A new CLT for additive functionals of Markov chains

Magda Peligrad

Department of Mathematical Sciences, University of Cincinnati, PO Box 210025, Cincinnati, Oh 45221-0025, USA.

email: peligrm@ucmail.uc.edu

Keywords: Markov chains, variance of partial sums, central limit theorem, projective criteria, absolute regularity.

Mathematical Subject Classification (2010): 60F05, 60G10, 60J05.

Abstract

In this paper we study the central limit theorem for additive functionals of stationary Markov chains with general state space by using a new idea involving conditioning with respect to both the past and future of the chain. Practically, we show that any stationary and ergodic Markov chain with \( \text{var}(S_n)/n \) uniformly bounded, satisfies a \( \sqrt{n} \)-central limit theorem with a random centering. We do not assume that the Markov chain is irreducible or aperiodic. However, the random centering is not needed if the Markov chain satisfies stronger forms of ergodicity.

1 Introduction

A basic result in probability theory is the central limit theorem. To go beyond the independent case, the dependence is often restricted by using projective criteria. For instance, the martingales are defined by using a projective condition with respect to the past sigma field. There also is an abundance of martingale-like conditions, which define classes of processes satisfying the CLT. Among them Gordin’s condition (Gordin, 1969), Heyde’s projective condition (Heyde, 1974), mixingales (McLeish, 1975), Maxwell and Woodroofe condition (2000), just to name a few. All of them have in common that the conditions are imposed to the conditional expectation of a variable with respect to the past sigma field.

There is, however, the following philosophical question. Note that a partial sum does not depend on the direction of time, i.e.

\[
S_n = X_1 + X_2 + \ldots + X_n = X_n + X_{n-1} + \ldots + X_1.
\]

However a condition of type ”martingale-like” depends on the direction of time. Therefore, in order to get results for \( S_n \), it is natural to also study projective conditions that are symmetric with respect to the direction of time. Furthermore, many mixing conditions (see Bradley, 2005, for a survey) and harnesses (see for instance Williams, 1973 and Mansuy and Yor, 2005) are independent of the direction of time.
In this paper we introduce a new idea, which involves conditioning with respect to both the past and the future of the process. By using this idea together with martingale approximation techniques, we shall prove that any Markov chain which is stationary and ergodic (in the ergodic theoretical sense) satisfies the CLT, provided that we use a random centering and we assume that $\text{var}(S_n)/n$ is uniformly bounded. In case when the stationary Markov chain satisfies stronger forms of ergodicity, the random centering is not needed. Among these classes are the absolutely regular Markov chains. For this class, our result gives as a corollary, a new interpretation of the limiting variance in the CLT in Theorem 2.3 of Chen (1999) and a totally different new approach.

Our paper is organized as follows. In Section 2 we present the results. Section 3 is dedicated to their proofs.

2 Results

Throughout the paper we assume that $(\xi_n)_{n \in \mathbb{Z}}$ is a stationary Markov chain defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in a measurable space $(S, \mathcal{A})$. Denote by $\mathcal{F}_n = \sigma(\xi_k, k \leq n)$ and by $\mathcal{F}^n = \sigma(\xi_k, k \geq n)$. The marginal distribution is denoted by $\pi(A) = \mathbb{P}(\xi_0 \in A)$. Next, let $L_0^2(\pi)$ be the set of measurable functions on $S$ such that $\int f^2 d\pi < \infty$ and $\int f d\pi = 0$. For a function $f \in L_0^2(\pi)$ let $X_i = f(\xi_i), S_n = \sum_{i=1}^n X_i$.

We denote by $||X||$ the norm in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Below, $\Rightarrow$ denotes the convergence in distribution and by $N(\mu, \sigma^2)$ we denote a normally distributed random variable with mean $\mu$ and variance $\sigma^2$.

2.1 Central limit theorem

We shall assume that the stationary Markov chain is ergodic in the ergodic theoretical sense, i.e. the invariant sigma field is trivial (see for instance Section 17.1 Meyn and Tweedie, 1996). We shall establish the following CLT.

**Theorem 1** Assume that

$$\sup_{n \geq 1} \frac{E(S_n^2)}{n} < \infty. \quad (1)$$

Then, the following limit exists

$$\lim_{n \to \infty} \frac{1}{n} ||S_n - E(S_n|\xi_0, \xi_n)||^2 = \sigma^2 \quad (2)$$

and

$$\frac{S_n - E(S_n|\xi_0, \xi_n)}{\sqrt{n}} \Rightarrow N(0, \sigma^2).$$

If the random centering is not present, we have the following result:
Corollary 2 Assume that (1) holds and in addition
\[
\frac{E(S_n|\xi_0, \xi_n)}{\sqrt{n}} \to p 0 \text{ as } n \to \infty.
\]
Then the following limit exists
\[
\lim_{n \to \infty} \frac{\pi}{2n} (E|S_n|)^2 = \sigma^2
\]
and
\[
\frac{S_n}{\sqrt{n}} \Rightarrow N(0, \sigma^2).
\] (3)

As an immediate consequence of Theorem (1) we give next necessary conditions for the CLT with the traditional limiting variance.

Corollary 3 Assume that (1) holds and in addition
\[
\lim_{n \to \infty} \frac{1}{n} ||E(S_n|\xi_0, \xi_n)||^2 = 0.
\] (4)
Then
\[
\lim_{n \to \infty} \frac{E(S_n^2)}{n} = \sigma^2
\] (5)
and
\[
\frac{S_n}{\sqrt{n}} \Rightarrow N(0, \sigma^2).
\] (6)

We can also give a sufficient condition for (4) in terms of individual random variables.

Proposition 4 Assume that (1) holds and the following condition is satisfied
\[
\lim_{n \to \infty} n ||E(X_0|\xi_0, \xi_n)||^2 = 0.
\] (7)
Then (3) and (6) hold.

In Subsection 2.2 in the context of absolutely regular Markov chains, we shall comment that condition (7) alone does not imply (6). However, a reinforced condition does:

Corollary 5 Assume that
\[
\sum_{k \geq 1} ||E(X_0|\xi_k, \xi_k)||^2 < \infty.
\] (8)
Then (3) and (6) hold.
2.2 Example: absolutely regular Markov chains

Given a stationary Markov chain \( \xi = (\xi_k)_{k \in \mathbb{Z}} \), with values in a Polish space, the coefficient of absolute regularity is defined as (see Proposition 3.22 (III,5) in Bradley, 2007)

\[
\beta_n = \beta_n(\xi) = E \left( \sup_{A \in \mathcal{B}} |P(\xi_n \in A| \xi_0) - P(\xi_0 \in A)| \right),
\]

where \( \mathcal{B} \) denotes the Borel sigma filed. The chain is called absolutely regular if \( \beta_n \to 0 \).

Let us mention that there are numerous examples of stationary absolutely regular Markov chains. For easy reference we refer to Section 3 in Bradley (2006) survey paper and to the references mentioned there. We know that a strictly stationary, countable state Markov chain is absolutely regular if and only if the chain is irreducible and aperiodic. Also, any strictly stationary Harris recurrent and aperiodic Markov chain is absolutely regular. It is also well-known that \( \beta_n \to 0 \) implies ergodicity in the ergodic theoretical sense. Also, in many situations these coefficients are tractable (see for instance Davydov, 1973, Doukhan et al. 1994, Bradley, 2007).

Due to their importance for the Monte Carlo simulations, the central limit theorem for Markov chains was intensively studied under the absolute regularity condition. In this direction we mention the books by Nummelin (1984), Meyn and Tweedie (1996) and Chen (1999) and we also refer to the survey paper by Jones (2004). By using regeneration techniques and partition in independent blocks (Nummelin’s splitting technique, 1978) it was proven that, in this setting, a necessary and sufficient condition for the CLT is that \( S_n / \sqrt{n} \) is stochastically bounded (Theorem 2.3 in Chen, 1999). However, the limit has a variance which is described in terms of the split chain and it is difficult to describe. Our next Corollary is obtained under a more general condition than in Theorem 3.1 in Chen (1999).

**Corollary 6** Assume that (1) holds and the sequence is absolutely regular. Then (3) holds.

To give a CLT where the limiting variance is \( \sigma^2 \) defined in (5), we shall verify condition (7) of Proposition 4. Denote by \( Q \) the quantile function of \( |X_0| \), i.e., the inverse function of \( t \mapsto \mathbb{P}(|X_0| > t) \). We obtain the following result:

**Corollary 7** Assume that (1) holds and the following condition is satisfied

\[
\lim_{n \to \infty} n \int_0^{\beta_n} Q^2(u) du = 0. \tag{9}
\]

Then (2) and (4) hold.

In terms of moments, by the Hölder inequality, (9) is implied by \( E(|X_0|^{2+\delta}) < \infty \) and \( n\beta_n^{2+\delta} \to 0 \), for some \( \delta > 0 \). If \( X_0 \) is bounded a.s., the mixing rate required for this corollary is \( n\beta_n \to 0 \).
Finally, condition (8) is verified if
\[
\sum_{n \geq 1} \int_0^{\beta_n} Q^2(u) \, du < \infty,
\]
and then (5) and (6) hold. Further reaching results could be found in Doukhan et al. (1994), where a larger class of processes was considered. According to Corollary 1 Doukhan et al. (1994), (10) alone is not enough for (6). Actually, the stronger condition (10) is a minimal condition for the CLT for \( S_n/\sqrt{n} \) in the following sense. In their Corollary 1, Doukhan et al. (1994) constructed a stationary absolutely regular Markov chain \((\xi_k)_{k \in \mathbb{Z}}\) and a function \(f \in L^2(\pi)\), which barely does not satisfies (10) and \( S_n/\sqrt{n} \) does not satisfy the CLT. For instance, this is the case when for an \( a > 1 \), \( \beta_n = cn^{-a} \) and \( Q^2(u) \) behaves as \( u^{-1+1/a} |\log u|^{-1} \) as \( u \to 0^+ \).

In this case
\[
\sum_{n \geq 1} \int_0^{\beta_n} Q^2(u) \, du = \int_0^1 \beta_n^{-1}(u)Q^2(u) \, du = \infty
\]
and, according to Corollary 1 in Doukhan et al. (1995), there exists a Markov chain with these specifications, such that \( S_n/\sqrt{n} \) does not satisfy the CLT. However, for these specifications (9) is satisfied. If in addition we know that (1) holds, then, by Corollary 7, the CLT holds for \( S_n/\sqrt{n} \).

Let us point out for instance, a situation where Corollary 7 is useful. Let \( Y = (Y_i) \) and \( Z = (Z_j) \) be two absolutely regular Markov chains of centered, bounded random variables, independent among them and satisfying the following conditions
\[
\sum_{n \geq 1} \beta_n(Y) < \infty \quad \text{and} \quad n \beta_n(Z) \to 0.
\]
If we define now the sequence \( X = (X_n) \), where for each \( n \) we set \( X_n = Y_n Z_n \), then, by Theorem 6.2 in Bradley (2007), we have \( \beta_n(X) \leq \beta_n(Y) + \beta_n(Z) \). As a consequence \( n \beta_n(X) \to 0 \) and condition (9) is satisfied. Certainly, this condition alone does not assure that the CLT holds. In order to apply Corollary 7 we have to verify that condition (1) holds. Conditioned by \( Z \) the partial sum of \((X_n)\) becomes a linear combination of the variables of \( Y \) and we can apply Corollary 7 in Peligrad and Utev (2006). It follows that there is a positive constant \( C \) such that
\[
E_Y \left( \sum_{k=1}^n Y_k Z_k \right)^2 \leq C \left( \sum_{k=1}^n Z_k^2 \right) \text{ a.s.,}
\]
where \( E_Y \) denotes the partial integral with respect to the variables of \( Y \). By the independence of the sequences \( Y \) and \( Z \) we obtain
\[
E \left( \sum_{k=1}^n Y_k Z_k \right)^2 \leq C n E(Z_0^2).
\]
Therefore condition (1) holds. By Corollary 7 we obtain that (5) and (6) hold for the sequence \((X_n)\).
3 Proofs

Proof of Theorem 1

The proof of this central limit theorem is based on the martingale approximation technique. Fix $m$ a positive integer and make consecutive blocks of size $m$. Denote by $Y_k$ the sum of variables in the $k$'th block. Let $u = u_n(m) = \lfloor n/m \rfloor$. So, for $k = 0, 1, ..., u - 1$, we have

$$ Y_k = Y_k(m) = (X_{km+1} + ... + X_{(k+1)m}). $$

Also denote

$$ Y_u = Y_u(m) = (X_{um+1} + ... + X_n). $$

For $k = 0, 1, ..., u - 1$ let us consider the random variables

$$ D_k = D_k(m) = \frac{1}{\sqrt{m}}(Y_k - E(Y_k|\xi_{km}, \xi_{(k+1)m})). $$

By the Markov property, conditioning by $\sigma(\xi_{km}, \xi_{(k+1)m})$ is equivalent to conditioning by $F_{km} \vee F_{(k+1)m}$. Note that $D_k$ is adapted to $F_{(k+1)m} = G_k$. Then we have $E(D_1|G_0) = 0$ a.s. and we have a stationary and ergodic sequence of square integrable martingale differences $(D_k, G_k)_{k \geq 0}$.

Therefore, by the classical central limit theorem for ergodic martingales, for every $m$, a fixed positive integer, we have

$$ \frac{1}{\sqrt{u}} M_u(m) := \frac{1}{\sqrt{u}} \sum_{k=0}^{u-1} D_k(m) \Rightarrow N_m, $$

where $N_m$ is a normally distributed random variable with mean $0$ and variance $m^{-1}||Y_0 - E(Y_0|\xi_0, \xi_m)||^2$. Now consider $(m')$ a subsequence of $N$ such that

$$ \lim_{m' \to \infty} \frac{1}{\sqrt{u}} ||Y_0 - E(Y_0|\xi_0, \xi_{m'})||^2 = \lim_{n \to \infty} \frac{1}{n} ||S_n - E(S_n|\xi_0, \xi_n)||^2 = \eta^2. \quad (11) $$

This means that

$$ N_{m'} \Rightarrow N(0, \eta^2) \text{ as } m' \to \infty. $$

Whence, according to Theorem 3.2 in Billingsley (1999), in order to establish the CLT for $\sum_{i=0}^{u-1} Y_i/\sqrt{n}$ we have only to show that

$$ \lim_{m' \to \infty} \limsup_{n \to \infty} \frac{1}{\sqrt{n}} (S_n - E(S_n|\xi_0, \xi_n)) - \frac{1}{\sqrt{u}} M_u(m') = 0. \quad (12) $$

Denote by $Z_k = m^{-1/2}E(Y_k|\xi_{km}, \xi_{(k+1)m})$ and let $R_u(m) = \sum_{k=0}^{u-1} Z_k$. Set

$$ S_u(m) = M_u(m) + R_u(m). \quad (13) $$
Let us show that $M_n(m)$ and $R_n(m)$ are orthogonal. We show this by analyzing
the expected value of all the terms of the product $M_n(m)R_n(m)$. Note that if
$j < k$, since $\mathcal{F}_{(j+1)m} \subset \mathcal{F}_{km}$, we have

$$E[(Y_k - E(Y_k|\xi_{km}, \xi_{(k+1)m}))E(Y_j|\xi_{jm}, \xi_{(j+1)m})] = E[E(Y_k - E(Y_k|\mathcal{F}_{km} \lor \mathcal{F}^{(k+1)m})|\mathcal{F}_{(j+1)m})E(Y_j|\xi_{jm}, \xi_{(j+1)m})] = 0.$$ 

On the other hand, if $j > k$, since $\mathcal{F}_m \subset \mathcal{F}^{(k+1)m}$ then

$$E[(Y_k - E(Y_k|\xi_{km}, \xi_{(k+1)m}))E(Y_j|\xi_{jm}, \xi_{(j+1)m})] = E[E(Y_k - E(Y_k|\mathcal{F}_{km} \lor \mathcal{F}^{(k+1)m})|\mathcal{F}_m)E(Y_j|\xi_{jm}, \xi_{(j+1)m})] = 0.$$ 

For $j = k$, by conditioning with respect to $\sigma(\xi_{km}, \xi_{(k+1)m})$, we note that

$$E[(Y_k - E(Y_k|\xi_{km}, \xi_{(k+1)m}))E(Y_k|\xi_{km}, \xi_{(k+1)m})] = 0.$$ 

Therefore $M_n(m)$ and $R_n(m)$ are indeed orthogonal. By using now the decom-
position (13), the fact that $M_n(m)$ and $R_n(m)$ are orthogonal and $M_n(m)$ is a
martingale, we obtain the identity

$$\frac{1}{u}||S_u(m)||^2 = \frac{1}{m}||S_m - E(S_m|\xi_0, \xi_m)||^2 + \frac{1}{u}||R_u(m)||^2. \quad (14)$$

Also, note that (1) and the definition of $Y_u$ imply that for some positive constant
$C$, we have $||Y_u|| \leq Cm$. Hence, by the properties of conditional expectations,
for every $m$, we have

$$||\frac{1}{\sqrt{n}}(S_n - E(S_n|\xi_0, \xi_n)) - \frac{1}{\sqrt{u}}(S_u(m) - E(S_u(m)|\xi_0, \xi_n))|| \to 0 \text{ as } n \to \infty.$$

(15)

Now, by using again the properties of conditional expectations and the identity
(14), for every $m$

$$\frac{1}{u}||S_u(m) - E(S_u(m)|\xi_0, \xi_n) - M_u(m)||^2 = \frac{1}{u}||R_u(m)||^2 - \frac{1}{u}||E(S_u(m)|\xi_0, \xi_n)||^2$$

$$= \frac{1}{u}(||S_u(m)||^2 - ||E(S_u(m)|\xi_0, \xi_n)||^2) - \frac{1}{u}||S_m - E(S_m|\xi_0, \xi_m)||^2$$

$$= \frac{1}{u}||S_u(m) - E(S_u(m)|\xi_0, \xi_n)||^2 - \frac{1}{u}||S_m - E(S_m|\xi_0, \xi_m)||^2.$$

By passing now to the limit in the last identity with $n \to \infty$, by (2) and (15) we obtain

$$\lim_{n \to \infty} \sup \frac{1}{\sqrt{n}}(S_n - E(S_n|\xi_0, \xi_n)) - \frac{1}{\sqrt{u}}M_u(m)||^2$$

$$= \eta^2 - \left(\frac{1}{m}||S_m - E(S_m|\xi_0, \xi_m)||^2\right).$$
By letting now $m' \to \infty$ on the subsequence defined in (11) and taking into account (15), we have that (12) follows. Therefore

\[
\frac{S_n - E(S_n|\xi_0, \xi_n)}{\sqrt{n}} \Rightarrow N(0, \eta^2).
\]

Then, by Skorohod’s representation theorem (i.e. Theorem 6.7 in Billingsley 1999) and by the Fatou lemma we get

\[
\limsup_{n \to \infty} \frac{E(S_n - E(S_n|\xi_0, \xi_n))^2}{n} \leq \liminf_{n \to \infty} \frac{E(S_n - E(S_n|\xi_0, \xi_n))^2}{n}.
\]

It follows that (2) holds as well as the CLT in Theorem 1. □

**Proof of Corollary 2**

From Theorem 1 and Theorem 3.1 in Billingsley (1999) we immediately obtain (3). Note that, by (1), we have that $|S_n|/\sqrt{n}$ is uniformly integrable and therefore, by (3) and the convergence of moments theorem (Theorem 3.5, Billingsley, 1999) we have that

\[
E\left|\frac{S_n}{\sqrt{n}}\right| \to \sqrt{2/\pi} \sigma.
\]

□

**Proof of Proposition 4**

This proposition follows by applying Corollary 3. Note that we have only to show that (7) implies (4).

We start the proof of this fact by fixing $0 < \varepsilon < 1$ and writing

\[
S_n = S_{[\varepsilon n]} + V_n(\varepsilon) + (S_n - S_{n-[\varepsilon n]}),
\]

where

\[
V_n(\varepsilon) = \sum_{j=[\varepsilon n]+1}^{n-[\varepsilon n]} X_j.
\]

Note that, by the triangle inequality, properties of the norm of the conditional expectation, condition (1) and stationarity, we easily get

\[
\limsup_{n \to \infty} \frac{1}{\sqrt{n}} E|S_n| \leq \limsup_{\varepsilon \to 0} \frac{1}{\sqrt{n}} E|V_n(\varepsilon)|. \quad (16)
\]

By the Cauchy-Schwartz inequality, for $1 \leq a \leq b \leq n$,

\[
E|\sum_{j=a}^{b} X_j| \leq n \sum_{j=a}^{b} E|X_j|.
\]

So, by stationarity

\[
\frac{1}{n} E(V_n(\varepsilon)|\xi_0, \xi_n)^2 \leq \sum_{j=[\varepsilon n]+1}^{n-[\varepsilon n]} E|X_0|\xi_{j-j}|^2.
\]

Since for $[\varepsilon n]+1 \leq j \leq n-[\varepsilon n]$ we have $F_{-j} \cap F_{n-j} \subset F_{[\varepsilon n]} \cap F^{[\varepsilon n]}$ it follows that

\[
\frac{1}{n} E(V_n(\varepsilon)|\xi_0, \xi_n)^2 \leq n E|X_0|\xi_{[\varepsilon n]}|\xi_{[\varepsilon n]}|^2.
\]
We obtain (4) by passing to the limit with $n \to \infty$ in the last inequality and taking into account (7) and (10). □

**Proof of Corollary 5** By the monotonicity of $||E(X_0|\xi_{-k}, \xi_k)||$, condition (8) implies (7). Condition (8) also implies the couple of conditions

$$\sum_{k \geq 1} ||E(X_0|\xi_{-k})||^2 < \infty$$

These two conditions lead to (5) by using Proposition 9.2 in Cuny and Lin (2016) combined with Lemma 6 in Zhao and Woodroofe (2008). It remains to apply Proposition 4. □

Before proving the corollaries in Subsection 2.2 we give a more general definition of the coefficient of absolute regularity $\beta$. As in relation (5) in Proposition 3.22 in Bradley, given two sigma algebra $A$ and $B$ with $B$ separable and for any $B \in B$ there is a regular conditional probability $P(B|A)$, then

$$\beta(A, B) = E(\sup_{B \in B} P(B|A) - P(B)).$$

We also need a technical lemma whose proof is given later.

**Lemma 8** Let $X, Z$ be two random variables on a probability space ($\Omega, \mathcal{K}, P$) with values in Polish space. Let $B \subset \mathcal{K}$ be a $\sigma-$algebra. Assume that $X$ and $Z$ are conditionally independent given $B$. Then

$$\beta(B, A \vee C) \leq \beta(A, B) + \beta(C, B) + \beta(A, C),$$

where $A = \sigma(X)$ and $C = \sigma(Z)$.

**Proof of Corollary 6** In order to apply Corollary 2 it is enough to verify that

$$\frac{E|E(S_n|\xi_0, \xi_n)|}{\sqrt{n}} \to 0.$$

Let $v$ be a positive integer. Then, by (11) we have

$$\lim_{n \to \infty} \frac{E|E(S_n|\xi_0, \xi_n)|}{\sqrt{n}} = \lim_{n \to \infty} \frac{E|E(\sum_{j=v+1}^{n} X_j|\xi_0, \xi_n)|}{\sqrt{n}}.$$ But, it is well-known that (see Ch.4 in Bradley 2007)

$$E|E(\sum_{j=v+1}^{n} X_j|\xi_0, \xi_n)| \leq 8\beta^{1/2}(\sigma(\xi_i; v \leq i \leq n - v), \sigma(\xi_0, \xi_n)) ||S_{n-2v}||_{2}.$$

By Lemma 8 applied with $A = \sigma(\xi_0)$, $B = \sigma(\xi_i; v \leq i \leq n - v)$, and $C = \sigma(\xi_n)$ we note that

$$\beta(\sigma(\xi_i; v \leq i \leq n - v), \sigma(\xi_0, \xi_n)) \leq 3\beta_v.$$
Therefore, for all $v \in N$

$$\lim_{n \to \infty} \frac{E|E(S_n|\xi_0, \xi_n)|}{\sqrt{n}} \leq 24\beta_v^{1/2}\sigma^2,$$

and the result follows by letting $v \to \infty$. □

**Proof of Corollary 7**

This corollary follows by verifying the conditions of Proposition 4. By Rio’s (1993) covariance inequality (see also Theorem 1.1 in Rio 2017) we know that

$$||E(X_0|\xi_{-n}, \xi_n)||^2 \leq 2\int_0^{\beta_n} Q^2(u)du,$$

where $\beta_n = \beta(\sigma(\xi_0), \sigma(\xi_{-n}, \xi_n))$.

But, according to Lemma 8 applied with $A = \sigma(\xi_{-n}), B = \sigma(\xi_0)$, and $C = \sigma(\xi_n)$ we obtain

$$\bar{\beta}_n = \beta(\sigma(\xi_0), \sigma(\xi_{-n}, \xi_n)) \leq 3\beta(\sigma(\xi_0), \sigma(\xi_n)) = 3\beta_n.$$

and the result follows. □

**Proof of Lemma 8** Denote the law of $X$ by $P_X$, the law of $Z$ by $P_Z$. Also by $P_{X|B}$ we denote the regular conditional distribution of $X$ given $B$ and by $P_{Z|B}$ the regular conditional distribution of $Z$ given $B$. By using the definition of $\beta(B, A \lor C)$ we have to evaluate the expression $I = E[\sup_H |P(H|B) - P(H)|]$, where the supremum is taken over all $H \subset A \lor C$. Denote by $I_H$ the indicator function of $H$. Since $X$ and $Z$ are conditionally independent given $B$ we have

$$P(H|B) = \iint I_H(x, z)P_{X|B}(dx)P_{Z|B}(dz) \text{ a.s.}$$

Also,

$$P(H) = \iint I_H(x, z)P_{X,Z}(dx, dz).$$

By the triangle inequality we can write $I \leq I_1 + I_2 + I_3$ where

$$I_1 = E \left( \sup_H \left| \iint I_H(x, z) (P_{X|B}(dx)P_{Z|B}(dz) - P_X(dx)P_{Z|B}(dz)) \right| \right),$$

$$I_2 = E \left( \sup_H \left| \iint I_H(x, z) (P_X(dx)P_{Z|B}(dz) - P_X(dx)P_Z(dz)) \right| \right)$$

and

$$I_3 = \sup_H \left| \iint I_H(x, z) (P_{X,Z}(dx, dz) - P_X(dx)P_Z(dz)) \right|.$$
Now, because $I_H$ is bounded by 1, we get

$$I_1 \leq E\left(\sup_{D \subseteq \mathcal{R}} |P_{X|B}(D) - P_X(D)|\right),$$

$$I_2 \leq E\left(\sup_{D \subseteq \mathcal{R}} |P_{Z|B}(D) - P_Z(D)|\right),$$

and

$$I_3 \leq \iint |P_{X,Z}(dx, dz) - P_{X,Z,*}|dxdz,$$

where $\mathcal{R}$ denote the Borel sigma field and $Z^*$ is a random variable distributed as $Z$ and independent of $X$.

The result follows by using the definition of $\beta$ and Theorem 3.29, both in Bradley (2007). □

Acknowledgement. This paper was partially supported by the NSF grant DMS-1832863. The author would like to thank Christophe Cuny, for discussions about the square root condition, to Wlodek Bryc for discussions about double tail sigma field and to Richard Bradley for correcting and suggesting a proof of Lemma 8.

References

[1] Billingsley, P. (1999). Convergence of probability measures. Second edition. Wiley, New York.

[2] Bradley, R.C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. Probability Surveys 2 107-144.

[3] Bradley, R.C. (2007). Introduction to strong mixing conditions 1, 2, 3. Kendrick Press, Heber City, UT.

[4] Chen, X. (1999). Limit theorems for functionals of ergodic Markov chains with general state space. Memoirs of the American Mathematical Society 139.

[5] Cuny, C. and Lin, M. (2016). Limit theorems for Markov chains by the symmetrization method. J. Math. Anal. Appl. 434 52-83.

[6] Davydov, Yu. A. (1973). Mixing conditions for Markov chains. Theory Probab. Appl. 18 312-328.

[7] Doukhan, P., Massart, P. and Rio, E. (1994). The functional central limit theorem for strongly mixing processes. Ann. Inst. H. Poincaré Probab. Statist. 30 63-82.

[8] Gordin, M. I. (1969). The central limit theorem for stationary processes, Soviet. Math. Dokl. 10 (5) 1174–1176.
[9] Heyde, C. C. (1974). On the central limit theorem for stationary processes. Z. Wahrsch. Verw. Gebiete 30 315-320.

[10] Jones, G. L. (2004). On the Markov chain central limit theorem. Probab. Surv. 1 299–320.

[11] Mansuy, R. and Yor, M. (2005). Harnesses, Lévy bridges and Monsieur Jourdain. Stochastic Process. Appl. 115 329–338.

[12] Maxwell, M. and Woodroofe, M. (2000). Central limit theorems for additive functionals of Markov chains. Ann. Probab. 28 713–724.

[13] McLeish, D. L. (1975). A generalization of martingales and mixing sequences. Adv. in Appl. Probab. 7 247-258.

[14] Meyn S.P. and Tweedie R.L. (2009). Markov chains and stochastic stability, 2nd edition. Cambridge University Press.

[15] Nummelin, E. (1978). A splitting technique for Harris recurrent chains. Z. Wahrs. verw Gebiete 43 309-318.

[16] Nummelin, E. (1984). General irreducible Markov chains and non-negative operators. Cambridge University Press, Cambridge, England.

[17] Peligrad, M. and Utev, S. (2006). Central limit theorem for stationary linear processes. Ann. Probab. 34 1608–1622.

[18] Rio, E. (1993). Covariance inequalities for strongly mixing processes. Ann. Inst. Henri Poincaré Probab. Stat. 29 587–597.

[19] Rio, E. (2017). Asymptotic theory of weakly dependent random processes. Springer.

[20] Williams, D. (1973). Some basic theorems on harnesses. Stochastic analysis (a tribute to the memory of Rollo Davidson), 349–363. Wiley, London.

[21] Zhao, O. and Woodroofe, M. (2008). On martingale approximations. Ann. Appl. Probab. 18 1831–1847.