

Markov Chain CLTs: Resolving Open Problems

Austin Brown^a, Jeffrey S. Rosenthal^b, and Quan Zhou^{a,1}

^aTexas A&M University and ^bUniversity of Toronto

Abstract

Markov chain central limit theorems (CLTs) and their associated variances are very important for implementing Markov chain Monte Carlo algorithms among other applications. Häggström and Rosenthal (2007) presented various results regarding the equality of different formulae for this variance and also posed seven open problems. We resolve all seven in this paper. For stationary, ergodic and reversible chains, we prove that whenever the normalized partial sums satisfy a $\sqrt{n}$ -CLT, the variance limit is finite if the function is square-integrable, otherwise undefined. Moreover, failure of the $\sqrt{n}$ -CLT forces the normalized partial sums to be non-tight. We also show that Roberts' holding-probability condition precludes a CLT even without assuming reversibility or square-integrability. Finally, we develop a general principle that expresses Fourier coefficients as the autocovariances of an ergodic nonreversible Markov chain.

1 Introduction

The central limit theorem (CLT) for i.i.d. finite-variance random variables is well known. For additive functionals of Markov chains, similar CLTs also exist, and the theory has been developed along several complementary lines. Spectral methods yield CLTs for reversible chains [14, 33]; martingale-based methods treat broader classes of chains through L^2 bounds on conditional expectations [11, 39, 55]; drift, mixing, and regeneration methods provide verifiable conditions on general state spaces [2, 29, 48]; and more recent works have also investigated criteria for chain-wide CLTs and heavy-tailed stationary distributions [5, 10]. Such CLT results are particularly useful for obtaining error bounds in Markov chain Monte Carlo algorithms [21, 23, 30], which often require estimating the asymptotic variance σ^2 associated with the corresponding CLT.

Häggström and Rosenthal [27] studied three variance formulae associated with a centered function h of a stationary Markov chain $(X_n)_{n \geq 0}$. In their notation, A is the limiting variance of $S_n/\sqrt{n}$, where S_n is the partial sum of $(h(X_k))_{k \geq 0}$, B is the corresponding infinite sum of lagged autocovariances, and for reversible chains with square-integrable h , C is the spectral representation of the asymptotic variance. Formal definitions are given in Section 2. Building on the classical result of Kipnis and Varadhan [33] on the relation between $\sqrt{n}$ -CLT and finiteness of C , Häggström and Rosenthal established conditions under which the quantities A, B, C are all equal to each other and to the variance σ^2 appearing in the CLT. The behavior of A has also been studied in many other works. For example, chain-level conditions ensuring finiteness of A are identified in Roberts and Rosenthal [49] and Derriennic and Lin [15], spectral characterizations are developed in Deligiannidis et al. [13] for normal (not necessarily reversible) transition operators, variational and Poisson-equation formulas for nonreversible chains are given in Huang and Mao [28], and sharper results are available for time-sampled chains, independence Metropolis–Hastings algorithms, alternating kernels and Gibbs samplers [12, 35, 38, 45]. Against this background, Häggström and Rosenthal [27] posed seven open problems (henceforth HR Problems) regarding the behavior of A, B, C under more general conditions. To our knowledge, those problems have remained unresolved, and this paper presents solutions to all of them.

¹Corresponding author. Email: quan@stat.tamu.edu

Behavior of h	Behavior of S_n	Behavior of A, B, C
$h \in L^2(\pi)$	$\sqrt{n}$ -CLT exists	$A = B = C = \sigma^2 < \infty$
$h \in L^2(\pi)$	$(S_n/\sqrt{n})$ not tight	$A = B = C = \infty$
$h \notin L^2(\pi)$	$\sqrt{n}$ -CLT exists	A, B are undefined; $\liminf A_n = \sigma^2$, $\limsup A_n = \infty$, and $\gamma_{2k} = \infty$, $\gamma_{2k+1} = -\infty$.
$h \notin L^2(\pi)$	$(S_n/\sqrt{n})$ not tight	Neither A nor B can be finite

Table 1: Four exhaustive scenarios for ergodic reversible Markov chains with stationary distribution π and centered $h \in L^1(\pi)$; $A_n = \text{Var}(S_n)/n$ and γ_k denotes the lag- k autocovariance.

HR Problems 2 and 6 ask, for reversible chains, whether A must be finite or undefined for a $\sqrt{n}$ -CLT to hold. Existing theories and examples suggest that the answer is affirmative [4, 19, 27, 37, 57]. We establish this conclusion assuming only tightness of $(S_n/\sqrt{n})_{n \geq 1}$ and obtain a sharp classification of the relation between the $\sqrt{n}$ -CLT and the variance quantities A and B . As shown in Table 1, every ergodic and reversible chain with centered $h \in L^1(\pi)$ satisfies the following dichotomy: either a $\sqrt{n}$ -CLT holds, or the sequence $(S_n/\sqrt{n})_{n \geq 1}$ is not even tight. This extends the same dichotomy result of Chen [8, Theorem II.3.1] for Harris recurrent aperiodic chains with $h \in L^2(\pi)$. When $h \in L^2(\pi)$, our additional conclusion is that tightness forces $A = B = C < \infty$; combined with classical results, this shows that $A < \infty$ is both necessary and sufficient for a $\sqrt{n}$ -CLT. When the CLT holds but $h \notin L^2(\pi)$, we prove that A, B must be undefined: A is undefined since the variance of $S_n/\sqrt{n}$ has oscillatory behavior with limit inferior σ^2 and limit superior ∞ , and B is undefined since lagged autocovariances (which exist) alternate between ∞ and $-\infty$. Thus, the qualitative behavior exhibited by Example 14 of Häggström and Rosenthal [27] is unavoidable. (And, C cannot even be defined when $h \notin L^2(\pi)$.) A key intermediate result is that, for reversible chains, tightness of $(S_n/\sqrt{n})_{n \geq 1}$ implies square-integrability of $h(X_0) + h(X_1)$, even when $h \notin L^2(\pi)$; in other words, the non-square-integrable contribution must cancel across adjacent observations. The proof involves a novel resolvent argument using characteristic functions. For $h \in L^2(\pi)$, we give an alternative regenerative proof that combines known results for cycle sums [2, 24] with a forward-backward martingale decomposition [54] to establish $C < \infty$, without assuming a finite second moment of the cycle length.

HR Problem 1 concerns the role of reversibility in such results. It is known that for nonreversible chains, the CLT behavior can be considerably less structured [26, 27]. Our solution to HR Problem 1 provides an explicit example where $h \in L^2(\pi)$, $A < \infty$ but B is undefined. To construct it, we develop a general Fourier-to-Markov covariance realization principle: from any strictly positive, continuous, even, 2π -periodic function f , we can construct a countable-state, irreducible, aperiodic, positive recurrent Markov chain with a centered $h \in L^2(\pi)$ whose lagged autocovariances coincide with the Fourier coefficients of f . Thus, the covariance behavior of nonreversible Markov chains is rich enough to encode the Fourier coefficients of every function in this broad class.

HR Problems 3, 5, and 7 concern a divergence condition introduced by Roberts [47], which has been used in the literature to establish the failure of $\sqrt{n}$ -CLT for independence Metropolis–Hastings samplers [47], although a complete proof has not previously appeared. We fill this gap by proving that, even without requiring reversibility or $h \in L^2(\pi)$, Roberts’ condition cannot coexist with a $\sqrt{n}$ -CLT. The proof involves deriving a lower bound on σ^2 using holding probabilities, which extends existing formulae for reversible chains with $h \in L^2(\pi)$ [12, 18].

One interesting aspect of this project is that all proofs were initially obtained by the GPT-5.6 Sol model [42]. Although we have completely re-written and verified its solutions, the essence of our results all originated there. Our experience adds to the emerging evidence that frontier

large language model software has now reached the point where it is competitive with human mathematicians in resolving challenging research questions, provided that their outputs are carefully checked and revised by humans [6], which has implications far beyond the current study [1].

AI Statement of Use. GPT-5.6 Sol was used to develop the proofs, generate the figures, and search the literature. In particular, the solution to HR Problem 2 involved a lengthy interactive discussion and investigation into counterexamples and techniques with GPT-5.6 Sol over the course of a week. The authors verified, substantially revised and reorganized all proof arguments, and they assume full responsibility for the accuracy and presentation of the paper.

2 Formal Definitions and HR Problems

2.1 Notation and Background

Throughout this paper, we let P be a time-homogeneous Markov transition kernel on a measurable space $(\mathcal{X}, \mathcal{F})$, with stationary probability distribution π . Let $\mathbb{E}_\mu, \text{Var}_\mu$ denote the expectation and variance under initial distribution μ , and we write $\mathbb{E}_x = \mathbb{E}_{\delta_x}$. Denote the n -step transition kernel by P^n . For every measurable function $f : \mathcal{X} \rightarrow \mathbb{R}$ and every probability measure μ on $(\mathcal{X}, \mathcal{F})$, define

$$(P^n f)(x) = \int_{\mathcal{X}} f(y) P^n(x, dy), \quad (\mu P^n)(E) = \int_{\mathcal{X}} P^n(x, E) \mu(dx),$$

for every $x \in \mathcal{X}$ and $E \in \mathcal{F}$. We also write $\pi(f) = \int_{\mathcal{X}} f(x) \pi(dx)$. For convenience, we assume that $\mathcal{X}$ is a Polish space and $\mathcal{F}$ is the Borel sigma-algebra so that the function $x \mapsto P(x, \{x\})$ is measurable and $\mathcal{F}$ is countably generated. We also assume that P is ergodic, which means that it is ϕ -irreducible and aperiodic. It follows that P is also recurrent [40, Proposition 10.1.1] and π -irreducible [40, Theorem 10.4.9], and $P^n(x, \cdot)$ converges to π in total variation distance for π -almost every x [48, Theorem 4]. Further, there exists an absorbing set G with $\pi(G) = 1$ such that P restricted to G is positive Harris recurrent [40, Theorem 9.0.1].

We always use $(X_n)_{n \geq 0}$ to denote a stationary Markov chain with kernel P . Fix some $h \in L_0^1(\pi)$; that is, h is integrable with $\pi(h) = 0$. Define for $n \geq 1$,

$$S_n = \sum_{k=0}^{n-1} h(X_k).$$

We say that a $\sqrt{n}$ -central limit theorem (CLT) holds if, for some $\sigma^2 \in [0, \infty)$,

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

where $\Rightarrow$ denotes weak convergence; we use σ^2 to denote this asymptotic variance throughout.

We say a family of real-valued random variables $(U_i)_{i \in \mathcal{I}}$, where $\mathcal{I}$ is an arbitrary index set, is bounded in probability or tight if

$$\lim_{N \rightarrow \infty} \sup_{i \in \mathcal{I}} \mathbb{P}(|U_i| \geq N) = 0.$$

If $S_n/\sqrt{n}$ converges in distribution or $\sup_{n \geq 1} n^{-1} \text{Var}_\pi(S_n) < \infty$, then the sequence $(S_n/\sqrt{n})_{n \geq 1}$ is bounded in probability.

2.2 Three Variance Expressions

Denote the lag- k autocovariance by

$$\gamma_k = \mathbb{E}_\pi[h(X_0)h(X_k)], \quad k \geq 0,$$

which is allowed to be infinite. When $h \in L^2(\pi)$, $|\gamma_k| < \infty$ for every k . For $n \geq 1$ and $m \geq 0$, let

$$A_n = \frac{1}{n} \text{Var}_\pi(S_n), \quad B_m = \gamma_0 + 2 \sum_{k=1}^m \gamma_k. \quad (2.1)$$

Note that B_m is well-defined provided that each γ_k is well-defined and the resulting sum does not contain the form $\infty - \infty$. When the limits exist, write

$$A = \lim_{n \rightarrow \infty} A_n, \quad B = \lim_{m \rightarrow \infty} B_m, \quad (2.2)$$

and both limits are allowed to be infinite. We have the elementary identity

$$A_n = \gamma_0 + 2 \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) \gamma_k = \frac{1}{n} \sum_{m=0}^{n-1} B_m, \quad (2.3)$$

provided that $(B_m)_{0 \leq m \leq n-1}$ are well-defined. Thus, A_n is the Cesàro mean of $B_0, \dots, B_{n-1}$. Consequently, if B exists, then A exists and $A = B$. We also record two useful observations about the behavior of $(A_n)_{n \geq 1}$.

Lemma 2.1. *If $A < \infty$, then $h \in L^2(\pi)$. If $h \notin L^2(\pi)$, then $\max\{A_n, A_{n+1}\} = \infty$ for every $n \geq 1$.*

Proof. Since $h(X_n) = S_{n+1} - S_n$ and $(X_n)_{n \geq 0}$ is stationary, $h \in L^2(\pi)$ whenever both S_n, S_{n+1} have finite variance for some $n \geq 1$. Similarly, if $h \notin L^2(\pi)$, then at least one of S_n, S_{n+1} has infinite variance for every $n \geq 1$. The two assertions then follow. $\square$

We say that Roberts' condition holds if

$$\lim_{n \rightarrow \infty} n \mathbb{E}_\pi[h(X_0)^2 r(X_0)^n] = \infty, \quad r(x) = P(x, \{x\}). \quad (2.4)$$

Lemma 2.2 (Roberts, 1999). *If Roberts' condition holds, then $A = \infty$.*

Proof. See Roberts [47]. In particular, $A_n \geq n \mathbb{E}_\pi[h(X_0)^2 r(X_0)^n]$ for each $n \geq 1$. $\square$

When P is reversible with respect to π and $h \in L^2(\pi)$, we define the spectral variance of h by

$$C = \int_{[-1,1]} \frac{1+\lambda}{1-\lambda} \mathcal{E}_h(d\lambda), \quad (2.5)$$

where $\mathcal{E}_h$ is the spectral measure for h ; that is, $\mathcal{E}_h(\cdot) = \langle h, \mathcal{E}(\cdot)h \rangle_{L^2(\pi)}$ where $\mathcal{E}$ is the unique spectral decomposition measure associated with P . If P is not reversible or $h \notin L^2(\pi)$, we leave C undefined. When $h \in L^2(\pi)$, several established results describe the relationship between the $\sqrt{n}$ -CLT and the quantities A, B, C , which we collect in the theorem below.

Theorem 2.3. *Let P be ergodic, $(X_n)_{n \geq 0}$ be stationary and $h \in L^2(\pi)$ with $\pi(h) = 0$.*

- (i) *Kipnis and Varadhan [33]: If P is reversible and $C < \infty$, then a $\sqrt{n}$ -CLT exists with $\sigma^2 = C$.*
- (ii) *Häggström and Rosenthal [27, Theorem 4]: If P is reversible, then $A = B = C \in [0, \infty]$.*
- (iii) *Peligrad [43, Theorem 1]: If $\sup_{n \geq 1} A_n < \infty$, then a $\sqrt{n}$ -CLT exists.*

Part (iii) is a direct consequence of Chen [8, Theorem II.3.1], which shows that for ergodic and stationary chains with $h \in L^2(\pi)$, $(S_n/\sqrt{n})_{n \geq 1}$ is tight if and only if a $\sqrt{n}$ -CLT exists.

2.3 HR Problems and Solutions

Häggström and Rosenthal [27, Section 5] posed seven open problems. Below we describe each of them and the corresponding solution we obtain. Throughout, we assume that P is ergodic. (Although this assumption is not explicitly stated in HR Problems 5 and 6, it appears to be implicit.)

HR Problem 1

Problem Statement. If the Markov chain is ergodic, but not necessarily reversible nor geometrically ergodic nor uniformly ergodic, does it necessarily follow that $A = B$ (allowing that they may both be infinite)? What if we assume that $P = P_1 P_2$ where each P_i is reversible?

We interpret the question as asking whether the conclusion holds for $h \in L^2(\pi)$, since Example 11 of Häggström and Rosenthal [27] already provides a counterexample where $A = \infty$ and B is undefined due to $\gamma_0 = \infty, \gamma_1 = -\infty$. Häggström and Rosenthal [27] proved that for $h \in L^2(\pi)$, $A = B$ if the chain is reversible and ergodic, or if the chain is uniformly ergodic (see their Remark 8). In particular, we have $A = B$ on any finite state space since ergodicity would imply uniform ergodicity. So HR Problem 1 is primarily motivated by the role of reversibility in this result. Since the convergence of B_n implies that A_n converges to the same limit, a counterexample must have a divergent sequence $(B_n)_{n \geq 0}$.

Resolution. We prove in Theorem 5.4 that both assertions fail with $h \in L^2(\pi)$, even if $P = P_1 P_2$ with P_i reversible. Counterexamples on countable spaces are constructed, where $A < \infty$ exists, $|\gamma_k| < \infty$ for every k , but B does not exist since $\liminf_{n \rightarrow \infty} B_n < \limsup_{n \rightarrow \infty} B_n = \infty$; the $\sqrt{n}$ -CLT also holds for our construction. The main technical contribution of our proof is a general device, presented in Theorem 5.1, for realizing the Fourier coefficients of a given function as the autocovariances of a regenerative Markov chain with finite excursions. Proposition 5.2 gives the construction of the chain, and Proposition 5.3 records the key analytic decomposition result needed to apply it. Once this realization device is established, the problem reduces to applying classical results and counterexamples in Fourier analysis.

HR Problem 2

Problem Statement. Is there a reversible, ergodic Markov chain where a $\sqrt{n}$ -CLT exists, but $A = \infty$?

For $h \in L^2(\pi)$ with $A = \infty$, several CLT results have also been established, but with a normalization different from $\sqrt{n}$ [37, 57]. Bradley [4] also provided multiple examples where $A = \infty$ but the $\sqrt{n}$ -CLT fails. For $h \notin L^2(\pi)$, Example 14 of Häggström and Rosenthal [27] has a $\sqrt{n}$ -CLT, while A is undefined rather than infinite [27, Remark 15]. A negative answer to HR Problem 2 would demonstrate that all these phenomena are not accidental. Compared to the existing results that use $\sup_{n \geq 1} A_n < \infty$ or $C < \infty$ to establish the $\sqrt{n}$ -CLT (see Theorem 2.3), the challenge is to proceed in the converse direction and extend the analysis beyond the condition $h \in L^2(\pi)$.

Resolution. We prove in Theorem 3.1 that the answer is negative. In particular, we show that

$$\sup_{n \geq 1} \frac{1}{2n} \text{Var}_\pi(S_{2n}) \leq 2\sigma^2,$$

which implies that A cannot be infinite. Combined with Lemma 2.1, this shows that A must be undefined if a $\sqrt{n}$ -CLT exists for $h \notin L^2(\pi)$. As in existing techniques to showing a Markov chain CLT [36, 54], we construct a function d such that for $k \geq 1$

$$h(X_{k-1}) + h(X_k) = d(X_{k-1}, X_k) + d(X_k, X_{k-1}), \quad (2.6)$$

and both $d(X_{k-1}, X_k)$ and $d(X_k, X_{k-1})$ are square-integrable martingale difference sequences. However, existing theories rely on square-integrability assumptions of h , while we require only $h \in L_0^1(\pi)$, making the construction of a function d a central technical challenge. Our key innovation is to construct d as the limit of a sequence of bounded functions with uniform L^2 control. The sequence is built using the resolvent of the characteristic function of S_n and the CLT is used to choose the correct scaling that yields (2.6); see Proposition 3.2. This decomposition reveals the hidden square integrability created by cancellation of adjacent observations. In fact, we show that this result can be proved under the strictly weaker assumption that $(S_n/\sqrt{n})_{n \geq 1}$ is tight.

HR Problem 3

Problem Statement. Is there a reversible, ergodic Markov chain where a $\sqrt{n}$ -CLT exists, but where Roberts' condition $\lim_{n \rightarrow \infty} n\mathbb{E}[h(X_0)^2 r(X_0)^n] = \infty$ holds?

By Lemma 2.2, $A = \infty$ whenever Roberts' condition holds. However, it does not necessarily preclude a $\sqrt{n}$ -CLT, since weak convergence alone does not imply convergence of moments.

Resolution. We prove in Theorem 4.1 that the answer is negative without assuming reversibility and with tightness of $(S_n/\sqrt{n})_{n \geq 1}$ in place of the $\sqrt{n}$ -CLT assumption. Hence, Roberts' condition does rule out a $\sqrt{n}$ -CLT for reversible and nonreversible ergodic chains. The key idea, which we isolate in Proposition 4.2, is to condition on the successive distinct states visited and integrate out the holding times at those states. These holding times are conditionally independent geometric random variables, and their fluctuations cause the modulus of the conditional characteristic function to decay multiplicatively. Under the CLT assumption, we obtain a lower bound on σ^2 in terms of the holding probabilities.

HR Problem 4

Problem Statement. If the chain is reversible and ergodic and a $\sqrt{n}$ -CLT exists, what further conditions (weaker than i.i.d. or uniformly ergodic) still imply that $A < \infty$?

This question is open-ended. Various sufficient conditions ensuring $A < \infty$ without assuming the CLT exist in the literature, including the variance-bounding property [49] and spectral criteria [13].

Resolution. Our solution to HR Problem 2, together with Theorem 4 of [27], implies that $h \in L^2(\pi)$ is a necessary and sufficient condition; see Corollary 3.7.

HR Problem 5

Problem Statement. In particular, if the chain is reversible (and ergodic) and a $\sqrt{n}$ -CLT exists, and there is $\delta > 0$ such that $r(x) \geq \delta$ for all $x \in \mathcal{X}$, does this imply that $A < \infty$?

Resolution. We give an affirmative answer in Corollary 4.3. This result follows readily from our solutions to HR Problems 2 and 3.

HR Problem 6

Problem Statement. If the Markov chain is reversible (and ergodic) and a $\sqrt{n}$ -CLT exists, and $\pi(h^2) < \infty$, does this imply that $A < \infty$?

Under the assumptions of the problem, we have $A = B = C$, though they may all be infinite [27, Theorem 4]. Hence, this problem essentially asks whether a converse to the Kipnis–Varadhan theorem holds. See the preceding discussion of HR Problem 2 for other related results.

Resolution. Our solution to HR Problem 2 shows that the answer is affirmative; see Corollary 3.6. Combined with Theorem 4 of Häggström and Rosenthal [27] and the Kipnis–Varadhan theorem, this shows that for ergodic and reversible Markov chains, if $h \in L^2(\pi)$, then either a $\sqrt{n}$ -CLT holds with $\sigma^2 = A = B = C < \infty$, or $(S_n/\sqrt{n})_{n \geq 1}$ is not tight and $A = B = C = \infty$. We give an alternative, more constructive proof in Section 3.4 that reverses the standard arguments used to establish Markov chain CLTs. In particular, we show that tightness forces the sum over a regeneration cycle to have a finite second moment, though the length of each cycle may not be square-integrable. This leads to an upper bound on the spectral variance C .

HR Problem 7

Problem Statement. Can the condition $\pi(h^2) < \infty$ be dropped from Corollaries 6 and 7?

For readers' convenience, here we include the relevant corollary statements from Häggström and Rosenthal [27].

HR Corollary 6. *If P is reversible and ergodic, and $\pi(h^2) < \infty$, and any one of A, B , and C is finite, then a $\sqrt{n}$ -CLT exists for h , with $\sigma^2 = A = B = C < \infty$.*

HR Corollary 7. *If P is reversible and ergodic, and $\lim_{n \rightarrow \infty} n\mathbb{E}_\pi[h^2(X_0)r(X_0)^n] = \infty$ where $\pi(h^2) < \infty$, then $A = B = C = \infty$.*

The spectral quantity C is defined only when $h \in L^2(\pi)$, and $B < \infty$ already implies $\gamma_0 = \pi(h^2) < \infty$. So for HR Corollary 6, dropping the L^2 assumption is nontrivial only when $A < \infty$.

For HR Corollary 7, Lemma 2.2 shows that Roberts' condition forces $A = \infty$. Since C is not defined outside $L^2(\pi)$, the remaining question is whether Roberts' condition also forces $B = \infty$. Because $\gamma_0 = \pi(h^2) = \infty$ when $h \notin L^2(\pi)$, a counterexample must have B undefined. This can occur either because some γ_k is undefined or because B_n involves $\infty - \infty$; we consider the latter, more substantive case. We also note that Roberts' condition holds trivially if $\mathbb{E}_\pi[h(X_0)^2 r(X_0)^n] = \infty$ for all large n . However, Roberts' condition may also hold when

$$\mathbb{E}_\pi[h(X_0)^2 r(X_0)^n] < \infty, \text{ for every } n \geq 1. \quad (2.7)$$

We impose (2.7) as an additional requirement.

Resolution. By Lemma 2.1, the condition $\pi(h^2) < \infty$ can be removed from HR Corollary 6. For HR Corollary 7, we prove in Theorem 4.5 that the condition $\pi(h^2) < \infty$ cannot be removed. A counterexample is given, where $h \in L_0^1(\pi) \setminus L^2(\pi)$, both (2.7) and Roberts' condition hold, $A = \infty$, but B is undefined due to $\gamma_0 = \infty$ and $\gamma_1 = -\infty$. The main idea is to introduce two transition mechanisms, where one mechanism produces $\gamma_0 = \infty, \gamma_1 = -\infty$, and the other ensures that Roberts' condition holds nontrivially. The parameters of the two mechanisms are chosen so that each mechanism generates its intended effect without interfering with the other.

2.4 Main Theorem

We summarize our main results and their consequences in the following theorem. Detailed proofs for HR Problems are presented in subsequent sections, and some standard results used in the proofs are collected in the Appendix.

Theorem 2.4. *Let $(X_n)_{n \geq 0}$ be a stationary, ergodic Markov chain with transition kernel P and invariant probability measure π . Let $h \in L_0^1(\pi)$, and define $S_n = \sum_{k=0}^{n-1} h(X_k)$. Let A, B, C be given by (2.1), (2.2) and (2.5). The following results hold.*

(i) If P is reversible, then exactly one of the following four scenarios occurs.

- (a) $h \in L^2(\pi)$, a $\sqrt{n}$ -CLT exists for h with $\sigma^2 = A = B = C < \infty$;
- (b) $h \in L^2(\pi)$, $(S_n/\sqrt{n})_{n \geq 1}$ is not tight, and $A = B = C = \infty$;
- (c) $h \notin L^2(\pi)$, a $\sqrt{n}$ -CLT exists for h , A is undefined since $\liminf_{n \rightarrow \infty} A_n = \sigma^2$ and $\limsup_{n \rightarrow \infty} A_n = \infty$, B is undefined since $\gamma_{2k} = \infty$ and $\gamma_{2k+1} = -\infty$ for each $k \geq 0$;
- (d) $h \notin L^2(\pi)$, $(S_n/\sqrt{n})_{n \geq 1}$ is not tight, and neither A nor B is finite.

(ii) If Roberts' condition (2.4) holds, then $A = \infty$ and $(S_n/\sqrt{n})_{n \geq 1}$ is not tight, but B may be undefined due to $\gamma_0 = \infty$ and $\gamma_1 = -\infty$.

(iii) There exist a nonreversible P and $h \in L^2(\pi)$ such that $A < \infty$, and B is undefined since $\liminf_{n \rightarrow \infty} B_n < \infty$ and $\limsup_{n \rightarrow \infty} B_n = \infty$.

Proof. If $h \in L^2(\pi)$, then $A = B = C \in [0, \infty]$, and if any one of them is finite, then a $\sqrt{n}$ -CLT holds with $\sigma^2 = A < \infty$ [27, Theorem 4, Corollary 6]. Further, by Theorem 3.1, $A = \infty$ implies that $(S_n/\sqrt{n})_{n \geq 1}$ is not tight, which proves the two possible scenarios when $h \in L^2(\pi)$. If $h \notin L^2(\pi)$, Lemma 2.1 implies that A cannot be finite, and since $\gamma_0 = \pi(h^2) = \infty$, B cannot be finite either. By Theorem 3.1, tightness of $(S_n/\sqrt{n})_{n \geq 1}$ implies a $\sqrt{n}$ -CLT, and the characterization of A, B when a $\sqrt{n}$ -CLT exists follows from Corollary 3.5. This proves part (i).

Part (ii) follows from Lemma 2.2, Theorem 4.1 and Theorem 4.5.

Part (iii) follows from Theorem 5.4. □

3 HR Problems 2, 4 and 6: Implications of a Reversible CLT

3.1 Main Result and Proof

Theorem 3.1 (Solution to HR Problem 2). *Let P be reversible and ergodic and $(X_n)_{n \geq 0}$ be stationary. Let $h \in L^1(\pi)$ with $\pi(h) = 0$. If $(S_n/\sqrt{n})_{n \geq 1}$ is bounded in probability, then $S_n/\sqrt{n} \Rightarrow \mathbf{N}(0, \sigma^2)$ for some $0 \leq \sigma^2 < \infty$, and*

$$\sup_{n \geq 1} A_{2n} \leq 2\sigma^2, \quad \lim_{n \rightarrow \infty} A_{2n} = \sigma^2, \quad (3.1)$$

where $A_n = \text{Var}_\pi(S_n)/n$. When $\sigma^2 = 0$, $S_{2n} = 0$ almost surely for every $n \geq 1$.

We aim to show that the non-square-integrable contribution must cancel within adjacent pairs. Our starting point is the forward-backward martingale decomposition that has been used previously to control $h(X_0) + h(X_1)$ [36, 54, 56]. But those works had different motivations: they typically start with the assumption $h \in L^2(\pi)$ to establish a CLT. Here we work in the converse direction, where a CLT or tightness is assumed to hold, and we want to recover the square integrability of $h(X_0) + h(X_1)$, the one-step change of S_{2n} . This is accomplished through bounded ‘‘characteristic resolvent’’ functions that capture the fluctuations of the entire Markovian dynamics and provide a uniformly bounded L^2 -approximation to $h(X_0) + h(X_1)$. Related characteristic resolvent constructions have also appeared in other contexts, for example in the large deviation theory [34].

Define a joint coupling Γ on $\mathcal{X} \times \mathcal{X}$ by

$$\Gamma(dx, dy) = \pi(dx)P(x, dy).$$

Since we assume $(X_n)_{n \geq 0}$ is stationary, $(X_0, X_1) \sim \Gamma$. The following two propositions are the key construction and novelty of our approach, which we prove in Section 3.3.

Proposition 3.2. *If $(S_n/\sqrt{n})_{n \geq 1}$ is bounded in probability, then there exists a real-valued function $d \in L^2(\Gamma)$ such that*

$$\begin{aligned} \mathbb{E}_\pi[d(X_0, X_1)^2] &< \infty, \\ \mathbb{E}[d(X_0, X_1) \mid X_0] &= 0 \quad \pi\text{-almost surely}, \\ h(x) + h(y) &= d(x, y) + d(y, x) \quad \text{for } \Gamma\text{-almost every } (x, y) \in \mathcal{X} \times \mathcal{X}. \end{aligned} \quad (3.2)$$

Proposition 3.3. *If $S_n/\sqrt{n} \Rightarrow \mathbf{N}(0, \sigma^2)$ for $0 \leq \sigma^2 < \infty$, then the conclusions of Proposition 3.2 hold with the following strengthening:*

$$\mathbb{E}_\pi[d(X_0, X_1)^2] \leq \sigma^2.$$

Proof of Theorem 3.1. Let d be given by Proposition 3.2. Define

$$M_k = d(X_{k-1}, X_k), \quad M_k^* = d(X_k, X_{k+1}).$$

Then $(M_k)_{k \geq 1}$ is a forward martingale-difference sequence with respect to the filtration $\mathcal{F}_k = \sigma(X_0, \dots, X_k)$, and $(M_k^*)_{k \geq 1}$ is a reverse martingale-difference sequence with respect to the reverse filtration $\mathcal{F}_k^* = \sigma(X_{k-1}, X_k, \dots)$. In particular, since P is reversible,

$$\mathbb{E}[M_k^* \mid \mathcal{F}_{k+1}^*] = \mathbb{E}[d(X_k, X_{k+1}) \mid X_k] = \mathbb{E}[d(X_k, X_{k+1}) \mid X_k] = 0.$$

See, e.g., Wu [54, Theorem 2.5] for a complete proof. Write $\|d\|_{L^2(\Gamma)}^2 = \mathbb{E}_\pi[d(X_0, X_1)^2]$. Then, $\mathbb{E}_\pi[M_k^2] = \mathbb{E}_\pi[(M_k^*)^2] = \|d\|_{L^2(\Gamma)}^2$, and by the pairwise L^2 -orthogonality within each sequence,

$$\mathbb{E}_\pi \left[\left(\sum_{k=0}^{n-1} M_{2k+1} \right)^2 \right] = \mathbb{E}_\pi \left[\left(\sum_{k=0}^{n-1} M_{2k+1}^* \right)^2 \right] = n \|d\|_{L^2(\Gamma)}^2. \quad (3.3)$$

By (3.2), the even sum subsequence yields the decomposition

$$S_{2n} = \sum_{k=0}^{n-1} [h(X_{2k}) + h(X_{2k+1})] = \sum_{k=0}^{n-1} M_{2k+1} + \sum_{k=0}^{n-1} M_{2k+1}^*. \quad (3.4)$$

Using $\pi(h) = 0$, $(a+b)^2 \leq 2a^2 + 2b^2$ for $a, b \in \mathbb{R}$ and (3.3), we have

$$\text{Var}_\pi(S_{2n}) = \mathbb{E}_\pi[S_{2n}^2] \leq 2\mathbb{E}_\pi \left[\left(\sum_{k=0}^{n-1} M_{2k+1} \right)^2 \right] + 2\mathbb{E}_\pi \left[\left(\sum_{k=0}^{n-1} M_{2k+1}^* \right)^2 \right] = 4n \|d\|_{L^2(\Gamma)}^2.$$

Since $\|d\|_{L^2(\Gamma)}^2 < \infty$, this shows that

$$\limsup_n \frac{1}{2n} \text{Var}_\pi(S_{2n}) \leq 2 \|d\|_{L^2(\Gamma)}^2 < \infty. \quad (3.5)$$

Let $V_k = (X_{2k}, X_{2k+1})$ for $k \geq 0$. Then $(V_k)_{k \geq 0}$ is a stationary and ergodic Markov chain, though it is not necessarily reversible. Let $H(x, y) = h(x) + h(y)$, which is in $L^2(\Gamma)$. By part (iii) of Theorem 2.3, condition (3.5) implies that a $\sqrt{n}$ -CLT holds for the partial sums of $(H(V_k))_{k \geq 0}$, i.e., $S_{2n}/\sqrt{2n} \Rightarrow \mathbf{N}(0, \sigma^2)$ for some $0 \leq \sigma^2 < \infty$. By Slutsky's theorem, the odd sum subsequence converges weakly to the same limit, and thus $S_n/\sqrt{n} \Rightarrow \mathbf{N}(0, \sigma^2)$. So by Proposition 3.3, we can further choose d such that $\|d\|_{L^2(\Gamma)}^2 \leq \sigma^2$, proving the bound (3.1).

To prove $A_{2n} \rightarrow \sigma^2$, we apply the CLT for martingale differences (see Lemma A.3) to sequences $(M_{2k+1})_{k \geq 0}$ and $(M_{2k+1}^*)_{k \geq 0}$, both of which are stationary and ergodic. This yields

$$J_n := \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} M_{2k+1} \Rightarrow \mathbf{N}\left(0, \|d\|_{L^2(\Gamma)}^2\right), \quad J_n^* := \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} M_{2k+1}^* \Rightarrow \mathbf{N}\left(0, \|d\|_{L^2(\Gamma)}^2\right).$$

By (3.3), this implies that both $(J_n^2)_{n \geq 1}$ and $((J_n^*)^2)_{n \geq 1}$ are uniformly integrable. Consequently, by (3.4), $(S_{2n}^2/n)_{n \geq 1}$ is also uniformly integrable, and thus A_{2n} converges to σ^2 .

Finally, letting $n = 1$, we get

$$\mathbb{E}_\pi[S_2^2] = \mathbb{E}_\pi[(h(X_0) + h(X_1))^2] \leq 4\sigma^2. \quad (3.6)$$

Hence, if $\sigma^2 = 0$, $h(X_0) + h(X_1) = 0$ almost surely, which completes the proof of Theorem 3.1. $\square$

Remark 3.1. Theorem 3.1 actually holds for a general state space $(\mathcal{X}, \mathcal{F})$. The current proof only requires that $\mathcal{F}$ be countably generated so that ergodicity can be used to obtain convergence in total variation distance; see Lemma 3.8 in Section 3.3. However, in Lemma A.1, we show that one can always restrict P on a general space to a countably generated sigma-algebra for the stationary process $(h(X_n))_{n \geq 0}$. In particular, Lemma A.1 implies that invariance, reversibility, ϕ -irreducibility, and aperiodicity can be preserved under this restriction.

3.2 Sharpness and Consequences of Theorem 3.1

By stationarity and (3.6), Theorem 3.1 implies that, for every $k \geq 1$, $h(X_k) + h(X_{k-1})$ is square-integrable with variance bounded by $4\sigma^2$. We give a simple example showing that the factor 4 in general cannot be improved.

Proposition 3.4. *Let $\mathcal{X} = \{-1, 1\}$ and P be the transition matrix with $P(x, x) = \delta$, $P(x, -x) = 1 - \delta$ for some $0 < \delta < 1$. Let $h(x) = x$. Then, $\text{Var}_\pi(S_2) = 4\delta$, while a $\sqrt{n}$ -CLT holds with*

$$\sigma^2 = \frac{\delta}{1 - \delta}.$$

Hence, $\sigma^{-2} \text{Var}_\pi(S_2) = 4(1 - \delta) \rightarrow 4$ as $\delta \downarrow 0$.

Proof. The second largest eigenvalue of P is $\lambda = 2\delta - 1$. Hence, the spectral formula yields $\sigma^2 = (1 + \lambda)/(1 - \lambda) = \delta/(1 - \delta)$. And a direct calculation gives $\text{Var}_\pi(S_2) = \mathbb{E}_\pi[S_2^2] = 4\delta$. $\square$

When $h \notin L^2(\pi)$ but a $\sqrt{n}$ -CLT holds, the fact that $h(X_k) + h(X_{k-1})$ is square-integrable can be used to precisely characterize the behavior of A, B .

Corollary 3.5. *Let P be reversible and ergodic and $(X_n)_{n \geq 0}$ be stationary. Let $h \in L_0^1(\pi) \setminus L^2(\pi)$, and assume a $\sqrt{n}$ -CLT holds for h with $0 \leq \sigma^2 < \infty$. Then,*

$$\liminf_{n \rightarrow \infty} A_n = \sigma^2, \quad \limsup_{n \rightarrow \infty} A_n = \infty, \quad \gamma_{2k} = \infty \text{ and } \gamma_{2k+1} = -\infty \text{ for } k \geq 0.$$

Proof. By Lemma 2.1, $h \notin L^2(\pi)$ yields $\limsup_{n \rightarrow \infty} A_n = \infty$. By Theorem 3.1, $A_{2n} \rightarrow \sigma^2$, which implies $\liminf_{n \rightarrow \infty} A_n \leq \sigma^2$. Moreover, since $u \mapsto u^2$ is lower semicontinuous and $S_n/\sqrt{n} \Rightarrow \mathbf{N}(0, \sigma^2)$, Fatou's lemma yields $\liminf_{n \rightarrow \infty} A_n \geq \sigma^2$, and thus $\liminf_{n \rightarrow \infty} A_n = \sigma^2$.

For the lag- k autocovariance, note that

$$h(X_0) - h(X_{2k}) = \sum_{j=0}^{2k-1} (-1)^j [h(X_j) + h(X_{j+1})],$$

where the sum of each adjacent pair, $h(X_j) + h(X_{j+1})$, is in L^2 . Hence, $h(X_0) - h(X_{2k}) \in L^2$. Since for any $a, b \in \mathbb{R}$, the negative part of ab , denoted by $(ab)^-$, satisfies $(ab)^- \leq (a - b)^2/4$,

$$\mathbb{E}_\pi [(h(X_0)h(X_{2k}))^-] \leq \frac{1}{4} \mathbb{E}_\pi [(h(X_0) - h(X_{2k}))^2] < \infty.$$

Meanwhile, because

$$h(X_0)^2 + h(X_{2k})^2 = (h(X_0) - h(X_{2k}))^2 + 2h(X_0)h(X_{2k}) \text{ is not integrable,}$$

the positive part of $h(X_0)h(X_{2k})$ cannot be integrable, which proves $\gamma_{2k} = \infty$. An analogous argument shows that $h(X_0) + h(X_{2k+1}) \in L^2$ and consequently $\gamma_{2k+1} = -\infty$. $\square$

Combining Theorem 3.1 with the results of Häggström and Rosenthal [27], we also obtain solutions to HR Problems 4 and 6.

Corollary 3.6 (Solution to HR Problem 6). *Let P be reversible and ergodic and $(X_n)_{n \geq 0}$ be stationary. Let $h \in L^2(\pi)$ with $\pi(h) = 0$ satisfy a $\sqrt{n}$ -CLT. Then, $A < \infty$.*

Proof. By Theorem 4 of Häggström and Rosenthal [27], if $h \in L^2(\pi)$, then A exists as an extended-real limit of the sequence A_n . By Theorem 3.1, the existence of a $\sqrt{n}$ -CLT implies that $A_{2n} \leq 2\sigma^2$ for each n , and thus $A < \infty$. $\square$

Corollary 3.7 (Solution to HR Problem 4). *Let P be reversible and ergodic and $(X_n)_{n \geq 0}$ be stationary. Let $h \in L^1(\pi)$ with $\pi(h) = 0$ satisfy a $\sqrt{n}$ -CLT. Then, $A < \infty$ if and only if $\pi(h^2) < \infty$.*

Proof. By Lemma 2.1, $A < \infty$ implies $h \in L^2(\pi)$. By Corollary 3.6, $h \in L^2(\pi)$ implies $A < \infty$. $\square$

3.3 Proofs of Propositions 3.2 and 3.3

For $0 < p < 1$, let N_p have a geometric distribution and be independent of $(X_n)_{n \geq 0}$. Thus,

$$\mathbb{P}(N_p = n) = p(1 - p)^{n-1}, \quad n = 1, 2, \dots \quad (3.7)$$

All expectations below include the randomness of N_p . Define $\tilde{N}_p = N_p - 1$. Let $\tilde{S}_n = \sum_{k=1}^n h(X_k)$ denote the sum starting at index 1, and define $\tilde{S}_0 = 0$. Pick $t \neq 0$ and define $\theta = t\sqrt{p}$. Define the “characteristic resolvent” functions by

$$f_p(x; t) = \mathbb{E}_x \left[e^{i\theta \tilde{S}_{\tilde{N}_p}} \right], \quad g_p(x; t) = \mathbb{E}_x \left[e^{i\theta S_{N_p}} \right] = e^{i\theta h(x)} f_p(x; t).$$

We will omit the dependence on t and simply write $f_p(x), g_p(x)$. Define

$$c_p := \pi(f_p) = \mathbb{E}_\pi \left[\exp \left(i\theta \tilde{S}_{\tilde{N}_p} \right) \right].$$

The proof of the propositions relies on two properties of f_p and g_p . The first shows that f_p becomes asymptotically independent of x as $p \downarrow 0$, and if the $\sqrt{n}$ -CLT holds, we can also find the limit of c_p . This is the only point in the proof that requires the use of the ergodicity assumption.

Lemma 3.8. For π -almost every x , $|f_p(x) - c_p| \rightarrow 0$ as $p \downarrow 0$. If $S_n/\sqrt{n} \Rightarrow \mathbf{N}(0, \sigma^2)$, then

$$\lim_{p \downarrow 0} c_p = \frac{1}{1 + t^2 \sigma^2 / 2}.$$

The second ingredient is a bound on the L^2 -norm of the normalized one-step difference. For $p \in (0, 1)$, define a complex-valued difference quotient

$$D_p(x, y) = \frac{g_p(y) - (Pg_p)(x)}{i\theta c_p}.$$

The centering by $(Pg_p)(x)$ ensures that $D_p(X_0, X_1)$ has conditional mean zero given X_0 .

Lemma 3.9. For any fixed $t \neq 0$ and $0 < p < 1$ such that $c_p \neq 0$,

$$\|D_p\|_{L^2(\Gamma)}^2 \leq \frac{2|1 - c_p|}{t^2(1 - p)^2|c_p|}.$$

We defer the proofs of both lemmas to the end of this subsection and first show how they imply the two propositions.

Proof of Proposition 3.2. Using $\tilde{S}_0 = 0$ and the Markov property,

$$f_p(x) = p + \sum_{k=1}^{\infty} p(1-p)^k \mathbb{E}_x \left[\exp \left(i\theta \tilde{S}_k \right) \right] = p + (1-p)(Pg_p)(x). \quad (3.8)$$

Taking expectation on both sides yields

$$c_p = p + (1-p)\pi(g_p) = p + (1-p)\phi_p(t), \quad \phi_p(t) = \mathbb{E}_\pi \left[e^{it\sqrt{p}S_{N_p}} \right]. \quad (3.9)$$

By Lemma A.2, the family $(\sqrt{p}S_{N_p})_{0 < p < 1}$ is also bounded in probability, and their characteristic functions are equicontinuous at zero. In particular, we can choose t such that $\sup_{0 < p < 1} |\phi_p(t) - 1| \leq 1/2$, and by (3.9), this implies

$$\sup_{0 < p < 1} |c_p - 1| \leq \frac{1}{2}. \quad (3.10)$$

Adding and subtracting $f_p(x)$ in the numerator of $D_p(x, y)$ yields

$$\begin{aligned} D_p(x, y) + D_p(y, x) &= \frac{[e^{i\theta h(y)} - 1]f_p(y)}{i\theta c_p} + \frac{[e^{i\theta h(x)} - 1]f_p(x)}{i\theta c_p} \\ &\quad + \frac{f_p(x) - (Pg_p)(x)}{i\theta c_p} + \frac{f_p(y) - (Pg_p)(y)}{i\theta c_p}. \end{aligned} \quad (3.11)$$

Since $|g_p(x)| \leq 1$ for every x and $\theta = t\sqrt{p}$, by (3.8),

$$\sup_{x \in \mathcal{X}} \left| \frac{f_p(x) - (Pg_p)(x)}{\theta} \right| \leq \frac{\sqrt{p}}{|t|} \sup_{x \in \mathcal{X}} |1 - (Pg_p)(x)| \leq \frac{2\sqrt{p}}{|t|}. \quad (3.12)$$

For the first two terms on the right-hand side of (3.11), we have

$$\lim_{p \downarrow 0} \frac{f_p(x)}{c_p} = 1, \quad \lim_{p \downarrow 0} \frac{e^{i\theta h(x)} - 1}{i\theta} = h(x). \quad (3.13)$$

The first claim follows from (3.10) and Lemma 3.8, and the second follows from $\theta = t\sqrt{p}$ and that $\lim_{|z| \rightarrow 0} (e^z - 1)/z = 1$ for complex z . So we obtain from (3.11), (3.12) and (3.13) that

$$\lim_{p \downarrow 0} [D_p(x, y) + D_p(y, x)] = h(x) + h(y), \text{ for } \Gamma\text{-almost every } (x, y) \in \mathcal{X} \times \mathcal{X}. \quad (3.14)$$

By (3.10) and Lemma 3.9, we have

$$\sup_{0 < p \leq 1/2} \|D_p\|_{L^2(\Gamma)}^2 \leq \sup_{0 < p \leq 1/2} \frac{2}{t^2(1-p)^2} \leq \frac{8}{t^2}. \quad (3.15)$$

Hence, we can choose a sequence $(p_j)_{j \geq 0}$ such that $p_j \downarrow 0$, and then the uniform upper bound (3.15) implies that $(D_{p_j})_{j \geq 0}$ is a bounded sequence in $L^2(\Gamma, \mathbb{C})$. Then, possibly reindexing, there exists a subsequence $(D_{p_j})_{j \geq 0}$ weakly convergent to a limit $D \in L^2(\Gamma, \mathbb{C})$, that is,

$$\int D_{p_j}(x, y) \overline{F(x, y)} \Gamma(dx, dy) \rightarrow \int D(x, y) \overline{F(x, y)} \Gamma(dx, dy) \text{ for every } F \in L^2(\Gamma, \mathbb{C});$$

see, for example, Kallenberg [31, Chapter 3]. Let $\mathcal{R}$ denote the reversal operator defined for functions $F : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$ by

$$(\mathcal{R}F)(x, y) = F(y, x).$$

By the reversibility of P , $G \in L^2(\Gamma, \mathbb{C})$ implies $\mathcal{R}G \in L^2(\Gamma, \mathbb{C})$, and $\int (\mathcal{R}F) \overline{G} d\Gamma = \int F \overline{\mathcal{R}G} d\Gamma$ (i.e., $\mathcal{R}$ is unitary). Hence, $\mathcal{R}D_{p_j} \rightarrow \mathcal{R}D$ weakly in $L^2(\Gamma, \mathbb{C})$, and thus

$$D_{p_j} + \mathcal{R}D_{p_j} \rightarrow D + \mathcal{R}D \text{ weakly in } L^2(\Gamma, \mathbb{C}). \quad (3.16)$$

Since $\mathcal{R}$ is unitary and using the bound (3.15),

$$\sup_{0 < p_j \leq 1/2} \|D_{p_j} + \mathcal{R}D_{p_j}\|_{L^2(\Gamma)}^2 \leq 4 \sup_{0 < p_j \leq 1/2} \|D_{p_j}\|_{L^2(\Gamma)}^2 \leq \frac{32}{t^2}. \quad (3.17)$$

Define $H : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ by $H(x, y) = h(x) + h(y)$. By (3.14), $D_{p_j} + \mathcal{R}D_{p_j} \rightarrow H$, Γ -almost everywhere. Combining (3.17) with Fatou's lemma implies $H \in L^2(\Gamma, \mathbb{C})$ and that

$$D_{p_j} + \mathcal{R}D_{p_j} \rightarrow H \text{ both strongly in } L^1(\Gamma, \mathbb{C}) \text{ and weakly in } L^2(\Gamma, \mathbb{C}). \quad (3.18)$$

Indeed, the strong convergence in $L^1(\Gamma, \mathbb{C})$ is due to uniform integrability, which, together with Hölder's inequality, implies that as $j \rightarrow \infty$, for any bounded measurable G ,

$$\int (D_{p_j} + \mathcal{R}D_{p_j}) \overline{G} d\Gamma \rightarrow \int H \overline{G} d\Gamma. \quad (3.19)$$

Since bounded measurable functions are dense in $L^2(\Gamma, \mathbb{C})$ and the uniform bound holds (3.17), (3.19) holds also for $G \in L^2(\Gamma, \mathbb{C})$, which proves the weak convergence in $L^2(\Gamma, \mathbb{C})$ stated in (3.18). Since weak limits are unique, (3.18) and (3.16) imply we have the identity

$$D + \mathcal{R}D = H. \quad (3.20)$$

Finally, define

$$d(x, y) = \operatorname{Re}(D(x, y)).$$

Then d is real-valued and belongs to $L^2(\Gamma)$. We verify that d satisfies the three conditions given in the proposition statement. First, (3.20) yields $d(x, y) + d(y, x) = h(x) + h(y)$ for Γ -almost every $(x, y) \in \mathcal{X} \times \mathcal{X}$. Second, by weak lower semicontinuity of the norm and (3.17),

$$\mathbb{E}_\pi[d(X_0, X_1)^2] = \|d\|_{L^2(\Gamma)}^2 \leq \liminf_{j \rightarrow \infty} \|D_{p_j}\|_{L^2(\Gamma)}^2 < \infty. \quad (3.21)$$

Third, let $\mathcal{K} = \{F \in L^2(\Gamma, \mathbb{C}) : \mathbb{E}[F(X_0, X_1) \mid X_0] = 0\}$, which is a closed linear subspace. Since $D_p \in \mathcal{K}$ for every p , the weak limit D also belongs to $\mathcal{K}$, which implies that $\mathbb{E}[d(X_0, X_1) \mid X_0] = 0$. This concludes the proof of Proposition 3.2. $\square$

Proof of Proposition 3.3. Since the $\sqrt{n}$ -CLT implies $(S_n/\sqrt{n})_{n \geq 1}$ is bounded in probability, all conclusions and proof arguments for Proposition 3.2 still hold. It only remains to prove the improved bound for $\|d\|_{L^2(\Gamma)}^2$. By Lemma 3.9 and (3.21),

$$\|d\|_{L^2(\Gamma)}^2 \leq \liminf_{j \rightarrow \infty} \|D_{p_j}\|_{L^2(\Gamma)}^2 \leq \liminf_{j \rightarrow \infty} \frac{2|1 - c_{p_j}|}{t^2(1 - p_j)^2|c_{p_j}|} = \sigma^2,$$

where the last step uses the limit of c_p given in Lemma 3.8. This concludes the proof. $\square$

Proof of Lemma 3.8. By the definition of f_p and the memoryless property of N_p ,

$$\begin{aligned} f_p(x) &= \mathbb{E}_x \left[e^{i\theta \tilde{S}_{\tilde{N}_p}} \mathbf{1}_{\{\tilde{N}_p < m\}} \right] + \mathbb{E}_x \left[e^{i\theta \tilde{S}_{\tilde{N}_p}} \mathbf{1}_{\{\tilde{N}_p \geq m\}} \right] \\ &= \mathbb{E}_x \left[e^{i\theta \tilde{S}_{\tilde{N}_p}} \mathbf{1}_{\{\tilde{N}_p < m\}} \right] + (1 - p)^m \mathbb{E}_x \left[e^{i\theta \tilde{S}_m} f_p(X_m) \right] \end{aligned}$$

for every fixed $m \geq 1$. Let $q = 1 - p$. Since $\mathbb{P}(\tilde{N}_p < m) = 1 - q^m$, we get

$$\begin{aligned} |f_p(x) - (P^m f_p)(x)| &\leq \left| \mathbb{E}_x \left[e^{i\theta \tilde{S}_{\tilde{N}_p}} \mathbf{1}_{\{\tilde{N}_p < m\}} \right] + q^m \mathbb{E}_x \left[e^{i\theta \tilde{S}_m} f_p(X_m) \right] - \mathbb{E}_x[f_p(X_m)] \right| \\ &\leq 2(1 - q^m) + q^m \left| \mathbb{E}_x \left[\left(e^{i\theta \tilde{S}_m} - 1 \right) f_p(X_m) \right] \right| \\ &\leq 2(1 - q^m) + q^m \mathbb{E}_x \left| e^{i\theta \tilde{S}_m} - 1 \right| \leq 4. \end{aligned}$$

With $\theta = t\sqrt{p}$, we apply the dominated convergence theorem to the preceding bound to obtain

$$\lim_{p \downarrow 0} |f_p(x) - (P^m f_p)(x)| = 0.$$

Meanwhile, since $|f_p(x)| \leq 1$ for every x and $c_p = \pi(f_p)$, by the variational characterization of the total variation distance,

$$|(P^m f_p)(x) - c_p| \leq 2\|P^m(x, \cdot) - \pi\|_{\text{TV}}.$$

Since P is ergodic, $\|P^m(x, \cdot) - \pi\|_{\text{TV}} \rightarrow 0$ for π -almost every x as $m \rightarrow \infty$. Applying the triangle inequality and letting $m \rightarrow \infty$ shows that $|f_p(x) - c_p|$ goes to 0.

Now assume the $\sqrt{n}$ -CLT exists. By Lemma A.2 and (3.9),

$$\lim_{p \downarrow 0} c_p = \lim_{p \downarrow 0} \phi_p(t) = \frac{1}{1 + t^2 \sigma^2 / 2},$$

which concludes the proof. $\square$

Proof of Lemma 3.9. Since $|g_p(x)| = |f_p(x)|$ and using (3.8), the conditional variance of D_p is

$$\begin{aligned} |c_p|^2 \theta^2 \mathbb{E} [|D_p(X_0, X_1)|^2 | X_0] &= \mathbb{E} [|g_p(X_1) - (Pg_p)(X_0)|^2 | X_0] \\ &= \mathbb{E} [|f_p(X_1)|^2 | X_0] - |(Pg_p)(X_0)|^2. \end{aligned}$$

Taking expectation over $X_0 \sim \pi$ yields

$$|c_p|^2 \theta^2 \|D_p\|_{L^2(\Gamma)}^2 = \|f_p\|_{L^2(\pi)}^2 - \|Pg_p\|_{L^2(\pi)}^2.$$

Using the identity (3.8) and $\|f_p - p\|_{L^2(\pi)}^2 = \|f_p\|_{L^2(\pi)}^2 - 2p \operatorname{Re}(\pi(f_p)) + p^2$, we get

$$\begin{aligned} \|D_p\|_{L^2(\Gamma)}^2 &= \frac{1}{|c_p|^2 \theta^2} \left[\|f_p\|_{L^2(\pi)}^2 - \frac{1}{(1-p)^2} \|f_p - p\|_{L^2(\pi)}^2 \right] \\ &= \frac{2 \operatorname{Re}(c_p) - \|f_p\|_{L^2(\pi)}^2 (2-p) - p}{t^2 (1-p)^2 |c_p|^2}. \end{aligned}$$

Since $\|f_p\|_{L^2(\pi)}^2 \geq |\pi(f_p)|^2 = |c_p|^2$, we have

$$\begin{aligned} 2 \operatorname{Re}(c_p) - \|f_p\|_{L^2(\pi)}^2 (2-p) - p &\leq 2 \operatorname{Re}(c_p) - |c_p|^2 (2-p) - p \\ &\leq 2(\operatorname{Re}(c_p) - |c_p|^2) \\ &= 2 \operatorname{Re}(\overline{c_p}(1 - c_p)) \\ &\leq 2|c_p||1 - c_p|. \end{aligned}$$

The claimed bound then follows. $\square$

3.4 An Alternative Proof for HR Problem 6

Although Theorem 3.1 readily resolves HR Problem 6, it does not provide any upper bound on the asymptotic variance σ^2 . But when $h \in L^2(\pi)$, additional spectral and regenerative tools become available. This motivates us to provide an alternative proof which yields an upper bound on σ^2 by reversing the standard arguments used to establish the $\sqrt{n}$ -CLT.

To explain our bound, we first recall the minorization condition and the resulting regenerative construction of the chain P . By Meyn and Tweedie [40, Theorem 5.2.1], we can choose an integer $m \geq 1$, a measurable set D with $\pi(D) > 0$, and a constant $\beta \in (0, 1]$ such that

$$P^m(x, E) \geq \beta \mathbf{1}_D(x) \pi_D(E), \quad x \in \mathcal{X}, E \in \mathcal{F},$$

where the probability measure π_D is defined by $\pi_D(\cdot) = \pi(\cdot \cap D)/\pi(D)$. Note that Meyn and Tweedie [40, Theorem 5.2.1] states the minorization condition as $P^m(x, E) \geq \delta \pi(E)$ for $x \in D$ and $E \subset D$, and we have renormalized the minorization constant by $\beta = \delta \pi(D)$. Define the residual kernel

$$L(x, E) = P^m(x, E) - \beta \pi_D(E) \mathbf{1}_D(x). \quad (3.22)$$

Define the skeleton chain $(Y_k)_{k \geq 0}$ by $Y_k = X_{km}$. Define the m -step block-sum function g and the corresponding partial sum by

$$g = \sum_{j=0}^{m-1} P^j h, \quad T_n = \sum_{k=0}^{n-1} g(Y_k).$$

The function g gives the expected sum of $h(X_n)$ over the next m steps conditional on the current state. Since $h \in L^2(\pi)$ and P is an $L^2(\pi)$ -contraction, $g \in L^2(\pi)$. We construct a split chain by lifting the skeleton chain up a dimension with Bernoulli random variables [40, Chapters 5 and 13]. Define $(I_k, Y_k)_{k \geq 0}$ with $I_k \in \{0, 1\}$ by the transition probabilities

$$\begin{aligned}\mathbb{P}(I_k = 1, Y_{k+1} \in E \mid Y_k = x) &= \beta \pi_D(E) \mathbf{1}_D(x) \\ \mathbb{P}(I_k = 0, Y_{k+1} \in E \mid Y_k = x) &= L(x, E).\end{aligned}$$

Marginally $(Y_k)_{k \geq 0}$ still has transition kernel P^m . Unless otherwise stated, the split chain is initialized with $Y_0 \sim \pi$, and $\mathbb{E}$ denotes expectation under the resulting split-chain law. When $Y_0 \sim \mu$, we write $\mathbb{E}_\mu$. Define regeneration times $1 \leq \tau_1 < \tau_2 < \dots$ by

$$\{\tau_i : i \geq 1\} = \{k + 1 : I_k = 1, k \geq 0\}.$$

For $i \geq 1$, define a sequence of i.i.d. random variables by

$$Z_i = \sum_{k=\tau_i}^{\tau_{i+1}-1} g(Y_k).$$

Each Z_i is the sum of $g(Y_k)$ accumulated over a regeneration cycle. The length of each cycle, $\tau_{i+1} - \tau_i$, has a finite mean, but its second moment is not necessarily finite. However, it can be proved that $\mathbb{E}[Z_1^2]$ is finite under the tightness assumption, which further yields an upper bound on σ^2 . The proof of the following theorem is deferred to the Appendix.

Theorem 3.10. *Let P be reversible and ergodic and $(X_n)_{n \geq 0}$ be stationary. Let $h \in L^2(\pi)$ with $\pi(h) = 0$ and $\beta, D, g, (Y_k)_{k \geq 0}, (T_n)_{n \geq 1}, (Z_i)_{i \geq 1}$ be as defined above. If $(S_n/\sqrt{n})_{n \geq 1}$ is bounded in probability, then*

$$\frac{S_n}{\sqrt{n}} \Rightarrow \mathbf{N}(0, \sigma^2), \quad \frac{T_n}{\sqrt{n}} \Rightarrow \mathbf{N}(0, \sigma_g^2),$$

where $\sigma^2 = A = B = C$ and

$$\sigma^2 \leq \sigma_g^2 = \beta \pi(D) \mathbb{E}[Z_1^2] < \infty.$$

4 HR Problems 3, 5 and 7: Implications of Roberts' Condition

4.1 A Holding-Time Lower Bound on σ^2

Theorem 4.1 (Solution to HR Problem 3). *Let P be ergodic and $(X_n)_{n \geq 0}$ be stationary. Let $h \in L^1(\pi)$ with $\pi(h) = 0$. Define $r(x) = P(x, \{x\})$ and $q(x) = 1 - r(x)$. If $(S_n/\sqrt{n})_{n \geq 1}$ is bounded in probability, then $h = 0$ on $\{q = 0\}$ π -almost surely, and*

$$\int_{\{q > 0\}} h(x)^2 \frac{r(x)}{q(x)} \pi(dx) < \infty. \quad (4.1)$$

Consequently,

$$\lim_{n \rightarrow \infty} n \mathbb{E}_\pi [h(X_0)^2 r(X_0)^n] = 0. \quad (4.2)$$

If $S_n/\sqrt{n} \Rightarrow \mathbf{N}(0, \sigma^2)$ for $0 \leq \sigma^2 < \infty$, then we also have

$$\int_{\{q > 0\}} h(x)^2 \frac{r(x)}{q(x)} \pi(dx) \leq \sigma^2. \quad (4.3)$$

Conditional on the sequence of distinct states visited by the chain, the holding times at each state are independent geometric random variables. The holding time L at state x has conditional variance $\text{Var}(L \mid x) = r(x)/q(x)^2$. Multiplying it by stationary frequency of new holding periods $q(x)\pi(dx)$ and the squared contribution $h(x)^2$ yields

$$q(x)\pi(dx) h(x)^2 \frac{r(x)}{q(x)^2} = h(x)^2 \frac{r(x)}{q(x)} \pi(dx),$$

which can be thought of as the variance per unit time generated purely by the geometric holding times. This is exactly the integrand in (4.1). Moreover, since these geometric fluctuations are conditionally independent, they cannot be canceled by the subsequent evolution of the jump chain. The main step of our proof is to turn this observation into an upper bound on $|\phi_p(t)|$, as formalized in the following proposition; see Section 4.2 for the proof. Since this inequality arises entirely from the holding times, it does not involve second moment of S_n .

As in the proof of Proposition 3.2, introduce the geometric random variable N_p independent of $(X_n)_{n \geq 0}$ with density given by (3.7). Define the characteristic function of $\sqrt{p}S_{N_p}$ by

$$\phi_p(t) = \mathbb{E}_\pi[e^{it\sqrt{p}S_{N_p}}].$$

For $m \geq 1$, define the truncation set Λ_m and bounded weight $a_m(x)$ by

$$\Lambda_m = \{x : q(x) \geq m^{-1}, |h(x)| \leq m\}, \quad a_m(x) = \frac{h(x)^2 r(x)}{q(x)^2} \mathbf{1}_{\Lambda_m}(x).$$

Proposition 4.2. *For every $m \geq 1$ and $t \neq 0$, we have*

$$\limsup_{p \downarrow 0} |\phi_p(t)| \leq \frac{1}{1 + t^2 G_m / 2},$$

where G_m is defined by

$$G_m := \int q(x) a_m(x) \pi(dx) = \int_{\Lambda_m} h(x)^2 \frac{r(x)}{q(x)} \pi(dx). \quad (4.4)$$

Proof of Theorem 4.1. By Lemma A.2, $(\sqrt{p}S_{N_p})_{0 < p < 1}$ is bounded in probability with equicontinuous characteristic functions, which implies that there is some $t_0 \neq 0$ such that

$$\inf_{0 < p < 1} |\phi_p(t_0)| \geq \frac{1}{2}.$$

It then follows from Proposition 4.2 that

$$\frac{1}{2} \leq \frac{1}{1 + t_0^2 G_m / 2}, \quad (4.5)$$

and thus $G_m \leq 2/t_0^2$. The sets (Λ_m) increase to $\{q > 0\}$, so by the monotone convergence theorem,

$$\int_{\{q > 0\}} h(x)^2 \frac{r(x)}{q(x)} \pi(dx) = \lim_{m \rightarrow \infty} G_m < \infty,$$

which proves (4.1).

Consider states with $q(x) = 0$. For every $\varepsilon > 0$, on the event $\{q(X_0) = 0, |h(X_0)| > \varepsilon\}$, we have $|S_n|/\sqrt{n} > \varepsilon\sqrt{n}$. Since $(S_n/\sqrt{n})_{n \geq 1}$ is bounded in probability, the probability of this event must

be zero. Letting ε range over the positive rationals shows that $h(x) = 0$ for π -almost every x with $r(x) = 1$. Now using the elementary inequality

$$nr^n \leq \sum_{k=1}^n r^k \leq \frac{r}{1-r}, \quad 0 \leq r < 1,$$

we get that

$$0 \leq nh(X_0)^2 r(X_0)^n \leq h(X_0)^2 \frac{r(X_0)}{1-r(X_0)} = h(X_0)^2 \frac{r(X_0)}{1-r(X_0)} \mathbf{1}_{\{r(X_0) < 1\}}.$$

The right-hand side is integrable due to (4.1). Since $\lim_{n \rightarrow \infty} nr^n = 0$ for each $r \in [0, 1)$, dominated convergence proves (4.2).

Finally, when the CLT assumption holds, for every $t \in \mathbb{R}$, Lemma A.2 yields

$$\lim_{p \downarrow 0} \phi_p(t) = \frac{1}{1 + \sigma^2 t^2 / 2}.$$

So the inequality (4.5) can be replaced by

$$\frac{1}{1 + \sigma^2 t^2 / 2} \leq \frac{1}{1 + t^2 G_m / 2},$$

which yields $G_m \leq \sigma^2$. Monotone convergence then proves (4.3). $\square$

Remark 4.1. The proof of Theorem 4.1 is closely related to the jump-chain representation, in which successive distinct states are separated by conditionally independent geometric holding times. This representation is standard for Metropolis–Hastings chains [16] and has been recorded for general Markov kernels by Vihola et al. [53, Proposition 24]. Moreover, an exact counterpart of (4.1) is present in the earlier theory for reversible chains with $h \in L^2(\pi)$; see Doucet et al. [18, Proposition 2] and Deligiannidis and Lee [12, Theorem 1]. Our contribution is to recover the sharp inequality directly from the CLT assumption, without assuming $h \in L^2(\pi)$, convergence of A_n , or reversibility.

Corollary 4.3 (Solution to HR Problem 5). *Let P be reversible and ergodic and $(X_n)_{n \geq 0}$ be stationary. Let $h \in L_0^1(\pi)$. If $(S_n / \sqrt{n})_{n \geq 1}$ is bounded in probability and $\inf_{x \in \mathcal{X}} P(x, \{x\}) \geq \delta > 0$, then $A < \infty$.*

Proof. By Theorem 4.1, the tightness of $(S_n / \sqrt{n})_{n \geq 1}$ implies that Roberts' condition fails and actually, $n \mathbb{E}_\pi [h(X_0)^2 r(X_0)^n] \rightarrow 0$. Hence, for sufficiently large n , we have

$$\infty > n \mathbb{E}_\pi [h(X_0)^2 r(X_0)^n] \geq n \delta^n \mathbb{E}_\pi [h(X_0)^2].$$

Since $\delta > 0$, this means $\pi(h^2) < \infty$. By Theorem 3.1, a $\sqrt{n}$ -CLT also holds, and then the conclusion follows from Corollary 3.6. $\square$

Remark 4.2. By Theorem 3.1, $S_n / \sqrt{n} \Rightarrow \mathbf{N}(0, \sigma^2)$ for some $0 \leq \sigma^2 < \infty$, and then one can use Theorem 4.1 to obtain a sharper bound on $\pi(h^2)$. Since $r(x)/(1-r(x)) \geq \delta/(1-\delta)$ on $\{r < 1\}$, and $h = 0$ on $\{r = 1\}$, we obtain that

$$\pi(h^2) \leq \frac{1-\delta}{\delta} \sigma^2 < \infty.$$

4.2 Proof of Proposition 4.2

Proof of Proposition 4.2. For the chain $(X_n)_{n \geq 0}$, the successive distinct states form a jump chain $(Y_k)_{k \geq 0}$. Explicitly, on $\{q > 0\}$, define the jump kernel

$$Q(x, E) = \frac{P(x, E) - r(x)\mathbf{1}_E(x)}{q(x)}.$$

Extend Q arbitrarily to $\{q = 0\}$. Then

$$P(x, dy) = r(x)\delta_x(dy) + q(x)Q(x, dy), \quad Q(x, \{x\}) = 0 \text{ whenever } q(x) > 0.$$

Then we realize N_p by independently killing the path after each recorded observation with probability p . Hence, if the current state is x , the three possible outcomes have probabilities

$$(1-p)r(x), \quad (1-p)q(x), \quad p,$$

corresponding respectively to a surviving self-transition, a surviving departure from x , and killing. The observed holding period terminates at each step with probability

$$d_p(x) = 1 - (1-p)r(x) = q(x) + pr(x), \quad (4.6)$$

and its length is geometric with parameter $d_p(x)$. Moreover, by the standard competing-risks property of geometric random variables, the observed holding time is independent of the terminal type, conditional on the current state. Conditional on departure, the next distinct state has law $Q(x, \cdot)$, again independently of the holding time.

Let $Y_0, \dots, Y_{K_p}$ be the $K_p + 1$ successive distinct states observed before killing, and let L_j be the observed holding time at Y_j . The final holding period is the one terminated by killing. Hence,

$$S_{N_p} = \sum_{j=0}^{K_p} L_j h(Y_j). \quad (4.7)$$

Let $\mathcal{G}_p = \sigma(K_p, Y_0, \dots, Y_{K_p})$ denote the observed jump chain. Our argument above shows that $L_0, \dots, L_{K_p}$ are conditionally independent given $\mathcal{G}_p$, and

$$L_j \mid \mathcal{G}_p \sim \text{Geom}(d_p(Y_j)), \quad 0 \leq j \leq K_p.$$

In particular, the final terminated period has the same conditional law as every preceding holding period, which is the primary motivation behind this geometric killing mechanism.

Define

$$\Psi_p(t) = \mathbb{E} \left[e^{it\sqrt{p}S_{N_p}} \mid \mathcal{G}_p \right].$$

Then, $\phi_p(t) = \mathbb{E}_\pi[\Psi_p(t)]$. To find an upper bound on $\Psi_p(t)$, we use (4.7) and conditional independence between $L_0, \dots, L_{K_p}$ to get

$$|\Psi_p(t)| = \prod_{j=0}^{K_p} \left| \mathbb{E} \left[e^{it\sqrt{p}h(Y_j)L_j} \mid \mathcal{G}_p \right] \right| = \prod_{j=0}^{K_p} e^{-b_{p,t}(Y_j)},$$

where we define, for $p \in (0, 1)$ and $t \in \mathbb{R}$,

$$b_{p,t}(x) = -\log \left| \mathbb{E} \left[e^{it\sqrt{p}h(x)L} \right] \right|, \quad L \sim \text{Geom}(d_p(x)).$$

Pick any $\eta \in (0, t^2/2)$. We prove in Lemma 4.4 below that every holding period at x produces an unavoidable decay proportional to $pa_m(x)$, yielding the following upper bound on $|\Psi_p(t)|$:

$$|\Psi_p(t)| \leq \exp \left[- \left(\frac{t^2}{2} - \eta \right) pR_{p,m} \right], \quad R_{p,m} = \sum_{j=0}^{K_p} a_m(Y_j),$$

for any $0 < p < p_0(m, t, \eta)$, where $p_0(m, t, \eta) > 0$ is constant independent of x . Jensen's inequality yields $|\phi_p(t)| \leq \mathbb{E}|\Psi_p(t)|$, and thus

$$|\phi_p(t)| \leq \mathbb{E} \left[\exp \left(- \frac{t^2 - 2\eta}{2} pR_{p,m} \right) \right]. \quad (4.8)$$

To convert (4.8) into a deterministic upper bound, we need to show that $pR_{p,m}$ converges. Consider the following pathwise representation of $R_{p,m}$:

$$R_{p,m} = a_m(X_0) + \sum_{k=1}^{N_p-1} W_{k,m}, \quad W_{k,m} = a_m(X_k) \mathbf{1}_{\{X_k \neq X_{k-1}\}}. \quad (4.9)$$

A routine calculation using stationarity gives $\mathbb{E}_\pi[W_{k,m}] = G_m$ where G_m is given by (4.4). Since $a_m \leq m^4$, the stationary sequence $(W_{k,m})_{k \geq 1}$ is bounded and ergodic. Birkhoff's ergodic theorem therefore gives

$$\frac{1}{n} \sum_{k=1}^n W_{k,m} \longrightarrow G_m \quad \text{almost surely.} \quad (4.10)$$

Because $N_p \rightarrow \infty$ in probability, the ergodic average in equation (4.10), evaluated at the independent random index $N_p - 1$, converges to G_m in probability. The initial term in equation (4.9) is bounded and therefore vanishes after multiplication by p . As $p \downarrow 0$, the rescaled horizon satisfies $pN_p \Rightarrow T$ with $T \sim \text{Exp}(1)$. So Slutsky's theorem gives

$$pR_{p,m} = pa_m(X_0) + \frac{p(N_p - 1)}{N_p - 1} \sum_{k=1}^{N_p-1} W_{k,m} \Rightarrow TG_m.$$

Since $u \mapsto e^{-cu}$ is bounded and continuous on $[0, \infty)$ for $c > 0$, it follows from (4.8) that

$$\limsup_{p \downarrow 0} |\phi_p(t)| \leq \mathbb{E} \left[\exp \left(- \frac{t^2 - 2\eta}{2} G_m T \right) \right] = \frac{1}{1 + (t^2/2 - \eta) G_m},$$

where the last step follows from the Laplace transform for $T \sim \text{Exp}(1)$. Letting $\eta \downarrow 0$ proves the proposition. $\square$

Lemma 4.4. Fix $m \geq 1$, $t \neq 0$, and $\eta \in (0, t^2/2)$. There exists $p_0 = p_0(m, t, \eta) > 0$ such that, for every $x \in \mathcal{X}$ and $0 < p < p_0$,

$$b_{p,t}(x) \geq p \left(\frac{t^2}{2} - \eta \right) a_m(x). \quad (4.11)$$

Proof. We first find an expression for $b_{p,t}(x)$. For $L \sim \text{Geom}(d)$ on $\{1, 2, \dots\}$, the characteristic function and its squared modulus are

$$\mathbb{E}[e^{izL}] = \frac{de^{iz}}{1 - (1-d)e^{iz}}, \quad |\mathbb{E}[e^{izL}]|^2 = \frac{d^2}{d^2 + 2(1-d)(1 - \cos z)}. \quad (4.12)$$

Using equations (4.6) and (4.12), we obtain

$$b_{p,t}(x) = \frac{1}{2} \log \left[1 + \frac{2(1-p)r(x)\{1 - \cos(t\sqrt{p}h(x))\}}{\{q(x) + pr(x)\}^2} \right]. \quad (4.13)$$

Now consider the bound (4.11). It is immediate outside Λ_m , because the right-hand side vanishes and $b_{p,t} \geq 0$. On Λ_m , the quantities $|t\sqrt{p}h(x)|$ converge to zero uniformly as $p \downarrow 0$. Using $1 - \cos u \sim u^2/2$ as $u \rightarrow 0$, we get

$$\frac{2\{1 - \cos(t\sqrt{p}h(x))\}}{pt^2h(x)^2} \rightarrow 1 \text{ uniformly on } \Lambda_m.$$

The ratio is interpreted as 1 when $h(x) = 0$. Since $q(x) \geq m^{-1}$ on Λ_m , we also have

$$\frac{(1-p)q(x)^2}{\{q(x) + pr(x)\}^2} \rightarrow 1 \text{ uniformly on } \Lambda_m.$$

Hence, using (4.13), Taylor expansion of the logarithm and $a_m \leq m^4$, we get

$$b_{p,t}(x) = \frac{pt^2}{2} a_m(x) \{1 + o(1)\} \text{ uniformly on } \Lambda_m.$$

The asserted lower bound follows for sufficiently small p . □

4.3 A Counterexample under Roberts' Condition

Theorem 4.5 (Solution to HR Problem 7). *There exist a reversible, irreducible, aperiodic countable-state Markov chain and $h \in L_0^1(\pi) \setminus L^2(\pi)$ such that*

$$\begin{aligned} \mathbb{E}_\pi[h(X_0)^2 r(X_0)^n] &< \infty \quad \text{for every } n \geq 1, \\ n\mathbb{E}_\pi[h(X_0)^2 r(X_0)^n] &\rightarrow \infty, \\ A &= \infty, \quad \gamma_1 = -\infty; \end{aligned}$$

consequently, B is undefined since $\gamma_0 = \pi(h^2) = \infty$.

Our construction involves two classes of states. The “oscillating states” are modeled in the same way as the sign-alternating construction in Example 14 of Häggström and Rosenthal [27], and they will supply the infinite second moment and the negative lag-one product. The other class consists of “sticky states” with long holding times, which will supply Roberts’s condition and contribute only a finite amount to the second moment. If one allows $\mathbb{E}_\pi[h(X_0)^2 r(X_0)^n] = \infty$ for all large n , then Roberts’ condition holds trivially, and simpler counterexamples can be constructed.

Proof of Theorem 4.5. Let the state space be

$$\mathcal{X} = \{0\} \cup \{o_{k,\varepsilon}, s_{k,\varepsilon} : k \geq 1, \varepsilon \in \{-1, 1\}\}.$$

The letters o and s stand for oscillating and sticky, respectively. Set

$$\kappa = \left(1 + 2 \sum_{k=1}^{\infty} 4^{-k} + 2 \sum_{k=1}^{\infty} 8^{-k} \right)^{-1} = \frac{21}{41},$$

and define a probability measure by

$$\pi(0) = \kappa, \quad \pi(o_{k,\varepsilon}) = \kappa 4^{-k}, \quad \pi(s_{k,\varepsilon}) = \kappa 8^{-k}.$$

From an oscillating state, the chain either changes sign or returns to the hub:

$$P(o_{k,\varepsilon}, o_{k,-\varepsilon}) = \frac{1}{2}, \quad P(o_{k,\varepsilon}, 0) = \frac{1}{2}. \quad (4.14)$$

From a sticky state, it either remains in place or returns to the hub:

$$P(s_{k,\varepsilon}, s_{k,\varepsilon}) = 1 - 4^{-k}, \quad P(s_{k,\varepsilon}, 0) = 4^{-k}.$$

From the hub, put

$$\begin{aligned} P(0, o_{k,\varepsilon}) &= \frac{1}{2} 4^{-k}, & P(0, s_{k,\varepsilon}) &= 32^{-k}, \\ P(0, 0) &= 1 - \sum_{k=1}^{\infty} 4^{-k} - 2 \sum_{k=1}^{\infty} 32^{-k} = \frac{56}{93}. \end{aligned} \quad (4.15)$$

All remaining transition probabilities are zero. Equations (4.14)–(4.15) define a Markov kernel. A routine calculation using the detailed balance condition verifies that P is π -reversible. Every state communicates with 0, and 0 reaches every state with positive probability; hence the chain is irreducible. The self-loop in (4.15) makes it aperiodic. Since the chain is irreducible with a countable state space and an invariant probability distribution, it is positive recurrent as well.

Define $h: \mathcal{X} \rightarrow \mathbb{R}$ by

$$h(0) = 0, \quad h(o_{k,\varepsilon}) = \varepsilon 2^k, \quad h(s_{k,\varepsilon}) = \varepsilon 2^k.$$

The sign symmetry gives $\pi(h) = 0$. Moreover,

$$\pi(|h|) = 2\kappa \sum_{k=1}^{\infty} (4^{-k} + 8^{-k}) 2^k = \frac{8}{3} \kappa < \infty.$$

On the other hand,

$$\pi(h^2) = 2\kappa \sum_{k=1}^{\infty} (4^{-k} + 8^{-k}) 4^k = 2\kappa \sum_{k=1}^{\infty} (1 + 2^{-k}) = \infty. \quad (4.16)$$

Therefore $h \in L_0^1(\pi) \setminus L^2(\pi)$. The divergence in (4.16) comes entirely from the oscillating family, each level of which contributes the same positive amount.

The holding probabilities satisfy $r(o_{k,\varepsilon}) = 0$, and $r(s_{k,\varepsilon}) = 1 - 4^{-k}$. The value of $r(0)$ is irrelevant because $h(0) = 0$. For $n \geq 1$,

$$\mathbb{E}_\pi[h(X_0)^2 r(X_0)^n] = 2\kappa \sum_{k=1}^{\infty} 8^{-k} 4^k (1 - 4^{-k})^n \leq 2\kappa \sum_{k=1}^{\infty} 2^{-k} = 2\kappa < \infty.$$

Let $K_n := \lceil \log_4(2n) \rceil$. For $k \geq K_n$, one has $n4^{-k} \leq 1/2$, and Bernoulli's inequality gives

$$(1 - 4^{-k})^n \geq 1 - n4^{-k} \geq \frac{1}{2}.$$

Consequently,

$$\mathbb{E}_\pi[h(X_0)^2 r(X_0)^n] \geq \kappa \sum_{k=K_n}^{\infty} 2^{-k} = \kappa 2^{1-K_n} \geq \kappa 2^{-\log_4(2n)} \geq \frac{\kappa}{\sqrt{2n}}.$$

It follows that

$$n\mathbb{E}_\pi[h(X_0)^2 r(X_0)^n] \geq \kappa \sqrt{\frac{n}{2}} \longrightarrow \infty,$$

which also implies $A = \infty$ by Lemma 2.2.

Consider γ_1 . The self-transitions at $s_{k,\varepsilon}$ yield

$$\mathbb{E}_\pi[(h(X_0)h(X_1))^+] = 2\kappa \sum_{k=1}^{\infty} 8^{-k} 4^k (1 - 4^{-k}) = 2\kappa \sum_{k=1}^{\infty} 2^{-k} (1 - 4^{-k}) < \infty,$$

while transitions from $o_{k,\varepsilon}$ to $o_{k,-\varepsilon}$ yield

$$\mathbb{E}_\pi[(h(X_0)h(X_1))^-] = 2\kappa \sum_{k=1}^{\infty} 4^{-k} 4^k \frac{1}{2} = \kappa \sum_{k=1}^{\infty} 1 = \infty.$$

So $\gamma_1 = -\infty$, while $\gamma_0 = \pi(h^2) = +\infty$. Consequently, B is not well-defined. $\square$

5 HR Problem 1: Nonreversible Counterexamples to $A = B$

5.1 A Fourier-to-Markov Representation Theorem

Our solution to HR Problem 1 relies on a new representation theorem that connects the Fourier transform of 2π -periodic functions to countable-space ergodic Markov chains. For an integrable 2π -periodic function v , write

$$\widehat{v}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} v(\theta) e^{-ik\theta} d\theta,$$

and define its symmetric Fourier partial sums and Fejér means at the origin by

$$(\mathcal{S}_N v)(0) = \sum_{|k| \leq N} \widehat{v}(k), \quad (\mathcal{M}_N v)(0) = \sum_{|k| \leq N} \left(1 - \frac{|k|}{N+1}\right) \widehat{v}(k). \quad (5.1)$$

Theorem 5.1. *Let f be a strictly positive, continuous, even, 2π -periodic function. There exist a countable-state, irreducible, aperiodic, positive recurrent Markov chain with stationary distribution π and a centered $h \in L^2(\pi)$ such that*

$$\gamma_k := \mathbb{E}_\pi[h(X_0)h(X_k)] = \frac{1}{2} \widehat{f}(k), \quad k \geq 0.$$

The proof separates into a probabilistic component and an analytic component. Proposition 5.2 provides the probabilistic one, where we construct an excursion-based Markov chain for realizing a given mixture of finite autocorrelations. See Section 5.2 for the proof.

Proposition 5.2. *For $q \geq 1$, let $a^{(q)} = (a_{q,0}, \dots, a_{q,\ell_q-1})$ be finite real vectors such that*

$$\sum_{q=1}^{\infty} \sum_{j=0}^{\ell_q-1} a_{q,j}^2 < \infty.$$

There exist a countable-state, irreducible, aperiodic, positive recurrent Markov chain P with stationary distribution π and a centered $h \in L^2(\pi)$ whose autocovariances satisfy

$$\gamma_k = \frac{1}{2} \sum_{q=1}^{\infty} \sum_{j=0}^{\ell_q-1-k} a_{q,j} a_{q,j+k}, \quad k \geq 0, \quad (5.2)$$

where the inner sum is zero when $k \geq \ell_q$. Moreover, the construction can be modified so that $P = P_1 P_2$, where each P_i is reversible with respect to π .

Proposition 5.2 reduces the probabilistic construction to finding finite vectors $(a^{(q)})_{q \geq 1}$. To connect this problem with Fourier analysis, let $a = (a_0, \dots, a_{\ell-1}) \in \mathbb{R}^{\ell}$ and define

$$\varphi_a(z) = \sum_{j=0}^{\ell-1} a_j z^j, \quad T_a(\theta) = \left| \varphi_a(e^{i\theta}) \right|^2.$$

A direct expansion gives

$$\widehat{T}_a(k) = \sum_{j=0}^{\ell-1-k} a_j a_{j+k}, \quad k \geq 0, \quad (5.3)$$

where the sum is zero when $k \geq \ell$. Thus, the autocorrelation sequence of a finite vector is the Fourier coefficient sequence of the squared modulus of a polynomial. Since Proposition 5.2 involves a sum of such autocorrelation blocks, we seek a corresponding decomposition into a sum of squared polynomial moduli. Proposition 5.3 provides this analytic result. See Section 5.3 for the proof.

Proposition 5.3. *Let f be a strictly positive, continuous, even, 2π -periodic function. Then there exist finite real vectors $a^{(q)} = (a_{q,0}, \dots, a_{q,\ell_q-1})$ for $q \geq 1$ such that, uniformly in θ ,*

$$f(\theta) = \sum_{q=1}^{\infty} \left| \sum_{j=0}^{\ell_q-1} a_{q,j} e^{ij\theta} \right|^2. \quad (5.4)$$

In particular, for every $k \geq 0$, $\widehat{f}(k) = 2\gamma_k$ where γ_k is given by (5.2). Moreover, $\widehat{f}(0) < \infty$.

Proof of Theorem 5.1. First, apply Proposition 5.3 to obtain the vectors $(a^{(q)})_{q \geq 1}$. Since $\widehat{f}(0) < \infty$, the vectors $(a^{(q)})_{q \geq 1}$ satisfy the assumption of Proposition 5.2, which proves Theorem 5.1. $\square$

Remark 5.1. Theorem 5.1 is similar in spirit to Herglotz's theorem, which can be used to show that the autocovariances of every stationary Gaussian process can be realized through the Fourier transform of a finite symmetric positive measure [44]. At an analytical level, there exist results similar to Proposition 5.3 which also decompose positive continuous functions into sums of positive trigonometric polynomials; see, for example, Rudin [51, Theorem 2.4.2 and Exercise 2.4.3] and the explicit discussion in Chen and Yuan [7, Section 4.1].

5.2 Proof of Proposition 5.2

We first provide a basic construction of P with the given expressions for $(\gamma_k)_{k \geq 0}$ and then show how to modify it so that $P = P_1 P_2$ with reversible P_1, P_2 .

Proof of Proposition 5.2 (Basic Construction). Let the state space be

$$\mathcal{X} = \{0\} \cup \{(q, \varepsilon, j) : q \geq 1, \varepsilon \in \{-1, 1\}, 0 \leq j < \ell_q\}.$$

We let 0 be the “regeneration state” of the chain. From 0, the chain stays at 0 with probability $1/2$; otherwise, it selects a block $q \geq 1$ and an independent symmetric sign ε , deterministically traverses the states $(q, \varepsilon, 0), \dots, (q, \varepsilon, \ell_q - 1)$ and then returns to 0. See Figure 1 for a graphical illustration. Let ρ_q denote the probability of selecting a block q at 0, which is given by

$$\rho_q = \frac{2^{-q}}{c_\rho \ell_q}, \quad c_\rho = \sum_{q=1}^{\infty} \frac{2^{-q}}{\ell_q}.$$

For later use, define

$$\bar{\ell} = \sum_{q=1}^{\infty} \rho_q \ell_q, \quad \alpha = \left(1 + \frac{\bar{\ell}}{2}\right)^{-1}.$$

Note that $\bar{\ell} < \infty$. The transition matrix can now be expressed by

$$\begin{aligned} P(0, 0) &= \frac{1}{2}, \\ P(0, (q, \varepsilon, 0)) &= \frac{\rho_q}{4} && \text{for } q \geq 1, \quad \varepsilon = \pm 1, \\ P((q, \varepsilon, j), (q, \varepsilon, j+1)) &= 1 && \text{for } q \geq 1, \quad \varepsilon = \pm 1, \quad 0 \leq j < \ell_q - 1, \\ P((q, \varepsilon, \ell_q - 1), 0) &= 1 && \text{for } q \geq 1, \quad \varepsilon = \pm 1. \end{aligned}$$

Every state communicates with 0, and the self-loop at 0 makes the chain aperiodic. The dynamics of the chain is visualized in Figure 1. The mean return time to 0 is

$$\frac{1}{2} + \sum_{q=1}^{\infty} \sum_{\varepsilon \in \{-1, 1\}} \frac{\rho_q}{4} (\ell_q + 1) = 1 + \frac{\bar{\ell}}{2} < \infty.$$

Hence the chain is positive recurrent. A routine calculation using renewal theory yields the stationary distribution:

$$\pi(0) = \alpha, \quad \pi(q, \varepsilon, j) = \frac{\alpha \rho_q}{4}. \quad (5.5)$$

Define the function h by

$$h(0) = 0, \quad h(q, \varepsilon, j) = \varepsilon \frac{a_{q,j}}{\sqrt{\alpha \rho_q}}.$$

The symmetric sign ε gives $\pi(h) = 0$, while

$$\pi(h^2) = \sum_{q=1}^{\infty} \sum_{\varepsilon \in \{-1, 1\}} \sum_{j=0}^{\ell_q-1} \frac{\alpha \rho_q}{4} \frac{a_{q,j}^2}{\alpha \rho_q} = \frac{1}{2} \sum_{q=1}^{\infty} \sum_{j=0}^{\ell_q-1} a_{q,j}^2 < \infty.$$

Starting from 0, every future nonzero excursion carries a fresh symmetric sign, and therefore

$$\mathbb{E}_0[h(X_t)] = 0, \quad t \geq 0. \quad (5.6)$$

If the chain starts at (q, ε, j) , its motion is deterministic until the current excursion ends; after the return to 0, the conditional mean is zero by (5.6). Thus

$$(P^k h)(q, \varepsilon, j) = \mathbf{1}_{\{j+k < \ell_q\}} \varepsilon \frac{a_{q,j+k}}{\sqrt{\alpha \rho_q}}. \quad (5.7)$$

Using stationarity and (5.5) and (5.7), we obtain

$$\begin{aligned}
\gamma_k &= \mathbb{E}_\pi[h(X_0)(P^k h)(X_0)] \\
&= \sum_{q=1}^{\infty} \sum_{\varepsilon \in \{-1, 1\}} \sum_{j=0}^{\ell_q-1-k} \frac{\alpha \rho_q}{4} \left(\varepsilon \frac{a_{q,j}}{\sqrt{\alpha \rho_q}} \right) \left(\varepsilon \frac{a_{q,j+k}}{\sqrt{\alpha \rho_q}} \right) \\
&= \frac{1}{2} \sum_{q=1}^{\infty} \sum_{j=0}^{\ell_q-1-k} a_{q,j} a_{q,j+k},
\end{aligned}$$

which completes the proof. $\square$

Remark 5.2. The finite excursion architecture used in our proof is a classical regenerative construction [52]. It was shown in Glynn [24] that covariance functions and spectral quantities of stationary regenerative processes can be represented in terms of quantities defined over individual excursions. A closely related construction already appears in Häggström and Rosenthal [27, Example 11 and Remark 13], where each visit to 0 initiates a deterministic finite excursion. Our construction allows an arbitrary finite sequence $a^{(q)}$ on each excursion and introduces an independent symmetric sign so that correlations across different excursions vanish.

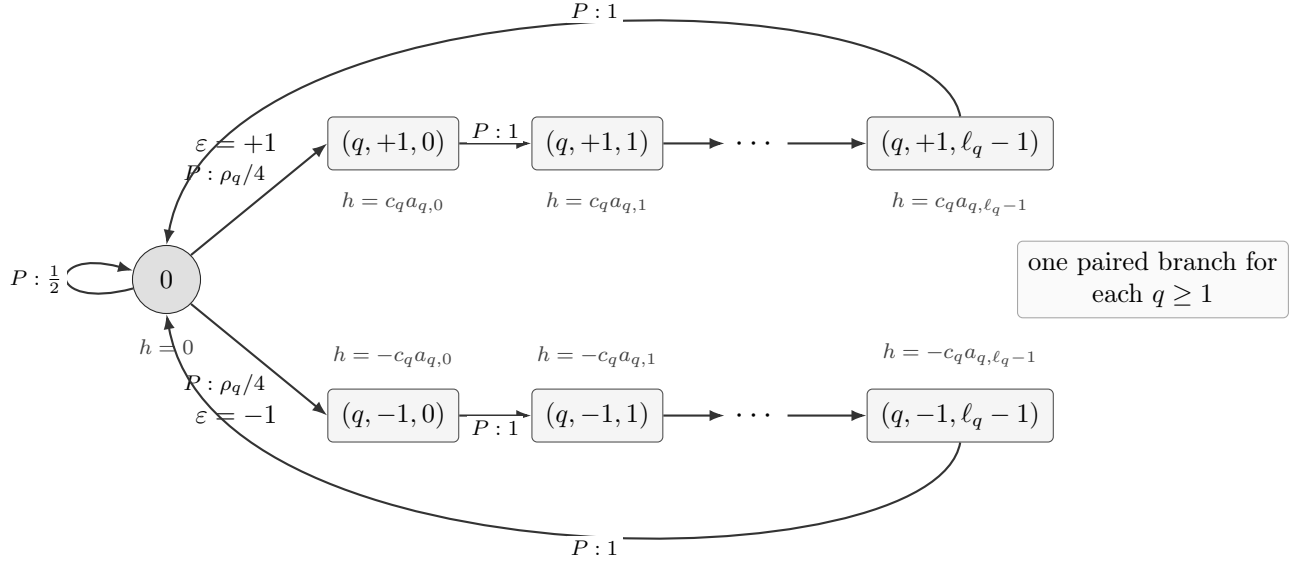


Figure 1: Dynamics of the basic excursion chain. Here $c_q = (\alpha \rho_q)^{-1/2}$.

We now give a factorized construction where $P = P_1 P_2$, and both P_1, P_2 are reversible. The key observation is that each deterministic forward step can be factorized into two reversible reflections: express $j \rightarrow j+1$ as $j \rightarrow -j \rightarrow 1 - (-j)$, with all operations performed modulo the cycle length L .

Proof of Proposition 5.2 (Factorized Construction). Let $\rho_q, \bar{\ell}, \alpha$ be defined as in the basic construction, and set $L_q = \ell_q + 1$. The basic construction uses one common regeneration state 0 for all excursions. Here we replace that common hub by a separate zero-valued anchor for each signed block, and we define the new state space by

$$\mathcal{X}^\circ = \{\star\} \cup \{(q, \varepsilon, j) : q \geq 1, \varepsilon \in \{-1, 1\}, j \in \mathbb{Z}_{L_q}\},$$

where $\mathbb{Z}_L = \{0, 1, \dots, L-1\}$ with addition and subtraction performed modulo L . The additional state $\star$ is regarded as a cycle of length one. For $j \in \mathbb{Z}_L$, introduce the two reflections

$$R_L(j) = -j \pmod{L}, \quad K_L(j) = 1 - j \pmod{L}.$$

They satisfy $K_L(R_L(j)) = j + 1 \pmod{L}$.

Let ν be the probability measure on the anchors given by

$$\nu(\star) = \frac{1}{2}, \quad \nu(q, \varepsilon, 0) = \frac{\rho_q}{4}.$$

Define P_1 to resample from ν at every anchor and to apply the first reflection elsewhere:

$$P_1(x, \cdot) = \begin{cases} \nu(\cdot), & x = \star \text{ or } (q, \varepsilon, 0), \\ \delta_{(q, \varepsilon, R_{L_q}(j))}(\cdot), & x = (q, \varepsilon, j), j \neq 0. \end{cases}$$

Define P_2 by fixing $\star$ and applying the second reflection on each nontrivial cycle:

$$P_2(\star, \star) = 1, \quad P_2((q, \varepsilon, j), (q, \varepsilon, K_{L_q}(j))) = 1.$$

Both P_1, P_2 are reversible with respect to the distribution π° defined by

$$\pi^\circ(\star) = \frac{\alpha}{2}, \quad \pi^\circ(q, \varepsilon, j) = \frac{\alpha \rho_q}{4}, \quad j \in \mathbb{Z}_{L_q}.$$

Indeed, on the anchor set one has $\pi^\circ = \alpha \nu$, so the refreshment step of P_1 satisfies detailed balance. At non-anchor states, both P_1 and P_2 are involutions within one cycle whose positions all have equal π° -mass, and thus they are also reversible.

Let $P = P_1 P_2$ (we use the convention that $P_1 P_2$ first applies P_1 and then P_2). At a non-anchor state, P moves from (q, ε, j) to $(q, \varepsilon, j + 1 \pmod{L_q})$ deterministically, while at an anchor P selects a new signed block according to ν and moves to its first position. So P has the same excursion dynamics as the basic construction, except that each signed block has its own anchor. We visualize its dynamics in Figure 2. The state $\star$ supplies a self-loop, so the chain is aperiodic; irreducibility and positive recurrence follow as before.

Finally, define the observable h° by letting it vanish at every anchor state:

$$h^\circ(\star) = h^\circ(q, \varepsilon, 0) = 0.$$

For non-anchor states, as in the basic construction, define

$$h^\circ(q, \varepsilon, j) = \varepsilon \frac{a_{q,j-1}}{\sqrt{\alpha \rho_q}}, \quad q \geq 1, \varepsilon \in \{-1, 1\}, 1 \leq j \leq \ell_q.$$

Then the expression for γ_k can be obtained by exactly the same calculations. In particular, sign symmetry gives $\pi^\circ(h^\circ) = 0$, and starting from any anchor, the next non-null cycle has a fresh symmetric sign, so all covariance across different cycles vanishes. $\square$

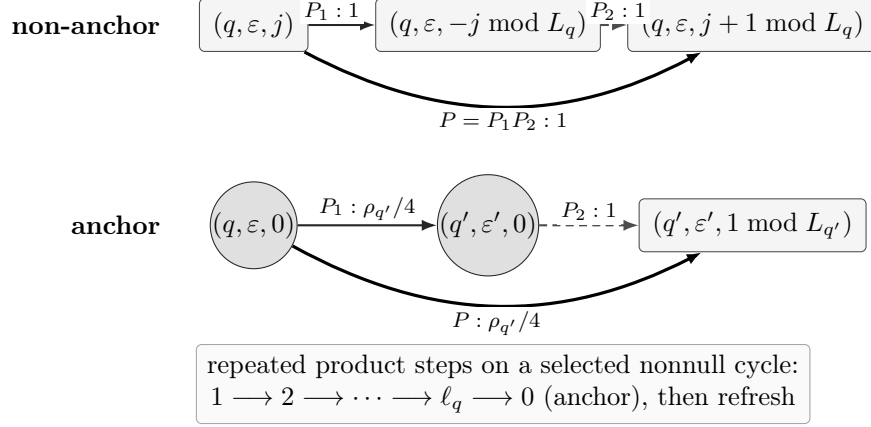


Figure 2: Dynamics of the chain in the factorized construction.

5.3 Proof of Proposition 5.3

Proof of Proposition 5.3. We first show that f can be presented as

$$f = \sum_{m=1}^{\infty} T_m, \quad (5.8)$$

where the convergence holds uniformly and every T_m is a nonnegative even trigonometric polynomial. For each $m \geq 1$, let

$$r_m = f - \sum_{k=1}^m T_k$$

denote the residual after the m -th approximation. Set $r_0 = f$.

We construct r_m inductively. Suppose that r_{m-1} is strictly positive, continuous, and even. By Fejér's theorem (see Lemma A.9), $\|\mathcal{M}_n r_{m-1} - r_{m-1}\|_{\infty} \rightarrow 0$, $\mathcal{M}_n r_{m-1}$ is nonnegative because the Fejér kernel is nonnegative, and it remains even [32, Chapter I]. Choose sufficiently large n_m such that

$$\|\mathcal{M}_{n_m} r_{m-1} - r_{m-1}\|_{\infty} < \frac{1}{2} \inf_{\theta \in [-\pi, \pi]} r_{m-1}(\theta).$$

Instead of subtracting all of this approximation, we subtract only half; that is, we let

$$T_m = \frac{1}{2} \mathcal{M}_{n_m} r_{m-1}, \quad r_m = r_{m-1} - T_m.$$

The approximation bound implies

$$\frac{1}{4} r_{m-1} \leq T_m \leq \frac{3}{4} r_{m-1}, \quad \frac{1}{4} r_{m-1} \leq r_m \leq \frac{3}{4} r_{m-1}.$$

Hence both the removed piece and the residual stay nonnegative, the residual remains strictly positive, and the residual decreases geometrically:

$$\|r_m\|_{\infty} \leq \left(\frac{3}{4}\right)^m \|f\|_{\infty}.$$

This proves that (5.8) holds uniformly.

By the Fejér–Riesz theorem [22] (see also Lemma A.10), for each m there is a polynomial

$$\Phi_m(z) = \sum_{j=0}^{d_m} c_{m,j} z^j \text{ such that } T_m(\theta) = |\Phi_m(e^{i\theta})|^2.$$

Write $c_{m,j} = x_{m,j} + iy_{m,j}$. For $k \geq 0$,

$$\widehat{T}_m(k) = \sum_{j=0}^{d_m-k} c_{m,j+k} \overline{c_{m,j}}.$$

Because T_m is even, this Fourier coefficient is real, and taking the real part gives

$$\widehat{T}_m(k) = \sum_{j=0}^{d_m-k} x_{m,j} x_{m,j+k} + \sum_{j=0}^{d_m-k} y_{m,j} y_{m,j+k}. \quad (5.9)$$

Thus each T_m contributes at most two finite real autocorrelation blocks, namely $(x_{m,j})_{j=0}^{d_m}$ and $(y_{m,j})_{j=0}^{d_m}$. Zero vectors may be discarded, and we index all remaining vectors by q . Recalling (5.3), we see that the two terms on the right-hand side of (5.9) are also Fourier coefficients of trigonometric polynomials, which gives the pointwise identity

$$T_m(\theta) = \left| \sum_{j=0}^{d_m} x_{m,j} e^{ij\theta} \right|^2 + \left| \sum_{j=0}^{d_m} y_{m,j} e^{ij\theta} \right|^2. \quad (5.10)$$

Combining (5.8) and (5.10) proves the uniform decomposition (5.4); taking Fourier coefficients gives the expression for $\widehat{f}(k)$. Finally, $\widehat{f}(0) = (2\pi)^{-1} \int_{-\pi}^{\pi} f(\theta) d\theta < \infty$. $\square$

5.4 Solution to HR Problem 1

Theorem 5.4 (Solution to HR Problem 1). *There exist a countable-state, irreducible, aperiodic, positive recurrent Markov transition kernel P with stationary distribution π and a function $h \in L^2(\pi)$ with $\pi(h) = 0$ such that*

$$A = \lim_{n \rightarrow \infty} A_n = 1, \quad \liminf_{n \rightarrow \infty} B_n = 1, \quad \limsup_{n \rightarrow \infty} B_n = \infty.$$

The construction can be modified so that $P = P_1 P_2$, where P_1 and P_2 are reversible Markov transition kernels. Moreover, h satisfies a $\sqrt{n}$ -CLT.

Proof of Theorem 5.4. Let f be a strictly positive, continuous, even, 2π -periodic function. By Theorem 5.1, we can construct a countable-space and ergodic P and a function h such that $\gamma_k = \widehat{f}(k)/2$. In particular, by Proposition 5.2, P can be expressed as $P = P_1 P_2$ with P_1, P_2 reversible with respect to π . Since f is even, $\widehat{f}(k) = \widehat{f}(-k)$, and by (2.3),

$$A_n = \frac{1}{2}(\mathcal{M}_{n-1}f)(0), \quad B_n = \frac{1}{2}(\mathcal{S}_n f)(0), \quad (5.11)$$

where $\mathcal{M}_n$ and $\mathcal{S}_n$ are defined by (5.1).

By Fejér's theorem, $(\mathcal{M}_N f)(0) \rightarrow f(0)$. Meanwhile, ordinary Fourier partial sums need not converge for continuous functions. Therefore, it only remains to find a function f such that

$$f(0) = 2, \quad \liminf_{N \rightarrow \infty} (\mathcal{S}_N f)(0) = 2, \quad \limsup_{N \rightarrow \infty} (\mathcal{S}_N f)(0) = \infty. \quad (5.12)$$

Here we simply use a classical example due to Fejér:

$$g(\theta) = \sum_{r=1}^{\infty} \frac{1}{r^2} \sin\left(\frac{2^{r^3} + 1}{2} |\theta|\right), \quad f(\theta) = 2 + g(\theta), \quad \theta \in [-\pi, \pi].$$

Note that f is strictly positive since

$$f(\theta) \geq 2 - \sum_{r=1}^{\infty} \frac{1}{r^2} = 2 - \frac{\pi^2}{6} > 0.$$

The well-known properties of the function g , which we review in Lemma A.11, imply that f satisfies (5.12). By (5.11), B_n does not converge, while $A_n \rightarrow 1$.

Finally, since $h \in L^2(\pi)$ and $A < \infty$, by part (iii) of Theorem 2.3, h satisfies a $\sqrt{n}$ -CLT. $\square$

Remark 5.3. For the factorized construction $P = P_1 P_2$ used in our proof, the transition kernels P_1, P_2 are reversible but are not conditional-expectation projections. Consequently, although Theorem 5.4 resolves HR Problem 1, it does not imply the same conclusion for two-component Gibbs samplers, which provided one of the motivations for this problem.

Acknowledgements

JSR was supported in part by NSERC of Canada discovery grant RGPIN-2019-04142. QZ was supported in part by NSF through grants DMS-2311307 and DMS-2245591.

A Appendix: Auxiliary results

Lemma A.1. *Let P be a Markov kernel on an arbitrary measurable space $(\mathcal{X}, \mathcal{F})$ and let $h : \mathcal{X} \rightarrow \mathbb{R}$ be measurable. There is a countably generated sigma-algebra $\mathcal{G} \subseteq \mathcal{F}$ such that*

1. h is $\mathcal{G}$ -measurable;
2. for every $E \in \mathcal{G}$, the function $x \mapsto P^n(x, E)$ is $\mathcal{G}$ -measurable for any integer $n \geq 1$.

In particular, the restricted Markov kernel $P_{\mathcal{G}}(x, E) := P(x, E)$, for $E \in \mathcal{G}$, is a Markov kernel on $(\mathcal{X}, \mathcal{G})$, and all finite-dimensional distributions of $(h(X_n))_{n \geq 0}$ for P coincide with using $P_{\mathcal{G}}$.

Proof. Exercise 1.17 of Revuz [46] shows that for any countable collection of sets in $\mathcal{F}$, there exists a sub-sigma-algebra $\mathcal{G}$ such that $\mathcal{G}$ is countably generated, and for every $E \in \mathcal{G}$, the function $x \mapsto P(x, E)$ is $\mathcal{G}$ -measurable. Hence, we can apply this result with the countable set

$$\mathcal{C} = \{\{x \in \mathcal{X} : h(x) < q\} : q \in \mathbb{Q}\},$$

where $\mathbb{Q}$ denotes the rationals. For completeness, here we explicitly show how $\mathcal{G}$ is constructed.

First, let $\mathcal{A}_0$ be the algebra generated by $\mathcal{C}$, which is still countable. Suppose $\mathcal{A}_r$ is a countable algebra contained in $\mathcal{F}$. Since P is a Markov kernel, $P(\cdot, E)$ is $\mathcal{F}$ -measurable for every $E \in \mathcal{A}_r$. Hence for each $E \in \mathcal{A}_r$ and $q \in \mathbb{Q}$, the set $\{x \in \mathcal{X} : P(x, E) > q\} \in \mathcal{F}$. Let $\mathcal{A}_{r+1}$ be the algebra generated by $\mathcal{A}_r$ and $\{\{x \in \mathcal{X} : P(x, E) > q\} : q \in \mathbb{Q}, E \in \mathcal{A}_r\}$. The algebra added at step r is

countable since both $\mathbb{Q}$ and $\mathcal{A}_r$ are countable. By induction, each $\mathcal{A}_r$ is a countable algebra, and the sequence $(\mathcal{A}_r)_{r \geq 0}$ is increasing. Define

$$\mathcal{A}_\infty = \bigcup_{r=0}^{\infty} \mathcal{A}_r,$$

and let $\mathcal{G} = \sigma(\mathcal{A}_\infty)$. Then $\mathcal{G}$ is countably generated, and h is $\mathcal{G}$ -measurable since $\mathcal{A}_0 \subseteq \mathcal{A}_\infty$.

Define $\mathcal{D} = \{E \in \mathcal{G} : P(\cdot, E) \text{ is } \mathcal{G}\text{-measurable}\}$. This is a λ -system by the properties of P . If $E \in \mathcal{A}_r$, then every rational strict superlevel set $\{x : P(x, E) > q\}$ is in $\mathcal{A}_{r+1} \subseteq \mathcal{G}$. Therefore, $P(\cdot, E)$ is $\mathcal{G}$ -measurable since a real-valued function is measurable whenever its rational superlevel sets are measurable. So $\mathcal{A}_\infty \subseteq \mathcal{D}$. Since $\mathcal{A}_\infty$ is a π -system, the π - λ theorem concludes $\mathcal{G} = \mathcal{D}$. This proves the second required condition with $n = 1$. An induction argument using

$$P_{\mathcal{G}}^{n+1}(x, E) = \int_{\mathcal{X}} P_{\mathcal{G}}^n(y, E) P_{\mathcal{G}}(x, dy)$$

extends this to every $n \geq 1$.

We have shown that $P_{\mathcal{G}}(x, E) = P(x, E)$. Assuming $P_{\mathcal{G}}^n(x, E) = P^n(x, E)$ and using the induction again, we find that $P_{\mathcal{G}}^{n+1}(x, E) = P^{n+1}(x, E)$. So the process $(h(X_n))_{n \geq 0}$ has the same finite-dimensional distributions under $P_{\mathcal{G}}$. $\square$

Lemma A.2. *Let $(U_n)_{n \geq 1}$ be real-valued random variables. For $0 < p < 1$, let $N_p \sim \text{Geom}(p)$ be independent of $(U_n)_{n \geq 1}$; in particular, $\mathbb{E}[N_p] = p^{-1}$.*

- (i) *If $(U_n)_{n \geq 1}$ is bounded in probability, then $(\sqrt{pN_p} U_{N_p})_{0 < p < 1}$ is also bounded in probability, and their characteristic functions are equicontinuous at zero.*
- (ii) *If $U_n \Rightarrow \mathcal{N}(0, \sigma^2)$ for some $0 \leq \sigma^2 < \infty$, then*

$$(pN_p, U_{N_p}) \Rightarrow (T, Z), \tag{A.1}$$

where $Z \sim \mathcal{N}(0, \sigma^2)$, $T \sim \text{Exp}(1)$, and T and Z are independent. Consequently,

$$\sqrt{pN_p} U_{N_p} \Longrightarrow \sqrt{T} Z,$$

and for every $t \in \mathbb{R}$,

$$\lim_{p \downarrow 0} \mathbb{E} \left[e^{it\sqrt{pN_p} U_{N_p}} \right] = \frac{1}{1 + \sigma^2 t^2 / 2}.$$

Proof. We first prove part (i). For every $\epsilon, R > 0$, we have

$$\begin{aligned} \mathbb{P}(|\sqrt{pN_p} U_{N_p}| > R) &\leq \mathbb{P}(|pN_p| > \epsilon^{-1}) + \mathbb{P}(|U_{N_p}| > R\sqrt{\epsilon}) \\ &\leq \epsilon + \sup_{n \geq 1} \mathbb{P}(|U_n| > R\sqrt{\epsilon}), \end{aligned}$$

since $\mathbb{E}[pN_p] = 1$. Since U_n is bounded in probability, for each fixed $\epsilon > 0$, we can choose R sufficiently large so that the supremum term is bounded by ϵ , and thus $\mathbb{P}(|\sqrt{pN_p} U_{N_p}| > R) \leq 2\epsilon$. This proves that $(\sqrt{pN_p} U_{N_p})_{0 < p < 1}$ is bounded in probability. The conclusion about characteristic functions follows from Lemma A.5.

Consider part (ii). Let f and g be bounded continuous functions. Independence between N_p and $(U_n)_{n \geq 1}$ gives

$$\mathbb{E} [f(U_{N_p}) g(pN_p)] = \sum_{n=1}^{\infty} \mathbb{P}(N_p = n) g(pn) \mathbb{E}[f(U_n)]. \tag{A.2}$$

Replacing $\mathbb{E}[f(U_n)]$ by $\mathbb{E}[f(Z)]$ in (A.2) introduces an error that can be bounded by

$$\sum_{n=1}^{\infty} \mathbb{P}(N_p = n) |g(pn)| \left| \mathbb{E}[f(U_n)] - \mathbb{E}[f(Z)] \right| \leq \|g\|_{\infty} \mathbb{E} \left| \mathbb{E}[f(U_{N_p}) | N_p] - \mathbb{E}[f(Z)] \right|. \quad (\text{A.3})$$

Since $N_p \rightarrow \infty$ in probability as $p \downarrow 0$ and $U_n \Rightarrow Z$, we have $\mathbb{E}[f(U_{N_p}) | N_p] \rightarrow \mathbb{E}[f(Z)]$ in probability. Further, $|\mathbb{E}[f(U_{N_p}) | N_p]| \leq \|f\|_{\infty}$, and thus the dominated convergence theorem implies that the right-hand side of (A.3) goes to zero as $p \downarrow 0$. Together with $pN_p \Rightarrow T$, this proves

$$\lim_{p \downarrow 0} \mathbb{E} [f(U_{N_p}) g(pN_p)] = \mathbb{E}[f(Z)] \lim_{p \downarrow 0} \mathbb{E}[g(pN_p)] = \mathbb{E}[f(Z)] \mathbb{E}[g(T)]. \quad (\text{A.4})$$

Taking $f(u) = e^{itu}$ and $g(v) = e^{isv}$ in (A.4) gives convergence of the joint characteristic functions to that of (T, Z) . Lévy's continuity theorem proves (A.1). The convergence of $\sqrt{pN_p} U_{N_p}$ follows from the continuous mapping theorem, which also implies the convergence of the characteristic function. The characteristic function limit is obtained by iterated expectation:

$$\mathbb{E} \left[e^{it\sqrt{pN_p} U_{N_p}} \right] \rightarrow \mathbb{E} \left[\mathbb{E}[e^{it\sqrt{T}Z} | T] \right] = \int_0^{\infty} e^{-xt^2\sigma^2/2} e^{-x} dx = \frac{1}{1 + t^2\sigma^2/2},$$

which concludes the proof. $\square$

Lemma A.3 (Billingsley, 1999, Theorem 18.3). *Let $(\xi_n)_{n \geq 0}$ be a stationary, ergodic martingale-difference sequence with respect to its natural filtration. If $s^2 := \mathbb{E}[\xi_0^2] < \infty$, then*

$$\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \xi_k \Longrightarrow \mathbf{N}(0, s^2).$$

Lemma A.4 (Bednorz et al., 2008, Theorem 3.4). *Let $(Z_k)_{k \geq 1}$ be independent and identically distributed real-valued random variables and $R_k = \sum_{i=1}^k Z_i$, with $R_0 = 0$. Let (J_n) be positive integer-valued random variables such that for some $\vartheta \in (0, \infty)$,*

$$\frac{J_n}{n} \rightarrow \vartheta, \text{ in probability.}$$

If $(R_{J_n}/\sqrt{n})_n$ is bounded in probability, then $\mathbb{E}(Z_1) = 0$ and $\mathbb{E}(Z_1^2) < \infty$.

Lemma A.5 (Kallenberg, 1997, Lemma 4.2). *A family of random variables $(Y_i)_{i \in \mathcal{I}}$ is bounded in probability if and only if their characteristic functions are equicontinuous at zero.*

Lemma A.6 (Kac occupation formula). *Let Q be an ergodic Markov kernel with invariant probability measure μ , and let E be measurable with $\mu(E) > 0$. Write*

$$\tau_E^+ = \inf\{n \geq 1 : X_n \in E\}.$$

Then, for every nonnegative measurable f ,

$$\mu(f) = \int_E \mu(dx) \mathbb{E}_x \left[\sum_{k=0}^{\tau_E^+ - 1} f(X_k) \right] = \mu(E) \mathbb{E}_{\mu(\cdot|E)} \left[\sum_{k=0}^{\tau_E^+ - 1} f(X_k) \right].$$

The identity extends to signed $f \in L^1(\mu)$ by applying it to f^+ and f^- .

Proof. The assumptions on Q imply that $\mathbb{P}_x(\tau_E^+ < \infty) = 1$ for μ -almost every x [17, Corollary 5.2.13]. The occupation identity is then exactly Douc et al. [17, Theorem 3.6.5]. $\square$

Lemma A.7 (Douc et al., 2018, Lemma 22.5.4). *Let Q be reversible with invariant probability μ . If $\|Q^n(x, \cdot) - \mu\|_{\text{TV}} \rightarrow 0$ for μ -almost every x , then, for every $h \in L^2(\mu)$ with $\mu(h) = 0$,*

$$\langle h, Q^n h \rangle_{L^2(\mu)} \rightarrow 0 \quad \text{and} \quad \mathcal{E}_h(\{-1, 1\}) = 0,$$

where $\mathcal{E}_h$ is the spectral measure of h .

Lemma A.8 (Glynn and Whitt, 2002, Theorem 2.1). *Let (X_n) be a positive-recurrent classically regenerative process. Let $T_0 < T_1 < \dots$ be regeneration epochs, suppose the noninitial cycles are i.i.d., and put $\mu_T = \mathbb{E}(T_2 - T_1) < \infty$. For a reward f define*

$$Z_i(\gamma) = \sum_{k=T_{i-1}}^{T_i-1} \{f(X_k) - \gamma\}.$$

Then

$$\frac{\sum_{k=0}^{n-1} f(X_k) - n\gamma}{\sqrt{n}} \Rightarrow \mathbf{N}(0, \sigma^2)$$

if and only if

$$\mathbb{E}Z_2(\gamma) = 0, \quad \mathbb{E}[Z_2(\gamma)^2] = \mu_T \sigma^2 < \infty.$$

Lemma A.9 (Fejér's theorem). *Let v be an integrable 2π -periodic function, and define*

$$(\mathcal{M}_N v)(x) = \sum_{|k| \leq N} \left(1 - \frac{|k|}{N+1}\right) \widehat{v}(k) e^{ikx}.$$

If v is continuous at x , then $(\mathcal{M}_N v)(x) \rightarrow v(x)$. If v is continuous everywhere, then the convergence is uniform. Moreover, if $v \geq 0$, then $\mathcal{M}_N v \geq 0$ for every N .

Lemma A.10 (Fejér–Riesz theorem). *Let $T: [-\pi, \pi] \rightarrow \mathbb{R}$ be given by*

$$T(\theta) = \sum_{k=-m}^m t_k e^{ik\theta}, \quad t_{-k} = \overline{t_k}.$$

If $T(\theta) > 0$ for every $\theta \in [-\pi, \pi]$, then there exists a polynomial

$$\varphi(z) = \sum_{j=0}^m c_j z^j, \quad c_j \in \mathbb{C},$$

such that $T(\theta) = |\varphi(e^{i\theta})|^2$ for $\theta \in [-\pi, \pi]$.

Lemma A.11 (Fejér's example). *Let $g: [-\pi, \pi] \rightarrow \mathbb{R}$ be defined by*

$$g(\theta) = \sum_{j=1}^{\infty} \frac{1}{j^2} \sin\left(\frac{2^{j^3} + 1}{2} |\theta|\right).$$

Define $\widehat{g}(k) = (2\pi)^{-1} \int_{-\pi}^{\pi} g(\theta) e^{-ik\theta} d\theta$, and

$$(\mathcal{S}_N g)(0) = \sum_{|k| \leq N} \widehat{g}(k), \quad (\mathcal{M}_N g)(0) = \sum_{|k| \leq N} \left(1 - \frac{|k|}{N+1}\right) \widehat{g}(k).$$

Then, g is continuous with $g(0) = 0$, $(\mathcal{S}_N g)(0) \geq 0$ for every $N \geq 0$, and

$$\limsup_{N \rightarrow \infty} (\mathcal{S}_N g)(0) = \infty, \quad \liminf_{N \rightarrow \infty} (\mathcal{S}_N g)(0) = 0, \quad \lim_{N \rightarrow \infty} (\mathcal{M}_N g)(0) = 0.$$

Proof. This is Fejér's classical example of a continuous function whose Fourier series diverges at 0 [20, Chapter 8.8]. The series defining g converges uniformly since

$$|g(\theta)| \leq \sum_{j=1}^{\infty} \frac{1}{j^2} = \frac{\pi^2}{6}.$$

It is shown in Duren [20, Chapter 8.8] that $\limsup_{N \rightarrow \infty} (\mathcal{S}_N g)(0) = \infty$ and

$$(\mathcal{S}_N g)(0) \geq 0, \text{ for every } N \geq 1.$$

On the other hand, Fejér's theorem and $g(0) = 0$ imply $(\mathcal{M}_N g)(0) \rightarrow 0$. Together with the nonnegativity of $(\mathcal{S}_N g)(0)$, this forces $\liminf_{N \rightarrow \infty} (\mathcal{S}_N g)(0) = 0$. $\square$

B Appendix: Proof of Theorem 3.10

B.1 Main Proof

The proof of Theorem 3.10 consists of three steps. The first step is to reduce the problem to the analysis of $\sum_k g(Y_k)$. In Proposition B.1, we show that tightness of $(S_n/\sqrt{n})_{n \geq 1}$ transfers to $(T_n/\sqrt{n})_{n \geq 1}$, and in the converse direction, to prove $\sigma^2 < \infty$, it suffices to control $\sup_n n^{-1} \text{Var}(T_n)$. Note that Proposition B.1 is independent of the minorization condition and holds for any integer $m \geq 1$.

Proposition B.1. *If $(S_n/\sqrt{n})_{n \geq 1}$ is bounded in probability, then $(T_n/\sqrt{n})_{n \geq 1}$ is also bounded in probability. Moreover, if*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \text{Var}(T_n) \leq M < \infty,$$

then $\sigma^2 = C \leq M < \infty$.

The second step is to show that $\mathbb{E}[Z_1^2] < \infty$. Our proof uses Bednorz et al. [2, Theorem 3.4].

Proposition B.2. *If $(T_n/\sqrt{n})_{n \geq 1}$ is bounded in probability, then $\mathbb{E}(Z_1^2) < \infty$.*

To convert $\mathbb{E}[Z_1^2] < \infty$ into a linear variance bound for T_n , we need to convert the cycle clock to one-step clock and handle the last incomplete cycle at a deterministic time n . In particular, the sum accumulated in the incomplete cycle is not necessarily square-integrable. The forward-backward martingale decomposition developed by Wu [54] enables us to overcome this difficulty. Let L be the residual kernel defined by (3.22), and let $\tau = \tau_1$ denote the first regeneration time. Define

$$u(x) = \sum_{j=0}^{\infty} (L^j g)(x) = \mathbb{E}_x \left[\sum_{j=0}^{\tau-1} g(Y_j) \right].$$

The standard theory of Nummelin [41] shows that u is well-defined π -almost everywhere and satisfies the Poisson equations. For completeness, we give a proof of this fact in the following lemma. An important point is that u is not necessarily in $L^2(\pi)$.

Lemma B.3. *The function u is well-defined and finite π -almost everywhere. Further,*

$$g = u - Lu = u - P^m u, \quad \pi\text{-almost everywhere,}$$

and $\pi_D(u) = 0$.

The forward and backward martingale differences are defined by

$$\xi_j = u(Y_j) - P^m u(Y_{j-1}), \quad \xi_j^* = u(Y_{j-1}) - P^m u(Y_j).$$

The final key result we need is the identity $\mathbb{E}[\xi_1^2] = \beta\pi(D)\mathbb{E}[Z_1^2]$, which says that the stationary quadratic variation per step equals the cycle-sum variance times the regeneration rate $\beta\pi(D)$. This identity is well known; see, for example, Meyn and Tweedie [40, Equations (17.13) and (17.44)]. However, usual formulations require $u \in L^2(\pi)$ or $\mathbb{E}[\tau^2] < \infty$, neither of which necessarily holds in our setting. Hence, we formally establish it in Proposition B.4, which only requires $\mathbb{E}[Z_1^2] < \infty$.

Proposition B.4. *If $\mathbb{E}[Z_1^2] < \infty$, then*

$$\mathbb{E}[\xi_1^2] = \beta\pi(D)\mathbb{E}[Z_1^2].$$

Proof of Theorem 3.10. The sequence $(\xi_j)_{j \geq 1}$ is a martingale-difference sequence since $(Y_k)_{k \geq 0}$ is a stationary Markov chain with transition kernel P^m and $\pi(g) = 0$. By reversibility and the stationarity of (Y_k) , $(\xi_j^*)_{j \geq 1}$ has the same properties with respect to the reverse filtration. The Poisson equation given in Lemma B.3 therefore gives, almost surely,

$$g(Y_{j-1}) + g(Y_j) = u(Y_{j-1}) - P^m u(Y_{j-1}) + u(Y_j) - P^m u(Y_j) = \xi_j^* + \xi_j.$$

This yields the decomposition of the sum

$$T_n = \sum_{j=0}^{n-1} g(Y_j) = \frac{1}{2} \sum_{j=1}^n (\xi_j + \xi_j^*) + \frac{1}{2} \{g(Y_0) - g(Y_n)\}. \quad (\text{B.1})$$

Compared to an ordinary forward Poisson decomposition

$$T_n = \sum_{j=1}^n \xi_j + u(Y_0) - u(Y_n), \quad (\text{B.2})$$

the endpoint terms in (B.1) only involve g instead of u .

Since ξ_j, ξ_j^* are martingale-difference sequences, within each sequence the differences are $L^2(\pi)$ -orthogonal to each other, and thus we obtain from (B.1) that

$$\begin{aligned} 2\sqrt{\text{Var}(T_n)} &\leq \left\| \sum_{j=1}^n \xi_j \right\|_{L^2(\pi)} + \left\| \sum_{j=1}^n \xi_j^* \right\|_{L^2(\pi)} + \|g(Y_0) - g(Y_n)\|_{L^2(\pi)} \\ &= \sqrt{n\mathbb{E}[\xi_j^2]} + \sqrt{n\mathbb{E}[(\xi_j^*)^2]} + \|g(Y_0) - g(Y_n)\|_{L^2(\pi)} \\ &\leq \sqrt{n\mathbb{E}[\xi_j^2]} + \sqrt{n\mathbb{E}[(\xi_j^*)^2]} + 2\sqrt{\pi(g^2)} \end{aligned}$$

By reversibility and since the process is stationary,

$$\mathbb{E}[\xi_j^2] = \mathbb{E}[(\xi_j^*)^2] = \int [u(y) - P^m u(x)]^2 P^m(x, dy) \pi(dx).$$

Propositions B.2 and B.4 yield $\mathbb{E}[\xi_1^2] = \beta\pi(\mathbf{D})\mathbb{E}[Z_1^2] < \infty$. Hence, using $g \in L^2(\pi)$, we get

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \text{Var}(T_n) \leq \limsup_{n \rightarrow \infty} \left(\sqrt{\mathbb{E}[\xi_1^2]} + \frac{\sqrt{\pi(g^2)}}{\sqrt{n}} \right)^2 \leq \beta\pi(\mathbf{D})\mathbb{E}[Z_1^2] < \infty.$$

By Proposition B.1, this implies $\sigma^2 = C \leq \beta\pi(\mathbf{D})\mathbb{E}[Z_1^2] < \infty$. Theorem 2.3 further yields that $\sigma^2 = A = B = C$ and $S_n/\sqrt{n} \Rightarrow \mathbf{N}(0, \sigma^2)$.

To conclude the proof, it remains to show that $\sigma_g^2 = \mathbb{E}[\xi_1^2]$. Since $\mathbb{E}[\xi_1^2] < \infty$, we can apply Lemma A.3 to obtain that $n^{-1/2} \sum_{j=1}^n \xi_j$ converges weakly to the Gaussian limit with variance $\mathbb{E}[\xi_1^2]$. By Lemma B.3, u is finite π -almost everywhere, and thus $n^{-1/2}[u(Y_0) - u(Y_n)] \rightarrow 0$ in probability. So the forward decomposition (B.2) shows that $T_n/\sqrt{n} \Rightarrow \mathbf{N}(0, \sigma_g^2)$ with $\sigma_g^2 = \mathbb{E}[\xi_1^2]$. $\square$

B.2 Proof of Proposition B.1

Proof. Define a filtration of the stationary Markov process by $\mathcal{H}_k = \sigma\{X_0, \dots, X_{km}\}$. Define

$$W_k = \sum_{j=0}^{m-1} h(X_{km+j})$$

so that

$$\sum_{k=0}^{n-1} W_k = \sum_{k=0}^{n-1} \sum_{j=0}^{m-1} h(X_{km+j}) = \sum_{i=0}^{nm-1} h(X_i) = S_{nm}.$$

By assumption and since m is constant, $n^{-1/2} \sum_{k=0}^{n-1} W_k$ is bounded in probability. Define

$$D_k = W_k - g(Y_k), \quad M_n = \sum_{k=0}^{n-1} D_k.$$

Then D_k is $\mathcal{H}_{k+1}$ -measurable and $\mathbb{E}[D_k \mid \mathcal{H}_k] = 0$, so $(M_n)_{n \geq 0}$ is a martingale with respect to $(\mathcal{H}_n)_{n \geq 0}$. This gives a decomposition

$$\frac{1}{\sqrt{n}} T_n = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} g(Y_k) = \frac{1}{\sqrt{n}} S_{mn} - \frac{1}{\sqrt{n}} M_n.$$

Since $(S_n/\sqrt{n})_{n \geq 1}$ is bounded in probability, it remains to show that $(M_n/\sqrt{n})_{n \geq 1}$ is bounded in probability. Since $h, g \in L^2(\pi)$, D_k is also square-integrable with $\mathbb{E}[D_j D_k] = 0$ for $j \neq k$. Using the stationarity, we get $n^{-1} \mathbb{E}(M_n^2) = \mathbb{E}(D_0^2) < \infty$. Chebyshev's inequality then yields the tightness of $(M_n/\sqrt{n})_{n \geq 1}$ and thus $(T_n/\sqrt{n})_{n \geq 1}$.

Now assume that $\limsup_{n \rightarrow \infty} n^{-1} \text{Var}(T_n) \leq M < \infty$. Let $\mathcal{E}_g^m$ denote the spectral measure of g with respect to P^m . By Rosenthal [50, Proposition 1],

$$\frac{1}{n} \text{Var}(T_n) \rightarrow \int_{[-1,1]} \frac{1+\lambda}{1-\lambda} \mathcal{E}_g^m(d\lambda),$$

with the convention $2/0 = \infty$. Hence, the limit on the right-hand side is bounded by M . Define $G(\lambda) = (1+\lambda)/(1-\lambda)$, and for $N \geq 0$, define $G_N = G \wedge N$. Define $\varphi_m(\lambda) = \sum_{j=0}^{m-1} \lambda^j$. Since G_N is bounded, we have by Conway [9, Theorem 2.3] that

$$\int_{\mathcal{X}} g(x) (G_N(P^m)g)(x) \pi(dx) = \int G_N(\lambda^m) \varphi_m(\lambda)^2 \mathcal{E}_h(d\lambda).$$

Here we used that φ_m is bounded on $[-1, 1]$. Passing to the limit with monotone convergence gives

$$\int_{[-1,1]} \frac{1+\lambda}{1-\lambda} \mathcal{E}_g^m(d\lambda) = \int_{[-1,1]} \frac{1+\lambda^m}{1-\lambda^m} \varphi_m(\lambda)^2 \mathcal{E}_h(d\lambda).$$

Now we have the lower bound for $-1 < \lambda < 1$ using $\varphi_m(\lambda) = \sum_{j=0}^{m-1} \lambda^j = (1 - \lambda^m)/(1 - \lambda)$ with

$$\begin{aligned} \frac{1+\lambda^m}{1-\lambda^m} \varphi_m(\lambda)^2 &= \frac{1+\lambda^m}{1-\lambda^m} \left(\frac{1-\lambda^m}{1-\lambda} \right)^2 \\ &= \frac{1+\lambda}{1-\lambda} \frac{1-\lambda^{2m}}{1-\lambda^2} \\ &= \frac{1+\lambda}{1-\lambda} \sum_{i=0}^{m-1} \lambda^{2i} \\ &\geq \frac{1+\lambda}{1-\lambda}. \end{aligned}$$

By Lemma A.7, $\mathcal{E}_h(\{1, -1\}) = 0$. Therefore, integration over the above lower bound with respect to $\mathcal{E}_h$ yields

$$M \geq \int_{[-1,1]} \frac{1+\lambda}{1-\lambda} \mathcal{E}_g^m(d\lambda) \geq \int_{[-1,1]} \frac{1+\lambda}{1-\lambda} \mathcal{E}_h(d\lambda) = C,$$

which completes the proof. $\square$

B.3 Proof of Proposition B.2

Proof. Let $\rho = \beta\pi(\mathbf{D}) > 0$. Let ω_n denote the number of regenerations by time n and $\tilde{\omega}_n$ the number of complete cycles by time n :

$$\omega_n = \sum_{k=0}^{n-1} I_k, \quad \tilde{\omega}_n = (\omega_n - 1)_+.$$

Since $\mathbb{E}[I_k | Y_k] = \beta \mathbf{1}_{\mathbf{D}}(Y_k)$, by the martingale strong law of large numbers and the ergodic theorem,

$$\frac{\omega_n}{n} \longrightarrow \beta\pi(\mathbf{D}) = \rho, \text{ almost surely.}$$

It also follows that $\tilde{\omega}_n/n \rightarrow \rho$ almost surely.

Decomposing the sum, for all sufficiently large n with $\tilde{\omega}_n \geq 1$,

$$T_n = \sum_{k=0}^{n-1} g(Y_k) = R^- + \sum_{i=1}^{\tilde{\omega}_n} Z_i + R_n^+,$$

where

$$R^- = \sum_{k=0}^{\tau_1-1} g(Y_k), \quad R_n^+ = \sum_{k=\tau_{\tilde{\omega}_n}+1}^{n-1} g(Y_k).$$

Since $\omega_n/n \rightarrow \rho$ almost surely, $\mathbb{P}(\tau_1 < \infty) = 1$. This implies $R^-/\sqrt{n} \rightarrow 0$ almost surely.

Define the most recent successful regeneration time not exceeding n by

$$\hat{\tau}_n = \max(\{0\} \cup \{k+1 \leq n : I_k = 1\}).$$

Define $E_n = n - \hat{\tau}_n$ so that E_n is the number of times after the most recent successful regeneration and before time n . If there have been no regenerations by time n , then $\hat{\tau}_n = 0$ and $E_n = n$. On the set $\{\tau_1 \leq n\}$, we have $\tau_{\tilde{\omega}_n+1} = \hat{\tau}_n$ and

$$R_n^+ = \sum_{k=\hat{\tau}_n}^{n-1} g(Y_k) = \sum_{j=1}^{E_n} g(Y_{n-j}).$$

Then for any $\epsilon > 0$, a union probability bound gives, for each fixed $M < \infty$,

$$\mathbb{P}\left(\frac{1}{\sqrt{n}}|R_n^+| > \epsilon\right) \leq \mathbb{P}(\tau_1 > n) + \mathbb{P}(E_n > M) + \mathbb{P}\left(\frac{1}{\sqrt{n}} \sum_{j=1}^M |g(Y_{n-j})| > \epsilon\right).$$

Taking the limit with respect to $n \rightarrow \infty$ gives

$$\limsup_{n \rightarrow \infty} \mathbb{P}\left(\frac{1}{\sqrt{n}}|R_n^+| > \epsilon\right) \leq \limsup_{n \rightarrow \infty} \mathbb{P}(E_n > M).$$

Now for $n \geq M$, the stationarity of the split chain implies

$$\begin{aligned} \mathbb{P}(E_n > M) &\leq \mathbb{P}(I_{n-M} = I_{n-M+1} = \dots = I_{n-1} = 0) \\ &= \mathbb{P}(I_0 = I_1 = \dots = I_{M-1} = 0), \end{aligned}$$

where the right-hand side goes to zero as $M \rightarrow \infty$ since $\omega_n/n \rightarrow \rho$ almost surely. Hence, $|R_n^+|/\sqrt{n} \rightarrow 0$ in probability. Moreover,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{\tilde{\omega}_n} Z_i = \frac{1}{\sqrt{n}} [T_n - R^- - R_n^+] \quad \text{on } \{\tau_1 \leq n\}.$$

Since $\mathbb{P}(\tau_1 > n) \rightarrow 0$, the left-hand side is bounded in probability.

To obtain a strictly positive clock, replace $\tilde{\omega}_n$ by $\tilde{\omega}_n \vee 1$; then $(\tilde{\omega}_n \vee 1)/n \rightarrow \rho$ almost surely. Since $\tilde{\omega}_n/n \rightarrow \rho$ in probability, $\mathbb{P}(\tilde{\omega}_n = 0) \rightarrow 0$. It follows that $n^{-1/2} \sum_{i=1}^{\tilde{\omega}_n \vee 1} Z_i$ is bounded in probability. Applying Lemma A.4 completes the proof. $\square$

Remark B.1. A simpler proof of Proposition B.2 is to first use Chen [8, Theorem II.3.1] to show that $T_n/\sqrt{n} \Rightarrow \mathcal{N}(0, \sigma^2)$ and then apply the result of Glynn and Whitt [25] (see Lemma A.8).

B.4 Proof of Lemma B.3

Proof. Define

$$R_\tau(g) = \sum_{j=0}^{\tau-1} g(Y_j). \tag{B.3}$$

We first verify that u is well-defined π -almost everywhere. Define

$$H(x) = \sum_{j=0}^{\infty} (L^j |g|)(x) = \mathbb{E}_x \left[\sum_{j=0}^{\tau-1} |g|(Y_j) \right], \quad \mathcal{N} = \{x \in \mathcal{X} : H(x) = \infty\}.$$

When the split chain is initialized with $Y_0 \sim \pi_{\mathbf{D}}$, the path $(Y_j)_{0 \leq j < \tau}$ has the law of a complete regeneration cycle. The stationary regeneration rate of the split chain is $\rho := \beta\pi(\mathbf{D})$.

Hence, applying Kac's occupation formula (see Lemma A.6) to the split chain yields

$$\pi(|g|) = \rho \mathbb{E}_{\pi_{\mathbf{D}}} \left[\sum_{j=0}^{\tau-1} |g|(Y_j) \right] = \rho \mathbb{E}_{\pi_{\mathbf{D}}} [H(Y_0)] = \rho \pi_{\mathbf{D}}(H). \quad (\text{B.4})$$

Since $g \in L^2(\pi)$, $\pi_{\mathbf{D}}(H) = \pi(|g|)/\rho < \infty$ and thus $\pi_{\mathbf{D}}(\mathcal{N}) = 0$. Since $H = |g| + LH$, we have $LH \leq H$, and by induction, $L^j H \leq H$ for each $j \geq 1$. Therefore, $\pi_{\mathbf{D}}(L^j H) \leq \pi_{\mathbf{D}}(H) < \infty$, and thus $(\pi_{\mathbf{D}} L^j)(\mathcal{N}) = 0$. Combining these results and applying Kac's occupation formula to $\mathbf{1}_{\mathcal{N}}$ yields

$$\pi(\mathcal{N}) = \pi(\mathbf{1}_{\mathcal{N}}) = \rho \mathbb{E}_{\pi_{\mathbf{D}}} \left[\sum_{j=0}^{\tau-1} \mathbf{1}_{\mathcal{N}}(Y_j) \right] = \rho \sum_{j=0}^{\infty} (\pi_{\mathbf{D}} L^j)(\mathcal{N}) = 0.$$

Therefore, the series defining u converges absolutely and u is well-defined and finite π -almost everywhere. Moreover, $u(x) = \mathbb{E}_x[R_{\tau}(g)]$ wherever the series converges absolutely. Since $u = g + Lu$, this yields the first Poisson equation $g = u - Lu$.

Using $\pi(g) = m\pi(h) = 0$ and replacing $|g|$ with g in Kac's occupation identity (B.4) yields

$$0 = \pi(g) = \rho \pi_{\mathbf{D}}(u).$$

Therefore, $\pi_{\mathbf{D}}(u) = 0$. Since

$$P^m(x, \cdot) = L(x, \cdot) + \beta \mathbf{1}_{\mathbf{D}}(x) \pi_{\mathbf{D}}(\cdot),$$

we get $P^m u = Lu$, π -almost everywhere. This proves the second claimed Poisson equation. $\square$

B.5 Proof of Proposition B.4

Proof. Write $\rho = \beta\pi(\mathbf{D})$. Let R_{τ} be defined by (B.3). By the definition of u and Jensen's inequality,

$$\pi_{\mathbf{D}}(u^2) = \int \mathbb{E}_x[R_{\tau}(g)]^2 \pi_{\mathbf{D}}(dx) \leq \mathbb{E}_{\pi_{\mathbf{D}}}[R_{\tau}(g)^2] = \mathbb{E}[Z_1^2] < \infty.$$

Here we used that, under $Y_0 \sim \pi_{\mathbf{D}}$, $R_{\tau}(g)$ has the same distribution as Z_1 . Since $P^m u = u - g$ and $g \in L^2(\pi_{\mathbf{D}})$, we also have $P^m u \in L^2(\pi_{\mathbf{D}})$. We have the decomposition

$$\mathbb{E}[\xi_1^2] = \int_{\mathcal{X} \times \mathcal{X}} [u(y) - P^m u(x)]^2 L(x, dy) \pi(dx) + \rho \int_{\mathcal{X}} \int_{\mathcal{X}} [u(y) - P^m u(x)]^2 \pi_{\mathbf{D}}(dy) \pi_{\mathbf{D}}(dx).$$

By Lemma B.3, $\pi_{\mathbf{D}}(u) = 0$, and thus

$$\int_{\mathcal{X}} \int_{\mathcal{X}} [u(y) - P^m u(x)]^2 \pi_{\mathbf{D}}(dy) \pi_{\mathbf{D}}(dx) = \pi_{\mathbf{D}}(u^2) + \pi_{\mathbf{D}}((P^m u)^2).$$

So we can express $\mathbb{E}[\xi_1^2]$ as

$$\mathbb{E}[\xi_1^2] = \pi(a) + \rho \pi_{\mathbf{D}}(u^2), \quad (\text{B.5})$$

where

$$a(x) = \int_{\mathcal{X}} [u(y) - P^m u(x)]^2 L(x, dy) + \beta \mathbf{1}_{\mathbf{D}}(x) (P^m u)(x)^2. \quad (\text{B.6})$$

To characterize $\mathbb{E}[Z_1^2]$, we use another martingale argument. For this martingale argument, initialize the split chain with $Y_0 \sim \pi_{\mathbb{D}}$ and define the filtration $\mathcal{G}_n = \sigma(Y_0, I_0, Y_1, \dots, I_{n-1}, Y_n)$. Define the martingale with respect to the filtration $\mathcal{G}_n$ and its difference by

$$M_n = \mathbb{E}_{\pi_{\mathbb{D}}}[R_{\tau}(g) \mid \mathcal{G}_n], \quad \Delta_{n+1} = M_{n+1} - M_n.$$

By Jensen's inequality,

$$\mathbb{E}_{\pi_{\mathbb{D}}}[M_n^2] \leq \mathbb{E}_{\pi_{\mathbb{D}}}[R_{\tau}(g)^2] = \mathbb{E}[Z_1^2] < \infty.$$

Therefore, M_n is an L^2 martingale, and by Lévy's upward convergence theorem,

$$M_n \xrightarrow{L^2} \mathbb{E}_{\pi_{\mathbb{D}}}[R_{\tau}(g) \mid \mathcal{G}_{\infty}] = R_{\tau}(g),$$

where the equality holds since $\tau < \infty$ almost surely. The convergence happens in L^2 so that

$$\lim_{n \rightarrow \infty} \mathbb{E}_{\pi_{\mathbb{D}}}[M_n^2] = \mathbb{E}_{\pi_{\mathbb{D}}}[R_{\tau}(g)^2] = \mathbb{E}[Z_1^2].$$

It remains to find an expression for $\lim_{n \rightarrow \infty} \mathbb{E}_{\pi_{\mathbb{D}}}[M_n^2]$. By construction, we can express M_n as

$$M_n = \sum_{k=0}^{(n \wedge \tau)-1} g(Y_k) + u(Y_n) \mathbf{1}_{\{n < \tau\}},$$

and in particular $M_0 = u(Y_0)$. It is also clear that $\Delta_{n+1} = 0$ if $n \geq \tau$. Hence, writing $M_n = M_0 + \sum_{j=0}^{n-1} \Delta_{j+1}$ and using the orthogonality of L^2 -martingale differences, we get

$$\mathbb{E}_{\pi_{\mathbb{D}}}[M_n^2] = \mathbb{E}_{\pi_{\mathbb{D}}}[M_0^2] + \mathbb{E}_{\pi_{\mathbb{D}}} \left[\sum_{j=0}^{n-1} \Delta_{j+1}^2 \right] = \pi_{\mathbb{D}}(u^2) + \mathbb{E}_{\pi_{\mathbb{D}}} \left[\sum_{j=0}^{(n \wedge \tau)-1} \Delta_{j+1}^2 \right]. \quad (\text{B.7})$$

Since $Y_0 \sim \pi_{\mathbb{D}}$, we can use the Poisson equation $g = u - P^m u$. If $j < \tau$ and $I_j = 0$, then

$$\Delta_{j+1} = g(Y_j) + u(Y_{j+1}) - u(Y_j) = u(Y_{j+1}) - P^m u(Y_j).$$

If $j < \tau$ and $I_j = 1$, then $\Delta_{j+1} = g(Y_j) - u(Y_j) = -P^m u(Y_j)$. So

$$\mathbb{E}_{\pi_{\mathbb{D}}}(\Delta_{j+1}^2 \mid \mathcal{G}_j) = \mathbf{1}_{\{j < \tau\}} a(Y_j),$$

where a is defined in (B.6). Since a is nonnegative, by Lemma A.6,

$$\pi(a) = \rho \mathbb{E}_{\pi_{\mathbb{D}}} \left[\sum_{j=0}^{\tau-1} a(Y_j) \right] = \rho \mathbb{E}_{\pi_{\mathbb{D}}} \left[\sum_{j=0}^{\tau-1} \mathbb{E}_{\pi_{\mathbb{D}}}(\Delta_{j+1}^2 \mid \mathcal{G}_j) \right] = \rho \mathbb{E}_{\pi_{\mathbb{D}}} \left[\sum_{j=0}^{\tau-1} \Delta_{j+1}^2 \right].$$

Then, by monotone convergence and (B.7),

$$\rho^{-1} \pi(a) = \mathbb{E}_{\pi_{\mathbb{D}}} \left[\sum_{j=0}^{\tau-1} \Delta_{j+1}^2 \right] = \lim_{n \rightarrow \infty} \mathbb{E}_{\pi_{\mathbb{D}}}[M_n^2] - \pi_{\mathbb{D}}(u^2) = \mathbb{E}[Z_1^2] - \pi_{\mathbb{D}}(u^2).$$

Comparing with (B.5) completes the proof. $\square$

References

- [1] Jarod Alper, Michael Barany, Alain Chavarri Villarelo, Sander Dahmen, Walter Dean, Karthik Ganapathy, Michael Harris, David Holmes, Mateja Jamnik, Steven Kelk, Bryna Kra, Ursula Martin, Bartosz Naskręcki, Rodrigo Ochigame, Jim Portegies, and Johannes Schmitt. Leiden declaration on artificial intelligence and mathematics, June 2026. URL <https://leidendeclaration.ai/>.
- [2] Witold Bednorz, Krzysztof Łatuszyński, and Rafał Łatała. A regeneration proof of the central limit theorem for uniformly ergodic Markov chains. *Electronic Communications in Probability*, 13:85–98, 2008. doi: 10.1214/ECP.v13-1354.
- [3] Patrick Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, 2nd edition, 1999. ISBN 978-0-471-19745-4. doi: 10.1002/9780470316962.
- [4] Richard C. Bradley. On the central limit question for strictly stationary, reversible Markov chains, 2025. URL <https://arxiv.org/abs/2508.12951>.
- [5] Miha Brešar, Aleksandar Mijatović, and Gareth O. Roberts. Central limit theorem for ergodic averages of Markov chains and the comparison of sampling algorithms for heavy-tailed distributions, 2025. URL <https://arxiv.org/abs/2512.18255>.
- [6] Sébastien Bubeck, Christian Coester, Ronen Eldan, Timothy Gowers, Yin Tat Lee, Alexandru Lupasca, Mehtaab Sawhney, Robert Scherrer, Mark Sellke, Brian K. Spears, Derya Unutmaz, Kevin Weil, Steven Yin, and Nikita Zhivotovskiy. Early science acceleration experiments with GPT-5, 2025. URL <https://arxiv.org/abs/2511.16072>.
- [7] Liwei Chen and Yuan Yuan. Hardy spaces and Szegő projection on quotient domains, 2023. URL <https://arxiv.org/abs/2310.12151>.
- [8] Xia Chen. *Limit Theorems for Functionals of Ergodic Markov Chains with General State Space*, volume 139, no. 664 of *Memoirs of the American Mathematical Society*. American Mathematical Society, Providence, RI, 1999. ISBN 0-8218-1060-X. doi: 10.1090/memo/0664.
- [9] John B. Conway. *A Course in Functional Analysis*, volume 96 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2nd edition, 1990. doi: 10.1007/978-1-4757-4383-8.
- [10] Christophe Cuny and Michael Lin. Global central limit theorems for Markov chains. *Stochastic Processes and their Applications*, 190:104709, 2025. doi: 10.1016/j.spa.2025.104709.
- [11] Christophe Cuny and Magda Peligrad. Central limit theorem started at a point for stationary processes and additive functionals of reversible Markov chains. *Journal of Theoretical Probability*, 25(1):171–188, 2012. doi: 10.1007/s10959-010-0321-8.
- [12] George Deligiannidis and Anthony Lee. Which ergodic averages have finite asymptotic variance? *The Annals of Applied Probability*, 28(4):2309–2334, 2018. doi: 10.1214/17-AAP1358.
- [13] George Deligiannidis, Magda Peligrad, and Sergey Utev. Asymptotic variance of stationary reversible and normal Markov processes. *Electronic Journal of Probability*, 20:1–26, 2015. doi: 10.1214/EJP.v20-3183.

- [14] Yves Derriennic and Michael Lin. The central limit theorem for Markov chains with normal transition operators, started at a point. *Probability Theory and Related Fields*, 119(4):508–528, 2001. doi: 10.1007/PL00008769.
- [15] Yves Derriennic and Michael Lin. Variance bounding Markov chains, L^2 -uniform mean ergodicity and the CLT. *Stochastics and Dynamics*, 11(1):81–94, 2011. doi: 10.1142/S0219493711003176.
- [16] Randal Douc and Christian P. Robert. A vanilla Rao–Blackwellization of Metropolis–Hastings algorithms. *The Annals of Statistics*, 39(1):261–277, 2011. doi: 10.1214/10-AOS838.
- [17] Randal Douc, Éric Moulines, Pierre Priouret, and Philippe Soulier. *Markov Chains*. Springer Series in Operations Research and Financial Engineering. Springer, Cham, 2018. ISBN 978-3-319-97704-1. doi: 10.1007/978-3-319-97704-1.
- [18] Arnaud Doucet, Michael K. Pitt, George Deligiannidis, and Robert Kohn. Efficient implementation of Markov chain Monte Carlo when using an unbiased likelihood estimator. *Biometrika*, 102(2):295–313, 2015. doi: 10.1093/biomet/asu075.
- [19] Paul Doukhan, Pascal Massart, and Emmanuel Rio. The functional central limit theorem for strongly mixing processes. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 30(1):63–82, 1994.
- [20] Peter L. Duren. *Invitation to Classical Analysis*, volume 17 of *Pure and Applied Undergraduate Texts*. American Mathematical Society, Providence, RI, 2012. ISBN 9780821869321.
- [21] James M. Flegal and Galin L. Jones. Batch means and spectral variance estimators in Markov chain Monte Carlo. *The Annals of Statistics*, 38(2):1034–1070, 2010. doi: 10.1214/09-AOS735.
- [22] Tryphon T. Georgiou and Anders Lindquist. On a Fejér–Riesz factorization of generalized trigonometric polynomials. *Communications in Information and Systems*, 21(3):371–384, 2021. doi: 10.4310/CIS.2021.v21.n3.a3.
- [23] Charles J. Geyer. Practical Markov chain Monte Carlo. *Statistical Science*, 7(4):473–483, 1992. doi: 10.1214/ss/1177011137.
- [24] Peter W. Glynn. The covariance function of a regenerative process. Technical report, Department of Engineering-Economic Systems and Operations Research, Stanford University, 1997. URL <https://web.stanford.edu/~glynn/papers/1997/G97.html>.
- [25] Peter W. Glynn and Ward Whitt. Necessary conditions in limit theorems for cumulative processes. *Stochastic Processes and their Applications*, 98(2):199–209, 2002. doi: 10.1016/S0304-4149(01)00146-6.
- [26] Olle Häggström. On the central limit theorem for geometrically ergodic Markov chains. *Probability Theory and Related Fields*, 132(1):74–82, 2005. doi: 10.1007/s00440-004-0390-7.
- [27] Olle Häggström and Jeffrey S. Rosenthal. On variance conditions for Markov chain CLTs. *Electronic Communications in Probability*, 12:454–464, 2007. doi: 10.1214/ECP.v12-1336.
- [28] Lu-Jing Huang and Yong-Hua Mao. Variational formulas for asymptotic variance of general discrete-time Markov chains. *Bernoulli*, 29(1):300–322, 2023. doi: 10.3150/21-BEJ1458.

- [29] Galin L. Jones. On the Markov chain central limit theorem. *Probability Surveys*, 1:299–320, 2004. doi: 10.1214/154957804100000051.
- [30] Galin L. Jones, Murali Haran, Brian S. Caffo, and Ronald Neath. Fixed-width output analysis for Markov chain Monte Carlo. *Journal of the American Statistical Association*, 101(476):1537–1547, 2006. doi: 10.1198/016214506000000492.
- [31] Olav Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer, New York, 1997. ISBN 9780387949574. doi: 10.1007/978-1-4757-4015-8.
- [32] Yitzhak Katznelson. *An Introduction to Harmonic Analysis*. Cambridge University Press, Cambridge, 3rd edition, 2004. doi: 10.1017/CBO9781139165372.
- [33] Claude Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Communications in Mathematical Physics*, 104(1):1–19, 1986. doi: 10.1007/BF01210789.
- [34] Ioannis Kontoyiannis and Sean P. Meyn. Spectral theory and limit theorems for geometrically ergodic Markov processes. *The Annals of Applied Probability*, 13(1):304–362, 2003. doi: 10.1214/aoap/1042765670.
- [35] Krzysztof Łatuszyński and Gareth O. Roberts. CLTs and asymptotic variance of time-sampled Markov chains. *Methodology and Computing in Applied Probability*, 15(1):237–247, 2013. doi: 10.1007/s11009-011-9237-8.
- [36] Martial Longla. Remarks on limit theorems for reversible Markov processes and their applications. *Journal of Statistical Planning and Inference*, 187:28–43, 2017. doi: 10.1016/j.jspi.2017.02.009.
- [37] Martial Longla, Costel Peligrad, and Magda Peligrad. On the functional central limit theorem for reversible Markov chains with nonlinear growth of the variance. *Journal of Applied Probability*, 49(4):1091–1105, 2012. doi: 10.1239/jap/1354716659.
- [38] Florian Maire, Randal Douc, and Jimmy Olsson. Comparison of asymptotic variances of inhomogeneous Markov chains with application to Markov chain Monte Carlo methods. *The Annals of Statistics*, 42(4):1483–1510, 2014. doi: 10.1214/14-AOS1209.
- [39] Michael Maxwell and Michael Woodroffe. Central limit theorems for additive functionals of Markov chains. *The Annals of Probability*, 28(2):713–724, 2000. doi: 10.1214/aop/1019160258.
- [40] Sean Meyn and Richard L. Tweedie. *Markov Chains and Stochastic Stability*. Cambridge Mathematical Library. Cambridge University Press, 2nd edition, 2009.
- [41] Esa Nummelin. On the Poisson equation in the potential theory of a single kernel. *Mathematica Scandinavica*, 68(1):59–82, 1991. doi: 10.7146/math.scand.a-12346.
- [42] OpenAI. GPT-5.6: Frontier intelligence that scales with your ambition, July 2026. URL <https://openai.com/index/gpt-5-6/>.
- [43] Magda Peligrad. On the CLT for stationary Markov chains with trivial tail sigma field. *Electronic Communications in Probability*, 28:1–7, 2023. doi: 10.1214/23-ECP509. URL <https://doi.org/10.1214/23-ECP509>.

- [44] Magda Peligrad and Wei Biao Wu. Central limit theorem for Fourier transforms of stationary processes. *The Annals of Probability*, 38(5):2009–2022, 2010. doi: 10.1214/10-AOP530.
- [45] Qian Qin. Analysis of two-component Gibbs samplers using the theory of two projections. *The Annals of Applied Probability*, 34(5):4310–4341, 2024. doi: 10.1214/24-AAP2066.
- [46] Daniel Revuz. *Markov Chains*, volume 11. Elsevier, 1st edition, 1984. ISBN 9780444864000.
- [47] Gareth O. Roberts. A note on acceptance rate criteria for CLTs for Metropolis–Hastings algorithms. *Journal of Applied Probability*, 36(4):1210–1217, 1999. ISSN 0021-9002. doi: 10.1239/jap/1032374766.
- [48] Gareth O. Roberts and Jeffrey S. Rosenthal. General state space Markov chains and MCMC algorithms. *Probability Surveys*, 1:20–71, 2004. doi: 10.1214/154957804100000024.
- [49] Gareth O. Roberts and Jeffrey S. Rosenthal. Variance bounding Markov chains. *The Annals of Applied Probability*, 18(3):1201–1214, 2008. doi: 10.1214/07-AAP486.
- [50] Jeffrey S. Rosenthal. Asymptotic variance and convergence rates of nearly-periodic Markov chain Monte Carlo algorithms. *Journal of the American Statistical Association*, 98(461):169–177, 2003. ISSN 0162-1459. doi: 10.1198/016214503388619193.
- [51] Walter Rudin. *Function Theory in Polydiscs*. W. A. Benjamin, New York, 1969.
- [52] W. L. Smith. Regenerative stochastic processes. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 232(1188):6–31, 1955. doi: 10.1098/rspa.1955.0198.
- [53] Matti Vihola, Jouni Helske, and Jordan Franks. Importance sampling type estimators based on approximate marginal Markov chain Monte Carlo. *Scandinavian Journal of Statistics*, 47(4):1339–1376, 2020. doi: 10.1111/sjos.12492.
- [54] Liming Wu. Forward-backward martingale decomposition and compactness results for additive functionals of stationary ergodic Markov processes. *Annales de l’Institut Henri Poincaré. Probabilités et Statistiques*, 35(2):121–141, 1999.
- [55] Wei Biao Wu and Michael Woodroffe. Martingale approximations for sums of stationary processes. *The Annals of Probability*, 32(2):1674–1690, 2004. doi: 10.1214/009117904000000351.
- [56] Ou Zhao and Michael Woodroffe. On martingale approximations. *The Annals of Applied Probability*, 18(5):1831–1847, 2008. doi: 10.1214/07-AAP505.
- [57] Ou Zhao, Michael Woodroffe, and Dalibor Volný. A central limit theorem for reversible processes with nonlinear growth of variance. *Journal of Applied Probability*, 47(4):1195–1202, 2010. doi: 10.1239/jap/1294170530.