition, linear processes are mixingales, but fail to be either martingales or mixing sequences (Withers (1981) [13]) if no further conditions are imposed. Moreover, McLeish's theorem doesn't assume the underlying process to be stationary, which makes it applicable in more general situations.

Verification of the validity of bootstrap and other resampling techniques often calls for a central limit theorem for triangular arrays of random variables. Instead of a central limit theorem, we prove an invariance principle, which is stronger. By appealing to theorem 19.4
of Billingsley [1], which employs the characterization of Brownian motion and other diffusion processes, we avoid going through a local central limit theorem.

A mixingale is a sequence of weakly dependent random variables, the size of whose conditional expectations with respect to distant predecessors are being restricted. To fix ideas, suppose we have a probability triple $(\Omega, \mathcal{F}, P)$. Let $\{\mathcal{F}^i, -\infty < i < \infty\}$ be a nondecreasing sequence of sub-$\sigma$-algebras of $\mathcal{F}$, $\mathcal{F}^{-\infty}$ the largest $\sigma$-algebra contained in all $\mathcal{F}^i$ and $\mathcal{F}^{\infty}$ the smallest $\sigma$-algebra which contains all $\mathcal{F}^i$. Denote $E_k U := E(U|\mathcal{F}^k)$. The following definition is taken from McLeish [11], $\| \cdot \|_2$ is the $L_2$ norm.

**Definition 2.1** The sequence $\{(X_n, \mathcal{F}^n)\}$ is a mixingale if there exists a nonincreasing positive sequence $\psi_k \to 0$ as $k \to \infty$ such that for all $i \geq 1, k \geq 0$,

$$\|E_{i-k}X_i\|_2 \leq \psi_k, \quad (1)$$

and

$$\|X_i - E_{i+k}X_i\|_2 \leq \psi_{k+1}. \quad (2)$$

Note $X_n$ need not be adapted to $\mathcal{F}^n$. If $\{(X_n, \mathcal{F}^n)\}$ is a mixingale, then $E_{-\infty}X_n = 0$ and $E_\infty X_n = X_n$ a.s. and if $X_n$ are adapted to $\mathcal{F}^n$, then inequalities (1) and (2) restrict the degree of dependence among $X_n$'s. In order that an invariance principle for $\{X_n\}$ holds, the controlling sequence $\psi_k$ has to decay to zero sufficiently fast. This is described in the following definition.

**Definition 2.2** (McLeish [11]) Call the sequence $\{\psi_k\}$ of size $-p$ if there exists a positive sequence $\{L(k)\}$ such that

(a) $\sum (nL(n))^{-1} < \infty$,

(b) $L(n) - L(n - 1) = O(L(n)/n)$,

(c) $L(n)$ is eventually non-decreasing and

(d) $\psi_n = o([n^{1/2}L(n)]^{-2p})$.

The definition of a sequence being of size $-p$ is closely related to regularly varying sequence of exponent $-p$, see Feller [8] p.275-276. $L(n) = (\log n)^\alpha$ for some $\alpha > 1$ would satisfy conditions (a)-(c) in the above definition.
We now give McLeish’s invariance principle for mixingales. Let \( \{X_i, i = 1, 2, 3, \ldots\} \) be a sequence of random variables in \( L_2(\Omega, \mathcal{F}, P) \) and set \( S_n = \sum_1^n X_i \). Suppose that

\[
E X_i = 0 \quad \text{for all } i, \quad \text{and} \quad \frac{E S_n^2}{n} \to \sigma^2 \quad \text{for some positive constant } \sigma^2. \tag{3}
\]

Define a random function on \( (\Omega, \mathcal{F}, P) \) by

\[
W_n(t) = \frac{S_{[nt]}}{\sqrt{nt}}, \tag{4}
\]

Then, the following theorem holds.

**Proposition 2.1** Let \( \{(X_i, \mathcal{F})\} \) be a mixingale satisfying (3) with \( \psi_k \) of size \(-1/2\). If \( \{X_i^2, i = 1, 2, \ldots\} \) is uniformly integrable, and if

\[
E \left\{ \frac{(S_{k+n} - S_k)^2}{n} \right\}^{\mathcal{F}^{k-m}} \to \sigma^2 \quad \text{in } L_1(\Omega) \text{ norm} \tag{5}
\]

as \( \min(m, k, n) \to \infty \), then \( \{W_n\} \) weakly converges to \( W \), a standard Brownian motion process on \( D = D[0, 1] \), the set of all functions on the interval \([0, 1]\) that are right continuous and have left hand limits at every point.

Here we conceive \( D \) to be endowed with the Skorohod metric, and the weak convergence is the weak convergence in a metric space, c.f. Billingsley [1].

According to McLeish [11], uniform integrability of the squared sequence and the size of \( \psi_k \) imply the tightness of \( W_n \), and the last condition is all that is additionally needed for the weak convergence. Also, \( \psi_k \) is of size \(-1/2\) if \( \psi_k = o(k^{-\epsilon}) \) for some \( \epsilon > 1/2 \). Condition (5) involves conditional expectation of partial sums with respect to the distant past, which basically suggests stability in a mean square sense. On occasions unconditional expectation would be adequate in (5) (see the next section, note \( E_{i-k} X_i \to E X_i \) as \( k \to \infty \)). As we have mentioned, the above invariance principle is applicable to mixing and linear processes with minimal assumptions. We will touch on this subject in the next section. But first, we shall extend this result to triangular arrays of dependent variables, again in a very general setting.

Now suppose we have a sequence of probability triples \( (\Omega_n, \mathcal{F}^n, P_n) \), each of which supports a sequence of mixingales \( \{(X_{n,j}, \mathcal{F}^{n,j})\}, j = 1, 2, \ldots\) \}. Denote \( E_{i-k} U = E(U|\mathcal{F}^{n,i-k}) \) if \( U \)
is measurable \((\Omega_n, \mathcal{F}^n, P_n)\). Suppose the array of mixingales \(\{(X_{n,i}, \mathcal{F}^n)\}\) satisfies
\[
||E_{i-k}X_{n,i}||_2 \leq \psi_k,
\]
and
\[
||X_{n,i} - E_{i+k}X_{n,i}||_2 \leq \psi_{k+1}
\]
for some nonincreasing sequence \(\psi_k\), for all \(i \geq 1, k \geq 0\), and all \(n = 1, 2, \ldots\) Assume
\[
EX_{n,j} = 0, \quad E\frac{S_{n,n}^2}{n} \rightarrow \sigma^2
\]
Let \(S_{n,k} = \sum_{j=1}^{k} X_{n,j}\), and
\[
W_n(t) = \frac{S_{n,\lfloor nt \rfloor}}{\sqrt{n}\sigma}.
\]
We have the following invariance principle for triangular array of mixingales:

**Theorem 2.1** Let \(\{(X_{n,j}, \mathcal{F}^n)\}\) be an array of mixingales satisfying (6) and (7) with \(\{\psi_k\}\) of size -1/2. If \(\{X_{n,j}, n, j = 1, 2, \ldots\}\) is uniformly integrable, then \(\{W_n\}\) is tight. If in addition that
\[
\left\| E_{k-m} \left[ \frac{(S_{n,k+d} - S_{n,k})^2}{d} \right] - \sigma^2 \right\|_1 \rightarrow 0,
\]
as \(\min(m, k, n, d) \rightarrow \infty\), then \(W_n\) converges weakly to \(W\) on \(D[0, 1]\).

**Proof.** See Appendix.

Again, \(D\) is endowed with the Skorohod metric.

In (6) and (7) we placed a uniform bound over all \(n, i\). (9) is a generalization of (5).

In the analysis of dependent variables, e.g. time series analysis, we are interested in the statistical properties of the process that generates the observations \(X_1, X_2, \ldots, X_n\). Since the problem is often of a complicated nature, our approach will be to use the bootstrap or other resampling schemes to approximate the behavior of \(\{X_i\}\) and reduce a complicated distribution issue to a simpler one by doing the analysis for the resampled series. We would like the array \(\{X_{n,i}\}\) generated by the resampling plan to be close to \(\{X_i\}\) (in a distributional way) asymptotically. Theorem 2.1 assists in justifying this.
3 Invariance Principle for Arrays of Mixing and Linear Processes

The concept of mixingale provides a unified approach to general dependent variables, including martingale difference processes, mixing processes, and linear processes. In this section, we shall engage in deducing invariance principles for mixing and linear processes as applications of Theorem 2.1.

3.1 Mixing

We first give invariance principles under strong (\(\alpha\)-) and uniform (\(\phi\)-) mixing conditions. Mixing coefficients are measures of dependence between \(\sigma\)-algebras. The uniform mixing (or \(\phi\)-mixing) coefficient and strong mixing (or \(\alpha\)-mixing) coefficient of two \(\sigma\)-algebras \(\mathcal{F}\) and \(\mathcal{U}\) are defined by

\[
\phi(\mathcal{F}, \mathcal{U}) = \sup\{|\Pr(G|F) - \Pr(G)| : F \in \mathcal{F}, G \in \mathcal{U}, \Pr(F) > 0\},
\]

and

\[
\alpha(\mathcal{F}, \mathcal{U}) = \sup\{|\Pr(FG) - \Pr(F)\Pr(G)| : F \in \mathcal{F}, G \in \mathcal{U}\}.
\]

We consider an array of doubly infinite sequences of random variables \(\{X_{n,i} : -\infty < i < \infty, n = 1, 2, \ldots\}\), with \(\{X_{n,i} : -\infty < i < \infty\}\) defined on \((\Omega_n, \mathcal{F}^n, P_n)\). Put \(\mathcal{F}^m_{n,k} = \sigma(X_{n,i} : k \leq i \leq m), R^m_{n,k} = \sigma(S_{n, m} - S_{n, k-1})\) for all \(k \leq m\). For each \(m \geq 0\), define the mixing coefficients as

\[
\hat{\phi}_{n,m} = \sup_k \phi(\mathcal{F}^k_{n,-\infty, \mathcal{F}^\infty_{n,k+m}}),
\]

\[
\hat{\alpha}_{n,m} = \sup_k \alpha(\mathcal{F}^k_{n,-\infty, \mathcal{F}^\infty_{n,k+m}}),
\]

\[
\phi_{n,m} = \sup_k \sup_{j \geq k+m} \phi(\mathcal{F}^k_{n,-\infty, R^j_{n,k+m}}),
\]

\[
\alpha_{n,m} = \sup_k \sup_{j \geq k+m} \alpha(\mathcal{F}^k_{n,-\infty, R^j_{n,k+m}}).
\]

Observe that \(\{X_{n,i} : -\infty < i < \infty\}\) is \(\phi\)-mixing if \(\hat{\phi}_{n,m} \to 0\) as \(m \to \infty\) and is strong mixing if \(\hat{\alpha}_{n,m} \to 0\) as \(m \to \infty\). In general it is true that \(\hat{\alpha}_{n,m} \leq \hat{\phi}_{n,m}\) and \(\alpha_{n,m} \leq \phi_{n,m} \leq \hat{\phi}_{n,m}\). We will treat only the two weaker conditions on \(\phi_{n,m}\) and \(\alpha_{n,m}\)'s.
Two σ-fields $\mathcal{F}$ and $\mathcal{U}$ are independent if and only if $\alpha(\mathcal{F},\mathcal{U}) = 0$ $(\phi(\mathcal{F},\mathcal{U}) = 0)$. If $\mathcal{F}$ and $\mathcal{U}$ are independent, and if $X$ is measurable $\mathcal{F}$ and $Y$ is measurable $\mathcal{U}$, and they both have finite moments, then $E(XY) - (EX)(EY) = 0$, and $E(X|\mathcal{U}) - EX = 0$. Nevertheless, if $\mathcal{F}$ and $\mathcal{U}$ are not independent, but $\alpha(\mathcal{F},\mathcal{U})$ $(\phi(\mathcal{F},\mathcal{U}))$ is small, then the discrepancies $E(XY) - (EX)(EY)$ and $E(X|\mathcal{U}) - EX$ are expected to be small compared to the moments (possibly higher moments) of $X$ and $Y$ (see Ibragimov and Linnik [10], chapter 17). To convince ourselves that mixing processes are mixingales under moment assumptions, we shall be interested in bounding the latter difference by mixing coefficients and moments. The following lemma is Lemma 3.5 of [11]. It contains exactly the inequalities appropriate for our purpose.

**Lemma 3.1** (McLeish [11]) *Suppose $X$ is a random variable measurable with respect to $\mathcal{U}$, and $1 \leq p \leq r \leq \infty$. Then*

\[
\|E(X|\mathcal{F}) - EX\|_p \leq 2[\phi(\mathcal{F},\mathcal{U})]^{1-1/r}\|X\|_r, \tag{10}
\]

*and*

\[
\|E(X|\mathcal{F}) - EX\|_p \leq 2(2^{1/p} + 1)[\alpha(\mathcal{F},\mathcal{U})]^{1/p-1/r}\|X\|_r. \tag{11}
\]

Now we shall apply Theorem 2.1 to mixing arrays. The case of an infinite sequence (not triangular array) is covered in McLeish [11]. The next theorem is an extension of Theorem 3.8 of [11]. We will place a uniform bound on each $\{\phi_{n,k} : n = 1, 2, \ldots\}$ and $\{\alpha_{n,k} : n = 1, 2, \ldots\}$, as we did in the uniform mixingale conditions (6) and (7) in Theorem 2.1. The proof of this theorem is almost identical to that of Theorem 3.8 of [11], and thus will be omitted.

**Theorem 3.1** *Let $\{X_{n,i} : -\infty < i < \infty\}$ be mixing with mixing coefficients $\phi_{n,k} \leq \phi_k$ and $\alpha_{n,k} \leq \alpha_k$ for all $n=1,2,\ldots$, and suppose $EX_{n,i} = 0, \forall n, i$, $E(S_{n,n}^2/n) \to \sigma^2$, and*

(a) $\{X_{n,i}^2 : n, i\}$ is uniformly integrable,
(b) $\sup_{n,i} \|X_{n,i}\|_{\beta} < \infty$, for some $2 \leq \beta \leq \infty$,
(c) $E((S_{n,k+m} - S_{n,k})^2/m) \to \sigma^2 > 0$ as $\min(n,k,m) \to \infty$,
(d) $\{\phi_k\}$ is of size $\beta/(2\beta - 2)$, or
(d') $\beta > 2$ and $\{\alpha_k\}$ is of size $\beta/(\beta - 2)$,

*then $W_n$ converges weakly to $W$ on $D$.**
Remark 3.1 (a) is implied by (b) when $\beta > 2$. If $\phi_k(\alpha_k)$ is decreasing, a sufficient condition for (d)((d')) is

$$\sum \phi_k^{1-1/\beta} < \infty \quad (\text{or} \quad \sum \alpha_k^{1-2/\beta} < \infty).$$

Taking $\beta = 2$ for $\phi_k$ and $\beta > 2$ for $\alpha_k$, we can link these two inequalities to the conditions of Ibragimov & Linnik [10], Theorem 18.5.2 and 18.5.3, which are among the weakest one can achieve.

If $\{X_{n,i}\}$ is an array of stationary processes, we have the following corollary.

**Corollary 3.1** Let $\{X_{n,i} : -\infty < i < \infty\}$ be stationary with mean 0. Suppose its mixing coefficients satisfy $\phi_{n,k} \leq \phi_k$ and $\alpha_{n,k} \leq \alpha_k$ for all $n=1,2,...$, and

(a) $\{X_{n,1}^2 : n\}$ is uniformly integrable,

(b) $\sup_n \|X_{n,1}\|_\beta < \infty$, for some $2 \leq \beta \leq \infty$,

(c) $E S_{n,m}^2 / m \to \sigma^2$ as $\min(n,m) \to \infty$,

(d) $\sum \phi_k^{1-1/\beta} < \infty$, or

(d') $\beta > 2$ and $\sum \alpha_k^{1-2/\beta} < \infty$,

then if $\sigma^2 > 0$, $W_n$ converges weakly to $W$ on $D$.

If $\{X_{n,i}\}$ are approximations of a stationary process $\{X_i\}$, and are in themselves stationary, the conditions in the last corollary will be true if the mixing coefficients of $\{X_{n,i}\}$ are dominated and converge to those of $\{X_i\}$, and if the dominating sequence satisfy d(d').

Condition (c) is implied by the mixing conditions.

In the case when $\sigma^2 = 0$, $S_{n,|nt|}/\sqrt{n}$ converges in probability to the zero function.

### 3.2 Linear process

Consider the linear process

$$X_{n,i} = \sum_{-\infty}^{\infty} \theta_{n,j} \xi_{n,i-j}$$

(12)

where $\xi_{n,j}, j = 0, \pm 1, \pm 2, ...$ are martingale differences with mean 0 and variance $\sigma_n^2$. Let $F_{n,i} = \sigma(\xi_{n,j} : j \leq i)$. Suppose $|\theta_{n,j}| \leq b_j$ and $\sigma_n^2 \leq a$ for all $n = 1, 2, ...$, then
Lemma 3.2 $\{(X_{n,i}, \mathcal{F}^{n,i})\}$ is a mixingale array, with uniform bound as in (6) and (7) given by $\psi_k^2 = a \sum_{|j| \geq k} b_j^2$.

$\psi_k$ is of size $-1/2$ if $\sum_{|j| \geq k} b_j^2 = O(k^{-\alpha})$ for some $\alpha > 1$. This last condition is in turn guaranteed if, for example, $b_j = O(|j|^{-\alpha})$.

Theorem 3.2 Let $\{X_{n,i}\}$ be a linear process as in (12). Suppose there exist positive real numbers $a$ and $b_j$, $j = 1, 2, ...$ such that

$$\sup_n \sigma_n^2 \leq a, \quad \sup_n |\theta_{n,j}| \leq b_j, j = 1, 2, ...$$

If

(a) $\sum_{|j| \geq k} b_j^2 = O(k^{-\alpha})$ for some $\alpha > 1$, and
(b) there exists the limit $\lim_{n \to \infty} \sigma_n^2 (\sum_j \theta_{n,j})^2 = \sigma^2 > 0$,

then $W_n$ converges weakly to $W$ on $D$.

To see how weak this condition is, imagine $b_j$ is related to a stationary linear process

$$X_i = \sum_{j=-\infty}^{\infty} b_j \xi_{i-j},$$

(13)

where $\xi_i$ are i.i.d. with mean 0 and variance 1. A central limit theorem for the sample mean of $\{X_n\}$, such as Theorem 7.1.2 of Brockwell & Davis [4] would require $\sum |b_j| < \infty$. As we have mentioned, if $b_j$ is nonincreasing in $|j|$, this leads to $b_j = o(|j|^{-1})$. Our convergence rate is slightly stronger than this. But if (13) is the AR($\infty$) representation of a causal, invertible stationary ARMA process, the most interesting linear process in practice, $b_j$ should dampen to 0 exponentially, and the polynomial rate trivially ensues. Withers [13] gives sufficient conditions for a linear process to be strong mixing. He also assumes the innovations to be independent. Corollary 2 and Corollary 1 of [13] require the coefficient to go to 0 at an exponential rate or a polynomial rate of order $> 2$. Our results are far more general.

4 Central Limit Theorems for Autocovariance Functions

Autocovariance and autocorrelation functions are of great interest in time series analysis. For example, the behavior of ACF helps one to identify a model to explain the data. Bartlett
(1955) showed the asymptotic normality of autocovariance estimates for stationary linear processes, which leads to asymptotic tests of a white noise series based on those estimates (see Box, Jenkins and Reinsel(1994) [3]). Tests based on spectral estimates are also related to the ACF’s. In this section, we shall develop central limit theorems for autocovariances under general conditions. We shall treat triangular arrays, leaving the classical central limit theorems as special cases.

Let \( \{X_{n,i}\} \) be an array of stationary dependent random variables. The autocovariance function of the nth row is \( R_n(k), k = 0, \pm 1, \ldots \). Fix an integer \( u \geq 0 \), let \( Y_{n,i} = X_{n,i}X_{n,i+u} - R_n(u) \), and \( T_{n,k} = \sum_{i=1}^{k} Y_{n,i} \). Also put \( V_n(t) = T_{n,|nt|}/n^{1/2} \). The sample autocovariance function is given by \( c_n(u) = n^{-1} \sum_{i=1}^{n} Y_{n,i} \).

4.1 Mixing

For a stationary mixing sequence \( \{X_{n,i} : i = 1, 2, \ldots\}, \{Y_{n,i} : i = 1, 2, \ldots\} \) is also mixing, with mixing coefficients being the translation of those of the original sequence by a lag \( u \). The conditions in Corollary 3.1 need only be revised by placing a stronger moment bound on \( X_{n,i} \) to ensure the corresponding bounds on \( Y_{n,i} \). An application of the same corollary leads to the following.

**Theorem 4.1** Let \( \{X_{n,i} : -\infty < i < \infty\} \) be stationary with mean 0. Suppose its mixing coefficients satisfy \( \phi_{n,k} \leq \phi_k \) and \( \alpha_{n,k} \leq \alpha_k \) for all \( n=1,2,\ldots \), and

(a) \( \{X_{n,1}^{4}\} \) is uniformly integrable,
(b) \( \sup_n \|X_{n,1}\|_{2\beta} < \infty \), for some \( 2 \leq \beta \leq \infty \),
(c) \( ET_{n,m}^{2}/m \to \tau^2 \) as \( \min(m,n) \to \infty \),
(d) \( \sum \phi_k^{1-1/\beta} < \infty \), or
(d') \( \beta > 2 \) and \( \sum \alpha_k^{1-2/\beta} < \infty \).

Then if \( \tau^2 > 0 \), \( V_n/\tau \) converges weakly to \( W \) on \( D \). In particular, \( \sqrt{n}(c_n(u) - R_n(u)) \to N(0, \tau^2) \).

Sometimes tapering is preferred before any statistical analysis is performed, which has an effect of reducing leakage when the sample ACF’s are used to estimate the spectrum (see Bloomfield [2]). Evidence has shown only a rescaling is needed in the asymptotic distribution.
(see Dahlhaus 1983 [5]). We shall consider tapering according to a function of bounded variation. Let \( h : [0, 1] \to [0, \infty) \) be a function of bounded variation. The ACF estimate at lag \( u \) is given by

\[
\hat{c}_n(u) = \frac{\sum_{i=1}^{n} h \left( \frac{i}{n} \right) h \left( \frac{i+u}{n} \right) X_{n,i} X_{n,i+u}}{\sum_{i=1}^{n} h \left( \frac{i}{n} \right) h \left( \frac{i+u}{n} \right)}.
\]

A simple application of the continuous mapping theorem and the invariance principle in the last theorem will yield the same limit distribution for \( \hat{c}_n(u) \). For a proof, see Theorem 2.2 of Dahlhaus [6].

**Corollary 4.1** Assume all conditions in Theorem 4.1 are met, and that \( h \) is of bounded variation. Then

\[
\sqrt{n}(\hat{c}_n(u) - R_n(u)) \to N \left( 0, \tau^2 \right).
\]

A joint distributional result can be obtained by employing the Cramér-Wold device:

**Corollary 4.2** Assume all conditions in Theorem 4.1 are met, and that \( h \) is of bounded variation. Then for any \( u_1, \ldots, u_t \),

\[
\sqrt{n}(\hat{c}_n(u_j) - R_n(u_j))_{j=1}^t
\]

converges weakly to a multivariate normal random variable with covariances equal to

\[
\lim_{n \to \infty} n \text{Cov} \left( \hat{c}_n(u_j) - R_n(u_j), \hat{c}_n(u_j') - R_n(u_j') \right).
\]

### 4.2 Linear process

Consider the linear process (12). Suppose \( \xi_{n,i} \) are i.i.d. with mean 0, variance \( \sigma_n^2 \), and fourth moment \( \gamma_n \). Again, we assume the uniform bound \( \sigma_n^2 \leq a \) and \( |\theta_{n,j}| \leq b_j \), and \( \gamma_n \leq c \) for all \( n = 1, 2, \ldots \). For this process, \( \{Y_{n,i}\} \) forms a mixingale array if some summability conditions are satisfied by \( \{b_j\} \):

**Lemma 4.1** If \( X_{n,i} \) is a linear process as described above, and if \( \sum_j b_j^2 < \infty \), then \( \{Y_{n,i}, \mathcal{F}^{n,j}\} \) is a mixingale array satisfying (6) and (7) with \( \psi_k = C \sum_{|j| \geq k-u} b_j^2 \), for some finite constant \( C \).
Additional assumptions, including $\sum_{|j|\geq k-u} b_j^2 = O(k^{-\alpha})$ for some $\alpha > 1$, are to be imposed for stability in a mean square sense. This is the content of the next theorem.

**Theorem 4.2** Let $\{X_{n,i}\}$ be the linear process (12), where $\xi_{n,i}, i = 1, 2, \ldots$ are martingale differences with mean 0, variance $\sigma_n^2$, and $E\xi_{n,i}^3 = 0$, $E\xi_{n,i}^4 = \gamma_n$. Suppose

$$
sup_n \sigma_n^2 \leq a, \quad sup_n \gamma_n \leq c \quad and \quad sup_n |\theta_{n,j}| \leq b_j, j = 1, 2, \ldots
$$

for some constants $a, c,$ and $b_j, j = 1, 2, \ldots$. If

(a) $\sum_{|j|\geq k} b_j^2 = O(k^{-\alpha})$ for some $\alpha > 1$, and

(b) there exists the limit

$$
\lim_{n \to \infty} (\gamma_n - 3\sigma_n^4) R_n(u) + \sum_i \left(R_n^2(i) + R_n(i-u)R_n(i+u)\right) = \tau^2 > 0,
$$

then $V_n/\tau$ converges weakly to $W$ on $D$.

With finite fourth moments, the summability condition (a) is comparable to Theorem 7.2.1 of Brockwell and Davis [4]. Again, we don’t need independence of $\xi_{n,i}$. See the discussion in the last section. A corollary for the joint distribution is ready to follow.

A Technical Proofs for Part II

**Proof of Theorem 2.1.** As in McLeish’s proof of Proposition 2.1 (Theorem 2.6 of [11]), we shall make use of Theorem 19.4 of Billingsley [1], which involves no local central limit theorems. By Theorem 8.4 of the same book, tightness of $\{W_n\}$ is equivalent to: $\forall \epsilon > 0$, $\exists \lambda > 1$ and $m_0$, $n_0$, such that

$$
Pr \left\{ \max_{i \leq m} |S_{n,k+i} - S_{n,k}| \geq \lambda \sigma \sqrt{m} \right\} \leq \epsilon / \lambda^2,
$$

$\forall m \geq m_0$, $n \geq n_0$, and $k \leq n$. A sufficient condition for this is:

\[
\left\{ \max_{i \leq m} \left( S_{n,k+i} - S_{n,k} \right)^2 / m : m, n, k \right\} \text{ is u.i.} \tag{14}
\]

By the assumption that $\{X_{n,i}^2 : n, i\}$ is u.i. and that $\psi_k$ is of size -1/2, we can show (14). The proof is a duplicate of Lemma 6.5 of McLeish [11]. The justification hinges on the
fact that the conditional expectations in (6) and (7) have uniform bounds \( \psi_k \), so that the maximal inequality (Lemma 6.2 of McLeish [11]) holds uniformly for all rows. The remaining conditions of Theorem 19.4 of [1] are checked just as in the proof of Theorem 2.6 of [11]. □

Proof of Lemma 3.2. Since \( E_{i-k} \xi_{n,i-j} = \xi_{n,i-j} \) if \( j \geq k \) and 0 otherwise for \( k \geq 0 \), it follows that

\[
E_{i-k}X_{n,i} = \sum_j \theta_{n,j} E_{i-k} \xi_{n,i-j} = \sum_{j \geq k} \theta_{n,j} \xi_{n,i-j}.
\]

So \( \|E_{i-k}X_{n,i}\|^2 \leq a \sum_{|j| \geq k} b_j^2 \). The corresponding bound for \( \|X_{n,i} - E_{i+k}X_{n,i}\|^2 \) can be derived in a similar manner. □

Proof of Theorem 3.2. It’s easy to see that (b) implies \( ES_{n,n}^2/n \to \sigma^2 \). So, we need only to show (9).

By stationarity, an equivalent condition is

\[
\left\| E_{-m} \frac{S_{n,d}}{d} - \sigma^2 \right\|_1 \to 0 \quad \text{as} \quad \min(m,n,d) \to \infty. \tag{15}
\]

With

\[
E_r(\xi_{n,l} \xi_{n,l'}) = \begin{cases} 
0, & \text{if } l \neq l' \text{ and } \max(l,l') > r \\
\sigma_n^2, & \text{if } l = l' > r \\
\xi_{n,l} \xi_{n,l'}, & \text{if } \max(l,l') \leq r
\end{cases}
\]

we have

\[
E_{-m}X_{n,i}X_{n,i'} = \sigma_n^2 \sum_{j > -m} \theta_{n,i-j} \theta_{n,i'-j} + \left( \sum_{j \leq -m} \theta_{n,i-j} \xi_{n,j} \right) \left( \sum_{|j'| \leq -m} \theta_{n,i'-j'} \xi_{n,j'} \right). \tag{16}
\]

Therefore

\[
E_{-m}(S_{n,d}^2/d) = I + II,
\]

where

\[
I = \sigma_n^2 \sum_{a > -m} d^{-1} \sum_{i=1}^d \sum_{i'=1}^d \theta_{n,i-a} \theta_{n,i'-a},
\]

\[
II = d^{-1} \left( \sum_{i=1}^d \sum_{j \leq -m} \theta_{n,i-j} \xi_{n,j} \right)^2.
\]
$I$ is approximately

$$
\sigma_n^2 \sum_{a=\infty}^{\infty} \frac{1}{d} \sum_{i=1}^{d} \sum_{i'=1}^{d} \theta_{n,i-a} \theta_{n,i'-a}
= \sigma_n^2 \sum_{j=-\infty}^{\infty} \sum_{\|l\|<d} \left(1 - \frac{|l|}{d}\right) \theta_{n,j+\theta_{n,j+l}}
\rightarrow \lim_{n \to \infty} \sigma_n^2 \left(\sum_{j=-\infty}^{\infty} \theta_{n,j}\right)^2
$$

as $\min(n, m, d) \to \infty$. Further, we have

$$
\|II\|_1 \leq \sum_{i=1}^{d} \left(\sum_{j=i+m}^{\infty} \theta_{n,j} \xi_{n,i-j}\right)^2.
$$

Let $B_k = \sum_{|l| \geq k} b_j^2$. Condition(a) implies $\sum_k B_k < \infty$, so that

$$
\|II\|_1 \leq \sum_{i=1}^{d} \left(\sum_{j=i+m}^{\infty} \theta_{n,j}^2 \right)
= O \left(\sum_{i=m+1}^{\infty} B_{i+m}\right) \to 0
$$

as $\min(n, m, d) \to \infty$. We thus have proved (15).

Proof of Lemma 4.1. From (16), we obtain the conditional expectations of $X_{n,i}X_{n,i+u}$:

$$
E_{i-k}X_{n,i}X_{n,i+u} = \sigma_n^2 \sum_{j<k} \theta_{n,j} \theta_{n,j+u} + \left(\sum_{j \geq k} \theta_{n,j} \xi_{n,i-j}\right) \left(\sum_{j \geq k+u} \theta_{n,j} \xi_{n,i+u-j}\right),
$$

$$
E_{i+k}X_{n,i}X_{n,i+u} = \sigma_n^2 \sum_{j<-k} \theta_{n,j} \theta_{n,j+u} + \left(\sum_{j \geq -k} \theta_{n,j} \xi_{n,i-j}\right) \left(\sum_{j \geq -k+u} \theta_{n,j} \xi_{n,i+u-j}\right).
$$

Therefore, noting that the acf for the nth row is $R_n(u) = \sigma_n^2 \sum_j \theta_{n,j} \theta_{n,j+u}$, the conditional expectations for $Y_{n,i}$ are

$$
E_{i-k}Y_{n,i} = -\sigma_n^2 \sum_{j \geq k} \theta_{n,j} \theta_{n,j+u} + \left(\sum_{j \geq k} \theta_{n,j} \xi_{n,i-j}\right) \left(\sum_{j \geq k+u} \theta_{n,j} \xi_{n,i+u-j}\right),
$$

$$
Y_{n,i} - E_{i+k}Y_{n,i} = X_{n,i}X_{n,i+u} - E_{i+k}X_{n,i}X_{n,i+u}
= \left(\sum_{j<-k} \theta_{n,j} \xi_{n,i-j}\right) \left(\sum_{j=-\infty}^{\infty} \theta_{n,j} \xi_{n,i+u-j}\right).
$$
\[
+ \left( \sum_{j \geq -k} \theta_{n,j} \xi_{n,i-j} \right) \left( \sum_{j < -k + u} \theta_{n,j} \xi_{n,i+u-j} \right) 
\]

\[
- \sigma_n^2 \left( \sum_{j < -k} \theta_{n,j} \theta_{n,j+u} \right). 
\]

Hölder's inequality implies \( \|XY\|_2 \leq \|X\|_4 \|Y\|_4 \). Thus with the uniform bounds on the moments and coefficients, the \( L_2 \) norms of the above are bounded by

\[
\| E_{i-k} Y_{n,i} \|_2 \leq 2 \sigma_n^2 \sum_{j \geq k} \theta_{n,j}^2 + \left\| \sum_{j \geq k} \theta_{n,j} \xi_{n,i-j} \right\|_4 \left\| \sum_{j \geq k+u} \theta_{n,j} \xi_{n,i-j} \right\|_4 
\]

\[
\leq 2 \sigma_n^2 \sum_{j \geq k} \theta_{n,j}^2 + \left( \gamma_n - 3 \sigma_n^4 \right) \sum_{j \geq k} \theta_{n,j}^4 + 3 \sigma_n^4 \left( \sum_{j \geq k} \theta_{n,j}^2 \right)^{1/2} 
\]

\[
\leq 4a \sum_{j \geq k} b_j^2 + c^{1/2} \left( \sum_{j \geq k} b_j^4 \right)^{1/2} 
\]

\[
\leq (4a + c^{1/2}) \sum_{j \geq k} b_j^2, 
\]

\[
\| Y_{n,i} - E_{i+k} Y_{n,i} \|_2 \leq \left\| \sum_{j < -k} \theta_{n,j} \xi_{n,i-j} \right\|_4 \left\| \sum_{j \geq k} \theta_{n,j} \xi_{n,i-j} \right\|_4 
\]

\[
+ \left\| \sum_{j < -k + u} \theta_{n,j} \xi_{n,i+u-j} \right\|_4 \left\| \sum_{j \geq -k} \theta_{n,j} \xi_{n,i-j} \right\|_4 
\]

\[
+ \sigma_n^2 \left\| \sum_{j < -k} \theta_{n,j} \theta_{n,j+u} \right\|_2 
\]

\[
\leq 2 \left\| \sum_{j} \theta_{n,j} \xi_{n,i-j} \right\|_4 \left\| \sum_{j \geq k+u} \theta_{n,j} \xi_{n,i-j} \right\|_4 + \sum_{j < -k + u} b_j^2 
\]

\[
\leq C \sum_{|j| > k-\mu} b_j^2. 
\]

C can be chosen to be, for example,

\[
C = 2(4a + c^{1/2}) \min \left( \sup_n \left\| \sum_{j} \theta_{n,j} \xi_{n,i-j} \right\|_4, 1 \right) + 1. 
\]

Hence \( \psi_k = C \sum_{|j| \geq k-\mu} b_j^2 \) can play the role of the uniform bound in (6) and (7). \( \square \)

**Proof of Theorem 4.2.** Because of condition (a), \( \psi_k \) is of size \(-1/2\). We still need conditions for \( \|E_m(T^{2}_{n,d}/d) - \tau^2\|_1 \to 0 \) as \( \min(m, n, d) \to \infty \). First, with the assumption \( E \xi_{n,j}^3 = 0 \) for
all $n, j$, if
\[
\lim_{n \to \infty} (\gamma_n - 3\sigma_n^4) R_n^2(u) + \sigma_n^4 \sum_i \left( R_n^2(i) + R_n(i - u) R_n(i + u) \right)
\]
exists, then $\tau^2$ is well defined and is equal to the above limit. Proceed as in the proof of Theorem 3.2. First of all, we have for $l_1 \leq l_2 \leq l_3 \leq l_4$ and $r,$
\[
E_r (\xi_{n, l_1}, \xi_{n, l_2}, \xi_{n, l_3}, \xi_{n, l_4}) = \begin{cases} 
\xi_{n, l_1}, \xi_{n, l_2}, \xi_{n, l_3}, \xi_{n, l_4}, & \text{if } l_4 \leq r, \\
\sigma_n^2 \xi_{n, l_1} \xi_{n, l_2}, & \text{if } l_2 \leq r < l_3 = l_4, \\
\sigma_n^4, & \text{if } r < l_1 = l_2 < l_3 = l_4, \\
\gamma_n, & \text{if } r < l_1 = l_2 = l_3 = l_4, \\
0, & \text{otherwise.}
\end{cases}
\]
From this we deduce
\[
E_{-m} X_{n,i'}X_{n,i+u}X_{n,i'}X_{n,i'+u} = \sum_{j_1} \sum_{j_2} \sum_{j_3} \sum_{j_4} \theta_{n,i-j_1} \theta_{n,i+u-j_2} \theta_{n,i'-j_3} \theta_{n,i'+u-j_4} E_{-m} \xi_{n,j_1} \xi_{n,j_2} \xi_{n,j_3} \xi_{n,j_4}
= I + II + III + IV,
\]
where
\[
I = \left( \sum_{j \leq -m} \theta_{n,i-j} \xi_{n,j} \right) \left( \sum_{j \leq -m} \theta_{n,i+u-j} \xi_{n,j} \right) \left( \sum_{j \leq -m} \theta_{n,i'-j} \xi_{n,j} \right) \left( \sum_{j \leq -m} \theta_{n,i'+u-j} \xi_{n,j} \right),
\]
\[
II = (\gamma_n - 3\sigma_n^4) \sum_{j > -m} \theta_{n,i-j} \theta_{n,i+u-j} \theta_{n,i'-j} \theta_{n,i'+u-j},
\]
\[
III = \sigma_n^4 \left[ \left( \sum_{j > -m} \theta_{n,i-j} \theta_{n,i+u-j} \right) \left( \sum_{j > -m} \theta_{n,i'-j} \theta_{n,i'+u-j} \right) + \text{two similar terms} \right],
\]
\[
IV = \sigma_n^2 \left[ \left( \sum_{j \leq -m} \theta_{i-j} \theta_{i+u-j} \right) \left( \sum_{j \leq -m} \theta_{n,i'-j} \xi_{n,j} \right) \left( \sum_{j \leq -m} \theta_{n,i'+u-j} \xi_{n,j} \right) + \text{five similar terms} \right].
\]
Under condition (a), $\sum_{ij \geq k} a_{ij}^2 = O(k^{-\alpha})$ for some $\alpha > 1,$ we have, by the same reasoning as for (17)-(19),
\[
\left\| \sum_{i=1}^{d} \left( \sum_{j \leq -m} \theta_{n,i-j} \xi_{n,j} \right) \left( \sum_{j \leq -m} \theta_{n,i+u-j} \xi_{n,j} \right) \right\|_1 = O \left( (1 + m)^{-\alpha+1} \right).
\]
17
which converges to 0 as $\min(m, n, d) \rightarrow \infty$. Similarly, the following can be proved:

$$
\left\| d^{-1} \sum_{i=1}^{d} \sum_{i'=1}^{d} I \right\|_{1} \rightarrow 0,
$$

$$
\left\| d^{-1} \sum_{i=1}^{d} \sum_{i'=1}^{d} \left( \gamma_n - 3\sigma_n^4 \right) R_n^2(u) \right\|_{1} \rightarrow 0,
$$

$$
\left\| d^{-1} \sum_{i=1}^{d} \sum_{i'=1}^{d} \left( III - \sum_{i} \left[ R_n^2(i) + R_n(i - u)R_n(i + u) \right] - dR_n^2(u) \right) \right\|_{1} \rightarrow 0,
$$

$$
\left\| d^{-1} \sum_{i=1}^{d} \sum_{i'=1}^{d} IV \right\|_{1} \rightarrow 0,
$$

and

$$
\left\| 2R_n(u) \sum_{i=1}^{d} E_{-m}(X_{n,i}X_{n,i+u}) - 2R_n^2(u) \right\|_{1} \rightarrow 0
$$

as $\min(m, n, d) \rightarrow \infty$. Therefore,

$$
\left\| E_{-m}(T_{n,d}^2/d) - \tau^2 \right\|_{1}
$$

$$
= \left\| d^{-1} \sum_{i=1}^{d} \sum_{i'=1}^{d} \left( I + II + III + IV \right) - 2R_n(u) \sum_{i=1}^{d} E_{-m}(X_{n,i}X_{n,i+u}) + dR_n^2(u) \right\|_{1}
$$

$$
\rightarrow 0
$$

as $\min(m, n, d) \rightarrow \infty$. \hfill \Box

### References


