



HAL
open science

Computable bounds of spectral gap for discrete Markov chains with band transition matrices

Loïc Hervé, James Ledoux

► **To cite this version:**

Loïc Hervé, James Ledoux. Computable bounds of spectral gap for discrete Markov chains with band transition matrices. 2015. hal-01117465v1

HAL Id: hal-01117465

<https://hal.science/hal-01117465v1>

Preprint submitted on 17 Feb 2015 (v1), last revised 13 Jan 2016 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Computable bounds of spectral gap for discrete Markov chains with band transition matrices

Loïc HERVÉ, and James LEDOUX *

February 17, 2015

Abstract

We analyse the $\ell^2(\pi)$ -convergence rate of irreducible and aperiodic Markov chains with N -band transition probability matrix P and with invariant distribution π . This analysis is heavily based on: first the study of the essential spectral radius $r_{ess}(P|_{\ell^2(\pi)})$ of $P|_{\ell^2(\pi)}$ derived from Hennion's quasi-compactness criteria; second the connection between the spectral gap property (SG₂) of P on $\ell^2(\pi)$ and the V -geometric ergodicity of P . Specifically, (SG₂) is shown to hold under the condition

$$\alpha_0 := \sum_{m=-N}^N \limsup_{i \rightarrow +\infty} \sqrt{P(i, i+m)P^*(i+m, i)} < 1.$$

Moreover $r_{ess}(P|_{\ell^2(\pi)}) \leq \alpha_0$. Simple conditions on asymptotic properties of P and of its invariant probability distribution π to ensure that $\alpha_0 < 1$ are given. In particular this allows us to obtain estimates of the $\ell^2(\pi)$ -geometric convergence rate of random walks with bounded increments. The specific case of reversible P is also addressed. Numerical bounds on the convergence rate can be provided via a truncation procedure. This is illustrated on the Metropolis-Hastings algorithm.

AMS subject classification : 60J10; 47B07

Keywords : Rate of convergence, ℓ^2 -spectral gap, V -geometric ergodicity, Essential spectral radius, Metropolis-Hastings algorithm.

1 Introduction

Let $P := (P(i, j))_{(i, j) \in \mathbb{X}^2}$ be a Markov kernel on a countable state space \mathbb{X} . For the sake of simplicity we suppose that $\mathbb{X} := \mathbb{N}$. Throughout the paper we assume that P is irreducible

*INSA de Rennes, IRMAR, F-35042, France; CNRS, UMR 6625, Rennes, F-35708, France; Université Européenne de Bretagne, France. {Loic.Herve,James.Ledoux}@insa-rennes.fr

and aperiodic, that P has a unique invariant probability measure denoted by $\pi := (\pi(i))_{i \in \mathbb{N}}$ (observe that $\forall i \in \mathbb{N}$, $\pi(i) > 0$ from irreducibility), and finally that

$$\exists i_0 \in \mathbb{N}, \exists N \in \mathbb{N}^*, \forall i \geq i_0 : |i - j| > N \implies P(i, j) = 0. \quad (\mathbf{AS1})$$

We denote by $(\ell^2(\pi), \|\cdot\|_2)$ the usual Hilbert space of sequences $(f(i))_{i \in \mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$ such that $\|f\|_2 := [\sum_{i \geq 0} |f(i)|^2 \pi(i)]^{1/2} < \infty$. It is well-known that P defines a linear contraction on $\ell^2(\pi)$, and that its adjoint operator P^* on $\ell^2(\pi)$ is defined by $P^*(i, j) := \pi(j) P(j, i) / \pi(i)$. The kernel P is said to have the spectral gap property on $\ell^2(\pi)$ at rate $\rho \in (0, 1)$ if there exists some positive constants $\rho \in (0, 1)$ and $C \in (0, +\infty)$ such that

$$\forall n \geq 1, \forall f \in \ell^2(\pi), \quad \|P^n f - \Pi f\|_2 \leq C \rho^n \|f\|_2 \quad \text{with} \quad \Pi f := \pi(f) 1_{\mathbb{N}}, \quad (\mathbf{SG}_2)$$

where $\pi(f) := \sum_{i \geq 0} f(i) \pi(i)$. A relevant and standard issue is to compute the value (or to find an upper bound) of

$$\varrho_2 := \inf\{\rho \in (0, 1) : (\mathbf{SG}_2) \text{ holds true}\}. \quad (1)$$

In this work we use the quasi-compactness criteria of [Hen93] to study (\mathbf{SG}_2) and to estimate ϱ_2 . In Section 2 it is proved that (\mathbf{SG}_2) holds when

$$\alpha_0 := \sum_{m=-N}^N \limsup_{i \rightarrow +\infty} \sqrt{P(i, i+m) P^*(i+m, i)} < 1. \quad (\mathbf{AS2})$$

Moreover $r_{ess}(P|_{\ell^2(\pi)}) \leq \alpha_0$. The main argument to obtain this result is the Doeblin-Fortet inequality in Lemma 2. We refer to [Hen93] for the definition of the essential spectral radius $r_{ess}(T)$ (related to quasi-compactness) of a bounded linear operator T on a Banach space. In Section 3, under the following assumptions

$$\forall m = -N, \dots, N, \quad P(i, i+m) \xrightarrow{i \rightarrow +\infty} a_m \in [0, 1]. \quad (\mathbf{AS3})$$

$$\frac{\pi(i+1)}{\pi(i)} \xrightarrow{i \rightarrow +\infty} \tau \in [0, 1] \quad (\mathbf{AS4})$$

$$\sum_{k=-N}^N k a_k < 0, \quad (\mathbf{NERI})$$

we establish that $(\mathbf{AS2})$ holds (hence (\mathbf{SG}_2)) and that α_0 can be explicitly computed in function of τ and the a_m 's. Observe that (\mathbf{NERI}) means that the expectation of the asymptotic random increments is negative. Moreover, using the inequality $r_{ess}(P|_{\ell^2(\pi)}) \leq \alpha_0$, Property (\mathbf{SG}_2) is proved to be connected to the so-called V -geometric ergodicity of P for $V := (\pi(n)^{-1/2})_{n \in \mathbb{N}}$, which corresponds to the spectral gap property on the usual weighted-supremum space \mathcal{B}_V associated with V . In particular, denoting the minimal V -geometrical ergodic rate by ϱ_V , it is proved that, either ϱ_2 and ϱ_V are both less than α_0 , or $\varrho_2 = \varrho_V$. As a result, an accurate bound of ϱ_2 is obtained for random walks (RW) with i.d. bounded increments using the results of [HL14b]. In the reversible case (Section 4) the previous results hold under Assumptions $(\mathbf{AS3})$ and $(\mathbf{AS4})$ provided that $a_m \neq a_{-m}$ for at least one m . A

first illustration to Birth-and-Death Markov chains is proposed in Subsection 4.1. The reversible case naturally contains the Markov kernels associated with the Metropolis-Hastings (M-H) Algorithm. In Subsection 4.2 we observe that, if the target distribution π and the proposal kernel $Q := (Q(i, j))_{(i, j) \in \mathbb{N}^2}$ satisfy **(AS1)**, **(AS3)** and **(AS4)**, then so is the associated reversible M-H kernel P , which then satisfies **(SG₂)**.

Estimating ϱ_2 is a difficult but relevant issue. This question is investigated in Section 5 where an accurate estimation of ϱ_2 is obtained by using the above mentioned link between ϱ_2 and ϱ_V and by applying the truncation procedure in [HL14a]. Numerical applications to discrete MCMC are presented at the end of Section 5. Bounding ϱ_2 in the reversible case is of special interest since **(SG₂)** holds in this case with $C = 1$ and $\rho = \varrho_2$.

2 **(SG₂)** under Assumption **(AS1)** on P

Theorem 1 *If Condition **(AS2)** holds, then P satisfies **(SG₂)**. Moreover $r_{ess}(P|_{\ell^2(\pi)}) \leq \alpha_0$.*

Proof. Let $\ell^1(\pi)$ denote the usual Banach space of sequences $(f(i))_{i \in \mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$ satisfying the following condition: $\|f\|_1 := \sum_{i \geq 0} |f(i)| \pi(i) < \infty$.

Lemma 1 *The identity map is compact from $\ell^2(\pi)$ into $\ell^1(\pi)$.*

Lemma 2 *For any $\alpha > \alpha_0$, there exists a positive constant $L \equiv L(\alpha)$ such that*

$$\forall f \in \ell^2(\pi), \quad \|Pf\|_2 \leq \alpha \|f\|_2 + L \|f\|_1.$$

It follows from these lemmas and from [Hen93] that P is quasi-compact on $\ell^2(\pi)$ with $r_{ess}(P|_{\ell^2(\pi)}) \leq \alpha$. Since α can be chosen arbitrarily close to α_0 , this gives $r_{ess}(P|_{\ell^2(\pi)}) \leq \alpha_0$. Then **(SG₂)** is deduced from aperiodicity and irreducibility assumptions. \square

Lemma 1 follows from the Cantor diagonal procedure.

Proof of Lemma 2. Under Assumption **(AS1)** we define

$$\forall i \geq i_0, \forall m = -N, \dots, N, \quad \beta_m(i) := \sqrt{P(i, i+m) P^*(i+m, i)}. \quad (2)$$

Let $\alpha > \alpha_0$, with α_0 given in **(AS2)**. Fix $\ell \equiv \ell(\alpha) \geq i_0$ such that $\sum_{m=-N}^N \sup_{i \geq \ell} \beta_m(i) \leq \alpha$. For $f \in \ell^2(\pi)$ we have from Minkowski's inequality and the band structure of P for $i \geq \ell$

$$\begin{aligned} \|Pf\|_2 &\leq \left[\sum_{i < \ell} |(Pf)(i)|^2 \pi(i) \right]^{1/2} + \left[\sum_{i \geq \ell} \left| \sum_{m=-N}^N P(i, i+m) f(i+m) \right|^2 \pi(i) \right]^{1/2} \\ &\leq C_\ell \sum_{i < \ell} |(Pf)(i)| \pi(i) + \left[\sum_{i \geq \ell} \left| \sum_{m=-N}^N P(i, i+m) f(i+m) \right|^2 \pi(i) \right]^{1/2} \end{aligned}$$

where $C_\ell > 0$ is derived from equivalent norms on the space \mathbb{C}^ℓ . Note that $\sum_{i < \ell} |(Pf)(i)| \pi(i) \leq \|Pf\|_1 \leq \|f\|_1$ so that setting $L := C_\ell$

$$\|Pf\|_2 \leq L\|f\|_1 + \left[\sum_{i \geq \ell} \left| \sum_{m=-N}^N P(i, i+m) f(i+m) \right|^2 \pi(i) \right]^{1/2}. \quad (3)$$

It remains to obtain the expected control of the second terms in the right hand side of (3). For $m = -N, \dots, N$, let us define $F_m = (F_m(i))_{i \in \mathbb{N}} \in \ell^2(\pi)$ by

$$F_m(i) := \begin{cases} 0 & \text{if } i < \ell \\ P(i, i+m) f(i+m) & \text{if } i \geq \ell. \end{cases}$$

Then

$$\begin{aligned} & \left[\sum_{i \geq \ell} \left| \sum_{m=-N}^N P(i, i+m) f(i+m) \right|^2 \pi(i) \right]^{1/2} = \left\| \sum_{m=-N}^N F_m \right\|_2 \\ & \leq \sum_{m=-N}^N \|F_m\|_2 = \sum_{m=-N}^N \left[\sum_{i \geq \ell} P(i, i+m)^2 |f(i+m)|^2 \pi(i) \right]^{1/2} \\ & = \sum_{m=-N}^N \left[\sum_{i \geq \ell} P(i, i+m) \frac{\pi(i) P(i, i+m)}{\pi_{i+m}} |f(i+m)|^2 \pi_{i+m} \right]^{1/2} \quad (\text{from the definition of } P^*) \\ & \leq \sum_{m=-N}^N \left(\sup_{i \geq \ell} \beta_m(i) \right) \left[\sum_{i \geq \ell} |f(i+m)|^2 \pi_{i+m} \right]^{1/2} \quad (\text{from (2)}) \\ & \leq \left(\sum_{m=-N}^N \sup_{i \geq \ell} \beta_m(i) \right) \|f\|_2. \end{aligned}$$

The statement in Lemma 2 can be deduced from the previous inequality and from (3). \square

3 (SG₂) and geometric ergodicity. Application to RWs with i.d. bounded increments

We specify Theorem 1 in terms of V -geometric ergodicity for $V := (\pi(n)^{-1/2})_{n \in \mathbb{N}}$. Let $(\mathcal{B}_V, \|\cdot\|_V)$ denote the weighted-supremum space of sequences $(g(n))_{n \in \mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$ such that $\|g\|_V := \sup_{n \in \mathbb{N}} V(n)^{-1} |g(n)| < \infty$. Recall that P is said to be V -geometrically ergodic if P satisfies the spectral gap property on \mathcal{B}_V , namely: there exists $C \in (0, +\infty)$ and $\rho \in (0, 1)$ such that

$$\forall n \geq 1, \forall f \in \mathcal{B}_V, \quad \|P^n f - \Pi f\|_V \leq C \rho^n \|f\|_V. \quad (\mathbf{SG}_V)$$

When this property holds, we define

$$\varrho_V := \inf\{\rho \in (0, 1) : (\mathbf{SG}_V) \text{ holds true}\}. \quad (4)$$

Remark 1 Under Assumptions **(AS3)** and **(AS4)**, we have

$$\alpha_0 := \sum_{m=-N}^N \limsup_{i \rightarrow +\infty} \sqrt{P(i, i+m) P^*(i+m, i)} = \begin{cases} \sum_{m=-N}^N a_m \tau^{-m/2} & \text{if } \tau \in (0, 1) \\ a_0 & \text{if } \tau = 0, \end{cases} \quad (5)$$

Indeed, if **(AS4)** holds with $\tau \in (0, 1)$, then the claimed formula follows from the definition of $P^*(\cdot, \cdot)$. If $\tau = 0$ in **(AS4)**, then $a_m = 0$ for every $m = 1, \dots, N$ since the invariance of π gives $\sum_{m=-N}^N P(i+m, i) \pi(i+m) / \pi(i) = 1$. Thus $a_{-m} = 0$ when $m < 0$. Hence $\alpha_0 = a_0$.

Proposition 1 If P and π satisfy Assumptions **(AS3)**, **(AS4)** and **(NERI)**, then P satisfies **(AS2)** (and $\alpha_0 < 1$ with α_0 given in (5)). Moreover P satisfies both **(SG₂)** and **(SG_V)**, we have $\max(r_{ess}(P|_{\mathcal{B}_V}), r_{ess}(P|_{\ell^2(\pi)})) \leq \alpha_0$, and the following assertions hold:

- (a) if $\varrho_V \leq \alpha_0$, then $\varrho_2 \leq \alpha_0$;
- (b) if $\varrho_V > \alpha_0$, then $\varrho_2 = \varrho_V$.

Proof. If $\tau = 0$ in **(AS4)**, then $\alpha_0 = a_0 < 1$ from (5) and **(NERI)**. Now assume that **(AS4)** holds with $\tau \in (0, 1)$. Then $\alpha_0 = \sum_{m=-N}^N a_m \tau^{-m/2} = \psi(\sqrt{\tau})$, where: $\forall t > 0$, $\psi(t) := \sum_{k=-N}^N a_k t^{-k}$. Moreover it easily follows from the invariance of π that $\psi(\tau) = 1$. Inequality $\alpha_0 = \psi(\sqrt{\tau}) < 1$ then follows from the following assertions: $\forall t \in (\tau, 1)$, $\psi(t) < 1$ and $\forall t \in (0, \tau) \cup (1, +\infty)$, $\psi(t) > 1$. To prove these properties, note that $\psi(\tau) = \psi(1) = 1$ and that ψ is convex on $(0, +\infty)$ since the second derivative of ψ is positive on $(0, +\infty)$. Moreover we have $\lim_{t \rightarrow +\infty} \psi(t) = +\infty$ since $a_k > 0$ for some $k < 0$ (use $\psi(\tau) = \psi(1) = 1$ and $\tau \in (0, 1)$). Similarly, $\lim_{t \rightarrow 0^+} \psi(t) = +\infty$ since $a_k > 0$ for some $k > 0$. This gives the desired properties on ψ since $\psi'(1) > 0$ from **(NERI)**.

(SG₂) and $r_{ess}(P|_{\ell^2(\pi)}) \leq \alpha_0$ follow from Theorem 1. Next **(SG_V)** is deduced from the well-known link (see [MT93]) between geometric ergodicity and the following drift inequality:

$$\forall \alpha \in (\alpha_0, 1), \exists L \equiv L_\alpha > 0, \quad PV \leq \alpha V + L 1_{\mathbb{N}}. \quad (6)$$

This inequality holds from

$$\frac{(PV)(i)}{V(i)} = \sum_{m=-N}^N P(i, i+m) \left(\frac{\pi(i)}{\pi(i+m)} \right)^{\frac{1}{2}} \xrightarrow{i \rightarrow +\infty} \alpha_0.$$

This gives (6), from which **(SG_V)** is derived using aperiodicity and irreducibility. It also follows from (6) that $r_{ess}(P|_{\mathcal{B}_V}) \leq \alpha$ (see [HL14b, Prop. 3.1]). Thus $r_{ess}(P|_{\mathcal{B}_V}) \leq \alpha_0$.

Now we prove (a) and (b) using the spectral properties of [HL14b, Prop. 2.1] of both $P|_{\ell^2(\pi)}$ and $P|_{\mathcal{B}_V}$ (due to quasi-compactness). We will also use the following obvious inclusion: $\ell^2(\pi) \subset \mathcal{B}_V$. In particular every eigenvalue of $P|_{\ell^2(\pi)}$ is also an eigenvalue for $P|_{\mathcal{B}_V}$. First assume that $\varrho_V \leq \alpha_0$. Then there is no eigenvalue for $P|_{\mathcal{B}_V}$ in the annulus $\Gamma := \{\lambda \in \mathbb{C} : \alpha_0 < |\lambda| < 1\}$ since $r_{ess}(P|_{\mathcal{B}_V}) \leq \alpha_0$. From $\ell^2(\pi) \subset \mathcal{B}_V$ it follows that there is also no eigenvalue for

$P|_{\ell^2(\pi)}$ in this annulus. Hence $\varrho_2 \leq \alpha_0$ since $r_{\text{ess}}(P|_{\ell^2(\pi)}) \leq \alpha_0$. Second assume that $\varrho_V > \alpha_0$. Then $P|_{\mathcal{B}_V}$ admits an eigenvalue $\lambda \in \mathbb{C}$ such that $|\lambda| = \varrho_V$. Let $f \in \mathcal{B}_V$, $f \neq 0$, such that $Pf = \lambda f$. We know from [HL14b, Prop. 2.2] that there exists some $\beta \equiv \beta_\lambda \in (0, 1)$ such that $|f(n)| = O(V(n)^\beta) = O(\pi(n)^{-\beta/2})$, so that $|f(n)|^2 \pi(n) = O(\pi(n)^{(1-\beta)})$, thus $f \in \ell^2(\pi)$ from (AS4). We have proved that $\varrho_2 \geq \varrho_V$. Finally the converse inequality is true since every eigenvalue of $P|_{\ell^2(\pi)}$ is an eigenvalue for $P|_{\mathcal{B}_V}$. Thus $\varrho_2 = \varrho_V$. \square

Example 1 (RWs with i.d. bounded increments) *Let P be defined as follows. There exist some positive integers $c, g, d \in \mathbb{N}^*$ such that*

$$\forall i \in \{0, \dots, g-1\}, \quad \sum_{j=0}^c P(i, j) = 1; \quad (7a)$$

$$\forall i \geq g, \forall j \in \mathbb{N}, \quad P(i, j) = \begin{cases} a_{j-i} & \text{if } i-g \leq j \leq i+d \\ 0 & \text{otherwise.} \end{cases} \quad (7b)$$

$$(a_{-g}, \dots, a_d) \in [0, 1]^{g+d+1} : a_{-g} > 0, a_d > 0, \sum_{k=-g}^d a_k = 1. \quad (7c)$$

We assume that P is aperiodic and irreducible, and that Assumption (NERI) holds, that is: $\sum_{k=-g}^d k a_k < 0$. Then P admits a unique invariant distribution π , and the conclusions of Proposition 1 hold. Moreover it can be derived from standard results of linear difference equation that $\pi(n) \sim c\tau^n$ when $n \rightarrow +\infty$, with $\tau \in (0, 1)$ defined by $\psi(\tau) = 1$, where $\psi(t) := \sum_{k=-N}^N a_k t^{-k}$. Thus, if $\gamma := \tau^{-1/2}$, then $\mathcal{B}_V = \{(g(n))_{n \in \mathbb{N}} \in \mathbb{C}^{\mathbb{N}}, \sup_{n \in \mathbb{N}} \gamma^{-n} |g(n)| < \infty\}$. Then we know from [HL14b, Prop. 3.2] that $r_{\text{ess}}(P|_{\mathcal{B}_V}) = \alpha_0$ with α_0 given in (5), and that ϱ_V can be computed from an algebraic polynomial elimination. More precisely, the procedure in [HL14b] developed for a special value $\hat{\gamma}$ can be applied for $\gamma := \tau^{-1/2}$ by considering $\Gamma := \{\lambda \in \mathbb{C} : \psi(\sqrt{\tau}) < |\lambda| < 1\}$. When Assertion (b) of Proposition 1 applies, we obtain the exact value of ϱ_2 (see Example 2). Property (SG₂) is proved in [Wüb12, Th. 2] under an extra weak reversibility assumption (with no explicit bound on ϱ_2). However, except in case $g = d = 1$ where reversibility is automatic, a RW with i.d. bounded increments is not reversible or even weak reversible in general. Note that no reversibility condition is required in Proposition 1.

Example 2 (Numerical examples in case $g = 2$ and $d = 1$) *Let P be defined by*

$$P(0, 0) = a \in (0, 1), \quad P(0, 1) = 1 - a, \quad P(1, 0) = b \in (0, 1), \quad P(1, 2) = 1 - b \quad (8)$$

$$\forall n \geq 2, \quad P(n, n-2) = 1/2, \quad P(n, n-1) = 1/3, \quad P(n, n) = 0, \quad P(n, n+1) = 1/6. \quad (9)$$

The form of boundary probabilities in (8) and the special values in (9) are chosen for convenience. Other (finitely many) boundary probabilities in (8) and other values in (9) could be considered provided that P is irreducible and aperiodic and that $(a_{-2}, a_{-1}, a_0, a_1)$ satisfies $a_{-2}, a_1 > 0$ and (NERI) i.e. $a_1 < 2a_{-2} + a_{-1}$. Here the fonction ψ is given by: $\psi(t) := t^2/2 + t/3 + 1/6t = 1 + (t-1)(t^2 - 5t/3 - 1/3)/2t$. Then function $\psi(\cdot) - 1$ has a unique zero over $(0, 1)$ which is $\tau = (\sqrt{37} - 5)/6 \approx 0.1805$ and $\alpha_0 = \psi(\sqrt{\tau}) \approx 0.6242$. Let $\gamma := 1/\sqrt{\tau} \approx 2.3540$ and $V := (\gamma^n)_{n \in \mathbb{N}}$. Using the procedure from [HL14b] and Proposition 1, we give in Table 1 the values of α_0 , ϱ_V and ϱ_2 for this instance.

(a, b)	α_0	ρ_V	ϱ_2
$(1/2, 1/2)$	0.624	0.624	≤ 0.624
$(1/10, 1/10)$	0.624	0.688	0.688
$(1/50, 1/50)$	0.624	0.757	0.757

Table 1: Convergence rate on $\ell^2(\pi)$ for different boundary transition probabilities (a, b)

Remark 2 If **(AS4)** in Proposition 1 is reinforced by the condition $\pi(n) \sim c\tau^n$ when $n \rightarrow +\infty$ with $\tau \in (0, 1)$ (e.g. see Example 1), then let us consider $\mathcal{B}_V = \{(g(n))_{n \in \mathbb{N}} \in \mathbb{C}^{\mathbb{N}}, \sup_{n \in \mathbb{N}} \gamma^{-n} |g(n)| < \infty\}$ with $\gamma := \tau^{-1/2}$. Then we deduce from [HL14b, Prop. 3.2] that $r_{\text{ess}}(P|_{\mathcal{B}_V}) = \alpha_0$ with α_0 given in (5), so that $\varrho_V \leq \alpha_0$ implies that $\varrho_V = \alpha_0$ since $\varrho_V \geq r_{\text{ess}}(P|_{\mathcal{B}_V})$. Then it follows from Proposition 1 that $\varrho_2 \leq \varrho_V$ and that this inequality is an equality when $\varrho_V > \alpha_0$. The passage from **(SG_V)** to **(SG₂)** and the inequality $\varrho_2 \leq \varrho_V$ was established in [RR97, Bar05] for general reversible V -geometrically ergodic Markov kernels. Again note that no reversibility condition is assumed in Proposition 1.

4 Applications to the reversible case

The reversible case corresponds to the condition $P = P^*$ (i.e. P is self-adjoint in $\ell^2(\pi)$), namely: $\forall (i, j) \in \mathbb{N}^2$, $\pi(i)P(i, j) = \pi(j)P(j, i)$ (detailed balance condition). Then **(SG₂)** is equivalent to the condition $\varrho_2 = \|P - \Pi\|_2 < 1$, where $\|\cdot\|_2$ denotes here the operator norm on $\ell^2(\pi)$. Thus, when **(SG₂)** holds in the reversible case, we have $C = 1$ and $\rho = \varrho_2$, that is

$$\forall n \geq 1, \forall f \in \ell^2(\pi), \quad \|P^n f - \pi(f)\mathbf{1}\|_2 \leq \varrho_2^n \|f\|_2. \quad (10)$$

Corollary 1 *If P is reversible, then:*

1. P satisfies **(SG₂)** and $r_{\text{ess}}(P|_{\ell^2(\pi)}) \leq \alpha_0$ under **(AS2)**, with:

$$\alpha_0 := \sum_{m=-N}^N \left(\limsup_{i \rightarrow +\infty} \sqrt{P(i, i+m)P(i+m, i)} \right) < 1.$$

2. If Condition **(AS3)** holds true, then

$$\alpha_0 = 1 - \sum_{m=1}^N (\sqrt{a_m} - \sqrt{a_{-m}})^2. \quad (11)$$

Consequently, if $a_m \neq a_{-m}$ for at least one $m \in \{1, \dots, N\}$, then P satisfies **(AS2)**.

3. If P satisfies **(AS3)** and if π satisfies **(AS4)** with $\tau \in (0, 1)$, then $a_m \neq a_{-m}$ for each $m \in \{1, \dots, N\}$ and the conclusions of Proposition 1 hold with α_0 given in (11).

4. If P satisfies **(AS3)** with $a_0 < 1$ and if π satisfies **(AS4)** with $\tau = 0$, then the conclusions of Proposition 1 hold.

Keep in mind that all our results are stated for positive recurrent Markov kernels. For instance, for Markov chain associated with $P(i, i-1) := p, P(i, i) := r, P(i, i+1) := q$ where $p+r+q = 1$, Formula (11) is $\alpha_0 = 1 - (\sqrt{q} - \sqrt{p})^2$, but the existence of π is only guaranteed when $p > q$.

Proof. The first statement follows from Theorem 1 and reversibility. Next **(AS3)** gives $\alpha_0 = \sum_{m=-N}^N \sqrt{a_m a_{-m}}$, hence Assertion 2. since $\sum_{m=-N}^N a_m = 1$. If moreover **(AS4)** holds with $\tau \in (0, 1)$, then $a_m \neq a_{-m}$ for every $m \in \{1, \dots, N\}$ since $\tau^m a_{-m} = a_m$ from the balance condition. Thus, under **(AS3)** and **(AS4)** with $\tau \in (0, 1)$, we obtain from Assertion 2. that $\alpha_0 < 1$. Moreover, since the real numbers α_0 given in (11) and in (5) are equal, all the spectral properties obtained in Proposition 1 remain valid. Idem for Assertion 4. from Remark 1. \square

4.1 Birth-and-Death Markov chains (BDMC)

The transition kernel $P := (P(i, j))_{(i, j) \in \mathbb{N}^2}$ of a Birth-and-Death Markov chains is defined by

$$P := \begin{pmatrix} r_0 & q_0 & 0 & \cdots & \cdots \\ p_1 & r_1 & q_1 & \ddots & \\ 0 & p_2 & r_2 & q_2 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}. \quad (12)$$

Recall that, under the following conditions

$$r_0 < 1, \quad \forall i \geq 1, \quad 0 < q_i, p_i < 1, \quad S := 1 + \sum_{i=1}^{\infty} \prod_{j=1}^i \frac{q_{j-1}}{p_j} < \infty, \quad (13)$$

P is irreducible, aperiodic and π (unique) is given by: $\pi(0) = 1/S$, $\pi(i) = (\prod_{j=1}^i \frac{q_{j-1}}{p_j})/S$. Moreover it is well-known that P is reversible w.r.t. π . Finally Condition **(AS2)** writes as:

$$\alpha_0 := \limsup_i \sqrt{p_i q_{i-1}} + \limsup_i r_i + \limsup_i \sqrt{q_i p_{i+1}} < 1. \quad (14)$$

Consequently, under Conditions (13) and (14), P satisfies **(SG₂)** and $r_{ess}(P|_{\ell^2(\pi)}) \leq \alpha_0$. In particular, if the sequences $(p_i)_{i \in \mathbb{N}^*}$, $(r_i)_{i \in \mathbb{N}}$ and $(q_i)_{i \in \mathbb{N}}$ in (12) admit a limit when $i \rightarrow +\infty$, say p, r, q , then **(SG₂)** holds provided that $p > q$. Moreover $r_{ess}(P|_{\ell^2(\pi)}) \leq 1 - (\sqrt{p} - \sqrt{q})^2$.

Example 3 (State-independent BDMC)

Let P given by (12) such that, for any $i \geq 1$, $p_i := p$, $r_i := r$, $q_i := q$, with $p, q, r \in [0, 1]$ such that $p + r + q = 1$ and $p > q > 0$. Let $r_0 \in (0, 1)$ and $\beta_0 := 1 - q - \sqrt{pq}$. The bounds for ϱ_V with $V(n) := (p/q)^{n/2}$ can be derived from [HL14b, Prop. 4.1], so that (Corollary 1):

- if $r_0 \in [\beta_0, 1)$, then $\varrho_2 \leq r + 2\sqrt{pq}$;

• if $r_0 \in (0, \beta_0]$, then :

(a) in case $2p \leq (1 - q + \sqrt{pq})^2$: $\varrho_2 \leq r + 2\sqrt{pq}$;

(b) in case $2p > (1 - q + \sqrt{pq})^2$, setting $\beta_1 := p - \sqrt{pq} - \sqrt{r(r + 2\sqrt{pq})}$:

$$\varrho_2 = \left| r_0 + \frac{p(1 - r_0)}{r_0 - 1 + q} \right| \quad \text{when } r_0 \in (0, \beta_1] \quad (15a)$$

$$\varrho_2 \leq r + 2\sqrt{pq} \quad \text{when } r_0 \in [\beta_1, \beta_0]. \quad (15b)$$

Remark 3 (Discussion on the $\ell^2(\pi)$ -spectral gap and the decay parameter)

Let P be a BDMC satisfying (13). It can be proved that the decay parameter of P , denoted by γ in [vDS95] but by γ_{DS} here to avoid confusion, equals to ϱ_2 , that is (from reversibility): $\gamma_{DS} = \varrho_2 = \|P - \Pi\|_2$. But note that γ_{DS} is only known for specific instances of BDMC from [vDS95] (see [Kov10] for a recent contribution). For a general Markov kernel P , we only have (see also [Pop77, Isa79]) $\gamma_{DS} \leq \varrho_2$. In particular, the decay parameter does not provide information on non-reversible RWs with i.d. bounded increments of Section 3.

4.2 The Metropolis-Hastings Algorithm

Let $\pi = (\pi(i))_{i \in \mathbb{N}}$ (target distribution) be a probability measure on \mathbb{N} known up to a multiplicative constant. Let $Q := (Q(i, j))_{(i, j) \in \mathbb{N}^2}$ (proposal kernel) be any transition kernel on \mathbb{N} . The associated Metropolis-Hastings (M-H) Markov kernel $P := (P(i, j))_{(i, j) \in \mathbb{N}^2}$ is defined by

$$P(i, j) := \begin{cases} \min \left(Q(i, j), \frac{\pi(j)Q(j, i)}{\pi(i)} \right) & \text{if } i \neq j \\ 1 - \sum_{\ell \neq i} P(i, \ell) & \text{if } i = j. \end{cases}$$

It is well-known that P is reversible with respect to π and that π is P -invariant.

Corollary 2 Assume that $\pi(i) > 0$ for every $i \in \mathbb{N}$ and that π satisfies **(AS4)** with $\tau \in (0, 1)$. Assume that Q is an aperiodic and irreducible Markov kernel on \mathbb{N} such that for every $(i, j) \in \mathbb{N}^2$, $Q(i, j) = 0 \Leftrightarrow Q(j, i) = 0$, satisfying **(AS1)** and the following condition (see **(AS3)**)

$$\forall m = -N, \dots, N, \quad q_m := \lim_{i \rightarrow +\infty} Q(i, i + m) \neq 0. \quad (16)$$

Then the associated M-H kernel P satisfies **(SG₂)** and $r_{ess}(P|_{\ell^2(\pi)}) \leq \alpha_0$ with

$$\alpha_0 := 1 - \sum_{m=1}^N (\sqrt{p_m} - \sqrt{p_{-m}})^2 \quad \text{where } p_k := \begin{cases} \min(q_k, \tau^k q_{-k}) & \text{if } k \neq 0 \\ 1 - \sum_{\ell=1}^N (p_\ell + p_{-\ell}) & \text{if } k = 0. \end{cases} \quad (17)$$

If **(AS4)** holds with $\tau = 0$, then $p_m = 0$ for every $m = 1, \dots, N$, and the above conclusions holds true with $\alpha_0 := p_0$.

Proof. It is well-known that P is irreducible and aperiodic under the basic assumptions on Q . If Q satisfies **(AS1)** for some N , then so is P (with the same N). Assumption **(AS3)** holds for P : $\lim_{i \rightarrow +\infty} P(i, i + m) = p_m$ with p_m defined in (17). Then apply Corollary 1. \square

Example 4 Assume that π (possibly known up to a multiplicative constant) is such that $\pi(i) > 0$ for every $i \in \mathbb{N}$ and satisfies **(AS4)**. Let Q be a transition kernel on \mathbb{N} satisfying

$$Q(0,0) := r < 1, \quad Q(0,1) := 1-r, \quad \forall i \geq 1, \quad Q(i,i-1) = q, \quad Q(i,i) = 1-2q, \quad Q(i,i+1) = q$$

for some $q \in (0, 1/2]$. The associated M-H Markov kernel $P^{(q)}$ is given by $P^{(q)}(0,1) = \min(1-r, q\pi(1)/\pi(0))$ and

$$\begin{aligned} \forall i \geq 1, \quad P^{(q)}(i,i-1) &= q \min\left(1, \frac{\pi(i-1)}{\pi(i)}\right) & P^{(q)}(i,i+1) &= q \min\left(1, \frac{\pi(i+1)}{\pi(i)}\right) \\ P^{(q)}(i,i) &:= 1 - \sum_{\ell \neq i} P^{(q)}(i,\ell). \end{aligned}$$

The conditions of Corollary 2 are trivially satisfied. Then $P^{(q)}$ satisfies **(SG₂)**. Next $\alpha_0 \equiv \alpha_0(q)$ in (17) is

$$\alpha_0(q) = 1 - q(1 - \sqrt{\tau})^2 \tag{18}$$

since the p_m 's in (17) are given by $p_{-1} = q$, $p_0 = 1 - q - q\tau$, $p_1 = q\tau$. When $q \in (0, 1/2]$, $\alpha_0(q)$ is minimal for $q = 1/2$, thus $q = 1/2$ provides the minimal bound for $r_{\text{ess}}(P_{|\ell^2(\pi)}^{(q)})$. The relevant question is to find $q \in (0, 1/2]$ providing the minimal value of $\varrho_2 \equiv \varrho_2(q)$ (See Example 6).

Example 5 (Simulation of Poisson distribution with parameter 1) Let π be the Poisson distribution with parameter $\lambda := 1$, defined by $\pi(i) := \exp(-1)/i!$. Then **(AS4)** holds with $\tau = 0$. Introduce the proposal kernel Q of Example 4 with $r := 1/2$ and $q := 1/2$. The associated M-H kernel P is given by $P^{(q)}(0,0) = P^{(q)}(0,1) = 1/2$ and

$$\forall i \geq 1, \quad P^{(q)}(i,i-1) = \frac{1}{2}, \quad P^{(q)}(i,i) = \frac{i}{2(i+1)}, \quad P^{(q)}(i,i+1) = \frac{1}{2(i+1)}.$$

We know from Example 4 that $P^{(q)}$ satisfies **(SG₂)** and $r_{\text{ess}}(P_{|\ell^2(\pi)}^{(q)}) \leq \alpha_0 = 1/2$. The rate of convergence $\varrho_2 \equiv \varrho_2(q)$ of $P^{(q)}$ is studied in Example 7.

5 Bound for ϱ_2 via truncation and numerical applications

Let us consider the following k -th truncated (and augmented) matrix P_k associated with P :

$$\forall (i,j) \in \{0, \dots, k-1\}^2, \quad P_k(i,j) := \begin{cases} P(i,j) & \text{if } 0 \leq i \leq k-1 \text{ and } 0 \leq j \leq k-2 \\ \sum_{\ell \geq k-1} P(i,\ell) & \text{if } 0 \leq i \leq k-1 \text{ and } j = k-1. \end{cases}$$

Let $\sigma(P_k)$ denote the set of eigenvalues of P_k , and define $\rho_k := \max\{|\lambda|, \lambda \in \sigma(P_k), |\lambda| < 1\}$.

Recall that $V(i) := \pi(i)^{-1/2}$ and that ϱ_V is defined in (4). The statement below follows from Proposition 1 and from the weak perturbation method in [HL14a] applied to $P|_{\mathcal{B}_V}$, for which the drift inequality (6) plays an important role.

Proposition 2 *If P satisfies (AS3), (AS4) and (NERI), then the following properties holds with α_0 given in (5):*

- (a) $\varrho_2 \leq \alpha_0 \iff \varrho_V \leq \alpha_0$, and in this case we have $\limsup_k \rho_k \leq \alpha_0$;
- (b) $\varrho_2 > \alpha_0 \iff \varrho_V > \alpha_0$, and in this case we have $\varrho_2 = \varrho_V = \lim_k \rho_k$.

Below the estimation of the convergence rate ϱ_2 for some Metropolis-Hastings Markov kernel P is derived from Proposition 2. Recall that Inequality (10) applies when P is reversible. The generic procedure for the following instances of Markov kernel P is as follows:

1. Compute α_0 given in (17) and choose a small $\varepsilon > 0$
2. $k := 2$
3. Consider the k -order truncated matrix P_k of the kernel P .
4. Compute the second highest eigenvalue ρ_k of P_k .
5. If $|\rho_k - \rho_{k-1}| > \varepsilon$ then ($k := k + 1$, return to step 3)
 else if $\rho_k > \alpha_0$ then $\varrho_2 \simeq \rho_k$ else $\varrho_2 \leq \alpha_0$.

It is clear that the control of the stabilization of the sequence $\{\rho_k\}_{k \geq 2}$ through the comparison between $|\rho_k - \rho_{k-1}|$ and ε only provides an estimation of ϱ_2 .

Example 6 (Example 4 (continued))

Let us consider the probability distribution π given by $\pi(i) := C(i+1)\tau^i$ for $n \in \mathbb{N}$ where C is a (possibly unknown) normalisation constant and $0 < \tau < 1$. Then (AS4) is satisfied. If we choose an RW as in Example 4 for the proposal kernel, the associated M-H kernel $P^{(q)}$ is defined by $P^{(q)}(0, 1) = \min(1 - r, 2q\tau)$ and

$$\forall i \geq 1, \quad P^{(q)}(i, i-1) = q \min\left(1, \frac{1}{\tau} \frac{i}{i+1}\right) \quad P^{(q)}(i, i+1) = q \min\left(1, \tau \frac{i+2}{i+1}\right)$$

$$P^{(q)}(i, i) := 1 - \sum_{\ell \neq i} P^{(q)}(i, \ell).$$

For $q \in (0, 1/2]$, $P^{(q)}$ satisfies (SG₂) with $r_{ess}(P_{|\ell^2(\pi)}^{(q)}) \leq \alpha_0(q) = 1 - q(1 - \sqrt{\tau})^2$ (see (18)). Table 2 based on Proposition 2 gives the estimate of $\varrho_2(q)$ of $P^{(q)}$.

Example 7 (Example 5 (continued)) *Table 3 based on Proposition 2 gives the estimation of $\varrho_2(q)$ of the M-H $P^{(q)}$ used in the simulation of the Poisson distribution of Example 5. Note that $q := 1/2$ gives the smallest value of $r_{ess}(P_{|\ell^2(\pi)}^{(q)}) = \alpha_0(q) = 0.5$, with α_0 given by (18). However $q := 1/2$ does not provide the minimal rate of convergence in $\ell^2(\pi)$ -norm (or in \mathcal{B}_V -norm). More precisely, for $q = 1/2$, the kernel $P^{(q)}$ admits some eigenvalues in the annulus $\Gamma := \{\lambda \in \mathbb{R} : 0.5 < |\lambda| < 1\}$, among which $\varrho_2(q) \approx 0.8090$ is the larger one in absolute value. Actually the minimal rate of convergence is achieved at $q \approx 0.38$ and note that every value $0.2 \leq q < 0.5$ in Table 3 provides a minimal rate than for $q := 1/2$. It could be conjectured from numerical evidence that for $q \leq q_0$ with $q_0 \approx 0.35$, $\varrho_2 = \alpha_0(q)$.*

		$\tau = 0.2$			$\tau = 0.5$		
q	$\alpha_0(q, \tau)$	$\rho_k(q)$	$\varrho_2(q)$	$\alpha_0(q, \tau)$	$\rho_k(q)$	$\varrho_2(q)$	
0.1	0.9694	$\rho_{27} \simeq 0.9710$	$\simeq 0.9710$	0.9914	$\rho_{39} \simeq 0.9921$	$\simeq 0.9921$	
0.2	0.9389	$\rho_{30} \simeq 0.9421$	$\simeq 0.9421$	0.9828	$\rho_{44} \simeq 0.9842$	$\simeq 0.9842$	
0.3	0.9083	$\rho_{31} \simeq 0.9131$	$\simeq 0.9131$	0.9743	$\rho_{47} \simeq 0.9763$	$\simeq 0.9763$	
0.4	0.8778	$\rho_{32} \simeq 0.8842$	$\simeq 0.8842$	0.9657	$\rho_{50} \simeq 0.9684$	$\simeq 0.9684$	
0.5	0.8472	$\rho_{33} \simeq 0.8552$	$\simeq 0.8552$	0.9571	$\rho_{51} \simeq 0.9605$	$\simeq 0.9605$	
		$\tau = 0.6$			$\tau = 0.8$		
q	$\alpha_0(q, \tau)$	$\rho_k(q)$	$\varrho_2(q)$	$\alpha_0(q, \tau)$	$\rho_k(q)$	$\varrho_2(q)$	
0.1	0.9949	$\rho_{44} \simeq 0.9953$	$\simeq 0.9953$	0.99889	$\rho_{55} \simeq 0.99883$	≤ 0.99889	
0.2	0.9898	$\rho_{51} \simeq 0.9906$	$\simeq 0.9906$	0.99777	$\rho_{66} \simeq 0.99781$	$\simeq 0.99781$	
0.3	0.9848	$\rho_{55} \simeq 0.9860$	$\simeq 0.9860$	0.99666	$\rho_{73} \simeq 0.9968$	$\simeq 0.9968$	
0.4	0.9797	$\rho_{58} \simeq 0.9814$	$\simeq 0.9814$	0.99554	$\rho_{79} \simeq 0.99579$	$\simeq 0.99579$	
0.5	0.9746	$\rho_{60} \simeq 0.9767$	$\simeq 0.9767$	0.99443	$\rho_{83} \simeq 0.9948$	$\simeq 0.9948$	

Table 2: Results for different values of τ with $\varepsilon = 10^{-5}$. The second eigenvalue $\rho_k \equiv \rho_k(q)$ of $P_k \equiv P^{(q)}_k$ is obtained from the observed empirical stabilization of ρ_k with respect to k .

q	$\alpha_0(q) \equiv r_{ess}(P^{(q)})$	$\rho_k(q)$	$\varrho_2(q)$
0.1	0.9	$\rho_{37} \simeq 0.9003$	$\simeq 0.9003$
0.2	0.8	$\rho_{83} \simeq 0.8008$	$\simeq 0.8008$
0.3	0.7	$\rho_{151} \simeq 0.7015$	$\simeq 0.7015$
0.38	0.62	$\rho_{61} \simeq 0.6301$	$\simeq 0.6301$
0.4	0.6	$\rho_{17} \simeq 0.6568$	$\simeq 0.6568$
0.5	0.5	$\rho_{14} \simeq 0.8090$	$\simeq 0.8090$

Table 3: $\rho_k \equiv \rho_k(q)$ is obtained from the observed empirical stabilization of ρ_k with $\varepsilon = 10^{-5}$.

References

- [Bax05] P. H. Baxendale. Renewal theory and computable convergence rates for geometrically ergodic Markov chains. *Ann. Appl. Probab.*, 15(1B):700–738, 2005.
- [Hen93] H. Hennion. Sur un théorème spectral et son application aux noyaux lipchitziens. *Proc. Amer. Math. Soc.*, 118:627–634, 1993.
- [HL14a] L. Hervé and J. Ledoux. Approximating Markov chains and V -geometric ergodicity via weak perturbation theory. *Stochastic Process. Appl.*, 124(1):613–638, 2014.
- [HL14b] L. Hervé and J. Ledoux. Spectral analysis of Markov kernels and application to the convergence rate of discrete random walks. *Adv. in Appl. Probab.*, 46(4):1036–1058, 2014.
- [Isa79] D. Isaacson. A characterization of geometric ergodicity. *Z. Wahrsch. Verw. Gebiete*, 49(3):267–273, 1979.
- [Kov10] Y. Kovchegov. Orthogonality and probability: mixing times. *Electron. Commun. Probab.*, 15:59–67, 2010.
- [MT93] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Springer-Verlag London Ltd., London, 1993.
- [Pop77] N. N. Popov. Geometric ergodicity conditions for countable Markov chains. *Dokl. Akad. Nauk SSSR*, 234(2):316–319, 1977.
- [RR97] G. O. Roberts and J. S. Rosenthal. Geometric ergodicity and hybrid Markov chains. *Elect. Comm. in Probab.*, 2:13–25, 1997.
- [vDS95] E. A. van Doorn and P. Schrijner. Geometric ergodicity and quasi-stationarity in discrete-time birth-death processes. *J. Austral. Math. Soc. Ser. B*, 37(2):121–144, 1995.
- [Wüb12] A. Wübker. Spectral theory for weakly reversible Markov chains. *J. Appl. Probab.*, 49(1):245–265, 2012.