

On the quenched CLT for stationary Markov chains.

Dedicated to Michael Lin's 80th birthday.

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, quenched limit theorems, central limit theorem, projective criteria.

Mathematical Subject Classification (2010): 60F05, 60F15, 60J05, 60G10.

Abstract. In this paper we give sufficient conditions for the almost sure central limit theorem started at a point, known under the name of quenched central limit theorem. This is achieved by using a new idea of conditioning with respect to both the past and the future of the Markov chain. As applications we provide a new sufficient projective condition for the quenched CLT.

1 Introduction and the main result

We assume that $(\xi_n)_{n \in \mathbb{Z}}$ is a stationary Markov chain, defined on a probability space $(\Omega, \mathcal{F}, P)$ with values in a Polish 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 on $\mathcal{A}$ is denoted by $\pi(A) = \mathbb{P}(\xi_0 \in A)$. We shall construct the Markov chain in a canonical way on $S^{\mathbb{Z}}$ from a kernel $Q(x, A)$, and we assume that an invariant distribution π exists.

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

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

Denote the regular conditional probability on $\mathcal{F}$, with respect to $\mathcal{F}_0$ by

$$P^0(\cdot)(\omega) = P(\cdot | \mathcal{F}_0)(\omega),$$

and the conditional expectation, $E^0(X) = E(X | \mathcal{F}_0)$. By the Markov property, if $A \in \mathcal{F}^0 = \sigma(\xi_i, i \geq 0)$, we have $P^0(A) = P(A | \xi_0)$, and for X measurable with respect to $\mathcal{F}^0$, $E^0(X) = E(X | \xi_0)$. We are studying the quenched central limit theorem for Markov chains, which can be stated in two equivalent ways: For P -almost all $\omega \in \Omega$

$$P^0(S_n / \sqrt{n} \leq t)(\omega) \rightarrow P(N(0, \sigma^2) \leq t) \text{ for any } t, \quad (2)$$

where $N(0, \sigma^2)$ is a normal random variable with mean 0 and variance σ^2 .

Another formulation is known under the name of the CLT started at a point. Let P^x be the probability associated to the Markov chain started from $x \in S$ and E^x be the corresponding conditional expectation. Then, for π -almost every $x \in S$,

$$P^x(S_n/\sqrt{n} \leq t) \rightarrow P(N(0, \sigma^2) \leq t) \text{ for any } t. \quad (3)$$

Clearly the quenched CLT implies that for any t

$$P(S_n/\sqrt{n} \leq t) \rightarrow P(N(0, \sigma^2) \leq t), \quad (4)$$

where $N(0, \sigma^2)$ is a normal random variable with mean 0 and variance σ^2 . This is called annealed CLT. On the other hand, there are numerous examples of processes satisfying the annealed CLT but failing to satisfy the quenched CLT. Some examples of this kind have been constructed by Volný and Woodroffe ([36], [38]). Therefore, some additional conditions are needed in order for the central limit theorem to hold in the quenched form.

The limit theorems started at a point are often encountered in evolutions in random media and they are of considerable importance in statistical mechanics. They are also useful for analyzing Markov chain Monte Carlo algorithms. Due to its importance, the problem was intensively studied in the literature. Two of the most influential papers are due to Derriennic and Lin ([14], [15]), which opened the way for many further results we shall mention throughout the paper. For a survey on quenched invariance principles under projective conditions we direct to [25].

The difficulty of obtaining quenched limit theorems consists in the fact that a Markov chain started at a point is no longer stationary. This is the reason this problem is very difficult to solve and there are still many open problems and long standing conjectures to be settled. Since stationary martingales satisfy the quenched CLT, the best technique to solve such a problem is to obtain a martingale approximation with a suitable rest. This technique was successfully used to get quenched CLT's for various classes of random variables in numerous papers, [14], [15], [39], [7], [36], [37], [21], [8], [12], [3], among others. The novelty here is that we use a martingale construction and approximation based on a new idea of conditioning with respect to both the past and the future of the Markov chain. This idea was introduced in [26], and [27]. In the annealed setting, if a stationary and ergodic Markov chain satisfies $E(S_n^2)/n$ is convergent, then the CLT holds (pending only a random centering) (see [26]). By using a similar martingale construction we obtain in this paper a new almost sure martingale approximation under P^x , for π -almost all starting points. This approximation will lead to the quenched CLT under the main condition that $E^x(S_n^2)/n$ is convergent π -almost surely. As application, we point out a new class of Markov chains satisfying the quenched CLT, defined by using projective conditions. In defining this class no assumption of irreducibility nor of aperiodicity is imposed. Under the additional assumptions that the Markov chain is irreducible, aperiodic and positively recurrent, Chen (Proposition 3.1., [6]) showed that if the CLT holds for the stationary Markov chain then the quenched CLT holds.

Here are some notations we shall use throughout the paper. We denote by $\|X\|$ the norm in $L^2(\Omega, \mathcal{F}, P)$. Unless otherwise specified, we shall assume the total ergodicity of the shift T of the sequence $(\xi_n)_{n \in \mathbb{Z}}$ with respect to P , i.e. T^m is ergodic for every $m \geq 1$. For the definition of the ergodicity of the shift we direct the reader to the subsection "A return to Ergodic Theory" in Billingsley [1] p. 494. Let us consider the operator Q induced by the kernel $Q(x, A)$ on bounded measurable functions on $(S, \mathcal{A})$ defined by $Qf(x) = \int_S f(y)Q(x, dy)$. By using Corollary 5 p. 97 in Rosenblatt [31], the shift of $(\xi_n)_{n \in \mathbb{Z}}$ is totally ergodic with respect to P if and only if the powers Q^m are ergodic with respect to π for all natural m (i.e. $Q^m f = f$ 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 [22].

Throughout the paper $\Rightarrow$ denotes the convergence in distribution. By the notation a.s. we understand P -almost surely. We shall also use the notation K for the conditional expectation operator on $L_1(P)$, namely

$$K(X) = E(X \circ T^{-1} | \xi_0), \quad K^n(X) = K(K^{n-1}(X)) = E(E(X \circ T^{-n}) | \xi_0).$$

The problem we address in Theorem 1 is to provide necessary and sufficient conditions for a quenched CLT for a class of Markov chains.

Theorem 1 *Assume (X_n) and (S_n) are defined by (1),*

$$\limsup_{n \rightarrow \infty} \frac{E(S_n^2)}{n} < \infty \quad (5)$$

and

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

Then there is $\sigma \geq 0$ such that

$$\frac{S_n}{\sqrt{n}} \Rightarrow N(0, \sigma^2) \text{ and } \frac{E(S_n^2)}{n} \rightarrow \sigma^2. \quad (7)$$

Furthermore, the following are equivalent:

$$(a) \quad \limsup_{n \rightarrow \infty} \frac{E^0(S_n^2)}{n} < \sigma^2 \text{ a.s.}$$

$$(b) \quad \frac{E^0(S_n^2)}{n} \text{ converges a.s.}$$

(c) *The quenched CLT in (2) holds and S_n^2/n is uniformly integrable under $P^0(\omega)$ for almost all ω .*

Remark 2 *Note that in condition (b) we do not have to specify the almost sure limit of $E^0(S_n^2)/n$. However, under our conditions it will always be σ^2 . In the sequel, when we say that the quenched limit theorem holds we understand that (c) of Theorem 1 holds.*

Relevant for the next Corollary is the notion of two-sided tail sigma field. We define the two-sided tail sigma field by

$$\mathcal{T}_d = \cap_{n \geq 1} (\mathcal{F}_{-n} \vee \mathcal{F}^n).$$

where $\mathcal{F}_{-n}$ and $\mathcal{F}^n$ are completed with all sets of probability 0. We say that $\mathcal{T}_d$ is trivial if for any $A \in \mathcal{T}_d$ we have $P(A) = 0$ or 1 .

Corollary 3 *Assume (X_n) and (S_n) are defined by (1), $\mathcal{T}_d$ is trivial and S_n^2/n is uniformly integrable. Then the Markov chain is totally ergodic, (7) holds and in addition (a), (b), and (c) of Theorem 1 are equivalent.*

Examples of stationary processes with trivial two-sided tail sigma field include absolutely regular Markov chains and interlaced mixing Markov chains. We refer to Subsection 2.5 in Bradley [4] for a survey and Bradley [5] for the proofs of the results in that survey. The definitions for these classes are also included in Subsection 2.1 of this paper.

In the next section we shall point out a sufficient condition for the quenched CLT by using projective criteria.

2 A sufficient condition for the quenched CLT

In this section we give a new sufficient condition for the quenched CLT based on the proof of Theorem 1. This condition arises in a computation of $E^0(S_n^2)$ by dyadic expansion.

We recall that the sequences (X_n) and (S_n) are defined by (1).

As shown in Theorem 2.7 in Cuny and Merlevède [8], it is known that the quenched CLT holds under a condition introduced by Maxwell and Woodroffe [20], namely

$$\sum_{n \geq 1} \frac{\|E(S_n|\xi_0)\|}{n^{3/2}} < \infty. \quad (8)$$

There are examples of Markov chains pointing out that, in general, condition (8) is as sharp as possible in some sense. Peligrad and Utev [24] constructed an example showing that for any sequence of positive constants (a_n) , $a_n \rightarrow 0$, there exists a stationary Markov chain such that

$$\sum_{n \geq 1} a_n \frac{\|E(S_n|\xi_0)\|}{n^{3/2}} < \infty$$

but $S_n/\sqrt{n}$ is not stochastically bounded. This example and other counterexamples provided by Volný [35], Dedecker [13] and Cuny and Lin [9], show that, in general, condition

$$\sum_{n \geq 1} \frac{\|E(S_n|\xi_0)\|^2}{n^2} < \infty \quad (9)$$

does not assure that $(S_n/\sqrt{n})$ is stochastically bounded. However, Corollary 3.5 in [27] contains a CLT under a reinforced form of (9). We provide next a quenched form of that result.

Theorem 4 *The quenched CLT holds under the condition*

$$\sum_{n \geq 1} \frac{\|E(S_n|\xi_0, \xi_n)\|^2}{n^2} < \infty. \quad (10)$$

As a corollary to Theorem 4, by Lemma 14 in [27] we have the following sufficient condition for (10) in terms of individual summands:

Corollary 5 *The quenched CLT holds under the condition*

$$\sum_{k \geq 1} \|E(X_0|\xi_{-k}, \xi_k)\|^2 < \infty. \quad (11)$$

We end this section by mentioning a conjecture due to Kipnis and Varadhan [19], which is unsolved. The conjecture asks if the quenched CLT and its functional form hold for stationary reversible and ergodic Markov chains ($Q = Q^*$ with Q^* the adjoint of Q) satisfying (9). For reversible Markov chains (9) is an equivalent formulation of $E(S_n^2)/n$ converges. This problem was investigated in several papers, [14], [7] where the quenched CLT for reversible Markov chains was obtained under various reinforcements of (9). Our condition (10) provides another reinforcement of (9), which in particular, gives a quenched CLT for reversible Markov chains.

2.1 Classes of stationary Markov Chains

There are several classes on stationary Markov chains for which $\|E(X_0|\xi_{-k}, \xi_k)\|$ is easy to estimate:

Absolutely regular Markov chains.

For a stationary Markov chain $\xi = (\xi_k)_{k \in \mathbb{Z}}$ with values in a separable Banach space endowed with the Borel sigma algebra $\mathcal{B}$, the coefficient of absolutely regularity is defined by (see Proposition 3.22 in Bradley [5])

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

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

Equivalently, (see Corollary 3.30 in Bradley [5])

$$\beta_n = \beta_n(\xi) = \beta(\xi_0, \xi_n) = \sup_{C \in \mathcal{B}^2} |P((\xi_0, \xi_n) \in C) - P((\xi_0, \xi_n^*) \in C)|,$$

where (ξ_0, ξ_n^*) are independent and identically distributed. This coefficient was introduced by Volkonskii and Rozanov [34] and was attributed there to Kolmogorov. If $\beta_n \rightarrow 0$, the Markov chain is called absolutely regular. Let us mention that there are numerous examples of stationary absolutely regular Markov

chains. 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. For easy reference we refer to Section 3 in Bradley [4] survey paper and to the references mentioned there. The condition $\beta_n \rightarrow 0$ implies total ergodicity.

In general, the CLT for this class requires the knowledge of the rate of convergence to 0 of the (β_n) coefficients (see for instance Doukhan et al. [16] and Peligrad [26] for a discussion on the CLT under $\beta_n \rightarrow 0$). Denote by Q the quantile function of $|X_0|$, i.e., the inverse function of $t \mapsto \mathbb{P}(|X_0| > t)$. According to Corollary 1 Doukhan et al. (1994), condition

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

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 $f \in \mathbb{L}_0^2(\pi)$, 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 $u \rightarrow 0^+$. 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. [12], there exists a Markov chain with these specifications, such that $S_n/\sqrt{n}$ does not satisfy the CLT.

If (12) holds, by Chen (Proposition 3.1., [6]), besides the CLT, the quenched CLT also holds.

It should be mentioned that a quenched CLT under (12) can be easily obtained from our Corollary 5. Indeed, by Rio's [29] covariance inequality (see also Theorem 1.1 in Rio [30]) 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 in Peligrad [26], 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 so, we recover the optimal quenched CLT for absolutely regular Markov chains under (12).

Perturbations of absolutely regular Markov chains.

Furthermore, let us point out another situation where Corollary 5 is useful. Let us consider two stationary Markov chains $\xi = (\xi_i)_{i \in \mathbb{Z}}$ and $\eta = (\eta_i)_{i \in \mathbb{Z}}$

independent among them. Then $(\xi_i, \eta_i)_{i \in \mathbb{Z}}$ is also a stationary Markov chain. Let f and g be two bounded functions, $|f| \leq C_1$ a.s. and $|g| \leq C_2$ a.s.

Define $Y_i = f(\xi_i) - E(f(\xi_i))$ and $Z_i = g(\eta_i) - E(g(\eta_i))$, and set $X_i = Y_i \cdot Z_i$. Assume that $\sum_{n \geq 1} \beta_n(\xi) < \infty$. Then, by independence,

$$\begin{aligned} \|E(X_0 | (\xi_n, \eta_n), (\xi_{-n}, \eta_{-n}))\| &= \|E(Y_0 | \xi_{-n}, \xi_n) E(Z_0 | \eta_{-n}, \eta_n)\|^2 \\ &\leq 6C_1^2 C_2^2 \beta_n(\xi), \end{aligned}$$

and, from Corollary 5, the quenched CLT follows for $(X_i)_{i \in \mathbb{Z}}$.

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

From Theorem 5 in Davydov (1973), we know that the β -mixing coefficients for ξ satisfy

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

where c is a positive constant. Construct the sequences $(Y_i)_{i \in \mathbb{Z}}$, $(Z_i)_{i \in \mathbb{Z}}$, and $(X_i)_{i \in \mathbb{Z}}$ as above. For this example,

$$\|E(Y_0 Z_0 | (\xi_n, \eta_n), (\xi_{-n}, \eta_{-n}))\| \leq c C_1^2 C_2^2 \frac{1}{n(\log(n + 1))^2},$$

and Corollary 5 applies to $(X_i)_{i \in \mathbb{Z}}$. Such examples cannot be treated by the results available so far in the literature.

Interlaced mixing Markov chains and their perturbations.

Another example where our results apply 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 $\rho_n^* \rightarrow 0$ as $n \rightarrow \infty$. The ρ^* -mixing condition goes back to Stein [33] and to Rosenblatt [32]. It is well-known that ρ^* -mixing implies total ergodicity. Also, there are known examples (see Example 7.16 in Bradley, [5]) of ρ^* -mixing sequences which are

not absolutely regular. Let X_0 as in Corollary 5. By the definition of ρ_n^* we immediately have

$$||E(X_0|\xi_{-n}, \xi_n)||^2 = |EX_0E(X_0|\xi_{-n}, \xi_n)| \leq \rho_n^* ||X_0|| \cdot ||E(X_0|\xi_0, \xi_n)||,$$

and so

$$||E(X_0|\xi_{-n}, \xi_n)|| \leq \rho_n^* ||X_0||.$$

Because for a ρ^* -mixing Markov chain, $\rho_n^* \rightarrow 0$ at least exponentially fast, (see Theorem 7.5 in Bradley [5]) by Corollary 5 we have that the quenched CLT holds for stationary ρ^* -mixing Markov chains. This result and further reaching results can be found in Cuny and Merlevède [8]. Although this result is not new for this class, the results in this paper provide a unified approach for different classes of Markov chains and also allow for other examples. As for the case of absolutely regular Markov chains, we can obtain a quenched CLT for certain functions of independent stationary Markov chains, where, only one of them is interlaced mixing and the other one is arbitrary.

3 Proofs

The starting point of the proofs is a new annealed CLT for Markov chains (see Theorem 1 in [26]):

Theorem 6 *Let $(X_n)_{n \in \mathbb{Z}}$ and $(S_n)_{n \geq 1}$ be as defined in (1), (ξ_n) is totally ergodic, and assume that (5) holds. Then, the following limit exists*

$$\lim_{n \rightarrow \infty} \frac{1}{n} ||S_n - E(S_n|\xi_0, \xi_n)||^2 = \sigma^2 \quad (13)$$

and

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

This result has the following consequence: (Corollary 5, [26]):

Theorem 7 *Assume that (5) and (6) holds. Then (7) holds.*

The main step of proving Theorem 1 is the following proposition:

Proposition 8 *If in addition to the conditions of Theorem 7 we assume that*

$$\limsup_{n \rightarrow \infty} \frac{E^0(S_n^2)}{n} \leq \sigma^2 \text{ a.s.} \quad (14)$$

then the quenched CLT in (2) holds.

Proof of Proposition 8

The proof of the quenched CLT is also based on the new idea to use a martingale approximation by conditioning with respect to past and future of the chain. We shall use the notations $E(X^2|\xi_0, \xi_n) = \|X\|_{0,n}^2$ and $E^0(X^2) = \|X\|_0^2$.

We start the proof by a decomposition in blocks of random variables, which is intended to weaken the dependence. 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 $u = u_n(m) = \lfloor n/m \rfloor$. So, for $k = 0, 1, \dots, u-1$, we have

$$Y_k = Y_k(m) = (X_{km+1} + \dots + X_{(k+1)m}). \quad (15)$$

Also denote

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

With this notations we write

$$\frac{1}{\sqrt{u}} S_u(m) := \frac{1}{\sqrt{u}} \sum_{k=0}^{u-1} \frac{1}{\sqrt{m}} Y_k(m) = \frac{1}{\sqrt{um}} S_{mu}.$$

In the first step of the proof we show that it is enough to prove that $S_u(m)/\sqrt{u}$ satisfies the quenched CLT. Let us show that the last block $Y_u(m)/\sqrt{n}$ has a negligible contribution to the convergence in distribution. With this aim, by Theorem 3.1 in Billingsley [1], it is enough to show that

$$E^0 \left(\frac{S_n - S_{mu}}{\sqrt{n}} \right)^2 = E^0 \left(\frac{Y_u(m)}{\sqrt{n}} \right)^2 \rightarrow 0 \text{ a.s. as } n \rightarrow \infty. \quad (16)$$

Note that the definition of $Y_u(m)$ and the Cauchy-Schwartz inequality imply that

$$E^0 \left(\frac{Y_u(m)}{\sqrt{n}} \right)^2 \leq m \frac{\max_{1 \leq j \leq n} E^0(X_j^2)}{n}.$$

Now, fix $M > 0$ and note that, for each $\varepsilon > 0$ and $n > M$,

$$\begin{aligned} \frac{\max_{1 \leq j \leq n} E^0(X_j^2)}{n} &\leq \varepsilon^2 + \frac{\sum_{j=1}^n E^0(X_j^2 I(|X_j| > \varepsilon \sqrt{n}))}{n} \\ &\leq \varepsilon^2 + \frac{\sum_{j=1}^n E^0(X_j^2 I(|X_j| > \varepsilon \sqrt{M}))}{n}. \end{aligned}$$

So, by Hopf's pointwise ergodic theorem for Dunford-Schwartz operators (Theorem 7.3 in Krengel [17])

$$\limsup_{n \rightarrow \infty} \frac{\max_{1 \leq j \leq n} E^0(X_j^2)}{n} \leq \varepsilon^2 + E(X_0^2 I(|X_0| > \varepsilon \sqrt{M})) \text{ a.s.}$$

and so, letting $\varepsilon \rightarrow 0$ and $M \rightarrow \infty$ we have

$$\limsup_{n \rightarrow \infty} \frac{\max_{1 \leq j \leq n} E^0(X_j^2)}{n} = 0 \text{ a.s.}$$

By the above arguments, we have proved that (16) holds for any m , and therefore $S_n/\sqrt{n}$ has the same limiting distribution as $S_{um}/\sqrt{n}$ under $P^0(\omega)$ for almost all ω . Since $um/n = [n/m](m/n) \rightarrow 1$ as $n \rightarrow \infty$, by Slutsky's theorem, $S_{um}/\sqrt{n}$ has the same limiting distribution as $S_u(m)/\sqrt{u}$. Furthermore, from (14) and (16) we easily derive that

$$\limsup_{u \rightarrow \infty} \frac{1}{u} \|S_u(m)\|_0^2 \leq \sigma^2 \text{ a.s.} \quad (17)$$

In the second step of the proof we construct the approximating martingale and mention its limiting properties.

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 $\mathcal{F}_{(k+1)m} = \mathcal{G}_k$. Also note that $\mathcal{G}_0 = \sigma(\xi_i, i \leq 0)$. Then we have $E(D_1 | \mathcal{G}_0) = 0$ a.s. Since we assumed that the shift T of the sequence $(\xi_n)_{n \in \mathbb{Z}}$ is totally ergodic, we deduce that for every m fixed, we have a stationary and ergodic sequence of square integrable martingale differences $(D_k, \mathcal{G}_k)_{k \geq 0}$.

Therefore, by the classical quenched central limit theorem for ergodic martingales, (see page 520 in Derriennic and Lin [14]) for every m , a fixed positive integer, we have for almost all $\omega \in \Omega$,

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

where N_m is a normally distributed random variable with mean 0 and variance $E(D_0^2) = m^{-1} \|S_m - E(S_m | \xi_0, \xi_m)\|^2$.

Since by (6) and (13),

$$m^{-1} \|S_m - E(S_m | \xi_0, \xi_m)\|^2 = m^{-1} (\|S_m\|^2 - \|E(S_m | \xi_0, \xi_m)\|^2) \rightarrow \sigma^2, \quad (18)$$

it follows that $N_m \Rightarrow N(0, \sigma^2)$. So, for almost all $\omega \in \Omega$,

$$\frac{1}{\sqrt{u}} M_u(m) \Rightarrow N_m \Rightarrow N(0, \sigma^2) \text{ under } P^0(\omega). \quad (19)$$

In the last step of the proof we shall approximate $S_u(m)$ by $M_u(m)$ in a suitable way, which will allow us to get the quenched limiting distribution $N(0, \sigma^2)$ also for $S_u(m)/\sqrt{u}$, completing the proof of this theorem. By using Theorem 3.2 in Billingsley [2] and taking into account (19), in order to establish the quenched CLT from Proposition 8, we have only to show that

$$\lim_{m \rightarrow \infty} \inf \limsup_{u \rightarrow \infty} E^0 \left(\frac{1}{\sqrt{u}} S_u(m) - \frac{1}{\sqrt{u}} M_u(m) \right)^2 = 0 \text{ a.s.} \quad (20)$$

Denote by

$$Z_k = m^{-1/2} E(Y_k | \xi_{km}, \xi_{(k+1)m}) \text{ and } R_u(m) = \sum_{k=0}^{u-1} Z_k. \quad (21)$$

With this notation we have:

$$S_u(m) = M_u(m) + R_u(m). \quad (22)$$

Let us show that $M_u(m)$ and $R_u(m)$ are orthogonal given $\mathcal{F}_0 \vee \mathcal{F}^n$. We show this property by analyzing the conditional expected value of all the terms of the product $M_u(m)R_u(m)$. For $m \leq n$, and $X \in \sigma(\xi_j, m \leq j \leq n)$ it is convenient to denote $E^{m,n}(X) = E(X | \mathcal{F}_m \vee \mathcal{F}^n) = E(X | \xi_m \vee \xi_n)$. Note that if $j < k$, since $\mathcal{F}_{(j+1)m} \subset \mathcal{F}_{km}$, and taking into account the Markov chain properties, we have that

$$\begin{aligned} & E^{0,n}[(Y_k - E(Y_k | \xi_{km}, \xi_{(k+1)m})) E(Y_j | \xi_{jm}, \xi_{(j+1)m})] \\ &= E^{0,n}[E^{(j+1)m,n}(Y_k - E(Y_k | \xi_{km}, \xi_{(k+1)m})) E(Y_j | \xi_{jm}, \xi_{(j+1)m})] \\ &= E^{0,n}[E^{(j+1)m,n}(Y_k - E(Y_k | \mathcal{F}_{km} \vee \mathcal{F}^{(k+1)m})) E(Y_j | \xi_{jm}, \xi_{(j+1)m})] = 0 \text{ a.s.} \end{aligned}$$

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

$$\begin{aligned} & E^{0,n}[(Y_k - E(Y_k | \xi_{km}, \xi_{(k+1)m})) E(Y_j | \xi_{jm}, \xi_{(j+1)m})] \\ &= E^{0,n}[E^{0,jm}(Y_k - E(Y_k | \xi_{km}, \xi_{(k+1)m})) E(Y_j | \xi_{jm}, \xi_{(j+1)m})] \\ &= E^{0,n}[E^{0,jm}(Y_k - E(Y_k | \mathcal{F}_{km} \vee \mathcal{F}^{(k+1)m})) E(Y_j | \xi_{jm}, \xi_{(j+1)m})] = 0 \text{ a.s.} \end{aligned}$$

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

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

Therefore $M_u(m)$ and $R_u(m)$ are indeed orthogonal under $E^{0,n}$ almost surely. By using now the decomposition (22), and the fact that $M_u(m)$ and $R_u(m)$ are orthogonal a.s. under $E^{0,n}$, we obtain the identity

$$\frac{1}{u} \|S_u(m)\|_{0,n}^2 = \frac{1}{u} \|M_n(m)\|_{0,n}^2 + \frac{1}{u} \|R_u(m)\|_{0,n}^2 \text{ a.s.} \quad (23)$$

By conditioning with respect to $\sigma(\xi_0)$ in (23), and taking into account the properties of conditional expectation, we also have

$$\frac{1}{u} \|S_u(m)\|_0^2 = \frac{1}{u} \|M_n(m)\|_0^2 + \frac{1}{u} \|R_u(m)\|_0^2.$$

By the definition of $M_u(m)$,

$$\frac{1}{u} \|M_n(m)\|_0^2 = \frac{1}{u} \sum_{k=0}^{u-1} E^0 D_k^2(m) = \frac{1}{u} \sum_{k=0}^{u-1} K^k (D_0^2(m)).$$

Now, by using the fact that (ξ_n) is totally ergodic along with Hopf's pointwise ergodic theorem for Dunford-Schwartz operators,

$$\lim_{u \rightarrow \infty} \frac{1}{u} \|M_u(m)\|_0^2 = \frac{1}{m} \|S_m - E(S_m | \xi_0, \xi_m)\|^2 \text{ a.s.}$$

So, by (18)

$$\lim_{m \rightarrow \infty} \lim_{u \rightarrow \infty} \frac{1}{u} \|M_u(m)\|_0^2 = \sigma^2.$$

By passing now to the limit in (23) and using (17) we obtain

$$\sigma^2 \geq \lim_{u \rightarrow \infty} \sup_u \frac{1}{u} \|S_u(m)\|_0^2 \geq \sigma^2 + \lim_{m \rightarrow \infty} \sup_m \lim_{u \rightarrow \infty} \sup_u \frac{1}{u} \|R_u(m)\|_0^2 \text{ a.s.}$$

Therefore,

$$\lim_{m \rightarrow \infty} \lim_{u \rightarrow \infty} \sup_u \frac{1}{u} \|R_u(m)\|_0^2 = 0 \text{ a.s.,}$$

which implies (20), and also the result follows. $\square$

In the next lemma we mention a property of the limit of $E^0(S_n^2)/n$. The idea of proof is borrowed from Dedecker and Merlevède [11], Subsection (3.2), where it was used in another context.

Lemma 9 *Assume that*

$$\frac{1}{n} E^0(S_n^2) \rightarrow \eta \text{ in } L_1.$$

Then η is measurable with respect to the invariant sigma field.

Proof. Recall the definition of shift T . Below, we denote by $TX = X \circ T^{-1}$. Clearly, η is $\mathcal{F}_0$ measurable. Then

$$E \left| E^0 \left(\frac{1}{n} S_n^2 - \eta \right) \right| \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (24)$$

Therefore, with the notation $E^1(\cdot) = E(\cdot | \mathcal{F}_1)$,

$$E \left| E^1 \left(\frac{1}{n} T S_n^2 - T \eta \right) \right| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Since $\mathcal{F}_0 \subset \mathcal{F}_1$, by the properties of conditional expectation, this implies

$$E \left| E^0 \left(\frac{1}{n} T S_n^2 - T \eta \right) \right| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

But, since the condition of this lemma implies that $E(S_n^2)/n$ is bounded,

$$\begin{aligned} \frac{1}{n} E |S_n^2 - T S_n^2| &\leq \frac{1}{n} E |(S_n^2 - (S_n - X_1 + X_{n+1})^2)| \\ &\leq \frac{1}{n} E |(X_1 - X_{n+1})(2S_n - X_1 + X_{n+1})| \\ &\leq \frac{4}{n} \|X_0\| \cdot (\|S_n\| + \|X_0\|) \rightarrow 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

So, by combining the last two limits, we also have

$$E \left| E^0 \left(\frac{1}{n} S_n^2 - T\eta \right) \right| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

By combining this limit with (24) we obtain

$$E | (\eta - E^0(T\eta)) | = 0,$$

implying that

$$\eta = E^0(T\eta) \text{ a.s.}$$

It remains to apply Lemma 3 from Dedecker and Merlevède [11], giving that $\eta = T\eta$ a.s. $\square$

Proof of Theorem 1

The first part in Theorem 1 is given in Theorem 7, so we have to prove only the second part of this theorem.

We argue first that (a) implies (c).

Since we assume (a) the quenched CLT holds by Proposition 8. Note that, by Theorem 25.11 in Billingsley [2], the quenched CLT implies that

$$\sigma^2 \leq \liminf_{n \rightarrow \infty} E^0(S_n^2)/n \text{ a.s.},$$

which combined with (a) gives $E^0(S_n^2)/n \rightarrow \sigma^2$ a.s. Now, because we have the quenched CLT and $E^0(S_n^2)/n \rightarrow \sigma^2$ a.s., by Theorem 3.6 in Billingsley [2], we have the uniform integrability of $(S_n^2/n)_n$ under $P^0(\omega)$ for almost all ω .

Clearly (c) implies (b) by the convergence of the moments in the CLT in Theorem 3.5 in Billingsley [2]. Actually (c) implies $E^0(S_n^2)/n \rightarrow \sigma^2$ a.s.

It remains to show that (b) implies (a).

We start from (b), which is: for some random variable η , $E^0(S_n^2)/n \rightarrow \eta$ a.s. By Theorem 7 the annealed CLT together with the convergence of the second moments hold. Furthermore, by Theorem 3.6 Billingsley [2] we have that S_n^2/n is uniformly integrable. This implies that $E^0(S_n^2)/n$ is also uniformly integrable, which, together with (b), implies the convergence $E^0(S_n^2)/n \rightarrow \eta$ in $L^1(\Omega, \mathcal{F}, P)$. By Lemma 9, the limit of $E^0(S_n^2)/n$ is measurable with respect to the trivial invariant sigma field, which means that it is constant. Because we assumed that $E(S_n^2)/n \rightarrow \sigma^2$ it follows that $\eta = \sigma^2$. $\square$

Proof of Corollary 3

The fact that $(\xi_k)_{k \in \mathbb{Z}}$ is totally ergodic follows by Proposition 2.12 in the Vol. 1 of Bradley [5]. Then, since we assume that (S_n^2/n) is uniformly integrable, by Lemma 4 in [28] we deduce that (6) holds. The result follows by Theorem 1. $\square$

We move now to prove Theorem 4. Relevant for the proof is the weak L_p space, define by

$$L^{p,w} = \{f \text{ measurable, } \sup_{\lambda>0} \lambda^p P(|f| \geq \lambda) < \infty\}.$$

Denote the norm in $L^{p,w}$ by $\|\cdot\|_{p,w}$. Below we also use the notation $\bar{S}_k = S_{2k} - S_k$.

The main step for proving Theorem 4 is the following upper bound concerning $E^0(S_n^2)/n$.

Lemma 10 *For any stationary and ergodic sequence (η_n) , not necessarily Markov, define (V_n) by $V_n = g(\eta_n)$ and $S_n = \sum_{k=1}^n V_k$. Assume V_0 is in L_2 and is centered at expectation. Let $\mathcal{K}_n = \sigma(\eta_j, j \leq n)$ and keep the notation $E^n(X) = E(X|\mathcal{K}_n)$. Then we have the following bound*

$$\|\sup_n \frac{1}{n} E^0(S_n^2)\|_{1,w} \leq 6E(V_0^2) + 12 \sum_{k \geq 0} \frac{1}{2^k} E|E^0(S_{2^k} \bar{S}_{2^k})|.$$

Proof. The proof follows the traditional technique of dyadic recurrence, initiated by Ibragimov [18] and further developed in [23], [5], [24], [8], among many others.

Let $2^{r-1} \leq n < 2^r$ and write its binary expansion:

$$n = \sum_{k=0}^{r-1} 2^k a_k \text{ where } a_{r-1} = 1 \text{ and } a_k \in \{0, 1\} \text{ for } k = 0, \dots, r-2.$$

Notice that

$$S_n = \sum_{i=0}^{r-1} a_i T_{2^i} \text{ where } T_{2^i} = \sum_{i=n_{i-1}+1}^{n_i} X_i, \quad n_i = \sum_{j=0}^i a_j 2^j \text{ and } n_{-1} = 0.$$

By the triangle inequality, (recall that $S_0 = 0$)

$$(E^0(S_n^2))^{1/2} = \|S_n\|_0 = \left\| \sum_{i=0}^{r-1} a_i E^0(S_{n_i} - S_{n_{i-1}}) \right\|_0 \leq \sum_{i=0}^{r-1} \|S_{n_i} - S_{n_{i-1}}\|_0.$$

Also, by stationarity and because $n_i - n_{i-1}$ is either 0 or 2^i , we obtain

$$\begin{aligned} E^0(S_{n_i} - S_{n_{i-1}})^2 &= E^0(E((S_{n_i} - S_{n_{i-1}})^2 | \mathcal{K}_{n_{i-1}})) = K^{n_{i-1}}(E^0(S_{n_i - n_{i-1}}^2)) \\ &\leq K^{n_{i-1}}(E^0(S_{2^i}^2)). \end{aligned}$$

It follows that

$$\frac{1}{n} E^0(S_n^2) \leq \frac{1}{n} \left(\sum_{i=0}^{r-1} [K^{n_{i-1}}(E^0(S_{2^i}^2))]^{1/2} \right)^2 \leq 6 \sup_i \frac{1}{2^i} [K^{n_{i-1}}(E^0(S_{2^i}^2))] \quad (25)$$

We fix $i \geq 1$ and evaluate the term in the right hand side of (25).

For each k and j , denote $S_{1,2^k} = S_{2^k}$, $S_{j,2^k} = S_{j2^k} - S_{(j-1)2^k}$. Clearly,

$$S_{2^i}^2 = S_{1,2^{i-1}}^2 + S_{2,2^{i-1}}^2 + 2S_{1,2^{i-1}}S_{2,2^{i-1}} = S_{1,2^{i-2}}^2 + S_{2,2^{i-2}}^2 + S_{3,2^{i-2}}^2 + S_{4,2^{i-2}}^2 \\ + 2(S_{1,2^{i-2}}S_{2,2^{i-2}} + S_{3,2^{i-2}}S_{4,2^{i-2}} + S_{1,2^{i-1}}S_{2,2^{i-1}}).$$

We continue the recurrence and get the representation:

$$S_{2^i}^2 = \sum_{j=1}^{2^i} V_j^2 + 2 \sum_{k=0}^{i-1} \sum_{j=1}^{2^{i-k-1}} S_{2j-1,2^k} S_{2j,2^k}.$$

Denoting by

$$g_k = E^0(S_{2^k} \bar{S}_{2^k}),$$

note that, by using the definition of the conditional expectation K ,

$$E^0(S_{2j-1,2^k} S_{2j,2^k}) = E^0(E(S_{2j-1,2^k} S_{2j,2^k} | \mathcal{K}_{(2j-2)2^k})) = K^{(j-1)2^{k+1}}(g_k).$$

By the above considerations,

$$\frac{1}{2^i} E^0(S_{2^i}^2) = \frac{1}{2^i} \sum_{j=1}^{2^i} K^j(V_0^2) + 2 \sum_{k=0}^{i-1} \frac{1}{2^{i-k}} \left(\sum_{j=1}^{2^{i-k-1}} K^{(j-1)2^{k+1}} \right) \left(\frac{1}{2^k} g_k \right).$$

So,

$$\frac{1}{2^i} K^{n_{i-1}}(E^0(S_{2^i}^2)) = \frac{1}{2^i} \sum_{j=1}^{2^i} K^j(K^{n_{i-1}}(V_0^2)) \\ + 2 \sum_{k=0}^{i-1} \frac{1}{2^{i-k}} \left(\sum_{j=1}^{2^{i-k-1}} K^{(j-1)2^{k+1}} \right) \left(\frac{1}{2^k} K^{n_{i-1}}(g_k) \right).$$

So, with the notation

$$\sup_n \frac{1}{n} \left(\sum_{j=0}^{n-1} K^{j2^{k+1}}(\cdot) \right) = \mathcal{M}_k(\cdot),$$

we obtain

$$\frac{1}{2^i} K^{n_{i-1}}(E^0 S_{2^i}^2) \leq \sup_n \frac{1}{n} \sum_{j=1}^n K^j(K^{n_{i-1}} V_0^2) + 2 \sum_{k=0}^{i-1} \frac{1}{2^k} \mathcal{M}_k(|K^{n_{i-1}}(g_k)|).$$

By using now Hopf's ergodic theorem (see, e.g., Krengel [17], Lemma 6.1, page 51, and Corollary 3.8, page 131),

$$\|\mathcal{M}_k\left(\left|K^{n_{i-1}}\left(\frac{g_k}{2^k}\right)\right|\right)\|_{1,w} \leq \frac{1}{2^k} \|K^{n_{i-1}} g_k\|_1 \leq \frac{1}{2^k} \|g_k\|_1.$$

and also

$$\left\| \sup_n \frac{1}{n} \sum_{j=1}^n K^j(K^{n_{i-1}}(V_0^2)) \right\|_{1,w} \leq E(K^{n_{i-1}} V_0^2) = E(V_0^2).$$

Therefore,

$$\left\| \sup_i \frac{1}{2^i} K^{n_{i-1}}(E^0(S_{2^i}^2)) \right\|_{1,w} \leq E(V_0^2) + 2 \sum_{k \geq 0} \frac{1}{2^k} E|E^0(S_{2^k} \bar{S}_{2^k})|.$$

To obtain the conclusion of this lemma we combine this last inequality with (25). $\square$

Based on this lemma we shall provide another bound needed for the proof of Theorem 4.

Lemma 11 *Assume in addition to the conditions of Lemma 10 that the sequence (η_n) has the Markov property. Then, for some universal constant C ,*

$$\| \sup_n \frac{1}{n} E^0(S_n^2) \|_{1,w} \leq C E(V_0^2) + C \sum_{n \geq 1} \frac{1}{n^2} E(E(S_n|\eta_0, \eta_n))^2. \quad (26)$$

Proof. This bound follows from Lemma 10. We start by noting that, by the properties of conditional expectations and the Markov property,

$$\begin{aligned} E(S_{2^k} \bar{S}_{2^k} | \eta_0) &= E(S_{2^k} E(\bar{S}_{2^k} | \eta_{2^k}) | \eta_0) = E(E(S_{2^k} E(\bar{S}_{2^k} | \eta_{2^k}) | \eta_0, \eta_{2^k}) | \eta_0) \\ &= E(E(S_{2^k} | \eta_0, \eta_{2^k}) E(\bar{S}_{2^k} | \eta_{2^k}) | \eta_0). \end{aligned}$$

So, by the Cauchy-Schwartz inequality,

$$\begin{aligned} E|E(S_{2^k} \bar{S}_{2^k} | \eta_0)| &\leq E|E(S_{2^k} | \eta_0, \eta_{2^k}) E(\bar{S}_{2^k} | \eta_{2^k})| \\ &\leq \frac{1}{2} E(E(S_{2^k} | \eta_0, \eta_{2^k}))^2 + \frac{1}{2} E(E(\bar{S}_{2^k} | \eta_{2^k}))^2 \\ &\leq E(E(S_{2^k} | \eta_0, \eta_{2^k}))^2. \end{aligned}$$

Therefore

$$\sum_{k \geq 0} \frac{1}{2^k} E|E((S_{2^k} \bar{S}_{2^k}) | \eta_0)| \leq \sum_{k \geq 0} \frac{1}{2^k} E(E(S_{2^k} | \eta_0, \eta_{2^k}))^2.$$

As proven in Lemmas 12 and 13 in [27], for some positive constant c ,

$$\sum_{k \geq 0} \frac{1}{2^k} E(E(S_{2^k} | \eta_0, \eta_{2^k}))^2 \leq c \sum_{n \geq 1} \frac{1}{n^2} E(E(S_n | \eta_0, \eta_n))^2.$$

It remains to apply Lemma 10 to obtain the desired result. $\square$

Proof of Theorem 4

The CLT and the convergence of moments under condition (10) are known (see Corollary 9 in [27]). The proof of the quenched CLT is based on the proof of Proposition 8 combined with Lemma 11.

For m fixed, we apply Lemma 11 with $\eta_{\ell+1} = (\xi_{\ell m}, \xi_{(\ell+1)m})$ and the sequence $V_{\ell+1}(m) = E(Y_{\ell} | \xi_{\ell m}, \xi_{(\ell+1)m}) / \sqrt{m}$ where $(Y_{\ell})_{\ell \in \mathbb{Z}}$ is the extension to a stationary sequence of Y_k defined in (15). It is easy to see that, by using the Markov property and the properties of the conditional expectation, we obtain for $k \geq 0$

$$E\left(\sum_{j=1}^{k+1} V_j | \eta_0, \eta_{k+1}\right) = \frac{1}{\sqrt{m}} E(S_{km} | \xi_0, \xi_{km}) + V_{k+1}.$$

It follows that

$$\|E(\sum_{j=1}^{k+1} V_j | \eta_0, \eta_{k+1})\|^2 \leq \frac{2}{m} \|E(S_{km} | \xi_0, \xi_{km})\|^2 + \frac{2}{m} \|E(S_m | \xi_0, \xi_m)\|^2.$$

So, for $R_u(m)$ defined in (21), $R_u(m) = \sum_{j=1}^u V_j(m)$, we obtain by Lemma 11, for some $C_1 > 0$,

$$\|\sup_u \frac{1}{u} E^0(R_u^2(m))\|_{1,w} \leq C_1 \sum_{k=1}^{\infty} \frac{1}{k^2 m} E(E(S_{km} | \xi_0, \xi_{km}))^2.$$

By the Cauchy-Schwartz inequality, and the properties of the conditional expectation,

$$\frac{1}{mk^2} E(E(S_{km} | \xi_0, \xi_{km}))^2 \leq \frac{1}{k^2} E(E(S_k | \xi_0, \xi_k))^2$$

and also

$$\sum_{k=1}^{\infty} \frac{1}{mk^2} E(E(S_{km} | \xi_0, \xi_{km}))^2 \leq \sum_{k=1}^{\infty} \frac{1}{k^2} E(E(S_k | \xi_0, \xi_k))^2 < \infty.$$

Since (10) implies that $\lim_{n \rightarrow \infty} E(E(S_n | \xi_0, \xi_n))^2/n = 0$, for any k fixed, we have that

$$\lim_{m \rightarrow \infty} \frac{1}{mk} E(E(S_{km} | \xi_0, \xi_{km}))^2 = 0.$$

So, by the dominated convergence theorem for discrete measures,

$$\lim_{m \rightarrow \infty} \sum_{k=1}^{\infty} \frac{1}{mk^2} E(E(S_{km} | \xi_0, \xi_{km}))^2 = 0.$$

It follows that

$$\lim_{m \rightarrow \infty} \|\sup_u \frac{1}{u} E^0(R_u^2(m))\|_{1,w} = 0.$$

By Theorem 4.1 in Billingsley [1], note that the Fatou Lemma also holds in the space $L^{1,w}$. Therefore,

$$\|\liminf_{m \rightarrow \infty} \sup_u \frac{1}{u} E^0(R_u^2(m))\|_{1,w} \leq \liminf_{m \rightarrow \infty} \|\sup_u \frac{1}{u} E^0(R_u^2(m))\|_{1,w} = 0.$$

and so

$$\liminf_{m \rightarrow \infty} \sup_u \frac{1}{u} E^0(R_u^2(m)) = 0 \text{ a.s.}$$

This proves that the martingale decomposition in (20) holds. The proof is now ended as in the proof of Proposition 8. $\square$

The author has no conflicts of interest to disclose.

The author confirms that the data supporting the findings of this study are available within the article.

Acknowledgement 12 *This paper was partially supported by the NSF grant DMS-2054598. It is dedicated to Michael Lin's 80th birthday. Many thanks for the referee's valuable comments and suggestions that improved the presentation of this paper.*

References

- [1] Billingsley, P. (1995). Probability and Measure. 3rd Edition, Wiley Series in Probability and Mathematical Statistics.
- [2] Billingsley, P. (1999). Convergence of probability measures. Second edition. Wiley, New York.
- [3] Barrera, D., Peligrad, C. and Peligrad, M. (2016). On the functional CLT for stationary Markov Chains started at a point. Stoc. Proc. Appl. 126 1885–1900.
- [4] Bradley, R.C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. Probability Surveys 2 107-144.
- [5] Bradley, R.C. (2007). Introduction to strong mixing conditions 1, 2, 3. Kendrick Press, Heber City, UT.
- [6] Chen, X. (1999). Limit theorems for functionals of ergodic Markov chains with general state space. Memoirs of the American Mathematical Society 139.
- [7] Cuny, C., and Peligrad, M. (2012). Central limit theorem started at a point for stationary processes and additive functional of reversible Markov Chains. Journal of Theoretical Probability 25 171-188.
- [8] Cuny, C. and Merlevède, F. (2014). On martingale approximations and the quenched weak invariance principle, Ann. Probab. 42 760–793.
- [9] Cuny, C. and Lin, M. (2016). Limit theorems for Markov chains by the symmetrization method. J. Math. Anal. Appl. 434 52–83.
- [10] Davydov, Yu. A. (1973). Mixing conditions for Markov chains. Theory Probab. Appl. 18 312-328.
- [11] Dedecker, J. and Merlevède, F. (2002). Necessary and sufficient conditions for the conditional central limit theorem. Ann. Probab. 30 1044–1081.
- [12] Dedecker, J., Merlevède, F. and Peligrad, M. (2014). A quenched weak invariance principle, Ann. Inst. H. Poincaré Probab. Statist. 50 872–898.
- [13] Dedecker, J. (2015). On the optimality of McLeish’s conditions for the central limit theorem. C. R. Math. Acad. Sci. Paris 353 557–561.
- [14] Derriennic, Y. and Lin, M. (2001). The central limit theorem for Markov chains with normal transition operators, started at a point. Probab. Theory Related Fields 119 508–528.
- [15] Derriennic, Y. and Lin, M. (2003). The central limit theorem for Markov chains started at a point. Probab. Theory Related Fields 125 73–76.

- [16] 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.
- [17] Krengel, U. (1985). *Ergodic Theorems*. de Gruyter Studies in Mathematics 6. de Gruyter, Berlin.
- [18] Ibragimov, I.A. (1975). A note on the central limit theorem for dependent random variables. *Theory Probab. Appl.* 20 135-141.
- [19] Kipnis, C. and Varadhan, S.R.S. (1986). Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.* 104 1–19.
- [20] Maxwell, M. and Woodroffe, M. (2000). Central limit theorems for additive functionals of Markov chains. *Ann. Probab.* 28 713–724.
- [21] Merlevède, F., Peligrad, C. and Peligrad, M. (2012). Almost sure invariance principles via martingale approximation. *Stochastic Process. Appl.* 122 170–190.
- [22] Quas, A. (2009). Ergodicity and mixing properties. in: R. E. Meyers (ed.), *Encyclopedia of Complexity and Systems Science* 2918–2933, Springer.
- [23] Peligrad, M (1982). Invariance principles for mixing sequences of random variables. *Ann. Probab.* 10 968-981.
- [24] Peligrad, M. and Utev, S. (2005). A new maximal inequality and invariance principle for stationary sequences. *Ann. Probab.* 33 798-815.
- [25] Peligrad, M. (2015). Quenched Invariance Principles via Martingale Approximation In: *Asymptotic Laws and Methods in Stochastics*. Volume in Honour of Miklos Csörgo. Fields Institute Communications Series, Springer-Verlag New York. 76 121-137.
- [26] Peligrad, M. (2020). A new CLT for additive functionals of Markov chains. *Stochastic Processes and their Applications*. 130 5695-5709.
- [27] Peligrad, M. (2020). On the CLT for additive functionals of Markov chains. *Electronic Communications in Probability*. 25. article number 40, 1-10.
- [28] Peligrad, M. (2022). On the CLT for stationary Markov chains with trivial tail sigma field. (2022). [arXiv:2208.14532](https://arxiv.org/abs/2208.14532). To appear in *Electronic Communications in Probability*.
- [29] Rio, E. (1993). Covariance inequalities for strongly mixing processes. *Ann. Inst. Henri Poincaré Probab. Stat.* 29 587–597.
- [30] Rio, E. (2017). *Asymptotic theory of weakly dependent random processes*. Springer.

- [31] Rosenblatt, M. (1971). Markov processes. Structure and asymptotic behavior. Springer, Berlin.
- [32] Rosenblatt, M. (1972). Central limit theorems for stationary processes. Proceedings of the Sixth Berkeley Symposium on Probability and Statistics, vol. 2, 551-561. University of California Press, Los Angeles.
- [33] Stein C. (1972). 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, 583-602. University of California Press, Los Angeles.
- [34] Volkonskii, V.A. and Rozanov Yu.A. (1959). Some limit theorems for random functions I. Theor. Probab. Appl. 4 178-197.
- [35] Volný, D. (2010). Martingale approximation and optimality of some conditions for the central limit theorem. J. Theoret. Probab. 23 888–903.
- [36] Volný, D. and Woodroffe, M. (2010). An example of non-quenched convergence in the conditional central limit theorem for partial sums of a linear process, in: Dependence in Analysis, Probability and Number Theory (The Phillipp Memorial Volume), Kendrick Press 317–323.
- [37] Volný, D. and Woodroffe, M. (2014). Quenched central limit theorems for sums of stationary processes, Statistics & Probability Letters 85 161-167.
- [38] Volný, D. and Woodroffe, M. (2017). Quenched central limit theorem for stationary linear processes. Statistica Sinica 27 519-533.
- [39] Wu, W.B. and Woodroffe M. (2004). Martingale approximations for sums of stationary processes. Ann. Probab. **32**, 1674–1690.