

A COARSENING OF THE STRONG MIXING CONDITION

BRENDAN K. BEARE

ABSTRACT. We consider a generalization of the α -mixing condition of Rosenblatt, which we term γ -mixing. Whereas α -mixing is defined in terms of entire σ -fields of sets generated by random variables in the distant past and future, γ -mixing is defined in terms of a more coarse collection of sets. We provide a Rosenthal inequality and central limit theorem for γ -mixing processes.

1. Introduction

Let $\{X_t : t \in \mathbb{Z}\}$ be a collection of random variables defined on some probability space (Ω, \mathcal{F}, P) . Mixing conditions provide one way to formalize the notion that these random variables are only weakly dependent on one another. There are many ways to define mixing; the monographs by Doukhan [8] and Bradley [5] list five classical definitions. The oldest and most general of these is the α -mixing condition of Rosenblatt [13, 4], also known as strong mixing. For any nonempty set of integers T , let $\mathcal{F}_T \subset \mathcal{F}$ denote the σ -field generated by the random variables $\{X_t : t \in T\}$. The α -mixing coefficients $\{\alpha_r : r \in \mathbb{N}\}$ associated with $\{X_t\}$ are given by

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

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

In this paper we investigate a generalization of α -mixing obtained by coarsening the families \mathcal{F}_S and \mathcal{F}_T appearing in (1.1). 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 define a sequence of γ -mixing coefficients $\{\gamma_r : r \in \mathbb{N}\}$ by

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

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 $\gamma_r \rightarrow 0$ as $r \rightarrow \infty$, we say that $\{X_t\}$ is γ -mixing.

Several other authors [12, 11, 7, 6] have investigated a coarsening of \mathcal{F}_S and \mathcal{F}_T in (1.1). The discussion in Dedecker and Prieur [7] is especially relevant. Those authors consider, among other dependence coefficients, a generalized α -mixing coefficient $\tilde{\alpha}_r$ proposed originally by [12]. This mixing coefficient is introduced in

Received 2014-7-8; Communicated by the editors.

2010 *Mathematics Subject Classification*. Primary 37A25; Secondary 60F05.

Key words and phrases. Mixing, weak dependence, Hardy-Krause variation.

Definition 2 of [7] using the notation $\alpha(r)$. After dividing by a constant factor of two, we may write $\tilde{\alpha}_r$ as

$$\tilde{\alpha}_r = \sup_{S, T} \sup_{A \in \mathcal{F}_S, B \in \mathcal{H}_T} |P(A \cap B) - P(A)P(B)|, \quad (1.3)$$

where this time the first supremum is taken over all nonempty finite sets of integers S, T such that $\min T - \max S \geq r$, and such that T is a singleton. Compared to (1.2), the set A in (1.3) is drawn from a larger collection of sets, while the set B is drawn from a smaller collection of sets. Clearly, $\tilde{\alpha}_r \leq \alpha_r$. We will shortly give an example of a process that is γ -mixing but not $\tilde{\alpha}$ -mixing, demonstrating that the γ -mixing property is more general than α -mixing, and distinct from $\tilde{\alpha}$ -mixing.

The main results of our paper are a Rosenthal inequality and central limit theorem for γ -mixing processes. The key to establishing them is a covariance inequality given in [3], which allows us to bound the covariance between two functions of a γ -mixing process by a quantity depending on the Hardy-Krause total variation norms of those functions. Our Rosenthal inequality represents a strict improvement over existing results for α -mixing processes: there is no cost to the coarsening of \mathcal{F}_S and \mathcal{F}_T that we adopt. The same cannot be said of our central limit theorem, which requires a much faster mixing rate than comparable results under α -mixing.

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

2. A Process That Is γ -mixing But Not $\tilde{\alpha}$ -mixing

Let $\{\varepsilon_t : t \in \mathbb{Z}\}$ be an iid 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 $\sum_{k=0}^{\infty} 2^{-k} \varepsilon_{t-k}$. One may 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 α -mixing by the construction of a set $A \in \sigma(X_0)$ and a sequence of sets $\{B_r\}$, $B_r \in \sigma(X_r)$, such that

$$|P(A \cap B_r) - P(A)P(B_r)| \geq 1/4 \quad (2.1)$$

for all $r \in \mathbb{N}$. Let $W_r = \{w_{r,1}, \dots, w_{r,m_r}\}$ denote the support of the random variable $X_r - 2^{-r}X_0$, and note that $m_r \leq 2^r$. Let $A = \{X_0 \leq 1/2\}$, and let

$$B_r = \left\{ X_r \in \bigcup_{k=1}^{m_r} [w_{r,k}, w_{r,k} + 2^{-r-1}] \right\}.$$

Now, since $X_0 \sim U(0, 1)$, we have $P(A) = 1/2$. And since $X_r = 2^{-r}X_0 + w_{r,k}$ for some $k = 1, \dots, m_r$, we have $A \subseteq B_r$. Consequently,

$$|P(A \cap B_r) - P(A)P(B_r)| = \frac{1}{2}(1 - P(B_r)).$$

But since $X_r \sim U(0, 1)$, we have $P(B_r) \leq m_r 2^{-r-1} \leq 1/2$. Thus (2.1) holds, and $\{X_t\}$ cannot be α -mixing.

Though $\{X_t\}$ is not α -mixing, it is $\tilde{\alpha}$ -mixing [7], with a geometric decay rate of $\tilde{\alpha}_r$. We can show that $\{X_t\}$ is also γ -mixing, with a geometric decay rate of γ_r .

Theorem 2.1. $\{X_t\}$ is γ -mixing, with $\gamma_r \leq 2^{1-r}$.

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

$$\begin{aligned} |P(A_x \cap B_y) - P(A_x)P(B_y)| &= \left| \int_{A_x} (P(B_y|\mathcal{F}_S) - P(B_y)) \, dP \right| \\ &\leq \frac{1}{2} E |P(B_y|\mathcal{F}_S) - P(B_y)|. \end{aligned}$$

Let \bar{s} denote the maximum element of S . Using the triangle inequality and the independence of $\cap_{t \in T} \{X_t - 2^{\bar{s}-t} X_{\bar{s}} \leq y_t\}$ and \mathcal{F}_S , we have

$$\begin{aligned} |P(B_y|\mathcal{F}_S) - P(B_y)| &\leq |P(\cap_{t \in T} \{X_t - 2^{\bar{s}-t} X_{\bar{s}} \leq y_t\} | \mathcal{F}_S) - P(B_y|\mathcal{F}_S)| \\ &\quad + |P(\cap_{t \in T} \{X_t - 2^{\bar{s}-t} X_{\bar{s}} \leq y_t\}) - P(B_y)|. \end{aligned}$$

Since $X_{\bar{s}}$ is nonnegative, we know that $B_y \subseteq \cap_{t \in T} \{X_t - 2^{\bar{s}-t} X_{\bar{s}} \leq y_t\}$, and so

$$\begin{aligned} |P(B_y|\mathcal{F}_S) - P(B_y)| &\leq P((\cap_{t \in T} \{X_t - 2^{\bar{s}-t} X_{\bar{s}} \leq y_t\}) \cap (\cup_{t \in T} \{X_t > y_t\}) | \mathcal{F}_S) \\ &\quad + P((\cap_{t \in T} \{X_t - 2^{\bar{s}-t} X_{\bar{s}} \leq y_t\}) \cap (\cup_{t \in T} \{X_t > y_t\})), \end{aligned}$$

from which it follows that

$$E |P(B_y|\mathcal{F}_S) - P(B_y)| \leq 2P((\cap_{t \in T} \{X_t - 2^{\bar{s}-t} X_{\bar{s}} \leq y_t\}) \cap (\cup_{t \in T} \{X_t > y_t\})).$$

The fact that $X_{\bar{s}} \leq 1$ now gives

$$\begin{aligned} E |P(B_y|\mathcal{F}_S) - P(B_y)| &\leq 2P((\cap_{t \in T} \{X_t - 2^{\bar{s}-t} \leq y_t\}) \cap (\cup_{t \in T} \{X_t > y_t\})) \\ &\leq 2P(\cup_{t \in T} \{y_t < X_t \leq y_t + 2^{\bar{s}-t}\}) \\ &\leq 2 \sum_{t \in T} P(y_t < X_t \leq y_t + 2^{\bar{s}-t}). \end{aligned}$$

The marginal distribution of each X_t is uniform on $[0, 1]$, and so

$$E |P(B_y|\mathcal{F}_S) - P(B_y)| \leq 2 \sum_{t \in T} 2^{\bar{s}-t} \leq 2 \sum_{t=\bar{s}+r}^{\infty} 2^{\bar{s}-t} = 2^{2-r}.$$

It follows that $\gamma_r \leq 2^{1-r}$ for all r . □

Theorem 2.1 demonstrates that $\{X_t\}$ is γ -mixing. But $\{X_t\}$ is also $\tilde{\alpha}$ -mixing, so we have yet to provide an example of a process that is γ -mixing but not $\tilde{\alpha}$ -mixing. In fact, this is now quite easy to achieve: we need merely consider the time reversed process $\{X_t^*\}$, where $X_t^* = X_{-t}$ for each $t \in \mathbb{Z}$. The time reversed process satisfies the dynamic equation $X_t^* = 2X_{t-1}^* \text{ mod}(1)$ a.s., and has been studied as an example of deterministic chaotic dynamics [2, 9, 14].

Theorem 2.2. $\{X_t^*\}$ is γ -mixing but not $\tilde{\alpha}$ -mixing, with $\gamma_r \leq 2^{1-r}$ and $\tilde{\alpha}_r \geq 1/4$.

Proof. $\gamma_r \leq 2^{1-r}$ follows from Theorem 2.1 and the invariance of γ_r under time reversal. $\tilde{\alpha}_r \geq 1/4$ follows by precisely the same argument used in [1] to show that $\{X_t\}$ is not α -mixing, repeated in the second paragraph of this section. Specifically, $B_r \in \sigma(X_{-r}^*)$ and $A = \{X_0^* \leq 1/2\}$, so from (1.3) we obtain $\tilde{\alpha}_r \geq |P(B_r \cap A) - P(B_r)P(A)| \geq 1/4$. \square

3. Covariance Inequalities

The following covariance inequality for a random process $\{X_t\}$ is well known [8, 5]: for any $r \in \mathbb{N}$, any nonempty finite sets of integers S and T such that $\min T - \max S \geq r \geq 1$, and any Borel measurable functions $f : \mathbb{R}^{|S|} \rightarrow \mathbb{R}$ and $g : \mathbb{R}^{|T|} \rightarrow \mathbb{R}$, we have

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

An inequality similar to (3.1) that involves γ -mixing coefficients rather than α -mixing coefficients has been proved in [3]. Before stating this inequality, we review the definitions of Vitali and Hardy-Krause variation for multivariate functions. For a more extensive discussion of these concepts, refer to [3, 10].

Let f be a real valued function defined on an n -dimensional rectangle $[a, b] = \{x \in \mathbb{R}^n : a \leq x \leq b\}$, and let $R = [c, d] \subseteq [a, b]$ be a smaller n -dimensional rectangle. Let

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

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

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

The Vitali variation of f is given by

$$\|f\|_V = \sup \sum_{R \in \mathcal{A}} |\Delta_R f|,$$

with the supremum taken over all finite collections of n -dimensional rectangles $\mathcal{A} = \{R_1, \dots, R_m\}$ 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 $f : [a, b] \rightarrow \mathbb{R}$, 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

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

Vitali variation and Hardy-Krause variation are equal when $n = 1$, but when $n \geq 2$ Hardy-Krause variation may be greater than Vitali variation.

Our covariance inequality for γ -mixing processes is as follows.

Theorem 3.1. *Suppose each X_t takes values in a bounded interval $[a_t, b_t]$. Let $r \in \mathbb{N}$, and let S and T denote nonempty finite sets of integers with $\min T - \max S \geq r$.*

Then for any functions $f : \prod_{s \in S} [a_s, b_s] \rightarrow \mathbb{R}$ and $g : \prod_{t \in T} [a_t, b_t] \rightarrow \mathbb{R}$ 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 [3], and the definition of γ_r . □

Theorem 3.1 is applicable to bounded random variables. Given a particular choice of f and g , it may be possible 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. Fix $r \in \mathbb{N}$, and let S and T be nonempty finite sets of integers with $\min T - \max S \geq r$. Let $A_1 = (3^{|S|} - 1)(3^{|T|} - 1)$ and $A_2 = 2|S| + 2|T|$. Then for $p, q \in [1, \infty]$ satisfying $\sup_{t \in S \cup T} \|X_t\|_p < \infty$ 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 $A = A_1$ if $q = 1$, or $A = A_1^{1/q} A_2^{(q-1)/q} q (q-1)^{(1-q)/q}$ if $q > 1$.

Proof. If $\gamma_r = 0$ then \mathcal{F}_S and \mathcal{F}_T must be independent, in which case the theorem is trivial. Assume $\gamma_r > 0$. Let $\bar{X}_t = \min \{ \max \{ X_t, -a_t \}, a_t \}$, where $a_t = \|X_t\|_p c^{-q/p} \gamma_r^{-1/p}$ for some constant $c > 0$. Begin by writing

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

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

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

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

We will use Theorem 3.1 to bound the first term on the right-hand side of (3.2). Clearly $\{\bar{X}_t\}$ is γ -mixing, with mixing coefficients bounded by those of $\{X_t\}$. Let the functions $f : \prod_{s \in S} [-a_s, a_s] \rightarrow \mathbb{R}$ and $g : \prod_{t \in T} [-a_t, a_t] \rightarrow \mathbb{R}$ be given by $f(x_s : s \in S) = \prod_{s \in S} x_s$ and $g(x_t : t \in T) = \prod_{t \in T} x_t$. For nonempty $I \subseteq S$ we have $f_I(x_s : s \in I) = \left(\prod_{s \in I} x_s \right) \left(\prod_{s \in S \setminus I} a_s \right)$. 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:

$$\|f_I\|_V = \int_{\prod_{s \in I} [-a_s, a_s]} \left(\prod_{s \in S \setminus I} a_s \right) \prod_{s \in I} dx_s = 2^{|I|} \left(\prod_{s \in S} a_s \right).$$

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

$$\|f\|_{\text{HK}} = \left(\prod_{s \in S} a_s \right) \left(\sum_{s=1}^{|S|} \frac{|S|!}{(|S| - s)!s!} 2^s \right) = \left(\prod_{s \in S} a_s \right) (3^{|S|} - 1),$$

and similarly $\|g\|_{\text{HK}} = \left(\prod_{t \in T} a_t \right) (3^{|T|} - 1)$. It now follows from Theorem 3.1 that

$$\left| \text{Cov} \left(\prod_{s \in S} \bar{X}_s, \prod_{t \in T} \bar{X}_t \right) \right| \leq A_1 \left(\prod_{t \in S \cup T} a_t \right) \gamma_r = A_1 \left(\prod_{t \in S \cup T} \|X_t\|_p \right) c^{1-q} \gamma_r^{1/q}. \tag{3.5}$$

Combining (3.2) through (3.5), we obtain

$$\left| \text{Cov} \left(\prod_{s \in S} X_s, \prod_{t \in T} X_t \right) \right| \leq (A_1 c^{1-q} + A_2 c) \left(\prod_{t \in S \cup T} \|X_t\|_p \right) \gamma_r^{1/q}.$$

Minimizing $A_1 c^{1-q} + A_2 c$ over c yields the constant A . □

Note that if we choose S and T to be singletons containing t and $t + r$ respectively, and set $q = 1$, then Theorem 3.2 states that

$$|\text{Cov}(X_t, X_{t+r})| \leq 4 \|X_t\|_\infty \|X_{t+r}\|_\infty \gamma_r. \tag{3.6}$$

If instead $q > 1$, then the constant term $A = 4q(q - 1)^{(1-q)/q}$ achieves a maximum value of 8 at $q = 2$, and so we have

$$|\text{Cov}(X_t, X_{t+r})| \leq 8 \|X_t\|_p \|X_{t+r}\|_p \gamma_r^{1/q}. \tag{3.7}$$

Inequalities (3.6) and (3.7) resemble the classic covariance inequalities for α -mixing processes [5, Theorems 1.11 and 3.7], achieving the familiar constant terms of 4 and 8 in the bounded and unbounded cases respectively. Since our inequalities involve γ -mixing coefficients rather than α -mixing, they constitute a refinement of the classic inequalities.

4. Rosenthal Inequality

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 $E|\sum_1^n X_t|^p \leq bD_n(p, \varepsilon, X)$ for all n . A Rosenthal inequality for α -mixing processes is given in [8].

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 α -mixing processes is proved in two steps. First, using a covariance inequality for α -mixing processes, the Rosenthal inequality is proved for any even integer p . Second, the so-called interpolation lemma [15, 8] is used to extend the inequality to all real $p > 1$.

To prove a Rosenthal inequality for γ -mixing processes, we modify the arguments used in the α -mixing case in the following way. First, in place of the covariance inequality for α -mixing processes, we employ Corollary 3.1 from above, which applies to γ -mixing processes. Second, we modify the interpolation lemma so that it is applicable under γ -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 γ -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 γ -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 γ -mixing coefficients are dominated by $\{\gamma_r\}$. Set

$$\begin{aligned} a &= V_n(p, \varepsilon, X)^{1/p}, \\ \bar{X}_t &= \min \{ \max \{ X_t, -a \}, a \}, \\ \underline{X}_t &= X_t - \bar{X}_t, \\ Y_t &= \bar{X}_t - E\bar{X}_t, \\ Z_t &= \underline{X}_t - E\underline{X}_t. \end{aligned}$$

Jensen's inequality allows us to bound $E|\sum_1^n X_t|^p$ by

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

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

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

We thus have

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

Y_t , ξ_t and ζ_t are all nondecreasing transformations of X_t , and therefore all have γ -mixing coefficients that are dominated by $\{\gamma_r\}$. Thus, under the hypothesis of

the lemma, there exists $b_1 < \infty$ such that

$$E \left| \sum_{t=1}^n X_t \right|^p \leq 2^{p-1} (b_1 V_n(k, \varepsilon, Y))^{p/k} + 2^{2p+k-3} b_1 V_n(k, \varepsilon, \xi) + 2^{2p+k-3} b_1 V_n(k, \varepsilon, \zeta) + 2^{2p+k-2} \left(\sum_{t=1}^n E |Z_t|^{p/k} \right)^k.$$

In [15, 8] it is shown that $V_n(k, \varepsilon, Y) \leq 2^k V_n(p, \varepsilon, X)^4 k/p$, that $V_n(k, \varepsilon, \xi) \leq 2^{k+p} V_n(p, \varepsilon, X)$, that $V_n(k, \varepsilon, \zeta) \leq 2^{k+p} V_n(p, \varepsilon, X)$, and, for $p \geq k - \varepsilon$, that $(\sum_{t=1}^n E |Z_t|^{p/k})^k \leq 2^p V_n(p, \varepsilon, X)$. We thus obtain $E |\sum_{t=1}^n X_t|^p \leq b_2 V_n(p, \varepsilon, X)$ 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 $p \in [0, k]$. \square

With Lemma 4.1 in hand, we may state our Rosenthal inequality for γ -mixing processes.

Theorem 4.2. *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 γ -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 ε such that, for all n ,*

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

Proof. The proof of this theorem differs from the proof under α -mixing – see e.g. [8, Section 1.4.1] – in only two respects. First, Theorem 3.2 is used in place of the covariance inequality for α -mixing processes. Second, Lemma 4.1 is used in place of the interpolation lemma [15, 8] for α -mixing processes. \square

Note that the only difference between Theorem 4.1 and the Rosenthal inequality for α -mixing processes stated in [8] is that we have replaced α -mixing coefficients with γ -mixing coefficients. Theorem 4.1 thus represents a strict refinement of that result.

5. Central Limit Theorem

In this section we prove a central limit theorem for stationary γ -mixing processes.

Theorem 5.1. *Suppose $\{X_t\}$ is stationary, and satisfies $EX_0 = 0$, $E|X_0|^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$, and $\gamma_r = O(\exp(-r^\delta))$ for some $\delta > (4 + \varepsilon)/(4 + 2\varepsilon)$ and all $r \in \mathbb{N}$. Then $\sum_{r=1}^\infty |EX_0 X_r| < \infty$, and if $\sigma^2 := EX_0^2 + 2 \sum_{r=1}^\infty EX_0 X_r > 0$, then $n^{-1/2} \sum_{t=1}^n X_t \rightarrow_d N(0, \sigma^2)$ as $n \rightarrow \infty$.*

Proof. Absolute convergence of $\sum_{r=1}^\infty EX_0 X_r$ follows from Theorem 3.2. Suppose $\sigma > 0$. We will show that $n^{-1/2} \sum_{t=1}^n X_t \rightarrow_d N(0, \sigma^2)$ using a lemma of Withers

[16, Lemma 3.1]. Split $\{X_t\}$ into k Bernstein blocks of length n_1 , separated by gaps of length n_2 , as follows:

$$\begin{aligned} \sum_{t=1}^n X_t &= \sum_{i=1}^k \eta_{in} + \sum_{i=1}^{k+1} \nu_{in}, & k &= \left\lceil \frac{n}{n_1+n_2} \right\rceil \\ \eta_{in} &= \sum_{t=(i-1)(n_1+n_2)+1}^{i(n_1+n_2)} X_t, & i &= 1, \dots, k \\ \nu_{in} &= \sum_{t=in_1+(i-1)n_2+1}^n X_t, & i &= 1, \dots, k \\ \nu_{k+1,n} &= \sum_{t=k(n_1+n_2)+1}^n X_t. \end{aligned}$$

The sequences $n_1(n)$ and $n_2(n)$ are chosen to satisfy $n_1 \sim n^\beta$ and $n_2 \sim n^\alpha$, where $0 < \alpha < \beta < 1$. Withers' lemma states that $n^{-1/2} \sum_{t=1}^n X_t \rightarrow_d N(0, \sigma^2)$ if the following four conditions are satisfied for ϕ, ψ being either sine or cosine functions:

$$\frac{1}{n} E \left(\sum_{i=1}^{k+1} \nu_{in} \right)^2 \rightarrow 0 \tag{5.1}$$

$$\frac{1}{n} \sum_{i=1}^k E \eta_{in}^2 \mathbf{1}(\eta_{in}^2 > n\epsilon) \rightarrow 0 \text{ for all } \epsilon > 0 \tag{5.2}$$

$$\frac{1}{n} \sum_{i=1}^{k-1} \sum_{j=i+1}^k \text{Cov}(\eta_{in}, \eta_{jn}) \rightarrow 0 \tag{5.3}$$

$$\sum_{j=2}^k \left| \text{Cov} \left(\phi \left(\omega n^{-1/2} \sum_{i=1}^{j-1} \eta_{in} \right), \psi \left(\omega n^{-1/2} \eta_{jn} \right) \right) \right| \rightarrow 0 \text{ for all } \omega > 0. \tag{5.4}$$

(We have simplified Withers' conditions by noting that $E(\sum_{t=1}^n X_t)^2 \sim n\sigma^2$; see e.g. [5, Prop. 8.3(IV)]). To verify (5.1), we note that Theorem 4.1 implies that $E(\sum_{i=1}^{k+1} \nu_{in})^2 = O(kn_2) = o(n)$. To verify (5.2), we use the inequalities of Hölder and Markov to obtain

$$E \eta_{in}^2 \mathbf{1}(\eta_{in}^2 > n\epsilon) \leq \|\eta_{in}\|_{2+\epsilon}^2 P(\eta_{in}^2 > n\epsilon)^{\epsilon/(2+\epsilon)} \leq (n\epsilon)^{-\epsilon/2} \|\eta_{in}\|_{2+\epsilon}^{2+\epsilon}.$$

Theorem 4.1 implies that $\|\eta_{in}\|_{2+\epsilon} = O(n_1^{1/2})$, and so the left-hand side of (5.2) is $O(k(n_1/n)^{1+\epsilon/2}) = O((n_1/n)^{\epsilon/2}) = o(1)$. To verify (5.3), we use (3.7) to obtain

$$|\text{Cov}(\eta_{in}, \eta_{jn})| \leq 8n_1^2 \|X_0\|_{2+\epsilon}^2 \gamma_{(j-i)n_2}^{\epsilon/(2+\epsilon)}$$

for $1 \leq i < j \leq k$. It follows that the left-hand side of (5.3) is $O(n\gamma_{n_2}^{\epsilon/(2+\epsilon)}) = o(1)$.

It remains to verify (5.4). Let $n_3 = n_3(n)$ be an increasing sequence satisfying $n_3 \sim n^\kappa$ for some $\kappa > 0$, and let $X_{tn} = \min\{n_3, \max\{X_t, -n_3\}\}$. For $j = 2, \dots, k$, define

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

Using Markov’s inequality and the boundedness of ϕ and ψ , we may show that

$$\begin{aligned} & \left| \text{Cov} \left(\phi \left(\omega n^{-1/2} \sum_{i=1}^{j-1} \eta_{in} \right), \psi \left(\omega n^{-1/2} \eta_{jn} \right) \right) \right| \\ & \leq \left| \text{Cov} \left(\phi \left(\omega n^{-1/2} \sum_{s \in S_j} X_{sn} \right), \psi \left(\omega n^{-1/2} \sum_{t \in T_j} X_{tn} \right) \right) \right| \\ & \quad + 4jn_1n_3^{-2-\varepsilon} \|X_0\|_{2+\varepsilon}^{2+\varepsilon}. \end{aligned} \tag{5.5}$$

We will use Theorem 3.1 to bound the first term on the right-hand side of (5.5). Let $f : [-n_3, n_3]^{|S_j|} \rightarrow \mathbb{R}$ and $g : [-n_3, n_3]^{|T_j|} \rightarrow \mathbb{R}$ be given by

$$\begin{aligned} f(x_s : s \in S_j) &= \phi \left(\omega n^{-1/2} \sum_{s \in S_j} x_s \right) \\ g(x_t : t \in T_j) &= \psi \left(\omega n^{-1/2} \sum_{t \in T_j} x_t \right). \end{aligned}$$

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

$$f_I(x_s : s \in I) = \phi \left(\omega n^{-1/2} \left(\sum_{s \in I} x_s + n_3 |S_j \setminus I| \right) \right).$$

The function obtained by differentiating f_I once with respect to each argument is bounded in absolute value by $(\omega n^{-1/2})^{|I|}$. Thus, $\|f_I\|_V \leq (2\omega n_3 n^{-1/2})^{|I|}$. Using the binomial theorem, we now have

$$\begin{aligned} \|f\|_{\text{HK}} &\leq \sum_{\emptyset \neq I \subseteq S_j} (2\omega n_3 n^{-1/2})^{|I|} \\ &= \sum_{s=1}^{|S_j|} \frac{|S_j|!}{(|S_j| - s)! s!} (2\omega n_3 n^{-1/2})^s \\ &= (1 + 2\omega n_3 n^{-1/2})^{|S_j|} - 1. \end{aligned}$$

We can show similarly that $\|g\|_{\text{HK}} \leq (1 + 2\omega n_3 n^{-1/2})^{|T_j|} - 1$. Thus, since the γ -mixing coefficients of $\{X_{tn}\}$ are dominated by those of $\{X_t\}$, Theorem 3.1 implies that the first term on the right-hand side of (5.5) is bounded by

$$\|f\|_{\text{HK}} \|g\|_{\text{HK}} \gamma_{n_2} \leq (1 + 2\omega n_3 n^{-1/2})^{|S_j|+|T_j|} \gamma_{n_2} = (1 + 2\omega n_3 n^{-1/2})^{jn_1} \gamma_{n_2}.$$

It follows that the quantity on the left-hand side of (5.4) is bounded by

$$k (1 + 2\omega n_3 n^{-1/2})^{kn_1} \gamma_{n_2} + 4k^2 n_1 n_3^{-2-\varepsilon} \|X_0\|_{2+\varepsilon}^{2+\varepsilon}.$$

Recall that $n_1 \sim n^\beta$, $n_2 \sim n^\alpha$, $n_3 \sim n^\kappa$ and $k \sim n^{1-\beta}$ for parameters α, β, κ satisfying $0 < \alpha < \beta < 1$ and $\kappa > 0$, and recall that $E|X_0|^{2+\varepsilon} < \infty$ and $\gamma_r =$

$O(\exp(-r^\delta))$ as $r \rightarrow \infty$. We therefore have

$$k \left(1 + 2\omega n_3 n^{-1/2}\right)^{kn_1} \gamma_{n_2} = O\left(n^{1-\beta} \left(1 + 2\omega n^{\kappa-1/2}\right)^n \exp(-n^{\alpha\delta})\right) \quad (5.6)$$

$$4k^2 n_1 n_3^{-2-\varepsilon} \|X_0\|_{2+\varepsilon}^{2+\varepsilon} = O\left(n^{2-\beta-2\kappa-\varepsilon\kappa}\right). \quad (5.7)$$

If we choose $\kappa < 1/2$, then $(1 + 2\omega n^{\kappa-1/2})^{n^{1/2-\kappa}} \sim \exp(2\omega)$, and the expression in (5.6) is $O(n^{1-\beta} \exp(n^{\kappa+1/2} - n^{\alpha\delta}))$. We may ensure that it vanishes by choosing α, κ to satisfy $\kappa < \alpha\delta - 1/2$. If, in addition, $\kappa > (2 - \beta)/(2 + \varepsilon)$, then the expression in (5.7) also vanishes, and (5.4) is satisfied. We can find κ to satisfy these conditions whenever α, β are such that $(2 - \beta)/(2 + \varepsilon) < 1/2$ and $(2 - \beta)/(2 + \varepsilon) < \alpha\delta - 1/2$. These two inequalities may be satisfied by choosing α, β sufficiently close to one, since the assumptions of our theorem imply that $1/(2 + \varepsilon) < \delta - 1/2$ \square

Note that the rate of γ -mixing required in Theorem 5.1 is substantially stronger than would be required under α -mixing. Using the central limit theorem given in [5, Theorem 10.7], we see that our γ -mixing condition may be replaced with the α -mixing condition $\sum_{r=1}^{\infty} \alpha_r^{\varepsilon/(2+\varepsilon)} < \infty$. Thus, in the case of bounded random variables, the memory condition $\alpha_r = O(r^{-\delta})$, $\delta > 1$, is sufficient for stationary α -mixing processes to satisfy a central limit theorem, whereas the analogous condition under Theorem 5.1 is $\gamma_r = O(\exp(-r^\delta))$, $\delta > 1/2$.

Acknowledgment. Parts of this paper derive from my doctoral research at Yale University. An earlier version was circulated under the title ‘‘A new mixing condition’’. I thank Donald Andrews, Yuichi Kitamura, Peter Phillips and Alessio Sancetta for helpful comments.

References

1. Andrews, D. W. K.: Non-strong mixing autoregressive processes, *J. Appl. Probab.* **21** (1984), no. 4, 930–934.
2. Bartlett, M. S.: Chance or chaos? *J. Roy. Statist. Soc. Ser. A* **153** (1990), no. 3, 321–347.
3. Beare, B. K.: A generalization of Hoeffding’s lemma and a new class of covariance inequalities, *Statist. Probab. Lett.* **79** (2009), no. 5, 637–642.
4. Blum, R. and Rosenblatt, M.: A class of stationary processes and a central limit theorem, *Proc. Nat. Ac. Sc. U.S.A.* **42** (1956), no. 7, 412–413.
5. Bradley, R. C.: *Introduction to Strong Mixing Conditions, vol. 1*, Kendrick Press, Heber City, 2007.
6. Dedecker, J., Doukhan, P., Lang, G., Le on, J. R., Louhichi, S. and Prieur, C.: *Weak Dependence: With Examples and Applications. Lecture Notes in Statistics vol. 190*, Springer, New York, 2007.
7. Dedecker, J. and Prieur, C.: New dependence coefficients. Examples and applications to statistics. *Probab. Theory Related Fields* **132** (2005), no. 2, 203–236.
8. Doukhan, P.: *Mixing: Properties and Examples. Lecture Notes in Statistics vol. 85*. Springer-Verlag, New York, 1994.
9. Lawrance, A. J.: Uniformly distributed first-order autoregressive time series models and multiplicative congruential random number generators, *J. Appl. Probab.* **29** (1992), no. 4, 896–903.
10. Owen, A. B.: Multidimensional variation for quasi-Monte Carlo. Technical report no. 2004-02, Department of Statistics, Stanford University, 2004.
11. Peligrad, M.: Some remarks on coupling of dependent random variables, *Statist. Probab. Lett.* **60** (2002), no. 2, 201–209.

12. Rio, E.: *Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants*, Springer-Verlag, Berlin, 2000.
13. Rosenblatt, M.: A central limit theorem and a strong mixing condition, *Proc. Nat. Ac. Sc. U.S.A.* **42** (1956), no. 1, 43–47.
14. Tong, H. and Chen, B.: A note on one-dimensional chaotic maps under time reversal, *Adv. Appl. Probab.* **24** (1992), no. 1, 219–220.
15. Utev, S.: 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, 1985.
16. Withers, C. S.: Central limit theorems for dependent variables, I, *Z. Wahrsch. Verw. Gebiete* **57** (1981), no. 4, 509–534.

BRENDAN K. BEARE: DEPARTMENT OF ECONOMICS, UNIVERSITY OF CALIFORNIA - SAN DIEGO,
LA JOLLA, CA 92101, USA
E-mail address: `bbeare@ucsd.edu`