SPA: 3663 Model 1 pp. 1–14 (col. fig: NIL)

ARTICLE IN PRESS

Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications xxx (xxxx) xxx

stochastic processes and their applications

www.elsevier.com/locate/spa

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

Received 25 August 2019; received in revised form 1 April 2020; accepted 10 April 2020
Available online xxxx

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 additive functionals of a stationary and totally ergodic Markov chain with $\mathrm{var}(S_n)/n$ uniformly bounded, satisfies a $\sqrt{n}-$central limit theorem with a random centering. We do not assume that the Markov chain is irreducible and aperiodic. However, the random centering is not needed if the Markov chain satisfies stronger forms of ergodicity. In absence of ergodicity the convergence in distribution still holds, but the limiting distribution might not be normal.

Keywords: Markov chains; Variance of partial sums; Central limit theorem; Projective criteria; Absolute regularity

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 [14], Gordin and Lifshits condition [15], Heyde's projective condition [17], [37], mixingales [23], Maxwell and Woodroofe condition [22], just to name a few. All of them have in common that the conditions are imposed on 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.

$$S_n = X_1 + X_2 + \cdots + X_n = X_n + X_{n-1} + \cdots + X_1.$$

E-mail address: peligrm@ucmail.uc.edu.

https://doi.org/10.1016/j.spa.2020.04.004
0304-4149/

Please cite this article as: M. Peligrad, A new CLT for additive functionals of Markov chains, Stochastic Processes and their Applications (2020), https://doi.org/10.1016/j.spa.2020.04.004.

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 [5], for a survey) and harnesses (see for instance [21,38]) are independent of the direction of time.

For additive functionals of reversible, stationary and ergodic Markov chains, with centered and square integrable variables, Kipnis and Varadhan [20] proved that if $E(S_n^2)/n$ converges to a finite limit, then the CLT holds. This is not true without assuming reversibility (see for instance [4] or [8], Prop. 9.5(ii), among other examples). On the other hand, for additive functionals of Harris recurrent and aperiodic Markov chains with centered and square integrable variables, Chen [7, Theorem II. 3.1] proved that if $S_n/\sqrt{n}$ is stochastically bounded, it satisfies the CLT.

These results suggest and motivate the study of limiting distribution for stationary Markov chains with additive functionals satisfying $\sup_n E(S_n^2)/n < \infty$. With this aim, 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 a blocking argument and martingale approximation techniques, we shall prove that functions of a Markov chain which is stationary and totally ergodic (in the ergodic theoretical sense) satisfy 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 II. 2.3 of Chen [7] and a totally different new approach. We also provide a new proof for the CLT for interlaced mixing Markov chains. We also point out that, when the Markov chain is stationary but not necessarily ergodic, the limiting distribution still exists and we express it as a mixture of distributions.

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 $\mathcal{F}_n = \sigma(\xi_k, k \le n)$ and by $\mathcal{F}^n = \sigma(\xi_k, k \ge n)$. The marginal distribution on $\mathcal{A}$ is denoted by $\pi(A) = \mathbb{P}(\xi_0 \in A)$. To settle the concern about the existence of a Markov chain with general state space, we shall construct the Markov chain from a kernel $P(x, A)$, we assume an invariant distribution π exists and invoke the Ionescu Tulcea [18] result.

Next, let $\mathbb{L}_0^2(\pi)$ 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 $f \in \mathbb{L}_0^2(\pi)$ let

$$X_i = f(\xi_i),\ S_n = \sum_{i=1}^{n} X_i. \tag{1}$$

We denote by $\|X\|$ the norm in $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$.

With the exception of Remark 3, in all the other results we shall assume the total ergodicity of the shift θ of the sequence $(\xi_n)_{n\in\mathbb{Z}}$ with respect to $\mathbb{P}$, i.e. θ^m is ergodic for every $m \ge 1$. For the definition of the ergodicity of the shift we direct the reader to the subsection "A return to Ergodic Theory" in [2] p. 494.

Let us consider the operator P induced by the kernel $P(x, A)$ on bounded measurable functions on $(S, \mathcal{A})$ defined by $Pf(x) = \int_S f(y)P(x, dy)$. By using Corollary 5 p. 97 in [33], the shift of $(\xi_n)_{n\in\mathbb{Z}}$ is totally ergodic with respect to $\mathbb{P}$ if and only if the powers P^m are ergodic with respect to π for all natural m (i.e. $P^m f = f$ for f bounded on $(S, \mathcal{A})$ implies f is constant π −a.e.). For more information on total ergodicity, we refer to the survey paper by Quas [30].

Below, $\Rightarrow$ denotes the convergence in distribution and by $N(\mu, \sigma^2)$ we denote a normally distributed random variable with mean μ and variance σ^2.

2.1. Central limit theorem

We shall establish the following CLT.

Theorem 1. *Assume that*

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

Then, the following limit exists

$$\lim_{n\to\infty} \frac{1}{n} \|S_n - E(S_n|\xi_0, \xi_n)\|^2 = \sigma^2 \tag{3}$$

and

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

Remark 2. It should be noted that, by condition (2), it follows that $0 \leq \sigma^2 < \infty$. When $\sigma^2 = 0$ then $(S_n - E(S_n|\xi_0, \xi_n))/\sqrt{n} \to^{\mathbb{P}} 0$ as $n \to \infty$. Also, for any stationary sequence, starting from the identity

$$E(S_n^2) = E(X_0^2) + 2\sum_{k=1}^{n-1}\sum_{j=1}^{k} E(X_0 X_j),$$

note that the convergence of sums of the covariances implies that $E(S_n^2)/n$ is convergent. Furthermore, if the sums of covariances are bounded by a constant then (2) holds.

We would like to mention that, as in the stationary martingale case, in the absence of ergodicity the limiting distribution still exists and it is a mixture of distributions.

Remark 3. If $(\xi_n)_{n\in\mathbb{Z}}$ is any stationary Markov chain, $(X_n)_{n\in\mathbb{Z}}$ and $(S_n)_{n\geq 1}$ are defined by (1) and (2) holds, then there is a random variable η^2 such that

$$\frac{S_n - E(S_n|\xi_0, \xi_n)}{\sqrt{n}} \Rightarrow \eta^2 N(0, 1), \tag{4}$$

where η^2 is independent of $N(0, 1)$.

If the random centering is not present, we have the following result:

Corollary 4. *Assume that* (2) *holds and in addition*

$$\frac{E(S_n|\xi_0, \xi_n)}{\sqrt{n}} \to^{\mathbb{P}} 0 \text{ as } n \to \infty.$$

Then the following limit exists

$$\lim_{n\to\infty} \frac{\pi}{2n} (E|S_n|)^2 = \sigma^2$$

and

$$\frac{S_n}{\sqrt{n}} \Rightarrow N(0, \sigma^2). \tag{5}$$

As an immediate consequence of Theorem 1, we give next sufficient conditions for the CLT with the traditional limiting variance.

Corollary 5. *Assume that (2) holds and in addition*

$$\lim_{n\to\infty} \frac{1}{n} \| E(S_n|\xi_0, \xi_n)\|^2 = 0. \tag{6}$$

Then

$$\lim_{n\to\infty} \frac{E(S_n^2)}{n} = \sigma^2 \tag{7}$$

and

$$\frac{S_n}{\sqrt{n}} \Rightarrow N(0, \sigma^2). \tag{8}$$

We can also give a sufficient condition for (6) in terms of individual random variables.

Proposition 6. *Assume that (2) holds and the following condition is satisfied*

$$\lim_{n\to\infty} n\| E(X_0|\xi_{-n}, \xi_n)\|^2 = 0. \tag{9}$$

Then (7) and (8) hold.

In Section 2.2, in the context of absolutely regular Markov chains, we shall comment that condition (9) alone does not imply (8). However, a reinforced condition does:

Corollary 7. *Assume that*

$$\sum_{k\geq 1} \| E(X_0|\xi_{-k}, \xi_k)\|^2 < \infty. \tag{10}$$

Then (7) and (8) hold with $\sigma^2 = \|X_0\|^2 + 2\sum_{k\geq 1} E(X_0 X_k)$.

2.2. Absolutely regular Markov chains

Relevant to this section is the coefficient of absolute regularity, which was introduced by Volkonskii and Rozanov [36] and was attributed there to Kolmogorov. For a stationary sequence $\xi = (\xi_k)_{k\in Z}$, not necessarily Markov, with values in a separable Banach space, the coefficient of absolute regularity is defined by

$$\beta_n = \beta_n(\xi) = E\left(\sup_{A\in\mathcal{F}^n} |\mathbb{P}(A|\mathcal{F}_0) - \mathbb{P}(A)| \right).$$

The chain is called absolutely regular if $\beta_n \to 0$. We can easily see from this definition that β_n is monotonic. Furthermore β_n is symmetric, in the sense that $\beta(\xi_0, \xi_n) = \beta(\xi_n, \xi_0)$. This fact can be easily seen by using the equivalent definition for β_n in terms of partitions (see Definitions 3.3 and 3.5 in [6]). For such sequences, Bradley [4] constructed an example of a stationary, pairwise independent, absolutely regular sequence for which a nondegenerate central limit theorem cannot hold.

For a Markov chain $\xi = (\xi_k)_{k\in Z}$, with values in a separable Banach space, the coefficient of absolute regularity is equal to (see Proposition 3.22 (III,5) in [6])

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

where $\mathcal{B}$ denotes the Borel sigma field.

Let us mention that there are numerous examples of stationary absolutely regular Markov chains. For easy reference we refer to Section 3 in [5] 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 total ergodicity in the measure theoretical sense. Also, in many situations these coefficients are tractable. The computation of the coefficients of absolute regularity is an area of intense research, with numerous applications to time series and statistics. There is a vast literature on this subject. See for instance [1,9,11–13,26], [6, Vol. 1,2,3] among others.

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 [28], Meyn and Tweedie [25] and Chen [7] and we also refer to the survey paper by Jones [19].

In the works mentioned above, the vast majority of results concerning the CLT for absolutely regular Markov chains require sufficient conditions in terms of moments and mixing rates. Some of them require rates which combine the tail distribution of a variable with the mixing coefficients.

By using regeneration techniques and partition in independent blocks (Nummelin's splitting technique, [27]) 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 II.2.3 in [7]). 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 II. 3.1 in [7], and sheds new light on the asymptotic variance in Chen's Theorem II.2.3. The advantage of these results is that no rate of convergence to zero of the mixing coefficients is required. However, some information about the variance of partial sums is needed.

Corollary 8. *Assume that (2) holds and the sequence is absolutely regular. Then (5) holds with* $\sigma^2 = \lim_{n\to\infty} \pi\,(E|S_n|)^2/2n$.

To give a CLT where the limiting variance is σ^2 defined in (7), we shall verify condition (9) of Proposition 6. 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 9. *Assume that (2) holds and the following condition is satisfied*

$$\lim_{n\to\infty} n \int_0^{\beta_n} Q^2(u)du = 0. \tag{11}$$

Then (7) and (8) hold.

In terms of moments, by Hölder's inequality, (11) is implied by $E(|X_0|^{2+\delta}) < \infty$ and $n\beta_n^{\delta/(2+\delta)} \to 0$, 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 (10) is verified if

$$\sum_{n\geq 1} \int_0^{\beta_n} Q^2(u)du < \infty, \tag{12}$$

and then (7) and (8) hold. Further reaching results could be found in [13], where a larger class of processes was considered. According to Corollary 1 Doukhan et al. [13], (11) alone is not enough for (8). Actually, the stronger condition (12) is a minimal condition for the CLT for $S_n/\sqrt{n}$ in the following sense. In their Corollary 1, Doukhan et al. [13] constructed a stationary absolutely regular Markov chain $(\xi_k)_{k\in Z}$ and a function $f \in \mathbb{L}_0^2(\pi)$, which barely does not satisfy (12) 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. [13], there exists a Markov chain with these specifications, such that $S_n/\sqrt{n}$ does not satisfy the CLT. However, for these specifications (11) is satisfied. If in addition we know that (2) holds, then, by Corollary 9, the CLT holds for $S_n/\sqrt{n}$.

Let us point out for instance, a situation where Corollary 9 is useful. Let $\mathbf{Y} = (Y_i)$ and $\mathbf{Z} = (Z_j)$ be two absolutely regular Markov chains of centered, bounded random variables, independent among them and satisfying the following conditions $\sum_{n\geq 1} \beta_n(Y) < \infty$ and $n\beta_n(Z) \to 0$. If we define now the sequence $\mathbf{X} = (X_n)$, where for each n we set $X_n = Y_n Z_n$, then, by Theorem 6.2 in [6], we have $\beta_n(X) \leq \beta_n(Y) + \beta_n(Z)$. As a consequence, by using the monotonicity of β_n, we have $n\beta_n(X) \to 0$ and condition (11) is satisfied. Certainly, this condition alone does not assure that the CLT holds. In order to apply Corollary 9 we have to verify that condition (2) holds. Conditioned by $\mathbf{Z}$ the partial sum of (X_n) becomes a linear combination of the variables of $\mathbf{Y}$ and we can apply Corollary 7 in [29]. It follows that there is a positive constant C such that

$$E_{\mathbf{Y}}(\sum_{k=1}^n Y_k Z_k)^2 \leq C(\sum_{k=1}^n Z_k^2) \text{ a.s.}, \tag{13}$$

where $E_{\mathbf{Y}}$ denotes the partial integral with respect to the variables of $\mathbf{Y}$. By the independence of the sequences $\mathbf{Y}$ and $\mathbf{Z}$ we obtain

$$E(\sum_{k=1}^n Y_k Z_k)^2 \leq CnE(Z_0^2). \tag{14}$$

Therefore condition (2) holds. By Corollary 9 we obtain that (7) and (8) hold for the sequence (X_n).

As a particular example of this kind let us consider two stationary renewal processes $\xi = (\xi_i)$ and $\eta = (\eta_i)$, with countable state space $\{0, 1, 2, \ldots\}$ and independent among them. For the transition probabilities of (ξ_i) we take for $i \geq 1$, $P(\xi_1 = i - 1|\xi_0 = i) = 1$, $p_i = P(\xi_1 = i|\xi_0 = 0) = \left(2i^3(\log(i+1))^2\right)^{-1}$, and $p_0 = P(\xi_1 = 0|\xi_0 = 0) = 1 - P(\xi_1 \geq 1|\xi_0 = 0)$.

Please cite this article as: M. Peligrad, A new CLT for additive functionals of Markov chains, Stochastic Processes and their Applications (2020), https://doi.org/10.1016/j.spa.2020.04.004.

For (η_i) we take for $i \geq 1$, $P(\eta_1 = i-1|\eta_0 = i) = 1$, $q_i = P(\eta_1 = i|\eta_0 = 0) = \big(2i^3(\log(i+1))\big)^{-1}$ and $q_0 = P(\eta_1 = 0|\eta_0 = 0) = 1 - P(\eta_1 \geq 1|\eta_0 = 0)$. From Theorem 5 in [9], we know that the $\beta-$mixing coefficients for these sequences are of orders

$$\beta_n(\xi) \leq c\frac{1}{n(\log(n+1))^2} \text{ and } \beta_n(\eta) \leq c\frac{1}{n\log(n+1)},$$

where c is a positive constant. Now, let f and g be two bounded function and define the sequences $\mathbf{Y}$ and $\mathbf{Z}$ by $Y_i = f(\xi_i) - E(f(\xi_i))$ and $Z_i = g(\eta_i) - E(g(\eta_i))$ and set $X_i = X_i Y_i$. Clearly, for this example $\sum_{n\geq 1}\beta_n(Y) < \infty$ and $n\beta_n(Z) \to 0$ and we can apply Corollary 9 for the sequence (X_n).

2.3. Interlaced mixing Markov chains

Another example where Corollary 5 applies is the class of interlaced mixing Markov chains. Let $\mathcal{A}, \mathcal{B}$ be two sub σ-algebras of $\mathcal{F}$. Define the maximal coefficient of correlation

$$\rho(\mathcal{A}, \mathcal{B}) = \sup_{f\in\mathbb{L}_0^2(\mathcal{A}),\ g\in\mathbb{L}_0^2(\mathcal{B})} \frac{|E(fg)|}{\|f\|\cdot\|g\|},$$

where $\mathbb{L}_0^2(\mathcal{A})$ ($\mathbb{L}_0^2(\mathcal{B})$) is the space of random variables that are $\mathcal{A}-$measurable (respectively $\mathcal{B}-$measurable), zero mean and square integrable. For a sequence of random variables, $(\xi_k)_{k\in\mathbb{Z}}$, we define

$$\rho_n^* = \sup \rho(\sigma(\xi_i, i \in S), \sigma(\xi_j, j \in T)) ,$$

where the supremum is taken over all pairs of disjoint sets, T and S or $\mathbb{R}$ such that $\min\{|t-s| : t \in T, s \in S\} \geq n$. We call the sequence ρ^*-mixing if $\rho_n^* \to 0$ as $n \to \infty$.

The ρ^*-mixing condition goes back to Stein [35] and to Rosenblatt [34]. It is well-known that ρ^*-mixing implies total ergodicity. Also, there are known examples (see Example 7.16 in [6]) of ρ^*-mixing sequences which are not absolutely regular.

Our next Corollary shows that our result provides an alternative proof of the CLT for interlaced ρ^*-mixing Markov chains. Although the result itself is not new, it provides another example where condition (6) is verified. For further reaching results see for instance Theorem 11.18 in [6] and Corollary 9.16 in [24].

Corollary 10. *Assume that $(\xi_k)_{k\in\mathbb{Z}}$ is a stationary ρ^*-mixing Markov chain. Then (6), (7) and (8) hold.*

3. Proofs

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

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

Also denote

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

For $k = 0, 1, \dots, u-1$ 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 $\mathcal{F}_{km} \vee \mathcal{F}^{(k+1)m}$. Note that D_k is adapted to $\mathcal{F}_{(k+1)m} = \mathcal{G}_k$. Then we have $E(D_1|\mathcal{G}_0) = 0$ a.s. Since we assumed that the shift θ of the sequence $(\xi_n)_{n\in\mathbb{Z}}$ is totally ergodic, we deduce that we have a stationary and ergodic sequence of square integrable martingale differences $(D_k, \mathcal{G}_k)_{k\geq 0}$.

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

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

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 $\mathbb{N}$ such that

$$\lim_{m'\to\infty} \frac{1}{m'}\|Y_0 - E(Y_0|\xi_0, \xi_{m'})\|^2 = \limsup_{n\to\infty} \frac{1}{n}\|S_n - E(S_n|\xi_0, \xi_n)\|^2 = \eta^2. \quad (15)$$

Note that by (2) it follows that $\eta^2 < \infty$. This means that

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

Whence, according to Theorem 3.2 in [3], in order to establish the CLT for $\sum_{i=0}^{u-1} Y_i/\sqrt{n}$ we have only to show that

$$\lim_{m'\to\infty} \limsup_{n\to\infty} \|\frac{1}{\sqrt{n}}(S_n - E(S_n|\xi_0, \xi_n)) - \frac{1}{\sqrt{u}} M_u(m')\|^2 = 0. \quad (16)$$

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

$$S_u(m) = M_u(m) + R_u(m). \quad (17)$$

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

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

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

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

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 decomposition (17), the fact that $M_n(m)$ and $R_n(m)$ are orthogonal and $M_n(m)$ is a martingale, we obtain the identity

$$\frac{1}{u}\|S_u(m)\|^2 = \frac{1}{m}\|S_m - E(S_m|\xi_0, \xi_m)\|^2 + \frac{1}{u}\|R_u(m)\|^2. \quad (18)$$

Also, note that (2) and the definition of Y_u imply that for some positive constant C, we have $\|Y_u\| \le Cm$. Hence, by the properties of conditional expectations, for every m fixed, we have

$$\|\frac{1}{\sqrt{n}}(S_n - E(S_n|\xi_0,\xi_n)) - \frac{1}{\sqrt{u}}(S_u(m) - E(S_u(m)|\xi_0,\xi_n))\| \to 0 \text{ as } n \to \infty. \tag{19}$$

Recall the definition of $M_u(m)$, which is orthogonal to $\mathcal{F}_0 \vee \mathcal{F}^{um}$. By using again the properties of conditional expectations and the identity (18), for every m we have

$$\begin{aligned}&\frac{1}{u}\|S_u(m) - E(S_u(m)|\xi_0,\xi_n) - M_u(m)\|^2 = \frac{1}{u}\|R_u(m)\|^2 - \frac{1}{u}\|E(S_u(m)|\xi_0,\xi_n)\|^2\\ &= \frac{1}{u}\left(\|S_u(m)\|^2 - \|E(S_u(m)|\xi_0,\xi_n)\|^2\right) - \frac{1}{m}\|S_m - E(S_m|\xi_0,\xi_m)\|^2\\ &= \frac{1}{u}\|(S_u(m) - E(S_u(m)|\xi_0,\xi_n))\|^2 - \frac{1}{m}\|S_m - E(S_m|\xi_0,\xi_m)\|^2.\end{aligned}$$

By passing now to the limit in the last identity with $n \to \infty$, by (15) and (19) we obtain

$$\begin{aligned}&\limsup_{n\to\infty} \|\frac{1}{\sqrt{n}}(S_n - E(S_n|\xi_0,\xi_n)) - \frac{1}{\sqrt{u}}M_u(m)\|^2\\ &= \eta^2 - \left(\frac{1}{m}\|S_m - E(S_m|\xi_0,\xi_m)\|^2\right).\end{aligned}$$

By letting now $m' \to \infty$ on the subsequence defined in (15) and taking into account (19), we have that (16) follows. Therefore

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

Then, by (15), Skorohod's representation theorem (i.e. Theorem 6.7 in [3]) and by Fatou's lemma we get

$$\limsup_{n\to\infty} \frac{E(S_n - E(S_n|\xi_0,\xi_n))^2}{n} = \eta^2 \le \liminf_{n\to\infty} \frac{E(S_n - E(S_n|\xi_0,\xi_n))^2}{n}.$$

It follows that (3) holds as well as the CLT in Theorem 1. □

Proof of Corollary 4. From Theorem 1 and Theorem 3.1 in [3] we immediately obtain (5). Note that, by (2), we have that $|S_n|/\sqrt{n}$ is uniformly integrable and therefore, by (5) and the convergence of moments theorem (Theorem 3.5, [3]) we have that $E|S_n|/\sqrt{n} \to \sqrt{2/\pi}\sigma$. □

Proof of Remark 3. The proof of this remark is based on two facts.

Fact 1. Raikov-type CLT for stationary martingale differences. (see Theorem 3.6 in [16]). If $(D_k)_{k\in\mathbb{Z}}$ is a square integrable sequence of martingale differences and $M_n = D_1 + \cdots + D_n$, then there is a random variable η^2 such that

$$\frac{M_n}{\sqrt{n}} \Rightarrow \eta^2 N(0,1),$$

where η^2 is independent on $N(0,1)$.

Fact 2. A variant of Theorem 3.2 in [3] for complete separable metric spaces (Theorem 2 in [10]). For random variables $(X_n(m'), Y_n)$ with $n \in \mathbb{N}$ and m' belonging to a subsequence of $\mathbb{N}$ which tends to ∞, assume that for every $\varepsilon > 0$

$$\lim_{m'\to\infty} \limsup_{n\to\infty} P(|X_n(m') - Y_n| > \varepsilon) = 0,$$

and for every m', $X_n(m') \Rightarrow Z(m')$ as $n \to \infty$. Then there is a random variable X such that $Z(m') \Rightarrow X$ as $m' \to \infty$ and $Y_n \Rightarrow X$ as $n \to \infty$.

To prove Remark 3, we define the subsequence (m') by (15) and start from relation (16). We apply next Fact 1 to the sequence of stationary martingale differences $(D_k(m'))_{k\geq 0}$ and obtain that

$$\frac{M_u(m')}{\sqrt{m'}} \Rightarrow \eta_{m'}^2 N(0,1) \text{ as } m' \to \infty,$$

where $\eta_{m'}^2$ are random variables independent on $N(0,1)$. Then, we apply Fact 2 and deduce that, for some random variable X, both $\eta_{m'}N(0,1) \Rightarrow X$ and $S_n/\sqrt{n} \Rightarrow X$. But the characteristic function of $\eta_{m'}^2 N(0,1)$ is $E\left(\exp(-t^2\eta_{m'}^2/2)\right)$ and therefore $\eta_{m'}^2$ converges in distribution to some random variable η implying (4). □

Proof of Proposition 6. This proposition follows by applying Corollary 5. Note that we have only to show that (9) implies (6).

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 (2) and stationarity, we easily get

$$\limsup_{n\to\infty} \frac{1}{\sqrt{n}} \|E(S_n|\xi_0,\xi_n)\| \leq \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{1}{\sqrt{n}} \|E(V_n(\varepsilon)|\xi_0,\xi_n)\|. \tag{20}$$

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

$$\|E(\sum_{j=a}^{b} X_j|\xi_0,\xi_n)\|^2 \leq n \sum_{j=a}^{b} \|E(X_j|\xi_0,\xi_n)\|^2.$$

So, by stationarity

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

Since for $[\varepsilon n]+1 \leq j \leq n-[\varepsilon n]$ we have $\mathcal{F}_{-j} \vee \mathcal{F}^{n-j} \subset \mathcal{F}_{-[\varepsilon n]} \vee \mathcal{F}^{[\varepsilon n]}$ 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 (6) by passing to the limit with $n \to \infty$ in the last inequality and taking into account (9) and (20). □

Proof of Corollary 7. By the monotonicity of $\|E(X_0|\xi_{-k},\xi_k)\|$, condition (10) implies (9). Condition (10) also implies the couple of conditions

$$\sum_{k\geq 1} \|E(X_0|\xi_{-k})\|^2 < \infty \text{ and } \sum_{k\geq 1} \|E(X_0|\xi_k)\|^2 < \infty. \tag{21}$$

Please cite this article as: M. Peligrad, A new CLT for additive functionals of Markov chains, Stochastic Processes and their Applications (2020), https://doi.org/10.1016/j.spa.2020.04.004.

Now note that by the properties of the conditional expectations, the Markov property and stationarity, for all $k \geq 1$ we easily obtain

$$|E(X_0 X_{2k})| = |E(X_0 E(X_{2k}|\xi_k))| = |E(E(X_0|\xi_k)E(X_{2k}|\xi_k))| \leq$$
$$\|E(X_0|\xi_k)\| \cdot \|E(X_0|\xi_{-k})\| \leq (\|E(X_0|\xi_{-k})\|^2 + \|E(X_0|\xi_k)\|^2)/2.$$

A similar relation holds for $|E(X_0 X_{2k+1})|$. Hence the two conditions in (21) lead to (2). The result follows by applying Proposition 6. □

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

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

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

Lemma 11. *Let X, Z be two random variables on a probability space $(\Omega, \mathcal{K}, P)$ with values in a separable Banach space. Let $\mathcal{B} \subset \mathcal{K}$ be a sub $\sigma-$algebra. Assume that X and Z are conditionally independent given $\mathcal{B}$. Then*

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

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

Proof of Corollary 8. In order to apply Corollary 4 it is enough to verify that

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

Let $v \leq n$ be a positive integer. Then, by (2) we have

$$\limsup_{n \to \infty} \frac{E|E(S_n|\xi_0, \xi_n)|}{\sqrt{n}} = \limsup_{n \to \infty} \frac{E|E(V_n(v)|\xi_0, \xi_n)|}{\sqrt{n}}, \tag{22}$$

where $V_n(v) = \sum_{j=v+1}^{n-v} X_j$. But, it is well-known that (see Ch.4 in [6])

$$E|E(V_n(v)|\xi_0, \xi_n)| \leq 8\beta^{1/2}(\sigma(\xi_i;\, v \leq i \leq n-v), \sigma(\xi_0, \xi_n))\|S_{n-2v}\|_2.$$

By Lemma 11, applied with $\mathcal{A} = \sigma(\xi_0)$, $\mathcal{B} = \sigma(\xi_i;\ v \leq i \leq n-v)$, and $\mathcal{C} = \sigma(\xi_n)$ and taking into account the properties β_v listed at the beginning of Section 2.2 along with stationarity, we obtain that

$$\beta(\sigma(\xi_i; m \leq i \leq n-m), \sigma(\xi_0, \xi_n)) \leq \beta(\xi_0, \xi_v) + \beta(\xi_n, \xi_{n-v}) + \beta(\xi_0, \xi_n) \leq 3\beta_v.$$

Therefore, for all $v \in \mathbb{N}$

$$\limsup_{n \to \infty} \frac{E|E(S_n|\xi_0, \xi_n)|}{\sqrt{n}} \leq 24\beta_v^{1/2} \sup_n \frac{1}{\sqrt{n}}\|S_n\|_2,$$

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

Proof of Corollary 9. This corollary follows by verifying the conditions of Proposition 6. By Rio's [31] covariance inequality (see also Theorem 1.1 in [32]) we know that

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

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

But, according to Lemma 11, applied with $\mathcal{A} = \sigma(\xi_{-n})$, $\mathcal{B} = \sigma(\xi_0)$, and $\mathcal{C} = \sigma(\xi_n)$ 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 Corollary 10. For this class of random variables it is well-known that condition (2) is satisfied (see for instance Lemma 8.23 in [6]). According to Corollary 5 we have only to verify condition (6). Note that by (22) it is enough to show that

$$\lim_{v\to\infty}\limsup_{n\to\infty}\frac{1}{\sqrt{n}}\|E(V_n(v)|\xi_0,\xi_n)\| = 0,$$

with $V_n(v) = \sum_{j=v+1}^{n-v} X_j$. By the definition of ρ_v^* we observe that

$$\|E(V_n(v)|\xi_0,\xi_n)\|^2 = |E(V_n(v)E(V_n(\varepsilon)|\xi_0,\xi_n))| \leq \rho_v^*\|V_n(v)\|\cdot\|E(V_n(v)|\xi_0,\xi_n)\|.$$

Whence,

$$\frac{1}{\sqrt{n}}\|E(V_n(v)|\xi_0,\xi_n)\| \leq \rho_v^*\frac{1}{\sqrt{n}}\|V_n(v)\|.$$

The result follows by (2). □

Proof of Lemma 11. Denote the law of X by P_X, the law of Z by P_Z. Also by $P_{X|\mathcal{B}}$ we denote the regular conditional distribution of X given $\mathcal{B}$ and by $P_{Z|\mathcal{B}}$ the regular conditional distribution of Z given $\mathcal{B}$. By using the definition of $\beta(\mathcal{B},\mathcal{A}\vee\mathcal{C})$ we have to evaluate the expression $I = E[\sup_H |P(H|\mathcal{B}) - P(H)|]$, where the supremum is taken over all $H \subset \mathcal{A}\vee\mathcal{C}$. Denote by I_H the indicator function of H. Since X and Z are conditionally independent given $\mathcal{B}$ we have

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

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

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

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

and

$$I_3 = \sup_H \left|\iint I_H(x,z)\left(P_{X,Z}(dx,dz) - P_X(dx)P_Z(dz)\right)\right|.$$

Now, because I_H is bounded by 1, we get

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

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

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 β and Theorem 3.29, both in [6]. □

Acknowledgments

This paper was partially supported by the NSF grant DMS-1811373. The author would like to thank Christophe Cuny for discussions about the variance of partial sums, to Wlodek Bryc for discussions about double tail sigma field and harnesses and to Richard Bradley for correcting and suggesting a proof of Lemma 11. Special thanks go to the two anonymous referees for carefully reading of the manuscript and for numerous suggestions, that improved the presentation of this paper.

References

[1] P. Ango Nze, Critères d'ergodicité géométrique ou arithmétique de modèles linéaires perturbés à représentation markovienne, C. R. Acad. Sci., Paris 326 (Série 1) (1998) 371–376.
[2] P. Billingsley, Probability and Measure, third ed., Wiley, New York, 1995.
[3] P. Billingsley, Convergence of Probability Measures, second ed., Wiley, New York, 1999.
[4] R.C. Bradley, Stationary, pairwise independent, absolutely regular sequence for which the central limit theorem fails, Probab. Theory Related Fields 81 (1989) 1–10.
[5] R.C. Bradley, Basic properties of strong mixing conditions. A survey and some open questions, Probab. Surv. 2 (2005) 107–144.
[6] R.C. Bradley, Introduction to Strong Mixing Conditions 1, 2, 3, Kendrick Press, Heber City, UT, 2007.
[7] X. Chen, Limit theorems for functionals of ergodic Markov chains with general state space, Mem. Amer. Math. Soc. (1999) 139.
[8] C. Cuny, M. Lin, Limit theorems for Markov chains by the symmetrization method, J. Math. Anal. Appl. 434 (2016) 52–83.
[9] Yu.A. Davydov, Mixing conditions for Markov chains, Theory Probab. Appl. 18 (1973) 312–328.
[10] H. Dehling, O. Durieu, D. Volný, New techniques for empirical processes of dependent data, Stochastic Process. Appl. 119 (2009) 3699–3718.
[11] R. Douc, E. Moulines, P. Soulier, Computable convergence rates for subgeometric ergodic Markov chains, Bernoulli 13 (2007) 831–848.
[12] P. Doukhan, Mixing: Properties and Examples, in: Lecture Notes in Statistics, vol. 85, Springer Verlag, 1994.
[13] P. Doukhan, P. Massart, E. Rio, The functional central limit theorem for strongly mixing processes, Ann. Inst. Henri Poincaré Probab. Stat. 30 (1994) 63–82.
[14] M.I. Gordin, The central limit theorem for stationary processes, Sov. Math. Dokl. 10 (1969) 1174–1176.
[15] M. Gordin, B. Lifshits, The central limit theorem for stationary Markov processes, Sov. Math. Dokl. 19 (1978) 392–393.
[16] P. Hall, C.C. Heyde, Martingale Limit Theory and its Application, Academic Press, New York, 1980.
[17] C.C. Heyde, On the central limit theorem for stationary processes, Z. Wahrscheinlichkeitstheor. Verwandte Geb. 30 (1974) 315–320.
[18] C.T. Ionescu Tulcea, Mesures dans les espaces produits, Atti Accad. Naz. Lincei Rend. 7 (1949) 208–211.

[19] G.L. Jones, On the Markov chain central limit theorem, Probab. Surv. 1 (2004) 299–320.
[20] C. Kipnis, S.R.S. Varadhan, Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions, Comm. Math. Phys. 104 (1986) 1–19.
[21] R. Mansuy, M. Yor, Harnesses, Lévy bridges and Monsieur Jourdain, Stochastic Process. Appl. 115 (2005) 329–338.
[22] M. Maxwell, M. Woodroofe, Central limit theorems for additive functionals of Markov chains, Ann. Probab. 28 (2000) 713–724.
[23] D.L. McLeish, A generalization of martingales and mixing sequences, Adv. Appl. Probab. 7 (1975) 247–258.
[24] F. Merlevède, M. Peligrad, S. Utev, Functional Gaussian Approximation for Dependent Structures, Oxford University Press, 2019.
[25] S.P. Meyn, R.L. Tweedie, Markov Chains and Stochastic Stability, second ed., Cambridge University Press, 2009.
[26] A. Mokkadem, Propriétés de mélange des modéles autorégressifs polynomiaux, Ann. Inst. Henri Poincare B 26 (2) (1990) 219–260.
[27] E. Nummelin, A splitting technique for harris recurrent chains, Z. Wahrscheinlichkeitstheor. Verwandte Geb. 43 (1978) 309–318.
[28] E. Nummelin, General Irreducible Markov Chains and Non-Negative Operators, Cambridge University Press, Cambridge, England, 1984.
[29] M. Peligrad, S. Utev, Central limit theorem for stationary linear processes, Ann. Probab. 34 (2006) 1608–1622.
[30] A. Quas, Ergodicity and mixing properties, in: R.E. Meyers (Ed.), Encyclopedia of Complexity and Systems Science, Springer, 2009, pp. 2918–2933.
[31] E. Rio, Covariance inequalities for strongly mixing processes, Ann. Inst. Henri Poincaré Probab. Stat. 29 (1993) 587–597.
[32] E. Rio, Asymptotic Theory of Weakly Dependent Random Processes, Springer, 2017.
[33] M. Rosenblatt, Markov Processes. Structure and Asymptotic Behavior, Springer, Berlin, 1971.
[34] M. Rosenblatt, Central limit theorems for stationary processes, in: Proceedings of the Sixth Berkeley Symposium on Probability and Statistics, Vol. 2, University of California Press, Los Angeles, 1972, pp. 551–561.
[35] C. Stein, A bound for the error in the normal approximation to the distribution of a sum of dependent random variables, in: Proceedings of the Sixth Berkeley Symposium on Probability and Statistics, Vol. 2, University of California Press, Los Angeles, 1972, pp. 583–602.
[36] V.A. Volkonskii, Yu.A. Rozanov, Some limit theorems for random functions I, Theory Probab. Appl. 4 (1959) 178–197.
[37] D. Volný, Approximating martingales and the central limit theorem for strictly stationary processes, Stochastic Process. Appl. 44 (1993) 41–74.
[38] D. Williams, Some basic theorems on harnesses, in: Stochastic Analysis (A Tribute to the Memory of Rollo Davidson), Wiley, London, 1973, pp. 349–363.

Please cite this article as: M. Peligrad, A new CLT for additive functionals of Markov chains, Stochastic Processes and their Applications (2020), https://doi.org/10.1016/j.spa.2020.04.004.