Geometric Ergodicity and the Spectral Gap of Non-Reversible Markov Chains

I. Kontoyiannis* S.P. Meyn†

June 27, 2009

Abstract

We argue that the spectral theory of non-reversible Markov chains may often be more effectively cast within the framework of the naturally associated weighted-L_∞ space L_∞^V, instead of the usual Hilbert space $L_2 = L_2(\pi)$, where π is the invariant measure of the chain. This observation is, in part, based on the following results. A discrete-time Markov chain with values in a general state space is geometrically ergodic if and only if its transition kernel admits a spectral gap in L_∞^V. If the chain is reversible, the same equivalence holds with L_2 in place of L_∞^V, but in the absence of reversibility it fails: There are (necessarily non-reversible, geometrically ergodic) chains that admit a spectral gap in L_∞^V but not in L_2. Moreover, if a chain admits a spectral gap in L_2, then for any $h \in L_2$ there exists a Lyapunov function $V_h \in L_1$ such that V_h dominates h and the chain admits a spectral gap in $L_\infty^{V_h}$. The relationship between the size of the spectral gap in L_∞^V or L_2, and the rate at which the chain converges to equilibrium is also briefly discussed.

Keywords: Markov chain, geometric ergodicity, spectral theory, stochastic Lyapunov function, reversibility, spectral gap.

2000 AMS Subject Classification: 60J05, 60J10, 37A30, 37A25

---

*Department of Informatics, Athens University of Economics and Business, Patission 76, Athens 10434, Greece. Email: `yiannis@aueb.gr`.
I.K. was supported in part by a Sloan Foundation Research Fellowship.
†Department of Electrical and Computer Engineering and the Coordinated Sciences Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. Email: `meyn@uiuc.edu`.
S.P.M. was supported in part by the National Science Foundation ECS-0523620. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

1 Introduction and Main Results

There is increasing interest in spectral theory and rates of convergence for Markov chains. Research is motivated by elegant mathematics as well as a range of applications. In particular, one of the most effective general methodologies used to establish bounds on the convergence rate of a geometrically ergodic chain is via an analysis of the spectrum of the chain's transition kernel. See, e.g., [16, 22, 21, 8, 5, 12, 11, 13, 23, 19, 3, 2, 17, 14, 7, 6], and the relevant references therein.

The word *spectrum* naturally invites techniques grounded in a Hilbert space framework. The majority of quantitative results on rates of convergence are obtained using such methods, within the Hilbert space $L_2 = L_2(\pi)$, where π denotes the stationary distribution of the Markov chain in question. Indeed, most successful studies have been carried out for Markov chains that are reversible, in which case a key to analysis is the fact that the transition kernel, viewed as a linear operator on L_2, is self-adjoint. In this paper we argue that, in the *absence* of reversibility, the Hilbert space framework may not be the appropriate setting for spectral analysis.

To be specific, let $\boldsymbol{X} = \{X(n) \,:\, n \geq 0\}$ denote a discrete-time Markov chain with values on a general state space X. We assume that X is equipped with a countably generated sigma-algebra $\mathcal{B}$. The distribution of $\boldsymbol{X}$ is described by its initial state $X(0) = x_0 \in \mathsf{X}$ and the transition semigroup $\{P^n \,:\, n \geq 0\}$, where, for each n,

$$P^n(x, A) := \Pr\{X(n) \in A \,|\, X(0) = x\}, \quad x \in \mathsf{X},\ A \in \mathcal{B}.$$

For simplicity we write P for the one-step kernel P^1. Recall that each P^n, like any (not necessarily probabilistic) kernel $Q(x, dy)$ acts on functions $F : \mathsf{X} \to \mathbb{C}$ and signed measures ν on $(\mathsf{X}, \mathcal{B})$, via,

$$QF(\cdot) = \int_{\mathsf{X}} Q(\cdot, dy)F(y) \quad \text{and} \quad \nu Q(\cdot) = \int_{\mathsf{X}} \nu(dx)Q(x, \cdot),$$

whenever the integrals exist. Throughout the paper, we assume that the chain $\boldsymbol{X} = \{X(n)\}$ is *ψ-irreducible and aperiodic*; cf. [15, 18]. This means that there is a σ-finite measure ψ on $(\mathsf{X}, \mathcal{B})$ such that, for any $A \in \mathcal{B}$ with $\psi(A) > 0$, and any $x \in \mathsf{X}$,

$$P^n(x, A) > 0, \quad \text{for all } n \text{ sufficiently large.}$$

Moreover, we assume that ψ is *maximal* in the sense that any other such ψ' is absolutely continuous with respect to ψ.

1.1 Geometric ergodicity

The natural class of chains to consider in the present context is that of *geometrically ergodic* chains, namely, chains with the property that there exists an invariant measure π on $(\mathsf{X}, \mathcal{B})$ and functions $\rho : \mathsf{X} \to (0, 1)$ and $C : \mathsf{X} \to [1, \infty)$, such that,

$$\|P^n(x, \cdot) - \pi\|_{\mathrm{TV}} \leq C(x)\rho(x)^n, \quad \text{for all } n \geq 0,\ \pi\text{-a.e. } x \in \mathsf{X},$$

where $\|\mu\|_{\mathrm{TV}} := \sup_{A\in\mathcal{B}} |\mu(A)|$ denotes the total variation norm on signed measures. Under ψ-irreducibility and aperiodicity this is equivalent [15, 19] to the seemingly stronger requirement that there is a single constant $\rho \in (0, 1)$, a constant $B < \infty$ and a π-a.e. finite function $V : \mathsf{X} \to [1, \infty]$, such that,

$$\|P^n(x, \cdot) - \pi\|_V \le B\, V(x)\rho^n, \quad \text{for all } n \ge 0,\ \pi\text{-a.e. } x \in \mathsf{X}, \tag{1}$$

where $\|\mu\|_V := \sup\{|\int F\, d\mu| \ :\ F \in L_\infty^V\}$ denotes the V-norm on signed measures, and where L_∞^V denotes the weighted-L_∞ space consisting of all measurable functions $F\colon \mathsf{X} \to \mathbb{C}$ with,

$$\|F\|_V := \sup_{x\in\mathsf{X}} \frac{|F(x)|}{V(x)} < \infty\,. \tag{2}$$

Another equivalent and operationally simpler definition of geometric ergodicity for a ψ-irreducible, aperiodic chain $\boldsymbol{X} = \{X(n)\}$, is that it satisfies the following drift criterion [15]:

$$\left.\begin{array}{l}\text{There is a function } V\colon \mathsf{X} \to [1, \infty], \text{ a small set } C \subset \mathsf{X},\\ \text{and constants } \delta > 0,\ b < \infty, \text{ such that:}\\[1em] \qquad\qquad PV \le (1 - \delta)V + b\mathbb{I}_C\,.\end{array}\right\} \qquad \textbf{(V4)}$$

We then say that the chain is *geometrically ergodic with Lyapunov function V*. In (V4) it is always assumed that the Lyapunov function V is finite for at least one x (and then it is necessarily finite ψ-a.e.). Also, recall that a set $C \in \mathcal{B}$ is *small* if there exist $n \ge 1$, $\epsilon > 0$ and a probability measure ν on $(\mathsf{X}, \mathcal{B})$ such that, $P^n(x, A) \ge \epsilon\mathbb{I}_C(x)\nu(A)$, for all $x \in \mathsf{X}$, $A \in \mathcal{B}$.

Our first result relates geometric ergodicity to the spectral properties of the kernel P. Its proof, given at the end of Section 3, is based on ideas from [12]. See Section 2 for more precise definitions.

Proposition 1.1. *A ψ-irreducible and aperiodic Markov chain $\boldsymbol{X} = \{X(n)\}$ is geometrically ergodic with Lyapunov function V if and only if P admits a spectral gap in L_∞^V.*

1.2 Reversibility

Recall that the chain $\boldsymbol{X} = \{X(n)\}$ is called *reversible* if there is a probability measure π on $(\mathsf{X}, \mathcal{B})$ satisfying the detailed balance conditions,

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

This is equivalent to saying that the linear operator P is self-adjoint on the space $L_2 = L_2(\pi)$ of (measurable) functions $F : \mathsf{X} \to \mathbb{C}$ that are square-integrable under π, endowed with the inner product $(F, G) = \int FG^*\, d\pi$, where '$*$' denotes the complex conjugate operation.

The following result is the natural analog of Proposition 1.1 for reversible chains. Its proof, given in Section 3, is partly based on results in [19].

Proposition 1.2. *A reversible, ψ-irreducible and aperiodic Markov chain $\boldsymbol{X} = \{X(n)\}$ is geometrically ergodic if and only if P admits a spectral gap in L_2.*

1.3 Spectral theory

The main question addressed in this paper is whether the reversibility assumption of Proposition 1.2 can be relaxed. In other words, whether the space L_2 can be used to characterize geometric ergodicity like L_∞^V was in Proposition 1.1. One direction is true without reversibility: A spectral gap in L_2 implies that the chain is "geometrically ergodic in L_2" [19][20], and this implies the existence of a Lyapunov function V satisfying (V4) [20]. Therefore, the chain is geometrically ergodic in the sense of [12], where it is also shown that it must admit a central gap in L_∞^V. A direct, explicit construction of a Lyapunov function V_h is given in our first main result stated next, where quantitative information about V_h is also obtained. It is proved in Section 3.

Theorem 1.3. *Suppose that a ψ-irreducible, aperiodic chain $\boldsymbol{X} = \{X(n)\}$ admits a spectral gap in L_2. Then, for any $h \in L_2$, there is π-integrable function V_h, such that the chain is geometrically ergodic with Lyapunov function V_h and $h \in L_\infty^{V_h}$.*

But the other direction may not hold in the absence of reversibility. Based on earlier counterexamples constructed by Häggström [9, 10] and Bradley [1], in Section 3 we prove the following:

Theorem 1.4. *There exists a ψ-irreducible, aperiodic Markov chain $\boldsymbol{X} = \{X(n)\}$ which is geometrically ergodic but does not admit a spectral gap in L_2.*

1.4 Convergence rates

The existence of a spectral gap is intimately connected to the exponential convergence rate for a ψ-irreducible, aperiodic Markov chain. For example, if the chain is reversible, we have the following well-known, quantitative bound. See Section 2 for detailed definitions; the result follows from the results in [19], combined with Lemma 2.2 given in Section 2.

Proposition 1.5. *Suppose that a reversible chain $\boldsymbol{X} = \{X(n)\}$ is ψ-irreducible, aperiodic, and has initial distribution μ. If the chain $\boldsymbol{X}$ admits a spectral gap $\delta_2 > 0$ in L_2, then,*

$$\|\mu P^n - \pi\|_{\mathrm{TV}} \le \frac{1}{2}\|\mu - \pi\|_2 (1-\delta_2)^n, \quad n \ge 1,$$

where the L_2-norm on signed measures ν is defined as the $L_2(\pi)$-norm of the density $d\nu/d\pi$ if it exists, and is set equal to infinity otherwise.

In the absence of reversibility, the size of the spectral gap in L_∞^V precisely determines the exponential convergence rate of *any* geometrically ergodic chain. The result of the following proposition is stated in Lemma 2.3 in Section 2.

Proposition 1.6. *Suppose that the chain $\boldsymbol{X} = \{X(n)\}$ is ψ-irreducible and aperiodic. If it admits a spectral gap $\delta_V > 0$ in L_∞^V, then, for π-a.e. x,*

$$\lim_{n\to\infty} \frac{1}{n} \log \|P^n(x,\cdot) - \pi\|_V = \log(1-\delta_V).$$

In fact, the convergence is uniform in that:

$$\lim_{n\to\infty} \frac{1}{n} \log \left(\sup_{x\in\mathsf{X},\, \|F\|_V = 1} \frac{|P^n F(x) - \int F\, d\pi|}{V(x)} \right) = \log(1-\delta_V). \tag{3}$$

Section 2 contains precise definitions regarding the spectrum and the spectral gap of the kernel P acting either on L_2 or the weighted-L_∞ space L_∞^V. Simple properties of the spectrum are also stated and proved. Section 3 contains the proofs of the first four results stated above.

2 Spectra and Geometric Ergodicity

We begin by giving precise definitions for the spectrum and spectral gap of the transition kernel P, viewed as a linear operator. The spectrum depends on the domain of P, for which we consider two possibilities:

(i) The Hilbert space $L_2 = L_2(\pi)$, equipped with the norm $\|F\|_2 = [\int |F|^2 \, d\pi]^{1/2}$.

(ii) The Banach space L_∞^V, with norm $\|\cdot\|_V$ defined in (2).

In either case, the spectrum is defined as the set of nonzero $\lambda \in \mathbb{C}$ for which the inverse $(I\lambda - P)^{-1}$ does not exist as a bounded linear operator on the domain of P. The transition kernel admits a *spectral gap* if there exists $\varepsilon_0 > 0$ such that $\mathcal{S} \cap \{z : |z| \geq 1 - \varepsilon_0\}$ is finite, and contains only poles of finite multiplicity; see [12, Section 4] for more details. The spectrum is denoted $\mathcal{S}_2$ when P is viewed as a linear operator on L_2, and it is denoted $\mathcal{S}_V$ when P is viewed as a linear operator on L_∞^V.

The induced operator norm of a linear operator $\widehat{P} \colon L_\infty^V \to L_\infty^V$ is defined as usual via,

$$\|\!|\widehat{P}|\!\|_V := \sup \frac{\|\widehat{P}F\|_V}{\|F\|_V},$$

where the supremum is over all $F \in L_\infty^V$ satisfying $\|F\|_V \neq 0$. An analogous definition gives the induced operator norm $\|\!|\widehat{P}|\!\|_2$ of a linear operator $\widehat{P}$ acting on L_2.

For a ψ-irreducible, aperiodic chain $\boldsymbol{X} = \{X(n)\}$, geometric ergodicity expressed in the form (1) implies that P^n converges to a rank-one operator, at a geometric rate: For some constants $B < \infty$, $\rho \in (0, 1)$,

$$\|\!|P^n - \mathbf{1} \otimes \pi|\!\|_V \leq B\,\rho^n\,, \qquad n \geq 0, \tag{4}$$

where the outer product $\mathbf{1} \otimes \pi$ denotes the kernel $\mathbf{1} \otimes \pi(x, dy) = \pi(dy)$. It follows that the inverse $[I\lambda - P + \mathbf{1} \otimes \pi]^{-1}$ exists as a bounded linear operator on L_∞^V, whenever $\lambda > \rho$. This in turn implies that P has a single isolated pole at $\lambda = 1$ in the set $\{\lambda \in \mathbb{C} : \lambda > \rho\}$, so that P admits a spectral gap.

In Lemma 2.1 we clarify the location of poles when the chain admits a spectral gap in L_2 or L_∞^V.

Lemma 2.1. *If a ψ-irreducible, aperiodic Markov chain admits a spectral gap in L_∞^V or L_2, then the only pole on the unit circle in $\mathbb{C}$ is $\lambda = 1$, and this pole has multiplicity one.*

Proof. We present the proof for L_∞^V; the proof in L_2 is identical.

We first note that the existence of a spectral gap implies ergodicity: There is a left eigenmeasure μ corresponding to the eigenvalue 1, satisfying $\mu P = \mu$ and $|\mu|(V) = \|\mu\|_V < \infty$. On letting $\pi(\,\cdot\,) = |\mu(\,\cdot\,)|/|\mu(\mathsf{X})|$ we conclude that π is super-invariant: $\pi P \geq \pi$. Since $\pi(\mathsf{X}) = 1$

we must have invariance. The ergodic theorem for positive recurrent Markov chains implies that $E[G(X(n)) \mid X(0) = x] \to \int G \, d\pi$, as $n \to \infty$, whenever $G \in L_1(\pi)$.

Ergodicity rules out the existence of multiple eigenfunctions corresponding to $\lambda = 1$. Hence, if this pole has multiplicity greater than one, then there is a generalized eigenfunction $h \in L_\infty^V$ satisfying,

$$Ph = h + 1.$$

Iterating gives $P^n h\,(x) = E[h(X(n)) \mid X(0) = x] = h(x) + n$ for $n \geq 1$. This rules out ergodicity, and proves that $\lambda = 1$ has multiplicity one.

We now show that if $\lambda \in \mathcal{S}_V$ with $|\lambda| = 1$, then $\lambda = 1$. To see this, let $h \in L_\infty^V$ denote an eigenfunction, $Ph = \lambda h$. Iterating, we obtain,

$$E[h(X(n)) \mid X(0) = x] = h(x)\lambda^n.$$

Then, letting $n \to \infty$, the right-hand-side converges to $\int h \, d\pi$ for a.e. x., so that $\lambda = 1$ and $h(x) = \int h \, d\pi$, π-a.e. □

Therefore, for a ψ-irreducible, aperiodic chain, the existence of a spectral gap in L_2 is equivalent to the existence of a single eigenvalue $\lambda = 1$ on the unit circle, which has multiplicity one. The spectral gap δ_2 is then defined as,

$$\delta_2 = 1 - \sup\{|\lambda| \ : \ \lambda \in \mathcal{S}_2, \, \lambda \neq 1\},$$

and similarly for δ_V.

Next we state two well-known, alternative expressions for the L_2-spectral gap δ_2 of a reversible chain. See, e.g., [19, Theorem 2.1] and [4, Proposition VIII.1.11].

Lemma 2.2. *Suppose $\boldsymbol{X}$ is a ψ-irreducible, aperiodic, reversible Markov chain. Then, its L_2-spectral gap δ_2 admits the alternative characterizations,*

$$\begin{aligned} \delta_2 &= 1 - \sup\Big\{ \frac{\|\nu P\|_2}{\|\nu\|_2} \ : \ \textit{signed measures } \nu \textit{ with } \nu(\mathsf{X}) = 0, \ \|\nu\|_2 \neq 0 \Big\} \\ &= 1 - \lim_{n\to\infty} \Big(|\!|\!| P^n - \mathbf{1} \otimes \pi |\!|\!|_2 \Big)^{1/n}, \end{aligned}$$

where the limit is the usual spectral radius of the semigroup $\{\widehat{P}^n\}$ generated by the kernel $\widehat{P} = P - \mathbf{1} \otimes \pi$, acting on functions in $L_2(\pi)$.

A similar result holds for δ_V, even in the absence of reversibility; see, e.g., [13].

Lemma 2.3. *Suppose $\boldsymbol{X}$ is a ψ-irreducible, aperiodic Markov chain. Then, its L_∞^V-spectral gap δ_V admits the following alternative characterization in terms of the spectral radius,*

$$\delta_V = 1 - \lim_{n\to\infty} \Big(|\!|\!| P^n - \mathbf{1} \otimes \pi |\!|\!|_V \Big)^{1/n}.$$

3 Proofs

First we prove Theorem 1.3. The following notation will be useful throughout this section.

For a Markov chain $\boldsymbol{X} = \{X(n)\}$, the first hitting time and first return time to a set $C \in \mathcal{B}$ are defined, respectively, by,

$$\begin{aligned} \sigma_C &:= \min\{n \geq 0 : X(n) \in C\}; \\ \tau_C &:= \min\{n \geq 1 : X(n) \in C\}. \end{aligned} \tag{5}$$

Conditional on $X(0) = x$, the expectation operator corresponding to the measure defining the distribution of the process $\boldsymbol{X} = \{X(n)\}$ is denoted $\mathsf{E}_x(\cdot)$, so that, for example, $P^n F(x) = E[F(X(n)) \,|\, X(0) = x] = \mathsf{E}_x[F(X(n))]$. For an arbitrary signed measure μ on $(\mathsf{X}, \mathcal{B})$, we write $\mu(F)$ for $\int F\, d\mu$, for any function $F : \mathsf{X} \to \mathbb{C}$ for which the integral exists.

Proof of Theorem 1.3. Since $\pi(h^2) < \infty$, and the chain is ψ-irreducible, it follows that there exists an increasing sequence of *h^2-regular sets* providing a π-a.e. covering of X [15, Theorem 14.2.5]. That is, there is a sequence of sets $\{S_r : r \in \mathbb{Z}_+\}$ such that $\pi(S_r) \to 1$ as $r \to \infty$, $S_r \subset S_{r+1}$ for each r, and the following bounds hold,

$$V_r(x) := \mathsf{E}_x\Big[\sum_{n=0}^{\tau_{S_r}} h^2(X(n))\Big] < \infty, \qquad \text{for } \pi\text{-a.e. } x$$

$$\sup_{x \in S_r} V_r(x) < \infty.$$

Since the chain admits a spectral gap in L_2, combining Theorem 2.1 of [19] with Lemma 2.2 and the results of [20], we have that it is geometrically ergodic. Hence, from [15, Theorem 15.4.2] it follows that there exists a sequence of *Kendall sets* providing a π-a.e. covering of X. That is, there is a sequence of sets $\{K_r : r \in \mathbb{Z}_+\}$ and positive constants $\{\theta_r : r \in \mathbb{Z}_+\}$ satisfying $\pi(K_r) \to 1$ as $r \to \infty$, $K_r \subset K_{r+1}$ for each r, and the following bounds hold,

$$U_r(x) := \mathsf{E}_x\big[\exp(\theta_r \tau_{K_r})\big] < \infty, \qquad \text{for } \pi\text{-a.e. } x$$

$$\sup_{x \in K_r} U_r(x) < \infty.$$

We also define another collection of sets,

$$C_{r,m} := \{x \in \mathsf{X} : U_r(x) + V_r(x) \leq m\}.$$

For each $r \geq 1$, these sets are non-decreasing in m, and $\pi(C_{r,m}) \to 1$ as $m \to \infty$. Moreoever, whenever $C_{r,m} \in \mathcal{B}^+$, this set is both an h^2-regular set and a Kendall set. This follows by combining Theorems 14.2.1 and 15.2.1 of [15]. Fix r_0 and m_0 so that $\pi(C_{r_0,m_0}) > 0$. We henceforth denote C_{r_0,m_0} by C, and let $\theta > 0$ denote a value satisfying the bound,

$$\mathsf{E}_x\big[\exp(\theta \tau_C)\big] < \infty, \qquad \text{for } \pi\text{-a.e. } x,$$

where the expectation is uniformly bounded over the Kendall set C.

The candidate Lyapunov function can now be defined as,

$$V_h(x) := \mathsf{E}_x\Big[\sum_{n=0}^{\sigma_C}\big(1 + |h(X(n))|\big)\exp\big(\tfrac{1}{2}\theta n\big)\Big]. \tag{6}$$

We first obtain a bound on this function. Writing,

$$V_h(x) = \mathsf{E}_x\Big[\sum_{n=0}^{\sigma_C}\exp\big(\tfrac{1}{2}\theta n\big)\Big] + \sum_{n=0}^{\infty}\mathsf{E}_x\Big[|h(X(n))|\exp\big(\tfrac{1}{2}\theta n\big)\mathbb{I}\{n \le \sigma_C\}\Big],$$

we see that the first term is finite π-a.e. by construction. The square of the second term is bounded above, using the Cauchy-Shwartz inequality, by,

$$\mathsf{E}_x\Big[\sum_{n=0}^{\infty}|h(X(n))|^2\mathbb{I}\{n \le \sigma_C\}\Big]\mathsf{E}_x\Big[\sum_{n=0}^{\infty}\exp\big(\theta n\big)\mathbb{I}\{n \le \sigma_C\}\Big] = U_{r_0}(x)\mathsf{E}_x\Big[\sum_{t=0}^{\sigma_C}\exp\big(\theta t\big)\Big],$$

so that V_h is finite π-a.e., and we also easily see that $|h| \le V_h$ so that $h \in L_\infty^{V_h}$.

Next we show that V_h satisfies (V4): First apply $e^{\frac{1}{2}\theta}P$ to the function V_h to obtain,

$$e^{\frac{1}{2}\theta}PV_h\,(x) = \mathsf{E}_x\Big[\sum_{n=1}^{\tau_C}\big(1 + |h(X(n+1))|\big)\exp\big(\tfrac{1}{2}\theta(t+1)\big)\Big] \tag{7}$$

We have $\tau_C = \sigma_C$ when $X(0) \in C^c$. This gives,

$$e^{\frac{1}{2}\theta}PV_h\,(x) = V_h(x) - \big(1 + |h(x)|\big), \qquad x \in C^c.$$

If $X(0) = x \in C$, then the previous arguments imply that the right-hand-side of (7) is finite, and in fact uniformly bounded over $x \in C$. Combining these results, we conclude that there exists a constant b_0 such that,

$$PV_h \le e^{-\frac{1}{2}\theta}V_h + b_0\mathbb{I}_C$$

Regular sets are necessarily small [15, Theorem 11.3.11] so that this is a version of the drift inequality (V4).

Finally note that, by the fact that (V4) implies the weaker drift condition (V3) of [15], the function V_h is π-integrable by [15, Theorem 14.0.1]. □

Theorem 1.3 states that (V4) holds for a Lyapounov function V_h with $h \in L_\infty^{V_h}$. If this could be strengthened to show that for every geometrically ergodic chain and any $h \in L_2$, the chain was geometrically ergodic with a Lyapunov function V_h that had $h^2 \in L_\infty^{V_h}$, then the central limit theorem would hold for the partial sums of $h(X(n))$ [15, Theorem 17.0.1]. But this is not generally possible:

Proposition 3.1. *There exists a geometrically ergodic Markov chain on a countable state space* X *and a function* $G \in L_2$ *with mean* $\pi(G) = 0$*, for which the central limit theorem fails in that the normalized partial sums,*

$$\frac{1}{\sqrt{n}}\sum_{i=0}^{n-1}G(X(i)), \quad n \ge 1, \tag{8}$$

converge neither to a normal distribution nor to a point mass.

The result of the proposition appears in [9, Theorem 1.3], and an earlier counterexample in [1] yields the same conclusion. Based on these counterexamples we now show that geometric ergodicity does not imply a spectral gap in the Hilbert space setting.

Proof of Theorem 1.4. Suppose that the Markov chain $\boldsymbol{X} = \{X(n)\}$ constructed in Proposition 3.1 does admit a spectral gap in L_2. Then its autocorrelation fuction decays geometrically fast, for any $h \in L_2$: Assuming without loss of generality that $\pi(h) = 0$, and letting $R_h(n) = \pi(hP^n h)$, for all n, we have the bound,

$$|R(n)| \leq \sqrt{\pi(h^2)\pi((P^n h)^2)}, \qquad n \geq 1.$$

Applying Theorem 1.3, we conclude that the right-hand-side decays geometrically fast as $n \to \infty$. Consequently, the sequence of normalized sums,

$$S_n := \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} h(X(i)), \qquad n \geq 1,$$

is uniformly bounded in L_2, i.e.,

$$\limsup_{n\to\infty} \mathsf{E}_\pi[S_n^2] \leq \sum_{n=-\infty}^{\infty} |R(n)|,$$

where $\mathsf{E}_\pi[\cdot]$ denotes the expectation operator corresponing to the stationary version of the chain. However, this is impossible for the choice of the function $h = G$ as in Proposition 3.1: In [9, p. 81] it is shown that the corresponding normalized sums in (8) fail to define a tight sequence of probability distributions. This is a consequence of [9, Lemma 3.2].

This contradiction establishes the claim that the Markov chain of Proposition 3.1 cannot admit a spectral gap in L_2. □

Finally we prove Propositions 1.1 and 1.2.

Proof of Proposition 1.1. The equivalence stated in the proposition is obtained on combining Lemma 2.1 with [12, Proposition 4.6]. To explain this, we introduce new terminology: The transition kernel is called V*-uniform* if $\lambda = 1$ is the only pole on the unit circle in $\mathbb{C}$, and this pole has multiplicity one. Proposition 4.6 of [12] states that geometric ergodicity with resepct to a Lyapunov function V is equivalent to V-uniformity of the kernel P. Consequently, the direct part of the proposition holds, since V-uniformity of P implies that it admits a spectral gap in L_∞^V.

Conversely, if the chain admits a spectral gap in L_∞^V, then Lemma 2.1 states that P is V-uniform. Applying Proposition 4.6 of [12] once more, we conclude that the chain is geometrically ergodic with the same Lyapunov function V. □

Proof of Proposition 1.2. The forward direction of the statement of the proposition is contained in [19] and [20].

The converse again follows from Lemma 2.1 and a minor modification of the arguments used in [12, Proposition 4.6]. If the chain admits a spectral gap in L_2, then the lemma states that $\lambda = 1$ has multiplicity one, and that this is the only pole on the unit circle in $\mathbb{C}$. It follows that for some $\rho < 1$, the inverse $[zI - (P - \mathbf{1} \otimes \pi)]^{-1}$ exists as a bounded linear operator on L_2, whenever $|z| \geq \rho$. Denote $b_\rho = \sup \{\|[zI - (P - \mathbf{1} \otimes \pi)]^{-1}\|_2 : |z| = \rho\}$, where $\| \cdot \|_2$ is the induced operator norm on L_2.

Following the proof of [12, Theorem 4.1], we conclude that finiteness of b_ρ implies a form of geometric ergodicity: For any $g \in L_2$,

$$\frac{1}{2\pi} \int_0^{2\pi} e^{in\phi} [\rho e^{in\phi} I - (P - \mathbf{1} \otimes \pi)]^{-1} g = \rho^{-n-1}(P^n g - \pi(g)).$$

Therefore, the L_2-norm of the left-hand-side is bounded by $b_\rho \|g\|_2$. This gives,

$$\|P^n g - \pi(g)\|_2 \leq b_\rho \|g\|_2 \rho^{n+1}, \qquad n \geq 1.$$

It follows from [15, Theorem 15.4.3] that the Markov chain is geometrically ergodic. ☐

References

[1] R.C. Bradley, Jr. Information regularity and the central limit question. *Rocky Mountain J. Math.*, 13(1):77–97, 1983.

[2] P. Brémaud. *Markov chains*, volume 31 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1999. Gibbs fields, Monte Carlo simulation, and queues.

[3] S.P. Brooks and G.O. Roberts. Convergence assessment techniques for Markov chain Monte Carlo. *Statistics and Computing*, 8:319–335, 1998.

[4] J.B. Conway. *A course in functional analysis*. Springer-Verlag, New York, second edition, 1990.

[5] Michael Dellnitz and Oliver Junge. On the approximation of complicated dynamical behavior. *SIAM J. on Numerical Analysis*, 36(2):491–515, 1999.

[6] P. Diaconis. The Markov chain Monte Carlo revolution. *Bull. Amer. Math. Soc. (N.S.)*, 46(2):179–205, 2009.

[7] P. Diaconis, K. Khare, and L. Saloff-Coste. Gibbs sampling, exponential families and orthogonal polynomials. *Statist. Sci.*, 23(2):151–200, 2008. With comments and a rejoinder by the authors.

[8] P. Diaconis and L. Saloff-Coste. What do we know about the Metropolis algorithm? *J. Comput. System Sci.*, 57(1):20–36, 1998. 27th Annual ACM Symposium on the Theory of Computing (STOC'95) (Las Vegas, NV).

[9] O. Häggström. On the central limit theorem for geometrically ergodic Markov chains. *Probab. Theory Related Fields*, 132(1):74–82, 2005.

[10] O. Häggström. Acknowledgement of priority concerning "On the central limit theorem for geometrically ergodic Markov chains" [Probab. Theory Related Fields **132** (2005), no. 1, 74–82.]. *Probab. Theory Related Fields*, 135(3):470, 2006.

[11] W. Huisinga, S.P. Meyn, and C. Schuette. Phase transitions and metastability in Markovian and molecular systems. *Ann. Appl. Probab.*, 2001. to appear.

[12] I. Kontoyiannis and S.P. Meyn. Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab.*, 13:304–362, February 2003.

[13] I. Kontoyiannis and S.P. Meyn. Large deviation asymptotics and the spectral theory of multiplicatively regular Markov processes. *Electron. J. Probab.*, 10(3):61–123, 2005.

[14] G. Lebeau and P. Diaconis. Métropolis: le jour où l'étoile probabilité entra dans le champ gravitationnel de la galaxie microlocale. In *Séminaire: Équations aux Dérivées Partielles. 2006–2007*, Sémin. Équ. Dériv. Partielles, pages Exp. No. XIV, 13. École Polytech., Palaiseau, 2007.

[15] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability.* Cambridge University Press, London, 2nd edition, 2009. Published in the Cambridge Mathematical Library. 1993 edition online: `http://black.csl.uiuc.edu/~meyn/pages/book.html`.

[16] S.P. Meyn and R.L. Tweedie. Computable bounds for geometric convergence rates of Markov chains. *Ann. Appl. Probab.*, 4(4):981–1011, 1994.

[17] R. Montenegro and P. Tetali. Mathematical aspects of mixing times in Markov chains. *Found. Trends Theor. Comput. Sci.*, 1(3), 2006.

[18] E. Nummelin. *General Irreducible Markov Chains and Nonnegative Operators.* Cambridge University Press, Cambridge, 1984.

[19] G.O. Roberts and J.S. Rosenthal. Geometric ergodicity and hybrid Markov chains. *Electron. Comm. Probab.*, 2:no. 2, 13–25 (electronic), 1997.

[20] G.O. Roberts and R.L. Tweedie. Geometric L^2 and L^1 convergence are equivalent for reversible Markov chains. *J. Appl. Probab.*, 38A:37–41, 2001. Probability, statistics and seismology.

[21] J.S. Rosenthal. Correction: "Minorization conditions and convergence rates for Markov chain Monte Carlo". *J. Amer. Statist. Assoc.*, 90(431):1136, 1995.

[22] J.S. Rosenthal. Minorization conditions and convergence rates for Markov chain Monte Carlo. *J. Amer. Statist. Assoc.*, 90(430):558–566, 1995.

[23] G. Winkler. *Image Analysis, Random Fields and Dynamic Monte Carlo Methods: A Mathematical Introduction.* Springer-Verlag, Berlin, 1995.