Elect. Comm. in Probab. **13** (2008), 85–98

ELECTRONIC COMMUNICATIONS in PROBABILITY

A REGENERATION PROOF OF THE CENTRAL LIMIT THEOREM FOR UNIFORMLY ERGODIC MARKOV CHAINS

WITOLD BEDNORZ
Institute of Mathematics, Warsaw University, 00-927 Warsaw, Poland
email: `wbednorz@mimuw.edu.pl`

KRZYSZTOF ŁATUSZYŃSKI
Institute of Econometrics, Warsaw School of Economics, 02-554 Warsaw, Poland
email: `latuch@gmail.com`

RAFAŁ LATAŁA
Institute of Mathematics, Warsaw University, 00-927 Warsaw, Poland
email: `rlatala@mimuw.edu.pl`

Submitted September 4, 2007, *accepted in final form* January 22, 2008

AMS 2000 Subject classification: 60J05, 60F05
Keywords: Markov chains, central limit theorems, regeneration, ergodicity, uniform ergodicity, Harris recurrence

Abstract

Central limit theorems for functionals of general state space Markov chains are of crucial importance in sensible implementation of Markov chain Monte Carlo algorithms as well as of vital theoretical interest. Different approaches to proving this type of results under diverse assumptions led to a large variety of CTL versions. However due to the recent development of the regeneration theory of Markov chains, many classical CLTs can be reproved using this intuitive probabilistic approach, avoiding technicalities of original proofs. In this paper we provide a characterization of CLTs for ergodic Markov chains via regeneration and then use the result to solve the open problem posed in [17]. We then discuss the difference between one-step and multiple-step small set condition.

1 Introduction

Let $(X_n)_{n\geqslant 0}$ be a time homogeneous, ergodic Markov chain on a measurable space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, with transition kernel $\mathbf{P}$ and a unique stationary measure π on $\mathcal{X}$. We remark that the ergodicity means that

$$\lim_{n\to\infty} \|\mathbf{P}^n(x,\cdot) - \pi\|_{tv} = 0, \qquad \text{for all } x \in \mathcal{X}, \tag{1}$$

where $\|\cdot\|_{tv}$ denotes the total variation distance. The process $(X_n)_{n\geqslant 0}$ may start from any initial distribution π_0. Let g be a real valued Borel function on $\mathcal{X}$, square integrable against the stationary measure π. We denote by $\bar{g}$ its centered version, namely $\bar{g} = g - \int g d\pi$ and for simplicity $S_n := \sum_{i=0}^{n-1} \bar{g}(X_i)$. We say that a $\sqrt{n}-$CLT holds for $(X_n)_{n\geqslant 0}$ and g if

$$S_n/\sqrt{n} \xrightarrow{d} N(0, \sigma_g^2), \qquad \text{as} \quad n \to \infty, \tag{2}$$

where $\sigma_g^2 < \infty$. First we aim to provide a general result, namely Theorem 4.1, that gives a necessary and sufficient condition for $\sqrt{n}$-CLTs for ergodic chains (which is a generalization of the well known Theorem 17.3.6 [11]). Assume for a moment that there exists a true atom $\alpha \in \mathcal{B}(\mathcal{X})$, i.e. such a set α that $\pi(\alpha) > 0$ and there exists a probability measure ν on $\mathcal{B}(\mathcal{X})$, such that $P(x, A) = \nu(A)$ for all $x \in \alpha$. Let τ_α be the first hitting time for α. In this simplistic case we can rephrase our Theorem 4.1 as follows:

Theorem 1.1. *Suppose that $(X_n)_{n\geqslant 0}$ is ergodic and possess a true atom α, then the $\sqrt{n}-$CLT holds if and only if*

$$\mathbf{E}_\alpha\left[\left(\sum_{k=1}^{\tau_\alpha} \bar{g}(X_k)\right)^2\right] < \infty. \tag{3}$$

Furthermore we have the following formula for the variance $\sigma_g^2 = \pi(\alpha)\mathbf{E}_\alpha\left[\left(\sum_{k=1}^{\tau_\alpha} \bar{g}(X_k)\right)^2\right]$.

Central limit theorems of this type are crucial for assessing the quality of Markov chain Monte Carlo estimation (see [10] and [5]) and are also of independent theoretical interest. Thus a large body of work on CLTs for functionals of Markov chains exists and a variety of results have been established under different assumptions and with different approaches (see [9] for a review). We discuss briefly the relation between two classical CLT formulations for geometrically ergodic and uniformly ergodic Markov chains. We say that a Markov chain $(X_n)_{n\geqslant 0}$ with transition kernel P and stationary distribution π is

- *geometrically ergodic,* if $\|\mathbf{P}^n(x,\cdot)-\pi(\cdot)\|_{tv} \leqslant M(x)\rho^n$, for some $\rho < 1$ and $M(x) < \infty \quad \pi$-a.e.,
- *uniformly ergodic,* if $\|\mathbf{P}^n(x,\cdot) - \pi(\cdot)\|_{tv} \leqslant M\rho^n$, for some $\rho < 1$ and $M < \infty$.

Recently the following CLT provided by [8] has been reproved in [17] using the intuitive regeneration approach and avoiding technicalities of the original proof (however see Section 6 for a commentary).

Theorem 1.2. *If a Markov chain $(X_n)_{n\geqslant 0}$ with stationary distribution π is geometrically ergodic, then a $\sqrt{n}-$CLT holds for $(X_n)_{n\geqslant 0}$ and g whenever $\pi(|g|^{2+\delta}) < \infty$ for some $\delta > 0$. Moreover $\sigma_g^2 := \int_{\mathcal{X}} \bar{g}^2 d\pi + 2\int_{\mathcal{X}} \sum_{n=1}^{\infty} \bar{g}(X_0)\bar{g}(X_n)d\pi$.*

Remark 1.3. *Note that for reversible chains the condition $\pi(|g|^{2+\delta}) < \infty$ for some $\delta > 0$ in Theorem 1.2 can be weakened to $\pi(g^2) < \infty$ as proved in [16], however this is not possible for the general case, see [2] or [6] for counterexamples.*

Roberts and Rosenthal posed an open problem, whether the following CLT version for uniformly ergodic Markov chains due to [4] can also be reproved using direct regeneration arguments.

Theorem 1.4. *If a Markov chain $(X_n)_{n\geqslant 0}$ with stationary distribution π is uniformly ergodic, then a $\sqrt{n}-$CLT holds for $(X_n)_{n\geqslant 0}$ and g whenever $\pi(g^2) < \infty$. Moreover $\sigma_g^2 := \int_{\mathcal{X}} \bar{g}^2 d\pi + 2\int_{\mathcal{X}} \sum_{n=1}^{\infty} \bar{g}(X_0)\bar{g}(X_n)d\pi$.*

The aim of this paper is to prove Theorem 4.1 and show how to derive from this general framework the regeneration proof of Theorem 1.4. The outline of the paper is as follows. In Section 2 we describe the regeneration construction, then in Section 3 we provide some preliminary results which may also be of independent interest. In Section 4 we detail the proof of Theorem 4.1, and derive Theorem 1.4 as a corollary in Section 5. Section 6 comprises a discussion of some difficulties of the regeneration approach.

2 Small Sets and the Split Chain

We remark that ergodicity as defined by (1) is equivalent to Harris recurrence and aperiodicity (see Proposition 6.3 in [13]). One of the main feature of Harris recurrent chains is that they are $\psi-$irreducible and admit the regeneration construction, discovered independently in [12] and [1], and which is now a well established technique. In particular such chains satisfy

Definition 2.1 (Minorization Condition)**.** *For some $\varepsilon > 0$, some $C \in \mathcal{B}^+(\mathcal{X}) := \{A \in \mathcal{B}(\mathcal{X}) : \psi(A) > 0\}$ and some probability measure ν_m with $\nu_m(C) = 1$ we have for all $x \in C$,*

$$\mathbf{P}^m(x, \cdot) \geqslant \varepsilon\nu_m(\cdot). \tag{4}$$

The minorization condition (4) enables constructing the split chain for $(X_n)_{n\geqslant 0}$ which is the central object of the approach (see Section 17.3 of [11] for a detailed description). The minorization condition allows to write P^m as a mixture of two distributions:

$$\mathbf{P}^m(x, \cdot) = \varepsilon\mathbb{I}_C(x)\nu_m(\cdot) + [1 - \varepsilon\mathbb{I}_C(x)]R(x, \cdot), \tag{5}$$

where $R(x, \cdot) = [1 - \varepsilon\mathbb{I}_C(x)]^{-1}[\mathbf{P}(x, \cdot) - \varepsilon\mathbb{I}_C(x)\nu_m(\cdot)]$. Now let $(X_{nm}, Y_n)_{n\geqslant 0}$ be the split chain of the $m-$skeleton i.e. let the random variable $Y_n \in \{0, 1\}$ be the level of the split $m-$skeleton at time nm. The split chain $(X_{nm}, Y_n)_{n\geqslant 0}$ is a Markov chain that obeys the following transition rule $\check{\mathbf{P}}$.

$$\begin{aligned}
\check{\mathbf{P}}(Y_n = 1, X_{(n+1)m} \in dy|Y_{n-1}, X_{nm} = x) &= \varepsilon\mathbb{I}_C(x)\nu_m(dy) && (6)\\
\check{\mathbf{P}}(Y_n = 0, X_{(n+1)m} \in dy|Y_{n-1}, X_{nm} = x) &= (1 - \varepsilon\mathbb{I}_C(x))R(x, dy), && (7)
\end{aligned}$$

and Y_n can be interpreted as a coin toss indicating whether $X_{(n+1)m}$ given $X_{nm} = x$ should be drawn from $\nu_m(\cdot)$ - with probability $\varepsilon\mathbb{I}_C(x)$ - or from $R(x, \cdot)$ - with probability $1 - \varepsilon\mathbb{I}_C(x)$.
One obtains the split chain $(X_k, Y_n)_{k\geqslant 0, n\geqslant 0}$ of the initial Markov chain $(X_n)_{n\geqslant 0}$ by defining appropriate conditional probabilities. To this end let $X_0^{nm} = \{X_0, \dots, X_{nm-1}\}$ and $Y_0^n = \{Y_0, \dots, Y_{n-1}\}$.

$$\begin{aligned}
&\check{\mathbf{P}}(Y_n = 1, X_{nm+1} \in dx_1, \dots, X_{(n+1)m-1} \in dx_{m-1}, X_{(n+1)m} \in dy|Y_0^n, X_0^{nm}; X_{nm} = x) = \\
&= \frac{\varepsilon\mathbb{I}_C(x)\nu_m(dy)}{\mathbf{P}^m(x, dy)}\mathbf{P}(x, dx_1)\cdots\mathbf{P}(x_{m-1}, dy), && (8)\\
&\check{\mathbf{P}}(Y_n = 0, X_{nm+1} \in dx_1, \dots, X_{(n+1)m-1} \in dx_{m-1}, X_{(n+1)m} \in dy|Y_0^n, X_0^{nm}; X_{nm} = x) = \\
&= \frac{(1 - \varepsilon\mathbb{I}_C(x))R(x, dy)}{\mathbf{P}^m(x, dy)}\mathbf{P}(x, dx_1)\cdots\mathbf{P}(x_{m-1}, dy). && (9)
\end{aligned}$$

Note that the marginal distribution of $(X_k)_{k\geqslant 0}$ in the split chain is that of the underlying Markov chain with transition kernel $\mathbf{P}$.

For a measure λ on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ let λ^* denote the measure on $\mathcal{X}\times\{0,1\}$ (with product $\sigma-$algebra) defined by $\lambda^*(B\times\{1\}) = \varepsilon\lambda(B\cap C)$ and $\lambda^*(B\times\{0\}) = (1-\varepsilon)\lambda(B\cap C) + \lambda(B\cap C^c)$. In the sequel we shall use ν_m^* for which $\nu_m^*(B\times\{1\}) = \varepsilon\nu_m(B)$ and $\nu_m^*(B\times\{0\}) = (1-\varepsilon)\nu_m(B)$ due to the fact that $\nu_m(C) = 1$.

Now integrate (8) over $x_1, \dots, x_{m-1}$ and then over y. This yields

$$\check{\mathbf{P}}(Y_n = 1, X_{(n+1)m} \in dy | Y_0^n, X_0^{nm}; X_{nm} = x) = \varepsilon \mathbb{I}_C(x)\nu_m(dy), \tag{10}$$

and

$$\check{\mathbf{P}}(Y_n = 1 | Y_0^n, X_0^{nm}; X_{nm} = x) = \varepsilon \mathbb{I}_C(x). \tag{11}$$

From the Bayes rule we obtain

$$\check{\mathbf{P}}(X_{(n+1)m} \in dy | Y_0^n, X_0^{nm}; Y_n = 1, X_{nm} = x) = \nu_m(dy), \tag{12}$$

and the crucial observation due to Meyn and Tweedie, emphasized here as Lemma 2.2 follows.

Lemma 2.2. *Conditional on $\{Y_n = 1\}$, the pre$-nm$ process $\{X_k, Y_i : k \leqslant nm, i \leqslant n\}$ and the post$-(n+1)m$ process $\{X_k, Y_i : k \geqslant (n+1)m, i \geqslant n+1\}$ are independent. Moreover, the post$-(n+1)m$ process has the same distribution as $\{X_k, Y_i : k \geqslant 0, i \geqslant 0\}$ with ν_m^* for the initial distribution of (X_0, Y_0).*

Next, let $\sigma_{\check\alpha}(n)$ denote entrance times of the split chain to the set $\check\alpha = C\times\{1\}$, i.e.

$$\sigma_{\check\alpha}(0) = \min\{k \geqslant 0 : Y_k = 1\}, \qquad \sigma_{\check\alpha}(n) = \min\{k > \sigma(n-1) : Y_k = 1\}, \quad n \geqslant 1,$$

whereas hitting times $\tau_{\check\alpha}(n)$ are defined as follows:

$$\tau_{\check\alpha}(1) = \min\{k \geqslant 1 : Y_k = 1\}, \qquad \tau_{\check\alpha}(n) = \min\{k > \tau_{\check\alpha}(n-1) : Y_k = 1\}, \quad n \geqslant 2.$$

We define also

$$s_i = s_i(\bar g) = \sum_{j=m(\sigma_{\check\alpha}(i)+1)}^{m(\sigma_{\check\alpha}(i+1)+1)-1} \bar g(X_j) = \sum_{j=\sigma_{\check\alpha}(i)+1}^{\sigma_{\check\alpha}(i+1)} Z_j(\bar g), \qquad \text{where} \quad Z_j(\bar g) = \sum_{k=0}^{m-1} \bar g(X_{jm+k}).$$

3 Tools and Preliminary Results

In this section we analyze the sequence $s_i(\bar g)$, $i \geqslant 0$. The basic result we often refer to is Theorem 17.3.1 in [11], which states that $(s_i)_{i\geqslant 0}$ is a sequence of 1-dependent, identically distributed r.v.'s with $\check{\mathbf{E}} s_i = 0$. In our approach we use the following decomposition: $s_i = \underline{s}_i + \overline{s}_i$, where

$$\underline{s}_i := \sum_{j=\sigma_{\check\alpha}(i)+1}^{\sigma_{\check\alpha}(i+1)-1} Z_j(\bar g) - \check{\mathbf{E}}_{\pi_0^*}\Bigg[\sum_{j=\sigma_{\check\alpha}(i)+1}^{\sigma_{\check\alpha}(i+1)-1} Z_j(\bar g)\Bigg], \qquad \overline{s}_i := Z_{\sigma_{\check\alpha}(i+1)}(\bar g) - \check{\mathbf{E}}_{\pi_0^*}\Big[Z_{\sigma_{\check\alpha}(i+1)}(\bar g)\Big].$$

A look into the proof of Lemma 3.3 later in this section clarifies that $\underline{s}_i$ and $\overline{s}_i$ are well defined.

Lemma 3.1. *The sequence $(\underline{s}_i)_{i\geqslant 0}$ consists of i.i.d. random variables.*

Proof. First note that $\underline{s}_i$ is a function of $\{X_{(\sigma_{\check{\alpha}}(i)+1)m}, X_{(\sigma_{\check{\alpha}}(i)+1)m+1}, \dots\}$ and that $Y_{\sigma_{\check{\alpha}}(i)} = 1$, hence by Lemma 2.2 $\underline{s}_0, \underline{s}_1, \underline{s}_2, \dots$ are identically distributed. Now focus on $\underline{s}_i, \underline{s}_{i+k}$ and $Y_{\sigma_{\check{\alpha}}(i+k)}$ for some $k \geqslant 1$. Obviously $Y_{\sigma_{\check{\alpha}}(i+k)} = 1$. Moreover $\underline{s}_i$ is a function of the pre$-\sigma_{\check{\alpha}}(i+k)m$ process and $\underline{s}_{i+k}$ is a function of the post$-(\sigma_{\check{\alpha}}(i+k)+1)m$ process. Thus $\underline{s}_i$ and $\underline{s}_{i+k}$ are independent again by Lemma 2.2 and for A_i, A_{i+k}, Borel subsets of R, we have

$$\check{\mathbf{P}}_{\pi_0^*}(\{\underline{s}_i \in A_i\} \cap \{\underline{s}_{i+k} \in A_{i+k}\}) = \check{\mathbf{P}}_{\pi_0^*}(\{\underline{s}_i \in A_i\})\check{\mathbf{P}}(\{\underline{s}_{i+k} \in A_{i+k}\}).$$

Let $0 \leqslant i_1 < i_2 < \cdots < i_l$. By the same pre- and post- process reasoning we obtain for $A_{i_1}, \dots, A_{i_l}$ Borel subsets of R that

$$\check{\mathbf{P}}_{\pi_0^*}(\{\underline{s}_{i_1} \in A_{i_1}\} \cap \cdots \cap \{\underline{s}_{i_l} \in A_{i_l}\}) = \check{\mathbf{P}}_{\pi_0^*}(\{\underline{s}_{i_1} \in A_{i_1}\} \cap \cdots \cap \{\underline{s}_{i_{l-1}} \in A_{i_{l-1}}\}) \cdot \check{\mathbf{P}}_{\pi_0^*}(\{\underline{s}_{i_l} \in A_{i_l}\}),$$

and the proof is complete by induction. □

Now we turn to prove the following lemma, which generalizes the conclusions drawn in [7] for uniformly ergodic Markov chains.

Lemma 3.2. *Let the Markov chain $(X_n)_{n\geqslant 0}$ be recurrent (and $(X_{nm})_{n\geqslant 0}$ be recurrent) and let the minorization condition (4) hold with $\pi(C) > 0$. Then*

$$\mathcal{L}(X_{\tau_{\check{\alpha}}(1)} | \{X_0, Y_0\} \in \check{\alpha}) = \mathcal{L}(X_{\sigma_{\check{\alpha}}(0)} | \{X_0, Y_0\} \sim \nu_m^*) = \pi_C(\cdot), \tag{13}$$

where $\pi_C(\cdot)$ is a probability measure proportional to π truncated to C, that is $\pi_C(B) = \pi(C)^{-1}\pi(B \cap C)$.

Proof. The first equation in (13) is a straightforward consequence of the split chain construction. To prove the second one we use Theorem 10.0.1 of [11] for the split $m-$skeleton with $A = \check{\alpha}$. Thus $\tau_A = \tau_{\check{\alpha}}(1)$ and $\check{\pi} := \pi^*$ is the invariant measure for the split $m-$skeleton. Let $C \supseteq B \in \mathcal{B}(\mathcal{X})$, and compute

$$\begin{aligned} \varepsilon\pi(B) &= \check{\pi}(B \times \{1\}) = \int_{\check{\alpha}} \check{\mathbf{E}}_{x,y}\left[\sum_{n=1}^{\tau_{\check{\alpha}}(1)} \mathbb{I}_{B\times\{1\}}(X_{nm}, Y_n)\right] \check{\pi}(dx, dy) \\ &= \check{\pi}(\check{\alpha})\check{\mathbf{E}}_{\nu_m^*}\left[\sum_{n=0}^{\sigma_{\check{\alpha}}(0)} \mathbb{I}_{B\times\{1\}}(X_{nm}, Y_n)\right] = \check{\pi}(\check{\alpha})\check{\mathbf{E}}_{\nu_m^*}\mathbb{I}_B(X_{\sigma_{\check{\alpha}}(0)}). \end{aligned}$$

This implies proportionality and the proof is complete. □

Lemma 3.3. $\check{\mathbf{E}}_{\pi_0^*}\overline{s}_i^2 \leq \frac{m^2\pi\bar{g}^2}{\varepsilon\pi(c)} < \infty$ *and $(\overline{s}_i)_{i\geqslant 0}$ are 1-dependent identically distributed r.v.'s.*

Proof. Recall that $\overline{s}_i = \sum_{k=0}^{m-1} \bar{g}(X_{\sigma_{\check{\alpha}}(i+1)m+k}) - \check{\mathbf{E}}_{\pi_0^*}\left(\sum_{k=0}^{m-1} \bar{g}(X_{\sigma_{\check{\alpha}}(i+1)m+k})\right)$ and is a function of the random variable

$$\{X_{\sigma_{\check{\alpha}}(i+1)m}, \dots, X_{\sigma_{\check{\alpha}}(i+1)m+m-1}\}. \tag{14}$$

By $\mu_i(\cdot)$ denote the distribution of (14) on $\mathcal{X}^m$. We will show that μ_i does not depend on i. From (8), (11) and the Bayes rule, for $x \in C$, we obtain

$$\begin{aligned} &\check{\mathbf{P}}(X_{nm+1} \in dx_1, \dots, X_{(n+1)m-1} \in dx_{m-1}, X_{(n+1)m} \in dy | Y_0^n, X_0^{nm}; Y_n = 1, X_{nm} = x) = \\ &= \frac{\nu_m(dy)}{P^m(x, dy)} P(x, dx_1) \cdots P(x_{m-1}, dy). \end{aligned} \tag{15}$$

Lemma 3.2 together with (15) yields

$$\begin{aligned}
\check{\mathbf{P}}(X_{nm} \in dx, X_{nm+1} \in dx_1, \ldots, X_{(n+1)m-1} \in dx_{m-1}, X_{(n+1)m} \in dy| \quad (16)\\
|Y_0^n, X_0^{nm}; Y_n = 1; \sigma_{\check{\alpha}}(0) < n) \;=\; \pi_C(dx)\frac{\nu_m(dy)}{P^m(x,dy)}P(x,dx_1)\cdots P(x_{m-1},dy).
\end{aligned}$$

Note that $\frac{\nu_m(dy)}{P^m(x,dy)}$ is just a Radon-Nykodym derivative and thus (16) is a well defined measure on $\mathcal{X}^{m+1}$, say $\mu(\cdot)$. It remains to notice, that $\mu_i(A) = \mu(A \times \mathcal{X})$ for any Borel $A \subset \mathcal{X}^m$. Thus μ_i, $i \geqslant 0$ are identical and hence $\overline{s}_i$, $i \geqslant 0$ have the same distribution. Due to Lemma 2.2 we obtain that $\overline{s}_i$, $i \geqslant 0$ are 1-dependent. To prove $\check{\mathbf{E}}_{\pi_0^*}\overline{s}_i^2 < \infty$, we first note that $\frac{\nu_m(dy)}{P^m(x,dy)} \leqslant 1/\varepsilon$ and also $\pi_C(\cdot) \leqslant \frac{1}{\pi(C)}\pi(\cdot)$. Hence

$$\mu_i(A) = \mu(A \times \mathcal{X}) \leqslant \frac{1}{\varepsilon\pi(C)}\mu_{\text{chain}}(A),$$

where μ_{chain} is defined by $\pi(dx)P(x,dx_1)\ldots P(x_{m-2},dx_{m-1})$. Thus

$$\left|\check{\mathbf{E}}_{\pi_0^*}\left(\sum_{k=0}^{m-1}\bar{g}(X_{\sigma_{\check{\alpha}}(i+1)m+k})\right)\right| \leq \frac{m\pi|\bar{g}|}{\varepsilon\pi(C)} < \infty.$$

Now let $\tilde{s}_i = \sum_{k=0}^{m-1}\bar{g}(X_{\sigma_{\check{\alpha}}(i+1)m+k})$ and proceed

$$\begin{aligned}
\check{\mathbf{E}}_{\pi_0^*}\overline{s}_i^2 &\leqslant \check{\mathbf{E}}_{\pi_0^*}\tilde{s}_i^2 \leqslant \frac{1}{\varepsilon\pi(c)}\mu_{\text{chain}}\tilde{s}_i^2 = \frac{1}{\varepsilon\pi(c)}\mathbf{E}_\pi\left(\sum_{k=0}^{m-1}\bar{g}(X_k)\right)^2\\
&\leqslant \frac{m}{\varepsilon\pi(c)}\mathbf{E}_\pi\left[\sum_{k=0}^{m-1}\bar{g}^2(X_k)\right] \leqslant \frac{m^2\pi\bar{g}^2}{\varepsilon\pi(c)}.
\end{aligned}$$

□

We need a result which gives the connection between stochastic boundedness and the existence of the second moment of $\underline{s}_i$. We state it in a general form.

Theorem 3.4. *Let $(X_n)_{n\geqslant 0}$ be a sequence of independent identically distributed random variables and $S_n = \sum_{k=0}^{n-1}X_k$. Suppose that (τ_n) is a sequence of positive, integer valued r.v.'s such that $\tau_n/n \to a \in (0,\infty)$ in probability when $n \to \infty$ and the sequence $(n^{-1/2}S_{\tau_n})$ is stochastically bounded. Then $\mathbf{E}X_0^2 < \infty$ and $\mathbf{E}X_0 = 0$.*

The proof of Theorem 3.4 is based on the following lemmas.

Lemma 3.5. *Let $\delta \in (0,1)$ and $t_0 := \sup\{t > 0\colon \sup_{0\leqslant k\leqslant n}\mathbf{P}(|S_k| \geqslant t) \geqslant \delta\}$. Then $\mathbf{P}(|S_{10n}| \geqslant 4t_0) \geqslant (1-\delta)(\delta/4)^{20}$ and $\mathbf{P}(\sup_{k\leqslant n}|S_k| \leqslant 3t_0) \geqslant 1-3\delta$.*

Proof. By the definition of t_0 there exists $0 \leqslant n_0 \leqslant n$ such that $\mathbf{P}(|S_{n_0}| \geqslant t_0) \geqslant \delta$. Then either $\mathbf{P}(|S_n| \geqslant t_0/2) \geqslant \delta/2$ or $\mathbf{P}(|S_n| \geqslant t_0/2) < \delta/2$ and consequently

$$\mathbf{P}(|S_{n-n_0}| \geqslant t_0/2) = \mathbf{P}(|S_n - S_{n_0}| \geqslant t_0/2) \geqslant \mathbf{P}(|S_{n_0}| \geqslant t_0) - \mathbf{P}(|S_n| \geqslant t_0/2) \geqslant \delta/2.$$

Thus there exists $n/2 \leqslant n_1 \leqslant n$ such that $\mathbf{P}(|S_{n_1}| \geqslant t_0/2) \geqslant \delta/2$. Let $10n = an_1 + b$ with $0 \leqslant b < n_1$, then $10 \leqslant a \leqslant 20$,

$$\begin{aligned}\mathbf{P}(|S_{an_1}| \geqslant 5t_0) &\geqslant \mathbf{P}(S_{an_1} \geqslant at_0/2) + \mathbf{P}(S_{an_1} \leqslant -at_0/2)\\ &\geqslant (\mathbf{P}(S_{n_1} \geqslant t_0/2))^a + (\mathbf{P}(S_{n_1} \leqslant -t_0/2))^a \geqslant (\delta/4)^a,\end{aligned}$$

hence

$$\mathbf{P}\big(|S_{10n}| \geqslant 4t_0\big) \geqslant \mathbf{P}\big(|S_{an_1}| \geqslant 5t_0\big)\mathbf{P}\big(|S_{10n} - S_{an_1}| \leqslant t_0\big) \geqslant (\delta/4)^a(1-\delta) \geqslant (1-\delta)(\delta/4)^{20}.$$

Finally by the Levy-Octaviani inequality we obtain

$$\mathbf{P}\Big(\sup_{k\leqslant n} |S_k| > 3t_0\Big) \leqslant 3\sup_{k\leqslant n}\mathbf{P}\big(|S_k| > t_0\big) \leqslant 3\delta.$$

□

Lemma 3.6. *Let $c^2 < \mathrm{Var}(X_1)$, then for sufficiently large n,* $\mathbf{P}(|S_n| \geqslant c\sqrt{n}/4) \geqslant 1/16$.

Proof. Let (X_i') be an independent copy of (X_i) and $S_k' = \sum_{i=1}^n X_i'$. Moreover let (ε_i) be a sequence of independent symmetric ± 1 r.v.'s, independent of (X_i) and (X_i'). For any reals (a_i) we get by the Paley-Zygmund inequality,

$$\begin{aligned}\mathbf{P}\Big(|\sum_{i=1}^n a_i\varepsilon_i| \geqslant \frac{1}{2}\Big(\sum_i a_i^2\Big)^{1/2}\Big) &= \mathbf{P}\Big(\Big|\sum_{i=1}^n a_i\varepsilon_i\Big|^2 \geqslant \frac{1}{4}\mathbf{E}\Big|\sum_{i=1}^n a_i\varepsilon_i\Big|^2\Big)\\ &\geqslant \Big(1-\frac{1}{4}\Big)^2\frac{\big(\mathbf{E}|\sum_{i=1}^n a_i\varepsilon_i|^2\big)^2}{\mathbf{E}|\sum_{i=1}^n a_i\varepsilon_i|^4} \geqslant \frac{3}{16}.\end{aligned}$$

Hence

$$\mathbf{P}\Big(|S_n - S_n'| \geqslant \frac{c}{2}\sqrt{n}\Big) = \mathbf{P}\Big(|\sum_{i=1}^n \varepsilon_i(X_i - X_i')| \geqslant \frac{c}{2}\sqrt{n}\Big) \geqslant \frac{3}{16}\mathbf{P}\Big(\sum_{i=1}^n (X_i - X_i')^2 \geqslant c^2 n\Big) \geqslant \frac{1}{8}$$

for sufficiently large n by the Weak LLN. Thus

$$\frac{1}{8} \leqslant \mathbf{P}\big(|S_n - S_n'| \geqslant \frac{c}{2}\sqrt{n}\big) \leqslant \mathbf{P}\big(|S_n| \geqslant \frac{c}{4}\sqrt{n}\big) + \mathbf{P}\big(|S_n'| \geqslant \frac{c}{4}\sqrt{n}\big) \leqslant 2\mathbf{P}\big(|S_n| \geqslant \frac{c}{4}\sqrt{n}\big).$$

□

Corollary 3.7. *Let $c^2 < \mathrm{Var}(X_1)$, then for sufficiently large n,* $\mathbf{P}(\inf_{10n\leqslant k\leqslant 11n} |S_k| \geqslant \frac{1}{4}c\sqrt{n}) \geqslant 2^{-121}$.

Proof. Let t_0 be as in Lemma 3.5 for $\delta = 1/16$, then

$$\begin{aligned}\mathbf{P}\Big(\inf_{10n\leqslant k\leqslant 11n} |S_k| \geqslant t_0\Big) &\geqslant \mathbf{P}\Big(|S_{10n}| \geqslant 4t_0, \sup_{10n\leqslant k\leqslant 11n} |S_k - S_{10n}| \leqslant 3t_0\Big)\\ &= \mathbf{P}\big(|S_{10n}| \geqslant 4t_0\big)\mathbf{P}\Big(\sup_{k\leqslant n}|S_k| \leqslant 3t_0\Big) \geqslant 2^{-121}.\end{aligned}$$

Hence by Lemma 3.5 we obtain $t_0 \geqslant c\sqrt{n}/4$ for large n. □

Proof of Theorem 3.4. By Corollary 3.7 for any $c^2 < \mathrm{Var}(X)$ we have,

$$\begin{aligned}\mathbf{P}\Big(|S_{\tau_n}| \geqslant \frac{c}{20}\sqrt{an}\Big) &\geqslant \mathbf{P}\Big(\big|\frac{\tau_n}{n} - a\big| \leqslant \frac{a}{21}, \inf_{\frac{20}{21}an \leqslant k \leqslant \frac{22}{21}an} |S_k| \geqslant \frac{c}{20}\sqrt{an}\Big) \geqslant \\ &\geqslant \mathbf{P}\Big(\inf_{\frac{20}{21}an \leqslant k \leqslant \frac{22}{21}an} |S_k| \geqslant \frac{c}{4}\sqrt{\frac{2an}{21}}\Big) - \mathbf{P}\Big(\big|\frac{\tau_n}{n} - a\big| > \frac{a}{21}\Big) \\ &\geqslant 2^{-121} - \mathbf{P}\Big(\big|\frac{\tau_n}{n} - a\big| > \frac{a}{21}\Big) \geqslant 2^{-122}\end{aligned}$$

for sufficiently large n. Since $(n^{-1/2}S_{\tau_n})$ is stochastically bounded, we immediately obtain $\mathrm{Var}(X_1) < \infty$. If $\mathbf{E}X_1 \neq 0$ then

$$\big|\frac{1}{\sqrt{n}} S_{\tau_n}\big| = \big|\frac{S_{\tau_n}}{\tau_n}\big|\big|\frac{\tau_n}{n}\big|\sqrt{n} \to \infty \qquad \text{in probability when} \quad n \to \infty.$$

$\square$

4 A Characterization of $\sqrt{n}$-CLTs

In this section we provide a generalization of Theorem 17.3.6 of [11]. We obtain an if and only if condition for the $\sqrt{n}$-CLT in terms of finiteness of the second moment of a centered excursion from $\check{\alpha}$.

Theorem 4.1. *Suppose that $(X_n)_{n\geqslant 0}$ is ergodic and $\pi(g^2) < \infty$. Let ν_m be the measure satisfying (4), then the $\sqrt{n}-$CLT holds if and only if*

$$\check{\mathbf{E}}_{\nu_m^*}\Big[\Big(\sum_{n=0}^{\sigma_{\check{\alpha}}(0)} Z_n(\bar{g})\Big)^2\Big] < \infty. \tag{17}$$

Furthermore we have the following formula for variance

$$\sigma_g^2 = \frac{\varepsilon\pi(C)}{m}\Big\{\check{\mathbf{E}}_{\nu_m^*}\Big[\Big(\sum_{n=0}^{\sigma_{\check{\alpha}}(0)} Z_n(\bar{g})\Big)^2\Big] + 2\check{\mathbf{E}}_{\nu_m^*}\Big[\Big(\sum_{n=0}^{\sigma_{\check{\alpha}}(0)} Z_n(\bar{g})\Big)\Big(\sum_{n=\sigma_{\check{\alpha}}(0)+1}^{\sigma_{\check{\alpha}}(1)} Z_n(\bar{g})\Big)\Big]\Big\}.$$

Proof. For $n \geqslant 0$ define

$$l_n := \max\{k \geqslant 1 : \; m(\sigma_{\check{\alpha}}(k) + 1) \leqslant n\}$$

and for completeness $l_n := 0$ if $m(\sigma_{\check{\alpha}}(0) + 1) \geqslant n$. First we are going to show that

$$\Big|\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} \bar{g}(X_j) - \frac{1}{\sqrt{n}}\sum_{j=0}^{l_n-1} s_j\Big| \to 0 \qquad \text{in probability.} \tag{18}$$

Thus we have to verify that the initial and final terms of the sum do not matter. First observe that by the Harris recurrence property of the chain $\sigma_{\check{\alpha}}(0) < \infty$, $\check{\mathbf{P}}_{\pi_0^*}$-a.s. and hence $\lim_{n\to\infty} \check{\mathbf{P}}_{\pi_0^*}(m\sigma_{\check{\alpha}}(0) \geqslant n) = 0$ and $\check{\mathbf{P}}_{\pi_0^*}(\sigma_{\check{\alpha}}(0) < \infty) = 1$. This yields

$$\Big|\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} \bar{g}(X_j) - \frac{1}{\sqrt{n}}\sum_{j=m(\sigma_{\check{\alpha}}(0)+1)}^{n-1} \bar{g}(X_j)\Big| \to 0, \qquad \check{\mathbf{P}} - \text{a.s.} \tag{19}$$

The second point is to provide a similar argument for the tail terms and to show that

$$\Big|\frac{1}{\sqrt{n}}\sum_{j=m(\sigma_{\check{\alpha}}(0)+1)}^{n-1}\bar{g}(X_j)-\frac{1}{\sqrt{n}}\sum_{j=m(\sigma_{\check{\alpha}}(0)+1)}^{m\sigma_{\check{\alpha}}(l_n)+m-1}\bar{g}(X_j)\Big|\to 0,\qquad \text{in probability.}\tag{20}$$

For $\varepsilon>0$ we have

$$\begin{aligned}\check{\mathbf{P}}_{\pi_0^*}\Big(\Big|\frac{1}{\sqrt{n}}\sum_{j=m(\sigma_{\check{\alpha}}(l_n)+1)}^{n-1}\bar{g}(X_j)\Big|>\varepsilon\Big) &\leqslant \check{\mathbf{P}}_{\pi_0^*}\Big(\frac{1}{\sqrt{n}}\sum_{j=\sigma_{\check{\alpha}}(l_n)+1}^{\sigma_{\check{\alpha}}(l_n+1)}Z_j(|\bar{g}|)>\varepsilon\Big)\\ &\leqslant \sum_{k=0}^{\infty}\check{\mathbf{P}}_{\check{\alpha}}\Big(\frac{1}{\sqrt{n}}\sum_{j=1}^{\tau_{\check{\alpha}}(1)}Z_j(|\bar{g}|)>\varepsilon,\ \tau_{\check{\alpha}}(1)\geqslant k\Big).\end{aligned}$$

Now since $\sum_{k=0}^{\infty}\check{\mathbf{P}}_{\check{\alpha}}(\tau_{\check{\alpha}}(1)\geqslant k)\leqslant\check{\mathbf{E}}_{\check{\alpha}}\tau_{\check{\alpha}}(1)<\infty$, where we use that $\check{\alpha}$ is an atom for the split chain, we deduce form the Lebesgue majorized convergence theorem that (20) holds. Obviously (19) and (20) yield (18).

We turn to prove that the condition (17) is sufficient for the CLT to hold. We will show that random numbers l_n can be replaced by their non-random equivalents. Namely we apply the LLN (Theorem 17.3.2 in [11])) to ensure that

$$\lim_{n\to\infty}\frac{l_n}{n}=\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{[n/m]-1}\mathbb{I}_{\{(X_{mk},Y_k)\in\check{\alpha}\}}=\frac{\check{\pi}(\check{\alpha})}{m},\qquad \check{\mathbf{P}}_{\pi_0^*}-\text{a.s.}\tag{21}$$

Let

$$n^*:=\lfloor\check{\pi}(\check{\alpha})nm^{-1}\rfloor,\qquad \underline{n}:=\lceil(1-\varepsilon)\check{\pi}(\check{\alpha})nm^{-1}\rceil,\qquad \overline{n}:=\lfloor(1+\varepsilon)\check{\pi}(\check{\alpha})nm^{-1}\rfloor.$$

Due to the LLN we know that for any $\varepsilon>0$, there exists n_0 such that for all $n\geqslant n_0$ we have $\check{\mathbf{P}}_{\pi_0^*}(\underline{n}\leqslant l_n\leqslant\overline{n})\geqslant 1-\varepsilon$. Consequently

$$\begin{aligned}\check{\mathbf{P}}_{\pi_0^*}\Big(\Big|\sum_{j=0}^{l_n-1}s_j-\sum_{j=0}^{n^*}s_j\Big|>\sqrt{n}\beta\Big) &\leqslant \varepsilon+\check{\mathbf{P}}_{\pi_0^*}\Big(\max_{\underline{n}\leqslant l\leqslant n^*}\Big|\sum_{j=l}^{n^*}s_j\Big|>\beta\sqrt{n}\Big)+\\ &\quad+\check{\mathbf{P}}_{\pi_0^*}\Big(\max_{n^*+1\leqslant l\leqslant\overline{n}}\Big|\sum_{j=n^*+1}^{l}s_j\Big|>\beta\sqrt{n}\Big).\end{aligned}\tag{22}$$

Since $(s_j)_{j\geqslant 0}$ are 1-dependent, $M_k:=\sum_{j=0}^{k}s_j$ is not necessarily a martingale. Thus to apply the classical Kolmogorov inequality we define $M_k^0=\sum_{j=0}^{\infty}s_{2j}\mathbb{I}_{\{2j\leq k\}}$ and $M_k^1=\sum_{j=0}^{\infty}s_{1+2j}\mathbb{I}_{\{1+2j\leq k\}}$, which are clearly square-integrable martingales (due to (17)). Hence

$$\begin{aligned}\check{\mathbf{P}}_{\pi_0^*}\big(\max_{\underline{n}\leqslant l\leqslant n^*}|M_{n^*}-M_l|>\beta\sqrt{n}\big) &\leqslant \check{\mathbf{P}}_{\pi_0^*}\Big(\max_{\underline{n}\leqslant l\leqslant n^*}|M_{n^*}^0-M_l^0|>\frac{\beta\sqrt{n}}{2}\Big)+\\ &\quad+\check{\mathbf{P}}_{\pi_0^*}\Big(\max_{\underline{n}\leqslant l\leqslant n^*}|M_{n^*}^1-M_l^1|>\frac{\beta\sqrt{n}}{2}\Big)\\ &\leqslant \frac{4}{n\beta^2}\sum_{k=0}^{1}\big(\check{\mathbf{E}}_{\pi_0^*}|M_{n^*}^k-M_{\underline{n}}^k|^2\big)\leqslant C\varepsilon\beta^{-2}\check{\mathbf{E}}_{\nu_m^*}(s_0^2),\end{aligned}\tag{23}$$

where C is a universal constant. In the same way we show that $\check{\mathbf{P}}(\max_{n^*+1\leqslant l\leqslant \overline{n}} |M_l - M_{n^*+1}| > \beta\sqrt{n}) \leqslant C\varepsilon\beta^{-2}\check{\mathbf{E}}_{\nu_m^*}(s_0^2)$, consequently, since ε is arbitrary, we obtain

$$\Big|\frac{1}{\sqrt{n}}\sum_{j=0}^{l_n-1} s_j - \frac{1}{\sqrt{n}}\sum_{j=0}^{n^*} s_j\Big| \to 0, \qquad \text{in probability.} \tag{24}$$

The last step is to provide an argument for the CLT for 1-dependent, identically distributed random variables. Namely, we have to prove that

$$\frac{1}{\sqrt{n}}\sum_{j=0}^{n} s_j \xrightarrow{d} \mathcal{N}(0,\bar{\sigma}^2), \ \text{ as } \ n\to\infty, \quad \text{where } \ \bar{\sigma}^2 := \check{\mathbf{E}}_{\nu_m^*}(s_0(\bar{g}))^2 + 2\check{\mathbf{E}}_{\nu_m^*}(s_0(\bar{g})s_1(\bar{g})). \tag{25}$$

Observe that (19), (20), (24) and (25) imply Theorem 4.1. We fix $k \geqslant 2$ and define $\xi_j := s_{kj+1}(\bar{g}) + ... + s_{kj+k-1}(\bar{g})$, consequently ξ_j are i.i.d. random variables and

$$\frac{1}{\sqrt{n}}\sum_{j=0}^{n} s_j = \frac{1}{\sqrt{n}}\sum_{j=0}^{\lfloor n/k\rfloor -1}\xi_j + \frac{1}{\sqrt{n}}\sum_{j=0}^{\lfloor n/k\rfloor} s_{kj}(\bar{g}) + \frac{1}{\sqrt{n}}\sum_{j=k[n/k]+1}^{n} s_j. \tag{26}$$

Obviously the last term converges to 0 in probability. Denoting

$$\sigma_k^2 := \check{\mathbf{E}}_{\pi_0^*}(\xi_j)^2 = (k-1)\check{\mathbf{E}}_{\nu_m^*}(s_0(\bar{g}))^2 + 2(k-2)\check{\mathbf{E}}_{\nu_m^*}(s_0(\bar{g})s_1(\bar{g})), \quad \text{and} \quad \sigma_s^2 := \check{\mathbf{E}}_{\nu_m^*}(s_0(\bar{g}))^2.$$

we use the classical CLT for i.i.d. random variables to see that

$$\frac{1}{\sqrt{n}}\sum_{j=0}^{\lfloor n/k\rfloor -1}\xi_j \xrightarrow{d} \mathcal{N}(0,k^{-1}\sigma_k^2), \qquad \text{and} \qquad \frac{1}{\sqrt{n}}\sum_{j=0}^{\lfloor n/k\rfloor} s_{kj}(\bar{g}) \xrightarrow{d} \mathcal{N}(0,k^{-1}\sigma_s^2). \tag{27}$$

Moreover

$$\lim_{n\to\infty}\Big[\frac{1}{\sqrt{n}}\sum_{j=0}^{\lfloor n/k\rfloor -1}\xi_j + \frac{1}{\sqrt{n}}\sum_{j=0}^{\lfloor n/k\rfloor} s_{kj}(\bar{g})\Big] \tag{28}$$

converges to $\mathcal{N}(0,\sigma_g^2)$, with $k\to\infty$. Since the weak convergence is metrizable we deduce from (26), (27) and (28) that (25) holds.

The remaining part is to prove that (17) is also necessary for the CLT to hold. Note that if $\sum_{k=0}^{n}\bar{g}(X_k)/\sqrt{n}$ verifies the CLT then $\sum_{j=0}^{l_n-1}s_j$ is stochastically bounded by (18). We use the decomposition $s_i = \underline{s}_i + \overline{s}_i$, $i\geqslant 0$ introduced in Section 3. By Lemma 3.3 we know that $\overline{s}_j$ is a sequence of 1-dependent random variables with the same distribution and finite second moment. Thus from the first part of the proof we deduce that $\sum_{j=0}^{l_n-1}\overline{s}_j/\sqrt{n}$ verifies a CLT and thus is stochastically bounded. Consequently the remaining sequence $\sum_{j=0}^{l_n-1}\underline{s}_j/\sqrt{n}$ also must be stochastically bounded. Lemma 3.1 states that $(\underline{s}_j)_{j\geqslant 0}$ is a sequence of i.i.d. r.v.'s, hence $\check{\mathbf{E}}[\underline{s}_j^2] < \infty$ by Theorem 3.4. Also $l_n/n \to \check{\pi}(\check{\alpha})m^{-1}$ by (21). Applying the inequality $(a+b)^2 \leqslant 2(a^2+b^2)$ we obtain

$$\check{\mathbf{E}}_{\pi_0^*}[s_j]^2 \leqslant 2(\check{\mathbf{E}}_{\pi_0^*}[\underline{s}_j^2] + \check{\mathbf{E}}_{\pi_0^*}[\overline{s}_j^2]) < \infty$$

which completes the proof. $\square$

Remark 4.2. *Note that in the case of $m=1$ we have $\bar{s}_i \equiv 0$ and for Theorem 4.1 to hold, it is enough to assume $\pi|g| < \infty$ instead of $\pi(g^2) < \infty$. In the case of $m>1$, assuming only $\pi|g| < \infty$ and (17) implies the $\sqrt{n}$-CLT, but the proof of the converse statement fails, and in fact the converse statement does not hold (one can easily provide an appropriate counterexample).*

5 Uniform Ergodicity

In view of Theorem 4.1 providing a regeneration proof of Theorem 1.4 amounts to establishing conditions (17) and checking the formula for the asymptotic variance. To this end we need some additional facts about small sets for uniformly ergodic Markov chains.

Theorem 5.1. *If $(X_n)_{n\geqslant 0}$, a Markov chain on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ with stationary distribution π is uniformly ergodic, then $\mathcal{X}$ is ν_m-small for some ν_m.*

Hence for uniformly ergodic chains (4) holds for all $x \in \mathcal{X}$. Theorem 5.1 is well known in literature, in particular it results from Theorems 5.2.1 and 5.2.4 in [11] with their $\psi = \pi$.

Theorem 5.1 implies that for uniformly ergodic Markov chains (5) can be rewritten as

$$\mathbf{P}^m(x,\cdot) = \varepsilon\nu_m(\cdot) + (1-\varepsilon)R(x,\cdot). \tag{29}$$

The following mixture representation of π will turn out very useful.

Lemma 5.2. *If $(X_n)_{n\geqslant 0}$ is an ergodic Markov chain with transition kernel P and (29) holds, then*

$$\pi = \varepsilon\mu := \varepsilon\sum_{n=0}^{\infty}\nu_m(1-\varepsilon)^n R^n. \tag{30}$$

Remark 5.3. *This can be easily extended to the more general setting than this of uniformly ergodic chains, namely let $\mathbf{P}^m(x,\cdot) = s(x)\nu_m(\cdot) + (1-s(x))R(x,\cdot)$, $s : \mathcal{X} \to [0,1]$, $\pi s > 0$. In this case $\pi = \pi s\sum_{n=0}^{\infty}\nu_m R_{\#}^n$, where $R_{\#}(x,\cdot) = (1-s(x))R(x,\cdot)$. Related decompositions under various assumptions can be found e.g. in [14], [7] and [3] and are closely related to perfect sampling algorithms, such as coupling form the past (CFTP) introduced in [15].*

Proof. First check that the measure in question is a probability measure.

$$\Big(\varepsilon\sum_{n=0}^{\infty}\nu_m(1-\varepsilon)^n R^n\Big)(\mathcal{X}) = \varepsilon\sum_{n=0}^{\infty}(1-\varepsilon)^n\big(\nu_m R^n\big)(\mathcal{X}) = 1.$$

It is also invariant for $\mathbf{P}^m$.

$$\begin{aligned}\Big(\sum_{n=0}^{\infty}\nu_m(1-\varepsilon)^n R^n\Big)\mathbf{P}^m &= \Big(\sum_{n=0}^{\infty}\nu_m(1-\varepsilon)^n R^n\Big)(\varepsilon\nu_m + (1-\varepsilon)R)\\ &= \varepsilon\mu\nu_m + \sum_{n=1}^{\infty}\nu_m(1-\varepsilon)^n R^n = \sum_{n=0}^{\infty}\nu_m(1-\varepsilon)^n R^n.\end{aligned}$$

Hence by ergodicity $\varepsilon\mu = \varepsilon\mu\mathbf{P}^{nm} \to \pi, \quad \text{as} \quad n\to\infty$. This completes the proof. □

Corollary 5.4. *The decomposition in Lemma 5.2 implies that*

$$(i) \qquad \check{\mathbf{E}}_{\nu_m^*}\Big(\sum_{n=0}^{\sigma(0)}\mathbb{I}_{\{X_{nm}\in A\}}\Big) = \check{\mathbf{E}}_{\nu_m^*}\Big(\sum_{n=0}^{\infty}\mathbb{I}_{\{X_{nm}\in A\}}\mathbb{I}_{\{Y_0=0,\ldots,Y_{n-1}=0\}}\Big) = \varepsilon^{-1}\pi(A),$$

$$\begin{aligned}(ii) \qquad &\check{\mathbf{E}}_{\nu_m^*}\Big(\sum_{n=0}^{\infty}f(X_{nm},X_{nm+1},\ldots;Y_n,Y_{n+1},\ldots)\mathbb{I}_{\{Y_0=0,\ldots,Y_{n-1}=0\}}\Big) =\\ &\qquad = \varepsilon^{-1}\check{\mathbf{E}}_{\pi^*}f(X_0,X_1,\ldots;Y_0,Y_1,\ldots).\end{aligned}$$

Proof. (i) is a direct consequence of (30). To see (ii) note that Y_n is a coin toss independent of $\{Y_0, \ldots, Y_{n-1}\}$ and X_{nm}, this allows for π^* instead of π on the RHS of (ii). Moreover the evolution of $\{X_{nm+1}, X_{nm+2}, \ldots; Y_{n+1}, Y_{n+2}, \ldots\}$ depends only (and explicitly by (8) and (9)) on X_{nm} and Y_n. Now use (i). $\square$

Our object of interest is

$$\begin{aligned} I &= \check{\mathbf{E}}_{\nu_m^*}\Big[\Big(\sum_{n=0}^{\sigma(0)} Z_n(\bar{g})\Big)^2\Big] = \check{\mathbf{E}}_{\nu_m^*}\Big[\Big(\sum_{n=0}^{\infty} Z_n(\bar{g})\mathbb{I}_{\{\sigma_{\check{\alpha}}(0)\geqslant n\}}\Big)^2\Big] \\ &= \check{\mathbf{E}}_{\nu_m^*}\Big[\sum_{n=0}^{\infty} Z_n(\bar{g})^2\mathbb{I}_{\{Y_0=0,\ldots,Y_{n-1}=0\}}\Big] + 2\check{\mathbf{E}}_{\nu_m^*}\Big[\sum_{n=0}^{\infty}\sum_{k=n+1}^{\infty} Z_n(\bar{g})\mathbb{I}_{\{\sigma(0)\geqslant n\}}Z_k(\bar{g})\mathbb{I}_{\{\sigma_{\check{\alpha}}(0)\geqslant k\}}\Big] \\ &= A + B \end{aligned} \tag{31}$$

Next we use Corollary 5.4 and then the inequality $2ab \leqslant a^2 + b^2$ to bound the term A in (31).

$$A = \varepsilon^{-1}\check{\mathbf{E}}_{\pi^*} Z_0(\bar{g})^2 = \varepsilon^{-1}\mathbf{E}_{\pi}\Big(\sum_{k=0}^{m-1}\bar{g}(X_k)\Big)^2 \leqslant \varepsilon^{-1} m \mathbf{E}_{\pi}\Big[\sum_{k=0}^{m-1}\bar{g}^2(X_k)\Big] \leqslant \varepsilon^{-1}m^2\pi\bar{g}^2 < \infty.$$

We proceed similarly with the term B

$$\begin{aligned} |B| &\leqslant 2\check{\mathbf{E}}_{\nu_m^*}\Big[\sum_{n=0}^{\infty}|Z_n(\bar{g})|\mathbb{I}_{\{\sigma_{\check{\alpha}}(0)\geqslant n\}}\sum_{k=1}^{\infty}|Z_{n+k}(\bar{g})|\mathbb{I}_{\{\sigma_{\check{\alpha}}(0)\geqslant n+k\}}\Big] \\ &= 2\varepsilon^{-1}\check{\mathbf{E}}_{\pi^*}\Big[|Z_0(\bar{g})|\sum_{k=1}^{\infty}|Z_k(\bar{g})|\mathbb{I}_{\{\sigma_{\check{\alpha}}(0)\geqslant k\}}\Big]. \end{aligned}$$

By Cauchy-Schwarz,

$$\begin{aligned} \check{\mathbf{E}}_{\pi^*}\big[\mathbb{I}_{\{\sigma_{\check{\alpha}}(0)\geqslant k\}}|Z_0(\bar{g})||Z_k(\bar{g})|\big] &\leqslant \sqrt{\check{\mathbf{E}}_{\pi^*}\big[\mathbb{I}_{\{\sigma_{\check{\alpha}}(0)\geqslant k\}}Z_0(\bar{g})^2\big]}\sqrt{\check{\mathbf{E}}_{\pi^*}Z_k(\bar{g})^2} \\ &= \sqrt{\check{\mathbf{E}}_{\pi^*}\big[\mathbb{I}_{\{Y_0=0\}}\mathbb{I}_{\{Y_1=0,\ldots,Y_{k-1}=0\}}Z_0(\bar{g})^2\big]}\sqrt{\check{\mathbf{E}}_{\pi^*}Z_0(\bar{g})^2}. \end{aligned}$$

Observe that $\{Y_1, \ldots, Y_{k-1}\}$ and $\{X_0, \ldots, X_{m-1}\}$ are independent. We drop $\mathbb{I}_{\{Y_0=0\}}$ to obtain

$$\check{\mathbf{E}}_{\pi^*}\big[\mathbb{I}_{\{\sigma_{\check{\alpha}}(0)\geqslant k\}}|Z_0(\bar{g})||Z_k(\bar{g})|\big] \leqslant (1-\varepsilon)^{\frac{k-1}{2}}\check{\mathbf{E}}_{\pi^*}Z_0(\bar{g})^2 \leqslant (1-\varepsilon)^{\frac{k-1}{2}}m^2\pi g^2.$$

Hence $|B| < \infty$, and the proof of (17) is complete. To get the variance formula note that the convergence we have established implies

$$I = \varepsilon^{-1}\check{\mathbf{E}}_{\pi^*}\Big[Z_0(\bar{g})\Big]^2 + 2\varepsilon^{-1}\check{\mathbf{E}}_{\pi^*}\Big[Z_0(\bar{g})\sum_{k=1}^{\infty}Z_k(\bar{g})\mathbb{I}_{\{\sigma_{\check{\alpha}}(0)\geqslant k\}}\Big].$$

Similarly we obtain

$$J := 2\check{\mathbf{E}}_{\nu_m^*}\Big[\big(\sum_{n=0}^{\sigma_{\check{\alpha}}(0)}Z_n(\bar{g})\big)\big(\sum_{n=\sigma_{\check{\alpha}}(0)+1}^{\sigma_{\check{\alpha}}(1)}Z_n(\bar{g})\big)\Big] = 2\varepsilon^{-1}\check{\mathbf{E}}_{\pi^*}\Big[Z_0(\bar{g})\sum_{k=\sigma_{\check{\alpha}}(0)+1}^{\infty}Z_k(\bar{g})\mathbb{I}_{\{\sigma_{\check{\alpha}}(1)\geqslant k\}}\Big].$$

Since $\pi(C) = 1$, we have $\sigma_g^2 = \varepsilon m^{-1}(I+J)$. Next we use Lemma 2.2 and $\check{\mathbf{E}}_{\pi^*} Z_0(\bar{g}) = 0$ to drop indicators and since for $f : \mathcal{X} \to R$, also $\check{\mathbf{E}}_{\pi^*} f = \mathbf{E}_\pi f$, we have

$$\varepsilon(I+J) = \check{\mathbf{E}}_{\pi^*}\Big[Z_0(\bar{g})\Big(Z_0(\bar{g}) + 2\sum_{k=1}^{\infty} Z_k(\bar{g})\Big)\Big] = \mathbf{E}_\pi\Big[Z_0(\bar{g})\Big(Z_0(\bar{g}) + 2\sum_{k=1}^{\infty} Z_k(\bar{g})\Big)\Big].$$

Now, since all the integrals are taken with respect to the stationary measure, we can for a moment assume that the chain runs in stationarity from $-\infty$ rather than starts at time 0 with $X_0 \sim \pi$. Thus

$$\begin{aligned}
\sigma_g^2 &= m^{-1}\mathbf{E}_\pi\Big[Z_0(\bar{g})\Big(\sum_{k=-\infty}^{\infty} Z_k(\bar{g})\Big)\Big] = m^{-1}\mathbf{E}_\pi\Big[\sum_{l=0}^{m-1}\bar{g}(X_l)\Big(\sum_{k=-\infty}^{\infty}\bar{g}(X_k)\Big)\Big] \\
&= \mathbf{E}_\pi\Big[\bar{g}(X_0)\sum_{k=-\infty}^{\infty}\bar{g}(X_k)\Big] = \int_{\mathcal{X}}\bar{g}^2 d\pi + 2\int_{\mathcal{X}}\sum_{n=1}^{\infty}\bar{g}(X_0)\bar{g}(X_n)d\pi.
\end{aligned}$$

6 The difference between $m = 1$ and $m \neq 1$

Assume the small set condition (4) holds and consider the split chain defined by (8) and (9). The following tours

$$\big\{\{X_{(\sigma(n)+1)m}, X_{(\sigma(n)+1)m+1}, \ldots, X_{(\sigma(n+1)+1)m-1}\}, n = 0, 1, \ldots\big\}$$

that start whenever $X_k \sim \nu_m$ are of crucial importance to the regeneration theory and are eagerly analyzed by researchers. In virtually every paper on the subject there is a claim these objects are independent identically distributed random variables. This claim is usually considered obvious and no proof is provided. However this is not true if $m > 1$.

In fact formulas (8) and (9) should be convincing enough, as $X_{mn+1}, \ldots, X_{(n+1)m}$ given $Y_n = 1$ and $X_{nm} = x$ are linked in a way described by $\mathbf{P}(x, dx_1)\cdots\mathbf{P}(x_{m-1}, dy)$. In particular consider a Markov chain on $\mathcal{X} = \{a, b, c, d, e\}$ with transition probabilities

$$\mathbf{P}(a,b) = \mathbf{P}(a,c) = \mathbf{P}(b,b) = \mathbf{P}(b,d) = \mathbf{P}(c,c) = \mathbf{P}(c,e) = 1/2, \;\; \mathbf{P}(d,a) = \mathbf{P}(e,a) = 1.$$

Let $\nu_4(d) = \nu_4(e) = 1/2$ and $\varepsilon = 1/8$. Clearly $\mathbf{P}^4(x,\cdot) \geqslant \varepsilon\nu_4(\cdot)$ for every $x \in \mathcal{X}$, hence we established (4) with $C = \mathcal{X}$. Note that for this simplistic example each tour can start with d or e. However if it starts with d or e the previous tour must have ended with b or c respectively. This makes them dependent. Similar examples with general state space $\mathcal{X}$ and $C \neq \mathcal{X}$ can be easily provided. Hence Theorem 4.1 is critical to providing regeneration proofs of CLTs and standard arguments that involve i.i.d. random variables are not valid.

References

[1] Athreya K. B. and Ney P. (1978). A new approach to the limit theory of recurrent Markov chains. *Trans. Amer. Math. Soc.* **245** 493–501. MR0511425

[2] Bradley, R. C. (1983) Information regularity and the central limit question. *Rocky Mountain Journal of Mathemathics* **13** 77–97. MR0692579

[3] Breyer L. A. and Roberts G. O. (2001). Catalytic perfect simulation. *Methodol. Comput. Appl. Probab.* **3** 161–177. MR1868568

[4] Cogburn, R. (1972). The Central Limit Theorem for Markov Processes. *In Le Cam, L. E., Neyman, J. & Scott, E. L. (Eds) Proc. Sixth Ann. Berkley Symp. Math. Sttist. and Prob.* **2** 458–512.

[5] Geyer C. J. (1992). Practical Markov Chain Monte Carlo. *Stat. Sci.* **7** 473–511.

[6] Häggström, O. (2005). On the Central Limit Theorem for Geometrically Ergodic Markov Chains. *Probability Theory and Related Fields* **132** 74–82. MR2136867

[7] Hobert J. P. and Robert C. P. (2004). A mixture representation of π with applications in Markov chain Monte Carlo and perfect smpling. *Ann. Appl. Probab.* **14** 1295–1305. MR2071424

[8] Ibragimov, I. A. and Linnik, Y. V. (1971). *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhof, Groningen. MR0322926

[9] Jones, G. L. (2005). On the Markov chain central limit theorem. *Probability Surveys* **1** 299–320. MR2068475

[10] Jones, G. L., Haran, M., Caffo, B. S. and Neath, R. (2006). Fixed-Width Output Analysis for Markov Chain Monte Carlo. *Journal of the American Statatistical Association.* **101** 1537–1547. MR2279478

[11] Meyn S. P. and Tweedie R. L. (1993). *Markov Chains and Stochastic Stability.* Springer-Verlag. MR1287609

[12] Nummelin E. (1978). A splitting technique for Harris recurrent chains. *Z. Wahrscheinlichkeitstheorie und Verw. Geb.* **43** 309–318. MR0501353

[13] Nummelin E. (1984). *General Irreducible Markov Chains and Nonnegative Operators.* Cambridge University Press, Cambridge. MR0776608

[14] Nummelin E. (2002). MC's for MCMC'ists. *International Statistical Review.* **70** 215–240.

[15] Propp, J. G. and Wilson, D. B. (1996). Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms.* **9** 223–252. MR1611693

[16] Roberts G. O. and Rosenthal J. S. (1997). Geometric Ergodicity and Hybrid Markov Chains. *Elec. Comm. Prob.* **2** 13–25 MR1448322

[17] Roberts G. O. and Rosenthal J. S. (2005). General state space Markov chains and MCMC algorithms. *Probability Surveys* **1** 20–71. MR2095565