On the invariance principle for reversible Markov chains

Magda Peligrad[1] and Sergey Utev [2]

Key words: reversible processes; Markov chains; functional central limit theorem.

Mathematical Subject Classification (2000): 60F05, 60F17, 60J05.

Abstract

In this paper, we investigate the functional central limit theorem for stochastic processes associated to partial sums of additive functionals of reversible Markov chains with general spate space, under the normalization standard deviation of partial sums. For this case, we show that the functional central limit theorem is equivalent to the fact that the variance of partial sums is regularly varying with exponent 1 and the partial sums satisfy the CLT. It is also equivalent to the conditional CLT.

1 Introduction and Result

Reversible Markov chains play a very important role in applications to infinite particle systems, random walks, processes in random media, Metropolis-Hastings algorithms. For instance, Kipnis and Varhadhan (1986) and Kipnis and Landim (1999) considered applications to interacting particle systems, Tierney (1994), Zhao et al. (2010), Longla et al. (2012) discussed the applications to Markov Chain Monte Carlo. Our paper is motivated by the functional limit theorem in the paper by Longla et al. (2012). Our result will bring further clarification on this subject. Without assuming aperiodicity or irreducibility properties, we shall show that for an additive functional of a stationary

[1]Department of Mathematical Sciences, University of Cincinnati, PO Box 210025, Cincinnati, Oh 45221-0025, USA. `peligrm@ucmail.uc.edu`

[2]Department of Mathematics, University of Leicester, University Road, LE1 7RH UK, `su35@leicester.ac.uk`

reversible Markov chain the functional CLT is equivalent to CLT plus the fact that the variance of partial sums is regularly varying with exponent 1.

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 general state space $(S, \mathcal{A})$. The marginal distribution is denoted by $\pi(A) = \mathbb{P}(\xi_0 \in A)$. Assume that there is a regular conditional distribution for ξ_1 given ξ_0 denoted by $Q(x, A) = \mathbb{P}(\xi_1 \in A | \, \xi_0 = x)$. Let Q also denote the Markov operator acting via $(Qf)(x) = \int_S f(s)Q(x, ds)$. 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 some function $f \in \mathbb{L}_0^2(\pi)$, let

$$X_i = f(\xi_i), \; S_n = \sum_{i=1}^{n} X_i, \; \sigma_n = (\mathbb{E}S_n^2)^{1/2}. \tag{1}$$

For any integrable random variable X we denote $\mathbb{E}_{\xi_n}(X) = \mathbb{E}(X|\xi_n)$. The symbol $\Rightarrow$ denotes convergence in distribution.

The Markov chain is called reversible if $Q = Q^*$, where Q^* is the adjoint operator of Q. The condition of reversibility is equivalent to requiring that (ξ_0, ξ_1) and (ξ_1, ξ_0) have the same distribution or

$$\int_A Q(\omega, B)\pi(d\omega) = \int_B Q(\omega, A)\pi(d\omega),$$

for all Borel sets $A, B \in \mathcal{A}$.

Gordin and Lifšic (1981) proved a CLT for functionals of a normal Markov chain. In particular this implies a CLT for functions of reversible Markov chains under the normalization $\sqrt{n}$. Kipnis and Varadhan (1986) provided a functional form of this result. They showed that for a stationary reversible and ergodic Markov chain the condition $var(S_n)/n \to \sigma_f^2$ implies the convergence of $S_{[ns]}/\sqrt{n}$ to the Brownian motion $|\sigma_f|W(s)$, (here, $0 \le s \le 1$, $[ns]$ is the integer part of ns and $W(s)$ is the standard Brownian motion). All these results used the normalization $\sqrt{n}$.

Zhao et al. (2010) addressed the central limit theorem question for reversible Markov chains under the weaker condition,

$$\sigma_n^2 = nh(n), \tag{2}$$

where h is a slowly varying function (i.e. $\lim_{n\to\infty} h(nu)/h(n) = 1$ for all $u > 0$). First, they proved that if $\liminf_n h(n) = 0$, then necessarily

$$2S_n = (1 + (-1)^{n-1})X_1 \quad a.s. \tag{3}$$

and if $\liminf_n h(n) = c^2 \neq 0$ then $\sigma_n^2/n \to c^2$. In the following, to avoid the trivial case, we shall assume that

$$\liminf \sigma_n^2/n > 0.$$

Zhao et al. (2010) also showed, by a class of examples, the surprising result that the distribution of S_n/σ_n needs not converge to the standard normal distribution under (2). Their example satisfies the central limit theorem, in the sense that

$$\frac{S_n}{\sigma_n} \Rightarrow |c|Z, \tag{4}$$

where Z is a standard normal variable, $N(0,1)$, and $|c| \neq 0, 1$. A large class of examples satisfying (4) is given in Deligiannidis et al. (2014). This paper also contains necessary and sufficient conditions for (2) in terms of the operator spectral measure.

A natural question is whether, in the context of reversible Markov chains, the central limit theorem in (4) implies the invariance principle namely

$$\frac{S_{[ns]}}{\sigma_n} \Rightarrow |c|W(s), \tag{5}$$

where $W(s)$, $s \geq 0$ is the standard Brownian motion. This question is interesting in itself, especially in the light of recent examples of stationary sequences which satisfy CLT but not its functional forms (Giraudo and Volný, 2014).

A step in this direction is Theorem 2 in Longla et al. (2012), showing that conditional CLT implies the functional CLT in (5) for functions of reversible Markov chains. By the conditional CLT we understand that some $c > 0$ and for all t,

$$\mathbb{E}_{\xi_0}(\exp it\frac{S_n}{\sigma_n}) \to \exp(-\frac{t^2}{2c^2}) \text{ in probability.} \tag{6}$$

This form of the conditional CLT was essentially used in Longla et al. (2012), in order to establish the convergence of finite dimensional distributions.

In this paper we show that actually (4) and (2) implies (5) for functionals of reversible Markov chains. More precisely we shall establish the following theorem:

Theorem 1 *Assume that* (X_n) *is defined by (1) and the Markov chain is stationary and reversible. Then, the following statements are equivalent:*

(a) The functional CLT in (5) holds.

(b) The CLT in (4) holds and the variance of partial sums is regularly varying with exponent 1 *(as in relation (2)).*

(c) The conditional CLT in (6) holds.

2 Proof of Theorem 1

The fact that (a) implies (b) follows by standard arguments in the following way. Clearly, since the partial sum is just a finite dimensional distribution for $s = 1$, the functional CLT in (5) implies the CLT in (4). Then, by (4), we have that $S_{[ns]}/\sigma_{[ns]} \Rightarrow |c|Z$, for every $s > 0$. On the other hand, the convergence in (5) implies $S_{[ns]}/\sigma_n \Rightarrow |c|sZ$ for all $s > 0$. By the theorem of types (see Theorem 14.2 in Billingsley, 1995), it follows that $\sigma^2_{[ns]}/\sigma^2_n \to s$. The fact that (c) implies (a) was established in Theorem 2 in Longla et al. (2012). It remains to prove that (b) implies (c). The idea of proof is to show that from any subsequence of $\mathbb{E}_{\xi_0}(\exp it(S_n/\sigma_n))$ we can extract one converging to $\exp(-t^2/2c^2)$ in probability. To achieve this goal, we need a technical lemma concerning conditional convergence.

Lemma 2 *Assume* (V_n, η) *is convergent in distribution to* (V, Y)*. Then, we can construct on the same probability space a sequence* (V'_n, η')*, where each vector is distributed as* (V_n, η) *and a vector* (V', η') *distributed as* (V, Y) *such that for all* t,

$$\mathbb{E}_{\eta'} \exp(itV'_n) \to \mathbb{E}_{\eta'} \exp(itV') \text{ in probability.}$$

Proof of Lemma 2. By the Skorohod theorem (see Skorohod, 1956), we can construct on the same probability space a sequence (V'_n, η'), where each vector is distributed as (V_n, η), and a vector (V', Y') distributed as (V, Y) such that $(V_n', \eta')(\omega) \to (V', Y')(\omega)$ for all ω. Clearly $\eta' = Y'$ and $V_n'(\omega) \to V'(\omega)$ for all ω. Then, by the mean value theorem, for any t and $\delta > 0$,

$$|\mathbb{E}_{\eta'} \exp(itV_n') - \mathbb{E}_{\eta'} \exp(itV')| \leq$$
$$\mathbb{E}_{\eta'} |\exp(itV_n') - \exp(itV')| I(|V_n' - V'| \leq \delta) + 2\mathbb{P}_{\eta'}(|V_n' - V'| > \delta)$$
$$\leq |t|\delta + 2\mathbb{P}_{\eta'}(|V_n' - V'| > \delta) \text{ a.s.}$$

By taking the expectation

$$\mathbb{E}|\mathbb{E}_{\eta'} \exp(itV_n') - \mathbb{E}_{\eta'} \exp(itV')| \leq |t|\delta + 2\mathbb{P}(|V_n' - V'| > \delta),$$

which tends to 0 as $n \to \infty$ and then $\delta \to 0$. $\diamond$

We continue to prove Theorem 1 by proving that (b) implies (c). We remind that we work under the assumption that $\liminf \sigma_n^2/n > 0$. Note that, by stationarity and the fact that X_0 is square integrable, it follows that $\max_{1 \leq i \leq n} |X_i|/\sqrt{n} \to 0$ *a.s.* and in $\mathbb{L}^2$ and therefore

$$\max_{1 \leq i \leq n} |X_i|/\sigma_n \to 0 \ a.s. \text{ and in } \mathbb{L}^2. \tag{7}$$

This property will allow us to adjust the sums by a few variables without changing the limiting distribution. We shall use some notations: $\bar{S}_n = X_{n+1} + ... + X_{2n}$; $\mathcal{P}_n = \sigma(\xi_i, i \leq n)$ is the past sigma field.

Step 1. As a preliminary computation, we show that for all t,

$$\mathbb{E}(\mathbb{E}_{\xi_0}(\exp \frac{itS_n}{\sigma_n}))^2 \to \exp(-\frac{t^2}{c^2}) \ . \tag{8}$$

By (7), the properties of conditional expectation and Markov property,

$$\mathbb{E}(\exp \frac{itS_{2n}}{\sigma_n}) = \mathbb{E}(\exp \frac{it(S_n + \bar{S}_n)}{\sigma_n}) = \mathbb{E}\left((\exp \frac{itS_n}{\sigma_n})\mathbb{E}(\exp \frac{it\bar{S}_n}{\sigma_n}|\mathcal{P}_n)\right)$$
$$= \mathbb{E}\left((\exp \frac{itS_n}{\sigma_n})\mathbb{E}_{\xi_n}(\exp \frac{it\bar{S}_n}{\sigma_n})\right) .$$

We write now $S_n = (X_0 + ... + X_{n-1}) - X_0 + X_n = S_{[0,n-1]} + (X_n - X_0)$.

Clearly by simple computations and (7),

$$|\mathbb{E}\left(\exp\frac{itS_n}{\sigma_n}\mathbb{E}_{\xi_n}(\exp\frac{it\bar{S}_n}{\sigma_n})\right)-\mathbb{E}\left(\exp\frac{itS_{[0,n-1]}}{\sigma_n}\mathbb{E}_{\xi_n}(\exp\frac{it\bar{S}_n}{\sigma_n})\right)|$$
$$\leq\mathbb{E}|\exp\frac{it(X_n-X_0)}{\sigma_n}-1|\to 0 \text{ as } n\to\infty.$$

Therefore, by combining these facts we obtain

$$\mathbb{E}(\exp\frac{itS_{2n}}{\sigma_n})-\mathbb{E}\left(\exp\frac{itS_{[0,n-1]}}{\sigma_n}\mathbb{E}_{\xi_n}(\exp\frac{it\bar{S}_n}{\sigma_n})\right)\to 0. \tag{9}$$

Note now that, by the properties of conditional expectation,

$$\mathbb{E}\left(\exp\frac{itS_{[0,n-1]}}{\sigma_n}\mathbb{E}_{\xi_n}(\exp\frac{it\bar{S}_n}{\sigma_n})\right)=\mathbb{E}\left(\mathbb{E}_{\xi_n}(\exp\frac{itS_{[0,n-1]}}{\sigma_n})\mathbb{E}_{\xi_n}(\exp\frac{it\bar{S}_n}{\sigma_n})\right).$$

By taking into account the reversibility of the process, we obtain

$$\mathbb{E}_{\xi_n}(\exp\frac{itS_{[0,n-1]}}{\sigma_n})=\mathbb{E}_{\xi_n}(\exp\frac{it\bar{S}_n}{\sigma_n}).$$

Therefore, by combining this identity with (9), it follows that

$$\mathbb{E}(\exp\frac{itS_{2n}}{\sigma_n})-\mathbb{E}\left(\mathbb{E}_{\xi_n}(\exp\frac{it\bar{S}_n}{\sigma_n})\right)^2\to 0.$$

So, by using now the stationary, the representation of the variance in (2) and the central limit theorem in (4), the convergence given in relation (8) follows.

Step 2. We show now that from any subsequence (n') we can extract one $(n")\subset(n')$ for which there is a pair of random variables (V,ξ), with ξ distributed as ξ_0 and V is normally distributed with mean 0 and variance c^2 (i.e. $N(0,c^2)$) and such that

$$\mathbb{E}_{\xi}\exp(itV'_{n"})\to\mathbb{E}_{\xi}\exp(itV) \text{ in probability,} \tag{10}$$

where $(V'_{n"},\xi)$ is distributed as $(S_{n"}/\sigma_{n"},\xi_0)$. In addition,

$$\mathbb{E}(\mathbb{E}_{\xi_0}(\exp\frac{itS_{n"}}{\sigma_{n"}}))^2\to\mathbb{E}(\mathbb{E}_{\xi}\exp(itV))^2. \tag{11}$$

Indeed, since $(S_n/\sigma_n,\xi_0)$ is tight, from any subsequence (n') we can extract one $(n")\subset(n')$ such that $(S_{n"}/\sigma_{n"},\xi_0)$ is convergent in distribution to (V,Y) where, by (4), V is

$N(0, c^2)$ and Y is distributed as ξ_0. By Lemma 2, applied to $V_{n''} = S_{n''}/\sigma_{n''}$, there exists two pairs of variables: $(V'_{n''}, \xi)$ distributed as $(S_{n''}/\sigma_{n''}, \xi'_0)$ and (V, ξ), with ξ distributed as ξ_0 and V centered normal with variance c^2, such that convergence in (10) holds. Now by starting from (10) and applying the Lebesgue dominated convergence theorem, (11) follows.

Step 3. In order to finish the proof of the theorem we shall show that the limit in (10) does not depend on the subsequence. As a matter of fact we shall show that

$$\mathbb{E}_\xi \exp(itV) = \exp(-\frac{t^2}{2c^2}) \text{ a.s.} \tag{12}$$

We shall use the fact that V is $N(0, c^2)$ and also, by Step 2 and by Step 1, we know that

$$\mathbb{E}(\mathbb{E}_\xi \exp(itV))^2 = \exp(-\frac{t^2}{c^2}). \tag{13}$$

Note that, in order to establish (12), it is enough to show that, for all integers $m \geq 0$,

$$\mathbb{E}_\xi(V^m) = \mathbb{E}(V^m) \; a.s. \tag{14}$$

With this aim, we redefine V on a larger probability space together with two independent variables uniformly distributed U and $\tilde{U}$ which are also independent on ξ and such that $V = f(\xi, U)$ and $\tilde{V} = f(\xi, \tilde{U})$. Note that V and $\tilde{V}$ are conditionally independent given ξ and (V, ξ) has the same distribution as $(\tilde{V}, \xi)$. In addition, let $N, \tilde{N}$ be i.i.d. random variables $N(0, c^2)$.

Then, by (13),

$$\mathbb{E}(\mathbb{E}_\xi \exp(itV))^2 = \mathbb{E}(\exp(it(V + \tilde{V})))^2$$
$$= \sum_{m=0}^{\infty} \mathbb{E}(V + \tilde{V})^m \frac{(it)^m}{m!} = \sum_{m=0}^{\infty} \mathbb{E}(N + \tilde{N})^m \frac{(it)^m}{m!}.$$

Hence, for all $n = 0, 1, 2, \ldots$

$$\mathbb{E}(V + \tilde{V})^n = \mathbb{E}(N + \tilde{N})^n. \tag{15}$$

Further, we proceed by induction to prove (14). Note that (14) obviously holds for $m = 0$. Assume (14) holds for $m \leq k$. To prove it for $(k + 1)$ we use (15) with $n = 2k + 2$ and develop the binomial

$$\sum_{\ell=0}^{2k+2} C_{2k+2}^{\ell} \mathbb{E} V^{\ell} \tilde{V}^{2k+2-\ell} = \sum_{\ell=0}^{2k+2} C_{2k+2}^{\ell} \mathbb{E} N^{\ell} \mathbb{E} N^{2k+2-\ell}. \tag{16}$$

By the induction hypothesis, conditional independence of V and $\tilde{V}$ and the properties of conditional expectation we obtain, for $0 \le \ell \le k$, and every integer m,

$$\mathbb{E} V^{\ell} \tilde{V}^{m} = \mathbb{E}(\mathbb{E}_{\xi}(V^{\ell} \tilde{V}^{m})) = \mathbb{E}(\mathbb{E}_{\xi} V^{\ell} \mathbb{E}_{\xi} V^{m}) = \mathbb{E}(V^{\ell}) \mathbb{E}(V^{m}) = \mathbb{E}(N^{\ell}) \mathbb{E}(N^{m})$$

and similarly,

$$\mathbb{E} \tilde{V}^{\ell} V^{m} = \mathbb{E}(N^{\ell}) \mathbb{E}(N^{m}).$$

By using both these estimates in (16), we observe that the terms in the sum from $0 \le \ell \le k$ and $k+2 \le \ell \le 2k+2$ all cancel, and it follows that

$$\mathbb{E} V^{k+1} \tilde{V}^{k+1} = (\mathbb{E} N^{k+1})^2 = (\mathbb{E} V^{k+1})^2.$$

Taking into account that

$$\mathbb{E} V^{k+1} \tilde{V}^{k+1} = \mathbb{E}(\mathbb{E}_{\xi} V^{k+1} \mathbb{E}_{\xi} \tilde{V}^{k+1}) = \mathbb{E}(\mathbb{E}_{\xi} V^{k+1})^2,$$

we obtain by the above arguments that

$$\mathbb{E}(\mathbb{E}_{\xi} V^{k+1})^2 = (\mathbb{E} V^{k+1})^2.$$

It follows that

$$\mathbb{E}(\mathbb{E}_{\xi} V^{k+1} - \mathbb{E} V^{k+1})^2 = 0,$$

which completes the proof of (14) and therefore of (12). Combining (12) with (10) we get

$$\mathbb{E}_{\xi} \exp(\frac{itS'_{n"}}{\sigma_{n"}}) \to \exp(-\frac{t^2}{2c^2}) \text{ in probability.}$$

Consequently,

$$\mathbb{E}_{\xi_0} \exp(\frac{itS_{n"}}{\sigma_{n"}}) \to \exp(-\frac{t^2}{2c^2}) \text{ in probability,}$$

completing the proof of the theorem.

Acknowledgement. This material was supported in part by a Charles Phelps Taft Memorial Fund grant at the University of Cincinnati, and the National Science Foundation grant number DMS-1208237. The authors would like to thank the referee for carefully reading the manuscript and suggestions which improved the presentation of the paper.

References

[1] Billingsley, P. (1995). *Probability and measure.* John Wiley & Sons, New York.

[2] Deligiannidis, G., Peligrad M. and Utev, S. (2014). Asymptotic variance of stationary reversible and normal Markov processes. http://arxiv.org/abs/1405.2411

[3] Giraudo, D. and Volný, D. (2014). A strictly stationary β-mixing process satisfying the central limit theorem but not the weak invariance principle. *Stoch. Proc. Appl.* **124** 3769-3781.

[4] Gordin, M. I. and Lifšic, B. (1981). A remark about a Markov process with normal transition operator, *Third Vilnius Conf. Proba. Stat., Akad. Nauk Litovsk*, (in Russian), Vilnius, **1** 147–148.

[5] 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.

[6] Kipnis, C. and Landim, C. (1999). *Scaling Limits of Interacting Particle Systems.* Springer, New York.

[7] Longla, M. Peligrad, C. and Peligrad, M. (2012). On the functional CLT for reversible Markov Chains with nonlinear growth of the variance. *J. Appl. Probab.* **49** 1091-1105.

[8] Skorohod, A.V. (1956). Limit theorems for stochastic processes. *Theory of Probability and Its Applications* **1** 261–290.

[9] Tierney, L. (1994). Markov chains for exploring posterior distribution (with discussion). *Ann. Statist.* **22** 1701-1762.

[10] Zhao, O., Woodroofe, M. and Volný, D. (2010). A central limit theorem for reversible processes with nonlinear growth of variance, *J. Appl. Probab.* **47** 1195-1202.