*ISSN 1471-0498*

# DEPARTMENT OF ECONOMICS

# DISCUSSION PAPER SERIES

## A NEW MIXING CONDITION

**Brendan K. Beare**

Number 348
September 2007

Manor Road Building, Oxford OX1 3UQ

# A New Mixing Condition

**Brendan K. Beare*,**

*Nuffield College*
*Oxford OX1 1NF*
*United Kingdom*
*e-mail:* brendan.beare@nuffield.ox.ac.uk

**Abstract:** In this paper a new mixing condition for sequences of random variables is considered. This mixing condition is termed $\gamma$-mixing. Whereas mixing conditions such as $\alpha$-mixing are typically defined in terms of entire $\sigma$-fields of sets generated by random variables in the distant past and future, $\gamma$-mixing is defined in terms of a smaller class of sets: the finite dimensional cylinder sets. This leads to a definition of mixing more general than those in current use. A Rosenthal inequality, law of large numbers, and functional central limit theorem are proved for $\gamma$-mixing processes.

**AMS 2000 subject classifications:** Primary 60F17, 60F25; secondary 62M10.
**Keywords and phrases:** mixing, weak dependence, Hardy-Krause variation.

*All email correspondence prior to November 2007 should be addressed to both brendan.beare@nuffield.ox.ac.uk and brendan.beare@yale.edu.

## 1. Introduction

Suppose $\{X_t : t \in \mathbb{Z}\}$ is a doubly infinite sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$. Mixing conditions provide one way to formalize the notion that the random variables $\{X_t\}$ may be only weakly dependent upon one another. There are many ways to define mixing; the monograph by Doukhan (5) lists five alternative definitions. Of those, the weakest is $\alpha$-mixing, also known as strong mixing. For any nonempty set of integers $T$, let $\mathcal{F}_T \subset \mathcal{F}$ denote the $\sigma$-field generated by the random variables $\{X_t : t \in T\}$. The $\alpha$-mixing coefficients $\{\alpha_r : r \in \mathbb{N}\}$ corresponding to $\{X_t\}$ are given by

$$\alpha_r = \sup_{S,T} \sup_{A \in \mathcal{F}_S, B \in \mathcal{F}_T} \left| P\left(A \cap B\right) - P\left(A\right) P\left(B\right) \right|, \tag{1.1}$$

where the first supremum is taken over all nonempty finite sets of integers $S, T$ such that $\min T - \max S \geq r$. If $\alpha_r \to 0$ as $r \to \infty$, then $\{X_t\}$ is said to be $\alpha$-mixing.

The subject of this paper is a new mixing condition more general than $\alpha$-mixing. For any nonempty set of integers $T$, let $\mathcal{H}_T \subset \mathcal{F}$ denote the class of sets of the form $\cap_{t \in T} \{X_t \leq x_t\}$, where each $x_t$ ranges over $\mathbb{R}$. We shall define the $\gamma$-mixing coefficients $\{\gamma_r : r \in \mathbb{N}\}$ corresponding to $\{X_t\}$ by

$$\gamma_r = \sup_{S,T} \sup_{A \in \mathcal{H}_S, B \in \mathcal{H}_T} \left| P\left(A \cap B\right) - P\left(A\right) P\left(B\right) \right|, \tag{1.2}$$

where, once again, the first supremum is taken over all nonempty finite sets of integers $S, T$ such that $\min T - \max S \geq r$. If $\gamma_r \to 0$ as $r \to \infty$, then we shall say that $\{X_t\}$ is $\gamma$-mixing.

The definition of $\alpha$-mixing given above is commonly attributed to Rosenblatt (12), but to the author's knowledge, the first definition of $\alpha$-mixing consistent with (1.1) was given by Volkonskii and Rozanov (14). Rosenblatt's definition of mixing was, in fact, much closer to $\gamma$-mixing than to $\alpha$-mixing: where we have defined $\gamma$-mixing in terms of sets of the form $\cap_{t \in T} \{X_t \leq x_t\}$, Rosenblatt's mixing condition was defined in terms of sets of the form $\cap_{t \in T} \{w_t < X_t \leq x_t\}$.

It will be shown in this paper that $\gamma$-mixing conditions may be used as a basis for constructing moment inequalities and invariance principles for weakly dependent processes. The key to establishing limit theorems under $\gamma$-mixing is a new covariance inequality developed in other work by the author (2), which allows us to bound the covariance between two functions of a $\gamma$-mixing process by a quantity depending on the Hardy-Krause total variation norms of those functions. This covariance inequality will be used to prove a Rosenthal inequality, law of large numbers, and functional central limit theorem for $\gamma$-mixing processes.

The paper is structured as follows. In section 2, an example of a process that is $\gamma$-mixing but not $\alpha$-mixing is given. Covariance inequalities applicable to $\gamma$-mixing processes are discussed in section 3. Our Rosenthal inequality and law of large numbers are proved in section 4, and our functional central limit theorem in section 5.

## 2. A process that is $\gamma$-mixing but not $\alpha$-mixing

In this section an example is provided of a process that is $\gamma$-mixing but not $\alpha$-mixing, thus demonstrating that the former mixing condition is strictly weaker than the latter. The process in question has been used in previous work (1; 16) to provide an example of a stable linear process that is not $\alpha$-mixing. Let $\{\varepsilon_t : t \in \mathbb{Z}\}$ be an independent sequence of random variables that each take the value 0 with probability 1/2 and the value 1/2 with probability 1/2. For $t \in \mathbb{Z}$, define $X_t$ as the limit in mean square of the series $\sum_{k=0}^{\infty} 2^{-k}\varepsilon_{t-k}$. It is easy to show that the marginal distribution of each $X_t$ is uniform on $[0, 1]$ by writing $X_t = (1/2) X_{t-1} + \varepsilon_t$ and using a simple argument with characteristic functions.

In (1) it is shown explicitly that $\{X_t\}$ is not $\alpha$-mixing by the construction of a set $A \in \mathcal{F}_0$ and a sequence of sets $\{B_r\}$, $B_r \in \mathcal{F}_r$, such that

$$|P(A \cap B_r) - P(A) P(B_r)| \geq 1/4$$

for all $r \in \mathbb{N}$. Let $W_r = \{w_{r,1}, \ldots, w_{r,m_r}\}$ denote the support of the random variable $X_r - 2^{-r}X_0$, and note that $m_r \leq 2^r$. The sets $A$, $\{B_r\}$ are defined as follows:

$$\begin{aligned} A &= \left\{X_0 \in \left(0, \frac{1}{2}\right)\right\} \\ B_r &= \left\{X_r \in \bigcup_{k=1}^{m_r} \left(w_{r,k}, w_{r,k} + 2^{-r-1}\right)\right\}. \end{aligned}$$

Now, since $X_0$ is distributed uniformly on $[0, 1]$, we have $P(A) = 1/2$. Furthermore, since $X_r = 2^{-r}X_0 + w_{r,k}$ for some $k = 1, \ldots, m_r$, we must have $A \subseteq B$. Consequently, $|P(A \cap B_r) - P(A) P(B_r)| = (1 - P(B_r))/2$. But, since $X_r$ is uniformly distributed on $[0, 1]$, we have $P(B_r) \leq m_r 2^{-r-1} \leq 1/2$. Thus, $|P(A \cap B_r) - P(A) P(B_r)| \geq 1/4$, and so $\{X_t\}$ cannot be $\alpha$-mixing.

**Theorem 2.1.** *$\{X_t\}$ is $\gamma$-mixing, with $\gamma_r \leq 2^{2-r}$ for all $r$.*

*Proof.* Fix two finite sets of integers $S$ and $T$ with $\min T - \max S \geq r$. For $x \in \mathbb{R}^{|S|}$ and $y \in \mathbb{R}^{|T|}$, let $A_x = \cap_{s\in S} \{X_s \leq x_s\}$ and $B_y = \cap_{t\in T} \{X_t \leq y_t\}$. Observe that

$$\begin{aligned} |P(A_x \cap B_y) - P(A_x) P(B_y)| &\leq |P(B_y|A_x) - P(B_y)| \\ &\leq E\,|P(B_y|\mathcal{F}_S) - P(B_y)|\,. \end{aligned}$$

Since $\{X_t\}$ is a Markov chain, it is clear that the term on the right-hand side of the above inequality depends only on the largest element of $S$. Without loss of generality, we may therefore assume that $S = \{s\}$, a singleton. Further, denoting the elements of $T$ in order from smallest to greatest by $t_1, \ldots, t_v$, the Markov property allows us to write

$$\begin{aligned}
& \left|P\left(B_y|\mathcal{F}_s\right)-P\left(B_y\right)\right| \\
= & \left|P\left(X_{t_v}\leq y_{t_v}|\mathcal{F}_{t_{v-1}}\right)P\left(X_{t_{v-1}}\leq y_{t_{v-1}}|\mathcal{F}_{t_{v-2}}\right)\cdots P\left(X_{t_1}\leq y_{t_1}|\mathcal{F}_s\right)\right. \\
& \left.-P\left(X_{t_v}\leq y_{t_v}|\mathcal{F}_{t_{v-1}}\right)P\left(X_{t_{v-1}}\leq y_{t_{v-1}}|\mathcal{F}_{t_{v-2}}\right)\cdots P\left(X_{t_1}\leq y_{t_1}\right)\right| \\
\leq & \left|P\left(X_{t_1}\leq y_{t_1}|\mathcal{F}_s\right)-P\left(X_{t_1}\leq y_{t_1}\right)\right|.
\end{aligned}$$

The term on the right-hand side of the above inequality depends only on the smallest element of $T$, so without loss of generality we shall set $T=\{t\}$, a singleton. We now have

$$\left|P\left(A_x\cap B_y\right)-P\left(A_x\right)P\left(B_y\right)\right|\leq E\left|P\left(X_t\leq y|\mathcal{F}_s\right)-P\left(X_t\leq y\right)\right|.$$

The triangle inequality, the independence of $X_t-2^{s-t}X_s$ and $X_s$, and the nonnegativity of $X_s$ respectively imply the following three relations:

$$\begin{aligned}
& \left|P\left(X_t\leq y|\mathcal{F}_s\right)-P\left(X_t\leq y\right)\right| \\
\leq & \left|P\left(X_t-2^{s-t}X_s\leq y|\mathcal{F}_s\right)-P\left(X_t\leq y|\mathcal{F}_s\right)\right| \\
& +\left|P\left(X_t-2^{s-t}X_s\leq y|\mathcal{F}_s\right)-P\left(X_t\leq y|\right)\right| \\
= & \left|P\left(X_t-2^{s-t}X_s\leq y|\mathcal{F}_s\right)-P\left(X_t\leq y|\mathcal{F}_s\right)\right| \\
& +\left|P\left(X_t-2^{s-t}X_s\leq y\right)-P\left(X_t\leq y|\right)\right| \\
\leq & P\left(X_t-2^{s-t}X_s\leq y<X_t|\mathcal{F}_s\right)+P\left(X_t-2^{s-t}X_s\leq y<X_t\right).
\end{aligned}$$

Consequently,

$$\begin{aligned}
E\left|P\left(X_t\leq y|\mathcal{F}_s\right)-P\left(X_t\leq y\right)\right| & \leq 2P\left(X_t-2^{s-t}X_s\leq y<X_t\right) \\
& \leq 2P\left(\left|X_t-y\right|<2^{s-t}X_s\right) \\
& \leq 2P\left(\left|X_t-y\right|<2^{-r}\right).
\end{aligned}$$

The quantity $P\left(\left|X_t-y\right|<2^{-r}\right)$ is maximized at $y=1/2$, at which we have $P\left(\left|X_t-1/2\right|<2^{-r}\right)=2^{1-r}$. It follows that $\gamma_r\leq 2^{2-r}$ for all $r$. □

The process discussed in this section provides one example of a stable linear process that is $\gamma$-mixing but not $\alpha$-mixing. General demonstrations of the $\alpha$-mixing property for linear processes (7; 15) require smoothness conditions on the distribution of innovations. It would be interesting to know whether a general demonstration of the $\gamma$-mixing property could be provided for linear processes without the imposition of smoothness conditions.

## 3. Covariance inequalities

Suppose $\{X_t\}$ has $\alpha$-mixing coefficients $\{\alpha_r\}$ and $\gamma$-mixing coefficients $\{\gamma_r\}$. The following covariance inequality is proved in (14): for any nonempty finite sets of integers $S$ and $T$ such that $\min T - \max S \geq r$, and any Borel measurable functions $f : \mathbb{R}^{|S|} \to \mathbb{R}$ and $g : \mathbb{R}^{|T|} \to \mathbb{R}$, we have

$$\left|\mathrm{Cov}\left(f\left(X_s : s \in S\right), g\left(X_t : t \in T\right)\right)\right| \leq 4 \left\|f\right\|_\infty \left\|g\right\|_\infty \alpha_r. \tag{3.1}$$

In (14), the inequality is stated with a constant term of 16, but it is clear that a constant of 4 is sufficient. To prove (3.1), suppose without loss of generality that $\|f\|_\infty = 1$ and $\|g\|_\infty = 1$, and define $f_S = f\left(X_s : s \in S\right)$, $g_T = g\left(X_t : t \in T\right)$. Since

$$\begin{aligned} \mathrm{Cov}\left(f_S, g_T\right) &= E\left(f_S\left(E\left(g_T|\mathcal{F}_S\right) - Eg_T\right)\right) \\ &= E\left(g_T\left(E\left(f_S|\mathcal{F}_T\right) - Ef_S\right)\right), \end{aligned}$$

the left-hand side of (3.1) can be maximized by choosing $f_S$ and $g_T$ that satisfy $f_S = 2 \cdot 1_A - 1$ and $g_T = 2 \cdot 1_B - 1$, where $A = \left\{E\left(g_T|\mathcal{F}_S\right) > Eg_T\right\} \in \mathcal{F}_S$ and $B = \left\{E\left(f_S|\mathcal{F}_T\right) > Ef_S\right\} \in \mathcal{F}_T$. But $\mathrm{Cov}\left(2 \cdot 1_A - 1, 2 \cdot 1_B - 1\right) = 4\left(P\left(A \cap B\right) - P\left(A\right)P\left(B\right)\right)$, so we may bound the left-hand side of (3.1) by $4\alpha_r$.

Inequality (3.1) is the foundation upon which moment inequalities and invariance principles for $\alpha$-mixing processes are built. To prove moment inequalities and invariance principles for $\gamma$-mixing processes, we require an inequality comparable to (3.1) that is expressed in terms of $\gamma$-mixing coefficients. Such an inequality has been proved in (2). Before stating this inequality, let us review the definitions of Vitali and Hardy-Krause variation for multivariate functions; these concepts are central to the results presented throughout the rest of the paper. For a more extensive discussion of Vitali and Hardy-Krause variation, refer to (2; 11).

Let $f$ be a real valued function defined on a bounded $n$-dimensional rectangle $[a, b] = \{x \in \mathbb{R}^n : a \leq x \leq b\}$, $n \in \mathbb{N}$. Suppose $R = [c, d]$ is an $n$-dimensional rectangle contained in $[a, b]$. Define

$$\Delta_R f = \sum_{I \subseteq \{1,\ldots,n\}} (-1)^{|I|} f\left(x_I\right),$$

where $x_I$ is the vector in $\mathbb{R}^n$ whose $i$th element is given by $c_i$ if $i \in I$, or by $d_i$ if $i \notin I$. For instance, if $n = 2$ then we have

$$\Delta_R f = f\left(d_1, d_2\right) - f\left(c_1, d_2\right) - f\left(d_1, c_2\right) + f\left(c_1, c_2\right).$$

The Vitali variation of $f$ is given by

$$\|f\|_{\mathrm{V}} = \sup \sum_{R \in \mathcal{A}} \left|\Delta_R f\right|,$$

where the supremum is taken over all finite collections of $n$-dimensional rectangles $\mathcal{A} = \{R_i : 1 \leq i \leq m\}$ such that $\bigcup_{i=1}^m R_i = [a, b]$, and the interiors of any two rectangles in $\mathcal{A}$ are disjoint.

Given a nonempty set $I \subseteq \{1,\dots,n\}$, and a function $f:[a,b]\to\mathbb{R}$, let $f_I$ denote the real valued function on $\prod_{i\in I}[a_i,b_i]$ obtained by setting the $i$th argument of $f$ equal to $b_i$ whenever $i\notin I$. The Hardy-Krause variation of $f$ is given by

$$\|f\|_{\mathrm{HK}} = \sum_{\emptyset\neq I\subseteq\{1,\dots,n\}} \|f_I\|_{\mathrm{V}}\,.$$

Evidently, both Vitali variation and Hardy-Krause variation reduce to the usual definition of total variation when $n=1$, but when $n\geq 2$ Hardy-Krause variation will typically be greater than Vitali variation.

If a function $f:[a,b]\to\mathbb{R}$ is of bounded Hardy-Krause variation, then for any $x\in(a,b]$ there exists a value, which we denote $f^-(x)$, such that $f(x_m)\to f^-(x)$ for any sequence of points $\{x_m\}$ in $[a,x)$ that converges to $x$. We will refer to $f^-$ as the left-hand limit of $f$, and set $f^-(x)=f(x)$ for $x\notin(a,b]$. If $f_I=f_I^-$ for all nonempty $I\subseteq\{1,\dots,n\}$, then we will say that $f$ is left-continuous.

Our fundamental covariance inequality for $\gamma$-mixing processes is as follows.

**Theorem 3.1.** *Suppose each $X_t$ takes values in a bounded interval $[a_t,b_t]$. Let $S$ and $T$ denote finite sets of integers with $\min T-\max S\geq r$. Then, for any functions $f:\prod_{s\in S}[a_s,b_s]\to\mathbb{R}$ and $g:\prod_{t\in T}[a_t,b_t]\to\mathbb{R}$ that are left-continuous and of bounded Hardy-Krause variation, we have*

$$|Cov(f(X_s:s\in S),g(X_t:t\in T))|\leq\|f\|_{HK}\|g\|_{HK}\gamma_r.$$

*Proof.* Immediate from Theorem 4.2 in (2), and the definition of $\gamma_r$. ☐

Theorem 3.1 reveals that bounded $\gamma$-mixing processes satisfy a version of $(\theta,\mathcal{F},\psi)$-weak dependence, a dependence condition introduced in (6). Given a sequence of positive real numbers $\theta=\{\theta_r\}$ tending to zero, a class of functions $\mathcal{F}$ each mapping $\mathbb{R}^u$ to $\mathbb{R}$ for some $u\geq 1$, and a function $\psi:\mathcal{F}^2\to\mathbb{R}$, a sequence of random variables is said to be $(\theta,\mathcal{F},\psi)$-weak dependent if, for any finite sets of integers $S,T$ with $\min T-\max S\geq r$, we have

$$|\mathrm{Cov}(f(X_s:s\in S),g(X_t:t\in T))|\leq\psi(h,k)\,\theta_r$$

for any functions $f,g\in\mathcal{F}$ with domains $\mathbb{R}^{|S|}$ and $\mathbb{R}^{|T|}$ respectively. Different choices of $\mathcal{F}$ and $\psi$ yield different forms of weak dependence. Theorem 3.1 tells us that, for stationary sequences of bounded random variables taking values in $[a,b]\subset\mathbb{R}$, $\gamma$-mixing implies $(\gamma,\mathcal{F},\psi)$-weak dependence with $\mathcal{F}$ equal to the set of all functions of $u$ variables ($u\geq 1$) that are left-continuous and of bounded Hardy-Krause variation on $[a,b]^u$, and $\psi(f,g)=\|f\|_{\mathrm{HK}}\|g\|_{\mathrm{HK}}$. By comparison, inequality (3.1) tells us that $\alpha$-mixing processes are $(\alpha,\mathcal{F},\psi)$-weak dependent with $\mathcal{F}$ equal to the set of all bounded measurable functions of $u$ variables ($u\geq 1$), and $\psi(f,g)=4\|f\|_\infty\|g\|_\infty$. The form of $(\theta,\mathcal{F},\psi)$-weak dependence studied in most depth in (6) has $\mathcal{F}$ equal to the set of bounded Lipschitz functions, and $\psi(f,g)$ depending on the Lipschitz norms of $f$ and $g$. The form of $(\theta,\mathcal{F},\psi)$-weak dependence implied by Theorem 3.1 does not appear to have been studied previously.

Theorem 3.1 provides a covariance inequality in terms of $\gamma$-mixing coefficients for processes whose marginal distributions are bounded. Given a particular choice of $f$ and $g$, it may be possible to use a truncation argument to extend Theorem 3.1 so that it is applicable to unbounded random variables. As an example, let us choose $f$ and $g$ to be product functions.

**Theorem 3.2.** *For any finite sets of integers $S$ and $T$ with $\min T - \max S \geq r$, and any $p, q \geq 1$ with $\sup_{t\in S\cup T} \|X_t\|_p < \infty$ and $(|S| + |T|)\, p^{-1} + q^{-1} = 1$, we have*

$$\left| Cov\left(\prod_{s\in S} X_s, \prod_{t\in T} X_t\right)\right| \leq A\left(\prod_{t\in S\cup T} \|X_t\|_p\right)\gamma_r^{1/q},$$

*where*

$$\begin{aligned} A &= \left(3^{|S|}-1\right)\left(3^{|T|}-1\right) + 2\,|S| + 2\,|T| + 2\,|S|^{\frac{|S|}{|S|+|T|}}\,|T|^{\frac{|T|}{|S|+|T|}} \\ &\leq 3^{|S|+|T|} + 4\,(|S|+|T|) - 5. \end{aligned}$$

*Proof.* If $\gamma_r = 0$ then the $\sigma$-fields $\mathcal{F}_S$ and $\mathcal{F}_T$ must be independent, in which case the theorem is trivial. Assume $\gamma_r > 0$. Define

$$A_1 = \left(3^{|S|}-1\right)\left(3^{|T|}-1\right)$$

and

$$A_2 = 2\,|S| + 2\,|T| + 2\,|S|^{\frac{|S|}{|S|+|T|}}\,|T|^{\frac{|T|}{|S|+|T|}},$$

so that $A = A_1 + A_2$. Let $\bar{X}_t = \min\{\max\{X_t, -a_t\}, a_t\}$, where, for some constant $c > 0$, we define $a_t = \|X_t\|_p\, c^{-q/p}\gamma_r^{-1/p}$. Begin by writing

$$\begin{aligned} &\left|\mathrm{Cov}\left(\prod_{s\in S} X_s, \prod_{t\in T} X_t\right)\right| \\ \leq\ &\left|\mathrm{Cov}\left(\prod_{s\in S} \bar{X}_s, \prod_{t\in T} \bar{X}_t\right)\right| + \left|\mathrm{Cov}\left(\prod_{s\in S} X_s - \prod_{s\in S}\bar{X}_s, \prod_{t\in T} \bar{X}_t\right)\right| \\ &+ \left|\mathrm{Cov}\left(\prod_{s\in S} \bar{X}_s, \prod_{t\in T} X_t - \prod_{t\in T}\bar{X}_t\right)\right| \\ &+ \left|\mathrm{Cov}\left(\prod_{s\in S} X_s - \prod_{s\in S}\bar{X}_s, \prod_{t\in T} X_t - \prod_{t\in T}\bar{X}_t\right)\right|. \end{aligned} \tag{3.2}$$

Using standard arguments with the inequalities of Hölder and Markov, we can bound the last three terms on the right-hand side of (3.2) as follows:

$$\left|\mathrm{Cov}\left(\prod_{s\in S} X_s - \prod_{s\in S}\bar{X}_s, \prod_{t\in T} \bar{X}_t\right)\right| \leq 2\,|S|\left(\prod_{t\in S\cup T}\|X_t\|_p\right) c\gamma_r^{1/q}, \tag{3.3}$$

$$\left|\mathrm{Cov}\left(\prod_{s\in S}\bar{X}_s, \prod_{t\in T}X_t - \prod_{t\in T}\bar{X}_t\right)\right| \le 2\left|T\right|\left(\prod_{t\in S\cup T}\left\|X_t\right\|_p\right)c\gamma_r^{1/q}, \tag{3.4}$$

and

$$\begin{aligned}&\left|\mathrm{Cov}\left(\prod_{s\in S}X_s - \prod_{s\in S}\bar{X}_s, \prod_{t\in T}X_t - \prod_{t\in T}\bar{X}_t\right)\right| \\ \le\ & 2\left|S\right|^{\frac{|S|}{|S|+|T|}}\left|T\right|^{\frac{|T|}{|S|+|T|}}\left(\prod_{t\in S\cup T}\left\|X_t\right\|_p\right)c\gamma_r^{1/q}.\end{aligned} \tag{3.5}$$

We will apply Theorem 3.1 to bound the first term on the right-hand side of (3.2). Observe that any set of the form $A = \cap_{s\in S}\left\{\bar{X}_s \le x_s\right\}$ can be written as $A = \cap_{s\in S}\left\{X_s \le x_s^*\right\}$, where $x_s^* = -\infty$ if $x_s < -a_s$, $x_s^* = x_s$ if $x_s < a_s$, and $x_s^* = \infty$ if $x_s \ge a_s$. It follows that $\left\{\bar{X}_t\right\}$ must be $\gamma$-mixing, with mixing coefficients bounded by those of $\left\{X_t\right\}$. Let the functions $f : \prod_{s\in S}\left[-a_s, a_s\right] \to \mathbb{R}$ and $g : \prod_{t\in T}\left[-a_t, a_t\right] \to \mathbb{R}$ be given by $f\left(x_s : s\in S\right) = \prod_{s\in S}x_s$ and $g\left(x_t : t\in T\right) = \prod_{t\in T}x_t$. Clearly, for nonempty $I\subseteq S$ we have $f_I\left(x_s : s\in I\right) = \left(\prod_{s\in I}x_s\right)\left(\prod_{s\in S\setminus I}a_s\right)$. The Vitali variation of $f_I$ is given by the $L_1$ norm of the mixed partial derivative obtained by differentiating $f_I$ once with respect to each argument:

$$\begin{aligned}\left\|f_I\right\|_{\mathrm{V}} &= \int_{\prod_{s\in I}\left[-a_s,a_s\right]}\left(\prod_{s\in S\setminus I}a_s\right)\prod_{s\in I}\mathrm{d}x_s \\ &= 2^{|I|}\left(\prod_{s\in S}a_s\right).\end{aligned}$$

Thus, using the binomial theorem, the Hardy-Krause variation of $f$ is given by

$$\begin{aligned}\left\|f\right\|_{\mathrm{HK}} &= \sum_{\emptyset\neq I\subseteq S}2^{|I|}\left(\prod_{s\in S}a_s\right) \\ &= \left(\prod_{s\in S}a_s\right)\left(\sum_{s=1}^{|S|}\frac{|S|!}{(|S|-s)!s!}2^s\right) \\ &= \left(\prod_{s\in S}a_s\right)\left(3^{|S|}-1\right),\end{aligned}$$

and similarly

$$\left\|g\right\|_{\mathrm{HK}} = \left(\prod_{t\in T}a_t\right)\left(3^{|T|}-1\right).$$

It now follows from Theorem 3.1 that

$$\begin{aligned}\left|\mathrm{Cov}\left(\prod_{s\in S}\bar{X}_s,\prod_{t\in T}\bar{X}_t\right)\right| &\leq \|f\|_{\mathrm{HK}}\|g\|_{\mathrm{HK}}\gamma_r \\ &= A_1\left(\prod_{t\in S\cup T}a_t\right)\gamma_r \\ &= A_1\left(\prod_{t\in S\cup T}\|X_t\|_p\right)c^{1-q}\gamma_r^{1/q}. \qquad (3.6)\end{aligned}$$

Combining (3.2) through (3.6), we obtain

$$\left|\mathrm{Cov}\left(\prod_{s\in S}X_s,\prod_{t\in T}X_t\right)\right| \leq \left(A_1c^{1-q}+A_2c\right)\left(\prod_{t\in S\cup T}\|X_t\|_p\right)\gamma_r^{1/q}.$$

The quantity $A_1c^{1-q}+A_2c$ is minimized by setting $c=\left(A_1(q-1)/A_2\right)^{1/q}$, at which $A_1c^{1-q}+A_2c=A_1^{1/q}A_2^{(q-1)/q}q(q-1)^{(1-q)/q}$. If we now choose $q$ to maximize this quantity, we obtain $A_1c^{1-q}+A_2c=A_1+A_2=A$ at $q=(A_1+A_2)/A_1$. This yields our desired bound. □

Note that, if we choose $S$ and $T$ to be singleton sets containing $t$ and $t+r$ respectively, then Theorem 3.2 states that

$$|\mathrm{Cov}(X_t,X_{t+r})| \leq 10\|X_t\|_p\|X_{t+r}\|_p\gamma_r^{1/q},$$

where $2p^{-1}+q^{-1}=1$. The analogous inequality under $\alpha$-mixing, given in (4), states that

$$|\mathrm{Cov}(X_t,X_{t+r})| \leq 10\|X_t\|_p\|X_{t+r}\|_p\alpha_r^{1/q}. \qquad (3.7)$$

Thus, in view of the fact that $\gamma_r\leq\alpha_r$, Theorem 3.2 provides a refinement of inequality (3.7). Note that (3.7) is frequently misstated in the literature as involving a constant term of 8 rather than 10.

Theorem 3.2 leads readily to the following corollary, which concerns covariances between polynomial functions of several random variables.

**Corollary 3.1.** *Fix positive integers $u,v,r$. For any integers $t_1\leq\cdots\leq t_{u+v}$ such that $t_u+r\leq t_{u+1}$, and any $p,q\geq 1$ with $\sup_{1\leq i\leq u+v}\|X_{t_i}\|_p<\infty$ and $(u+v)p^{-1}+q^{-1}=1$, we have*

$$\left|Cov\left(\prod_{i=1}^{u}X_{t_i},\prod_{i=u+1}^{u+v}X_{t_i}\right)\right| \leq A\left(\prod_{i=1}^{u+v}\|X_{t_i}\|_p\right)\gamma_r^{1/q},$$

*where*

$$\begin{aligned}A &= (3^u-1)(3^v-1)+2u+2v+2u^{\frac{u}{u+v}}v^{\frac{v}{u+v}} \\ &\leq 3^{u+v}+4(u+v)-5.\end{aligned}$$

*Proof.* This result differs from Theorem 3.2 only if some of the indices $t_i$ are equal to one another. In this case, we may construct a new set of variables, $X_1^*, \dots, X_{u+v}^*$, such that $X_i^* = X_{t_i}$ for each $i$. The $\gamma$-mixing coefficient associated with these variables is clearly the same as for $\{X_{t_i} : 1 \leq i \leq u+v\}$, and so our desired result now follows immediately from Theorem 3.2. □

Corollary 3.1 will be used in the next section to prove our Rosenthal inequality for $\gamma$-mixing processes.

## 4. Rosenthal inequality and law of large numbers

Given constants $p \geq 0$ and $\varepsilon > 0$, and a sequence of random variables $X = \{X_t\}$, define $W_n(p, \varepsilon, X)$ and $D_n(p, \varepsilon, X)$ as follows:

$$\begin{aligned} W_n(p,\varepsilon,X) &= \sum_{t=1}^{n} \|X_t\|_{p+\varepsilon}^p \\ D_n(p,\varepsilon,X) &= W_n(p,0,X) \text{ for } p \leq 1 \\ &= W_n(p,\varepsilon,X) \text{ for } 1 < p \leq 2 \\ &= \max\left\{W_n(p,\varepsilon,X), (W_n(2,\varepsilon,X))^{p/2}\right\} \text{ for } p \geq 2. \end{aligned}$$

The random variables $\{X_t\}$ are said to satisfy a Rosenthal inequality if there exists a constant $b < \infty$ such that $E\left|\sum_1^n X_t\right|^p \leq bD_n(p,\varepsilon,X)$ for all $n$. A Rosenthal inequality for $\alpha$-mixing processes is given in (5).

When $p \leq 1$, the Rosenthal inequality is a trivial consequence of the inequality $(a+b)^p \leq a^p + b^p$, which holds for any positive $a, b$. When $p > 1$, the Rosenthal inequality for $\alpha$-mixing processes is proved in two steps. First, using the covariance inequality for $\alpha$-mixing processes in (4), reproduced above as inequality (3.7), the Rosenthal inequality is proved for any even integer $p$. Second, the so-called interpolation lemma (13; 5) is used to extend the inequality to all real $p > 1$.

To prove a Rosenthal inequality for $\gamma$-mixing processes, we modify the arguments used in the $\alpha$-mixing case in the following way. First, in place of the inequality in (4) for $\alpha$-mixing processes, we employ Corollary 3.1 from above, which applies to $\gamma$-mixing processes. Second, we modify the interpolation lemma so that it is applicable under $\gamma$-mixing. The following lemma provides this modification. We will say that one sequence of numbers $\{\gamma_r\}$ dominates another sequence $\{\gamma'_r\}$ if $\gamma'_r \leq \gamma_r$ for all $r$.

**Lemma 4.1.** *Fix $k \geq 0$, $\varepsilon > 0$, and a sequence of nonnegative real numbers $\{\gamma_r\}$. Suppose there exists a constant $b < \infty$ such that any centered sequence of random variables $X = \{X_t\}$ whose $\gamma$-mixing coefficients are dominated by $\{\gamma_r\}$ satisfies*

$$E\left|\sum_{t=1}^{n} X_t\right|^k \leq bV_n(k,\varepsilon,X)$$

*for all n, where*

$$\begin{aligned} V_n(k,\varepsilon,X) &= W_n(k,\varepsilon,X) \text{ for } k\leq 2 \\ &= \max\left\{W_n(k,\varepsilon,X),(W_n(2,\varepsilon,X))^{k/2}\right\} \text{ for } k\geq 2. \end{aligned}$$

*Then, for any $p\in[0,k]$ there exists a constant $b'<\infty$ such that any centered sequence of random variables $X=\{X_t\}$ whose $\gamma$-mixing coefficients are dominated by $\{\gamma_r\}$ satisfies*

$$E\left|\sum_{t=1}^n X_t\right|^p \leq b' V_n(p,\varepsilon,X)$$

*for all n.*

*Proof.* The lemma is trivial for $p\leq 1$, so we assume $k,p\geq 1$. Suppose $X=\{X_t\}$ is a centered sequence of r.v.s whose $\gamma$-mixing coefficients are dominated by $\{\gamma_r\}$. Set $a=V_n(p,\varepsilon,X)^{1/p}$, $\bar{X}_t=\min\{\max\{X_t,-a\},a\}$, $\underline{X}_t=X_t-\bar{X}_t$, $Y_t=\bar{X}_t-E\bar{X}_t$, and $Z_t=\underline{X}_t-E\underline{X}_t$. Jensen's inequality allows us to write

$$\begin{aligned} &E\left|\sum_{t=1}^n X_t\right|^p \\ \leq\ & 2^{p-1}\left(E\left|\sum_{t=1}^n Y_t\right|^p+E\left|\sum_{t=1}^n Z_t\right|^p\right) \\ \leq\ & 2^{p-1}\left(E\left|\sum_{t=1}^n Y_t\right|^p+2^{p-1}\left(E\left|\sum_{t=1}^n Z_t 1_{\{Z_t\geq 0\}}\right|^p+E\left|\sum_{t=1}^n Z_t 1_{\{Z_t<0\}}\right|^p\right)\right) \\ \leq\ & 2^{p-1}\left(E\left|\sum_{t=1}^n Y_t\right|^k\right)^{p/k}+2^{2p-2}E\left(\sum_{t=1}^n |Z_t|^{p/k} 1_{\{Z_t\geq 0\}}\right)^k \\ & +2^{2p-2}E\left(\sum_{t=1}^n |Z_t|^{p/k} 1_{\{Z_t<0\}}\right)^k. \end{aligned}$$

Define the random variables

$$\begin{aligned} \xi_t &= |Z_t|^{p/k} 1_{\{Z_t\geq 0\}}-E\left(|Z_t|^{p/k} 1_{\{Z_t\geq 0\}}\right) \\ \zeta_t &= -|Z_t|^{p/k} 1_{\{Z_t<0\}}+E\left(|Z_t|^{p/k} 1_{\{Z_t<0\}}\right), \end{aligned}$$

and observe that

$$\begin{aligned} E\left(\sum_{t=1}^n |Z_t|^{p/k} 1_{\{Z_t\geq 0\}}\right)^k &= E\left(\sum_{t=1}^n \xi_t+\sum_{t=1}^n E\left(|Z_t|^{p/k} 1_{\{Z_t\geq 0\}}\right)\right)^k \\ &\leq 2^{k-1}\left(E\left|\sum_{t=1}^n \xi_t\right|^k+\left(\sum_{t=1}^n E|Z_t|^{p/k}\right)^k\right) \end{aligned}$$

and

$$E\left(\sum_{t=1}^{n}|Z_t|^{p/k}1_{\{Z_t<0\}}\right)^k = E\left(-\sum_{t=1}^{n}\zeta_t+\sum_{t=1}^{n}E\left(|Z_t|^{p/k}1_{\{Z_t<0\}}\right)\right)^k$$
$$\leq 2^{k-1}\left(E\left|\sum_{t=1}^{n}\zeta_t\right|^k+\left(\sum_{t=1}^{n}E|Z_t|^{p/k}\right)^k\right).$$

We thus have

$$E\left|\sum_{t=1}^{n}X_t\right|^p \leq 2^{p-1}\left(E\left|\sum_{t=1}^{n}Y_t\right|^k\right)^{p/k}+2^{2p+k-3}E\left|\sum_{t=1}^{n}\xi_t\right|^k$$
$$+2^{2p+k-3}E\left|\sum_{t=1}^{n}\zeta_t\right|^k+2^{2p+k-2}\left(\sum_{t=1}^{n}E|Z_t|^{p/k}\right)^k.$$

$Y_t$, $\xi_t$ and $\zeta_t$ are all nondecreasing transformations of $X_t$, and therefore all have $\gamma$-mixing coefficients that are dominated by $\{\gamma_r\}$. Thus, under the hypothesis of the lemma, there exists $b_1<\infty$ such that

$$E\left|\sum_{t=1}^{n}X_t\right|^p \leq 2^{p-1}\left(b_1V_n(k,\varepsilon,Y)\right)^{p/k}+2^{2p+k-3}b_1V_n(k,\varepsilon,\xi)$$
$$+2^{2p+k-3}b_1V_n(k,\varepsilon,\zeta)+2^{2p+k-2}\left(\sum_{t=1}^{n}E|Z_t|^{p/k}\right)^k.$$

Using arguments identical to those in (13; 5), we can show that

$$V_n(k,\varepsilon,Y)\leq 2^kV_n(p,\varepsilon,X)^{k/p},$$
$$V_n(k,\varepsilon,\xi)\leq 2^{k+p}V_n(p,\varepsilon,X),$$
$$V_n(k,\varepsilon,\zeta)\leq 2^{k+p}V_n(p,\varepsilon,X),$$

and, provided that $p\geq k-\varepsilon$,

$$\left(\sum_{t=1}^{n}E|Z_t|^{p/k}\right)^k\leq 2^pV_n(p,\varepsilon,X).$$

We thus obtain

$$E\left|\sum_{t=1}^{n}X_t\right|^p\leq b_2V_n(p,\varepsilon,X)$$

for some $b_2\geq 0$ not depending on $n$ or $X$. This completes the proof for the case where $p\geq k-\varepsilon$. But if the theorem is true for $p\geq k-\varepsilon$, then it must also be true for $p\geq k-2\varepsilon$, and so on for all $p\in[0,k]$. □

With Lemma 4.1 in hand, we are now in a position to state our Rosenthal inequality for $\gamma$-mixing processes.

**Theorem 4.1.** *Fix $p \geq 0$ and $\varepsilon > 0$, and let $k$ denote the smallest even integer equal to or greater than $p$. Let $\{X_t\}$ satisfy $EX_t = 0$ and $E|X_t|^{p+\varepsilon} < \infty$ for each $t$, and have $\gamma$-mixing coefficients satisfying $\sum_{r=1}^{\infty}(r+1)^{k-2}\gamma_r^{\varepsilon/(k+\varepsilon)} < \infty$. Then there exists a constant $b < \infty$ not depending on $\varepsilon$ such that, for all $n$,*

$$E\left|\sum_{t=1}^{n} X_t\right|^p \leq bD_n(p, \varepsilon, X).$$

*Proof.* The proof of this theorem differs from the proof under $\alpha$-mixing - see e.g. section 1.4.1 in (5) - in only two respects. First, Corollary 3.1 is used in place of the inequality in (4) for $\alpha$-mixing processes. Second, Lemma 4.1 is used in place of the interpolation lemma (13; 5) for $\alpha$-mixing processes. □

Note that the only difference between Theorem 4.1 and the Rosenthal inequality for $\alpha$-mixing processes stated in (5) is that we have replaced $\alpha$-mixing coefficients with $\gamma$-mixing coefficients. Theorem 4.1 thus represents a strict refinement of existing results under $\alpha$-mixing.

As an application of Theorem 4.1, we now present a law of large numbers for $\gamma$-mixing processes. Conditions are provided under which the sample average converges in $L_1$. A weak law of large numbers is thus an obvious corollary. Theorem 4.1 will be used again in the next section to demonstrate the tightness of a partial sum process under $\gamma$-mixing.

**Theorem 4.2.** *Suppose $\sup_t E|X_t|^{1+\varepsilon} < \infty$ and $\gamma_r = O\left(r^{-(2+\delta)/\delta}\right)$ for some $\varepsilon > \delta > 0$. Then*

$$E\left|\frac{1}{n}\sum_{t=1}^{n}(X_t - EX_t)\right| \to 0$$

*as $n \to \infty$.*

*Proof.* For $p \in (1, \min\{1+\varepsilon, 2\})$, Theorem 4.1 implies that

$$E\left|\sum_{t=1}^{n}(X_t - EX_t)\right|^p = O(n),$$

provided that

$$\sum_{r=1}^{\infty}\gamma_r^{\frac{1+\varepsilon-p}{3+\varepsilon-p}} < \infty.$$

Since $\gamma_r = O\left(r^{-(2+\delta)/\delta}\right)$, this condition will hold if

$$\left(\frac{2+\delta}{\delta}\right)\left(\frac{1+\varepsilon-p}{3+\varepsilon-p}\right) > 1,$$

which can be obtained by choosing $p$ sufficiently close to 1. We then have $E\left|n^{-1}\sum_1^n(X_t - EX_t)\right|^p = O\left(n^{1-p}\right) = o(1)$. □

## 5. Functional central limit theorem

In this section we prove a functional central limit theorem for $\gamma$-mixing processes. The main result we employ is Lemma 3.1 of Withers (16), which provides basic conditions under which a heterogeneous, dependent sequence of random variables will satisfy a central limit theorem. These conditions have also been considered in earlier work (8). Tightness of the sequence of partial sum processes follows easily from the Rosenthal inequality proved in the previous section.

Throughout this section, we assume stationarity of $\{X_t\}$.

**Theorem 5.1.** *Let $\{X_t\}$ be stationary. Suppose $EX_t = 0$, $E|X_t|^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$, and $\gamma_r = O\left(\exp\left(-r^{\delta}\right)\right)$ for some $\delta > (2+\varepsilon)/(2+2\varepsilon)$. Then the series*

$$\sigma^2 = EX_0^2 + 2\sum_{r=1}^{\infty} EX_0X_r$$

*is absolutely convergent; if $\sigma > 0$ then*

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[n\cdot]}(X_t - EX_t) \Rightarrow \sigma B$$

*as $n \to \infty$, where $B$ denotes a standard Brownian motion on $[0,1]$.*

*Proof.* Absolute convergence of $\sigma^2$ follows easily from Theorem 3.2. Suppose $\sigma > 0$, and define $B_n(r) = \frac{1}{\sigma\sqrt{n}}\sum_1^{[nr]} X_t$. We will first show finite dimensional convergence of $B_n$ to $B$, and then show tightness of $B_n$. Finite dimensional convergence will be shown using the Cramér-Wold device. That is, we will show that

$$\sum_{j=1}^{m}\lambda_j\left(B_n(r_j) - B_n(r_{j-1})\right) \to_d N\left(0, \sum_{j=1}^{m}\lambda_j^2(r_j - r_{j-1})\right)$$

as $n \to \infty$, for any $0 = r_0 < r_1 < \cdots < r_m = 1$ and any $\lambda \in \mathbb{R}^m \setminus \{0\}$, $m \in \mathbb{N}$. Observe that

$$\sum_{j=1}^{m}\lambda_j\left(B_n(r_j) - B_n(r_{j-1})\right) = \frac{1}{\sigma\sqrt{n}}\sum_{t=1}^{n} X_{tn},$$

where $X_{tn} = \lambda_{tn}X_t$, and $\lambda_{tn} = \lambda_{\min\{j:[nr_j]\geq t\}}$. We will show that

$$n^{-1/2}\sum_{t=1}^{n} X_{tn} \to_d N\left(0, \sigma^2\sum_{j=1}^{m}\lambda_j^2(r_j - r_{j-1})\right) \tag{5.1}$$

using parts (b) and (d) of Lemma 3.1 of Withers (16), a central limit theorem for dependent heterogeneous random arrays. Split $\{X_{tn} : 1 \leq t \leq n\}$ into $k$

Bernstein blocks of length $n_1$, separated by gaps of length $n_2$, as follows:

$$\begin{aligned}
\sum_{t=1}^{n} X_{tn} &= \sum_{i=1}^{k} \eta_{in} + \sum_{i=1}^{k+1} \nu_{in}, \quad k = \left[\frac{n}{n_1+n_2}\right] \\
\eta_{in} &= \sum_{t=(i-1)(n_1+n_2)+1}^{in_1+(i-1)n_2} X_{tn}, \; i = 1, \ldots, k \\
\nu_{in} &= \sum_{t=in_1+(i-1)n_2+1}^{i(n_1+n_2)} X_{tn}, \quad i = 1, \ldots, k \\
\nu_{k+1,n} &= \sum_{t=k(n_1+n_2)+1}^{[nr_m]} X_{tn}.
\end{aligned}$$

The sequences $n_1(n)$ and $n_2(n)$ are assumed to satisfy $n_1 \to \infty$, $n_2 \to \infty$, $n_1 = o(n)$ and $n_2 = o(n_1)$ as $n \to \infty$. Let $\tilde{\sigma}_n^2 = \sum_{i=1}^{k} E\eta_{in}^2$, and observe that

$$n^{-1}\tilde{\sigma}_n^2 \to \sigma^2 \sum_{j=1}^{m} \lambda_j^2 (r_j - r_{j-1}). \tag{5.2}$$

Withers' lemma states that $\tilde{\sigma}_n^{-1} \sum_1^n X_{tn} \to_d N(0,1)$ if the following three conditions are satisfied for $\phi, \psi$ being either sine or cosine functions:

$$\frac{1}{\tilde{\sigma}_n^2} E\left(\sum_{i=1}^{k+1} \nu_{in}\right)^2 \to 0 \tag{5.3}$$

$$\frac{1}{\tilde{\sigma}_n^2} \sum_{i=1}^{k} E\eta_{in}^2 1(|\eta_{in}| \geq \epsilon\tilde{\sigma}_n) \to 0 \text{ for all } \epsilon > 0 \tag{5.4}$$

$$\sum_{j=2}^{k} \left|\mathrm{Cov}\left(\phi\left(\frac{\omega}{\tilde{\sigma}_n}\sum_{i=1}^{j-1}\eta_{in}\right), \psi\left(\frac{\omega}{\tilde{\sigma}_n}\eta_{jn}\right)\right)\right| \to 0 \text{ for all } \omega > 0. \tag{5.5}$$

To see that (5.3) holds, observe that Theorem 4.1 implies that, for some constant $b < \infty$,

$$\begin{aligned}
E\left(\sum_{i=1}^{k+1} \nu_{in}\right)^2 &\leq b\sum_{i=1}^{k+1} \sum_{t=in_1+(i-1)n_2+1}^{i(n_1+n_2)} \|X_{tn}\|_{2+\varepsilon}^2 \\
&= O(kn_2).
\end{aligned}$$

It follows from (5.2) that $\tilde{\sigma}_n^{-2} E\left(\sum_1^{k+1} \nu_{in}\right)^2 = O\left(n^{-1}kn_2\right) = o(1)$. Next, observe that

$$n_1^{-1} E\eta_{in}^2 1(|\eta_{in}| \geq \epsilon\tilde{\sigma}_n) = E\left|n_1^{-1/2}\eta_{in}\right|^2 1\left(\left|n_1^{-1/2}\eta_{0n}\right| \geq \epsilon n_1^{-1/2}\tilde{\sigma}_n\right).$$

By Theorem 4.1, we have

$$\begin{aligned} E\left|n_1^{-1/2}\eta_{in}\right|^2 &\leq bn_1^{-1}\sum_{(i-1)(n_1+n_2)+1}^{in_1+(i-1)n_2}\|X_{tn}\|_{2+\varepsilon}^2 \\ &\leq b\bar{\lambda}^2\|X_t\|_{2+\varepsilon}^2\,, \end{aligned}$$

where $\bar{\lambda}=\max_j|\lambda_j|$. Dominated convergence thus gives

$$n_1^{-1}E\eta_{in}^2 1\left(|\eta_{in}|\geq\epsilon\tilde{\sigma}_n\right)=o\left(1\right),$$

and (5.4) follows.

It remains to confirm (5.5). Fix $a>0$, and let $\bar{X}_t=\min\{a,\max\{X_t,-a\}\}$, $\underline{X}_t=X_t-\bar{X}_t$, $\bar{X}_{tn}=\lambda_{tn}\bar{X}_t$ and $\underline{X}_{tn}=X_{tn}-\bar{X}_{tn}$. For $j=2,\ldots,k$, define

$$\begin{aligned} S_j &= \cup_{i=1}^{j-1}\left\{(i-1)\left(n_1+n_2\right)+1,\ldots,in_1+(i-1)\,n_2\right\} \\ T_j &= \left\{(j-1)\left(n_1+n_2\right)+1,\ldots,jn_1+(j-1)\,n_2\right\}. \end{aligned}$$

The mean value theorem can be used to show that

$$\begin{aligned} &\left|\mathrm{Cov}\left(\phi\left(\frac{\omega}{\tilde{\sigma}_n}\sum_{i=1}^{j-1}\eta_{in}\right),\psi\left(\frac{\omega}{\tilde{\sigma}_n}\eta_{jn}\right)\right)\right| \\ \leq &\left|\mathrm{Cov}\left(\phi\left(\frac{\omega}{\tilde{\sigma}_n}\sum_{s\in S_j}\bar{X}_{sn}\right),\psi\left(\frac{\omega}{\tilde{\sigma}_n}\sum_{t\in T_j}\bar{X}_{tn}\right)\right)\right| \\ &+\frac{2\omega}{\tilde{\sigma}_n}\sum_{s\in S_j}E\left|\underline{X}_{sn}\right|+\frac{2\omega}{\tilde{\sigma}_n}\sum_{t\in T_j}E\left|\underline{X}_{tn}\right|. \end{aligned}$$

Markov's inequality gives $E\left|\underline{X}_{tn}\right|\leq\bar{\lambda}^{2+\epsilon}a^{-1-\varepsilon}E\left|X_t\right|^{2+\varepsilon}$, so we have

$$\begin{aligned} &\left|\mathrm{Cov}\left(\phi\left(\frac{\omega}{\tilde{\sigma}_n}\sum_{i=1}^{j-1}\eta_{in}\right),\psi\left(\frac{\omega}{\tilde{\sigma}_n}\eta_{jn}\right)\right)\right| \\ \leq &\left|\mathrm{Cov}\left(\phi\left(\frac{\omega}{\tilde{\sigma}_n}\sum_{s\in S_j}\bar{X}_{sn}\right),\psi\left(\frac{\omega}{\tilde{\sigma}_n}\sum_{t\in T_j}\bar{X}_{tn}\right)\right)\right| \\ &+jn_1\frac{2\omega\bar{\lambda}^{2+\epsilon}}{\tilde{\sigma}_n}a^{-1-\varepsilon}E\left|X_t\right|^{2+\varepsilon}. \end{aligned} \tag{5.6}$$

We will use Theorem 3.1 to bound the first term on the right-hand side of inequality (5.6). Let $f:[-a,a]^{|S_j|}\to\mathbb{R}$ and $g:[-a,a]^{|T_j|}\to\mathbb{R}$ be given by

$$\begin{aligned} f\left(x_s:s\in S_j\right) &= \phi\left(\frac{\omega}{\tilde{\sigma}_n}\sum_{s\in S_j}\lambda_{sn}x_s\right) \\ g\left(x_t:t\in T_j\right) &= \psi\left(\frac{\omega}{\tilde{\sigma}_n}\sum_{t\in T_j}\lambda_{tn}x_t\right). \end{aligned}$$

Clearly, for nonempty $I \subseteq S_j$, we have

$$f_I\left(x_s : s \in I\right) = \phi\left(\frac{\omega}{\tilde{\sigma}_n}\left(\sum_{s\in I}\lambda_{sn}x_s + a\sum_{s\in S_j\setminus I}\lambda_{sn}\right)\right).$$

The function obtained by differentiating $f_I$ once with respect to each argument is bounded in absolute value by $\left(\omega\bar{\lambda}/\tilde{\sigma}_n\right)^{|I|}$. Thus,

$$\|f_I\|_{\mathrm{V}} \leq \left(\frac{2a\omega\bar{\lambda}}{\tilde{\sigma}_n}\right)^{|I|}.$$

Using the binomial theorem, we now have

$$\begin{aligned}
\|f\|_{\mathrm{HK}} &\leq \sum_{\emptyset\neq I\subseteq S_j}\left(\frac{2a\omega\bar{\lambda}}{\tilde{\sigma}_n}\right)^{|I|} \\
&= \sum_{s=1}^{|S_j|}\frac{|S_j|!}{(|S_j|-s)!s!}\left(\frac{2a\omega\bar{\lambda}}{\tilde{\sigma}_n}\right)^{s} \\
&= \left(1+\frac{2a\omega\bar{\lambda}}{\tilde{\sigma}_n}\right)^{|S_j|} - 1.
\end{aligned}$$

We can show similarly that

$$\|g\|_{\mathrm{HK}} \leq \left(1+\frac{2a\omega\bar{\lambda}}{\tilde{\sigma}_n}\right)^{|T_j|} - 1.$$

Thus, since the $\gamma$-mixing coefficients of $\left\{\bar{X}_t\right\}$ are dominated by those of $\{X_t\}$, Theorem 3.1 implies that

$$\begin{aligned}
&\left|\mathrm{Cov}\left(\phi\left(\frac{\omega}{\tilde{\sigma}_n}\sum_{s\in S_j}\bar{X}_{sn}\right),\psi\left(\frac{\omega}{\tilde{\sigma}_n}\sum_{t\in T_j}\bar{X}_{tn}\right)\right)\right| \\
= \;& \left|\mathrm{Cov}\left(f\left(\bar{X}_s : s\in S_j\right), g\left(\bar{X}_t : t\in T_j\right)\right)\right| \\
\leq \;& \|f\|_{\mathrm{HK}}\,\|g\|_{\mathrm{HK}}\,\gamma_{n_2} \\
\leq \;& \left(1+\frac{2a\omega\bar{\lambda}}{\tilde{\sigma}_n}\right)^{|S_j|+|T_j|}\gamma_{n_2} \\
= \;& \left(1+\frac{2a\omega\bar{\lambda}}{\tilde{\sigma}_n}\right)^{jn_1}\gamma_{n_2}.
\end{aligned}$$

From (5.6), we thus obtain the bound

$$\begin{aligned}
&\sum_{j=2}^{k}\left|\mathrm{Cov}\left(\phi\left(\frac{\omega}{\tilde{\sigma}_n}\sum_{i=1}^{j-1}\eta_{in}\right),\psi\left(\frac{\omega}{\tilde{\sigma}_n}\eta_{jn}\right)\right)\right| \\
\leq \;& k\left(1+\frac{2a\omega\bar{\lambda}}{\tilde{\sigma}_n}\right)^{kn_1}\gamma_{n_2} + k^2 n_1\frac{2\omega\bar{\lambda}^{2+\epsilon}}{\tilde{\sigma}_n}a^{-1-\varepsilon}E\left|X_t\right|^{2+\varepsilon}.
\end{aligned}$$

Fix $\kappa \in (0, 1/2)$ and set $a = a_n = n^\kappa$. Since $\kappa < 1/2$, we have

$$\left(1 + 2a_n\omega\bar{\lambda}/\tilde{\sigma}_n\right)^{\tilde{\sigma}_n/a_n} = \exp\left(2\omega\bar{\lambda}\right) + o\left(1\right),$$

and so

$$\begin{aligned}&\sum_{j=2}^{k}\left|\mathrm{Cov}\left(\phi\left(\frac{\omega}{\tilde{\sigma}_n}\sum_{i=1}^{j-1}\eta_{in}\right), \psi\left(\frac{\omega}{\tilde{\sigma}_n}\eta_{jn}\right)\right)\right| \\ = \quad & O\left(k\exp\left(\frac{2\omega\bar{\lambda}a_n k n_1}{\tilde{\sigma}_n} - n_2^\delta\right)\right) + O\left(k^2 n^{-1/2} n_1 a_n^{-1-\varepsilon}\right).\end{aligned}$$

Let $n_1 = \left[n^\beta\right]$ and $n_2 = [n^\alpha]$, with $0 < \alpha < \beta < 1$. We now have

$$\begin{aligned}&\sum_{j=2}^{k}\left|\mathrm{Cov}\left(\phi\left(\frac{\omega}{\tilde{\sigma}_n}\sum_{i=1}^{j-1}\eta_{in}\right), \psi\left(\frac{\omega}{\tilde{\sigma}_n}\eta_{jn}\right)\right)\right| \\ = \quad & O\left(n^{1-\beta}\exp\left(O\left(n^{\gamma+1/2}\right) - [n^\alpha]^\delta\right)\right) + O\left(n^{3/2-\beta-\kappa-\varepsilon\kappa}\right).\end{aligned}$$

Therefore, in order to satisfy (5.5), we require the following two inequalities to hold:

$$\begin{aligned}\kappa &< \alpha\delta - \frac{1}{2} \\ \kappa &> \frac{3-2\beta}{2+2\varepsilon}.\end{aligned}$$

By choosing $\alpha$ and $\beta$ close to one, we can reduce the inequalities to

$$\begin{aligned}\kappa &< \delta - \frac{1}{2} \\ \kappa &> \frac{1}{2+2\varepsilon}.\end{aligned}$$

We can find $\kappa \in (0, 1/2)$ satisfying these inequalities provided that

$$\frac{1}{2+2\varepsilon} < \delta - \frac{1}{2},$$

which is true since $\delta > (2+\varepsilon)/(2+2\varepsilon)$. Thus (5.5) holds, and so Withers' lemma yields $\tilde{\sigma}_n^{-1}\sum_1^n X_{tn} \to_d N(0,1)$. In view of (5.2), we conclude that (5.1) holds. This demonstrates the finite dimensional convergence of $B_n$ to $B$.

To prove tightness of $B_n$, it suffices for us to show that, for any $\varepsilon > 0$, there exists $\lambda > 1$ such that

$$\lambda^2 P\left\{\max_{1\le m\le n}\left|\sum_{t=1}^{m} X_t\right| \ge \lambda\sigma\sqrt{n}\right\} \le \varepsilon$$

for all $n$; see ch. 3 in (3). For $\delta \in (0, \varepsilon)$, we have from Theorem 4.1 that $E\left|\sum_1^n X_t\right|^{2+\delta} \le bn^{1+\delta/2}$ for some constant $b < \infty$. Thus, Theorem 3.1 in (9) implies that $E\max_{1\le m\le n}\left|\sum_1^m X_t\right|^{2+\delta} \le b' n^{1+\delta/2}$ for some constant $b' < \infty$. Markov's inequality thus gives us

$$\begin{aligned}\lambda^2 P\left\{\max_{1\le m\le n}\left|\sum_{t=1}^m X_t\right| \ge \lambda\sigma\sqrt{n}\right\} &\le \lambda^2 \frac{E\max_{1\le m\le n}\left|\sum_{t=1}^m X_t\right|^{2+\delta}}{\lambda^{2+\delta}\sigma^{2+\delta}n^{1+\delta/2}} \\ &\le \frac{b'}{\lambda^\delta\sigma^{2+\delta}},\end{aligned}$$

which can be made arbitrarily small by choosing $\lambda$ sufficiently large. This completes the proof of tightness, and therefore proves weak convergence of $B_n$ to $B$. □

It is interesting to note that the rate of decay of mixing coefficients required in Theorem 5.1 is substantially stronger than that required by invariance principles for $\alpha$-mixing processes. For instance, the functional central limit theorem for $\alpha$-mixing processes in (10) requires that $\sum_1^\infty \alpha_r^{\varepsilon/(\varepsilon+2)} < \infty$ and $E\left|X_t\right|^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$. Thus, in the case of bounded random variables, the memory condition $\alpha_r = O\left(r^{-\delta}\right)$, $\delta > 1$, would be sufficient for $\alpha$-mixing processes to satisfy a functional central limit theorem, whereas the analogous condition for bounded $\gamma$-mixing processes is $\gamma_r = O\left(\exp\left(r^{-\delta}\right)\right)$, $\delta > 1/2$.

This difference in memory conditions can be understood by comparing the covariance inequality in (14) for bounded $\alpha$-mixing processes (restated above as inequality (3.1)) with our own covariance inequality for bounded $\gamma$-mixing processes (Theorem 3.1). A key step in the Bernstein block approach to proving central limit theory for dependent random variables is the verification that

$$\sum_{j=1}^{k-1}\left|\mathrm{Cov}\left(\exp\left(\frac{i\omega}{\sigma_n}\sum_{s=1}^{j-1}\eta_s\right), \exp\left(\frac{i\omega}{\sigma_n}\eta_k\right)\right)\right| \to 0 \tag{5.7}$$

for any $\omega > 0$. Here, $\sigma_n^2 = E\left(\sum_1^n X_t\right)^2$, and $\{X_t : 1 \le t \le n\}$ is divided into $k$ Bernstein blocks $\eta_s = \sum_{(s-1)(n_1+n_2)+1}^{sn_1+(s-1)n_2} X_t$ of length $n_1$ separated by gaps of length $n_2$. We let $n_1, n_2 \to \infty$, $n_1 = o(n)$, $n_2 = o(n_1)$. Let $\phi$ and $\psi$ denote sine or cosine functions. Evidently, (5.7) can be true only if

$$\mathrm{Cov}\left(\phi\left(\frac{\omega}{\sigma_n}\sum_{s=1}^{k-1}\eta_s\right), \psi\left(\frac{\omega}{\sigma_n}\eta_k\right)\right) \to 0. \tag{5.8}$$

In the case of bounded $\alpha$-mixing processes, the r.v.s $\phi\left(\left(\omega/\sigma_n\right)\sum_{s=1}^{k-1}\eta_s\right)$ and $\psi\left(\left(\omega/\sigma_n\right)\eta_k\right)$ are measurable functions of finite subsets of $\{X_t\}$ separated by a gap of length $n_2$, and so inequality (3.1) yields

$$\left|\mathrm{Cov}\left(\phi\left(\frac{\omega}{\sigma_n}\sum_{s=1}^{k-1}\eta_s\right), \psi\left(\frac{\omega}{\sigma_n}\eta_k\right)\right)\right| \le 4\alpha_{n_2}.$$

(5.8) will thus be achieved under any rate of decay of $\alpha$-mixing coefficients. Theorem 3.1 is much less powerful in this context, providing (via the binomial theorem) the bound

$$\left|\mathrm{Cov}\left(\phi\left(\frac{\omega}{\sigma_n}\sum_{s=1}^{k-1}\eta_s\right),\psi\left(\frac{\omega}{\sigma_n}\eta_k\right)\right)\right|\leq\left(1+c\frac{\sigma}{\sigma_n}\right)^{kn_1}\gamma_{n_2},$$

where $c=2a\omega/\sigma$, $\sigma=\lim n^{-1/2}\sigma_n$, and each variable $X_t$ is bounded in absolute value by $a<\infty$. As $n\to\infty$, the quantity $(1+c\sigma/\sigma_n)^{kn_1}$ grows like $\exp\left(cn^{1/2}\right)$. We are thus lead to assume $\gamma_r=O\left(\exp\left(r^{-\delta}\right)\right)$, $\delta>1/2$, in order to achieve (5.8).

**Acknowledgements**

This paper is a version of the second chapter in my doctoral dissertation, written at Yale University. I thank my advisor Peter Phillips and committee members Donald Andrews and Yuichi Kitamura for their support and advice.

Financial support from the Cowles Foundation under a Carl Arvid Anderson Prize Fellowship is gratefully acknowledged.

## References

[1] Andrews, D. W. K. (1984). Non strong mixing autoregressive processes. *J. Appl. Probab.* **21** 930–934.
[2] Beare, B. K. (2007). A generalization of Hoeffding's lemma, and a new class of covariance inequalities. Unpublished manuscript.
[3] Billingsley, P. (1968). *Weak Convergence of Probability Measures.* Wiley, New York.
[4] Davydov, Y. A. (1970). The invariance principle for stationary processes. *Theory Probab. Appl.* **15** 487–498.
[5] Doukhan, P. (1994). *Mixing: Properties and Examples. Lecture Notes in Statistics* **85**. Springer-Verlag, New York.
[6] Doukhan, P. and Louhichi, S. (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications* **84** 313–342.
[7] Gorodetskii, V. V. (1977). On the strong mixing property for linear sequences. *Theory Probab. Appl.* **22** 411–413.
[8] Ibragimov, I. A. and Linnik, I. V. (1971). *Independent and Stationary Sequences of Random Variables.* Walter Nordhoof Pub., Groningen.
[9] Moricz, F. A., Serfling, R. J., and Stout, W. F. (1982). Moment and probability bounds with quasi super-additive structure for the maximum partial sum. *Ann. Probab.* **10** 1032–1040.
[10] Oodaira, H. and Yoshihara, K. I. (1972). Functional central limit theorems for strictly stationary processes satisfying the strong condition. *Kodai Math. Sem. Rep.* textbf24 259–269.

[11] Owen, A. B. (2004). Multidimensional variation for quasi-Monte Carlo. Technical report no. 2004-02, Department of Statistics, Stanford University.

[12] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Ac. Sc. U.S.A.* **42** 43–47.

[13] Utev, S. (1985). Inequalities for sums of weakly dependent random variables and estimates of the convergence rate in the invariance principle. In Borovkov, A. A., ed. *Advances in Probability Theory: Limit Theorems for Sums of Random Variables.* Optimization Software Inc., New York.

[14] Volkonskii, V. A. and Rozanov, Y. A. (1959). Some limit theorems for random functions, Part I. *Theory Probab. Appl.* **4** 178–197.

[15] Withers, C. S. (1981). Conditions for linear processes to be strong-mixing. *Z. Wahrsch. Verw. Gebiete* **57** 477–480.

[16] Withers, C. S. (1981). Central limit theorems for dependent variables. I. *Z. Wahrsch. Verw. Gebiete* **57** 509–534.