Quasi-compactness of Markov kernels on weighted-supremum spaces and geometrical ergodicity

Denis Guibourg, Loïc Hervé, James Ledoux

Quasi-compactness of Markov kernels on weighted-supremum spaces and geometrical ergodicity

Denis GUIBOURG, Loïc HERVÉ, and James LEDOUX *

14 Octobre 2011

Abstract

Let P be a Markov kernel on a measurable space $\mathbb{X}$ and let $V : \mathbb{X} \to [1, +\infty)$. We provide various assumptions, based on drift conditions, under which P is quasi-compact on the weighted-supremum Banach space $(\mathcal{B}_V, \|\cdot\|_V)$ of all the measurable functions $f : \mathbb{X} \to \mathbb{C}$ such that $\|f\|_V := \sup_{x\in\mathbb{X}} |f(x)|/V(x) < \infty$. Furthermore we give bounds for the essential spectral radius of P. Under additional assumptions, these results allow us to derive the convergence rate of P on $\mathcal{B}_V$, that is the geometric rate of convergence of the iterates P^n to the stationary distribution in operator norm. Applications to discrete Markov kernels and to iterated function systems are presented.

AMS subject classification : 60J10; 47B07

Keywords : Markov chain, drift condition, essential spectral radius, convergence rate.

Contents

*Université Européenne de Bretagne, I.R.M.A.R. (UMR-CNRS 6625), Institut National des Sciences Appliquées de Rennes. Denis.Guibourg@ens.insa-rennes.fr, Loic.Herve@insa-rennes.fr, James.Ledoux@insa-rennes.fr

1 Introduction

Let P be a Markov kernel on a measurable space $(\mathbb{X}, \mathcal{X})$. Let us introduce the weighted-supremum Banach space $(\mathcal{B}_V, \|\cdot\|_V)$ composed of measurable functions $f : \mathbb{X} \to \mathbb{C}$ such that

$$\|f\|_V := \sup_{x \in \mathbb{X}} \frac{|f(x)|}{V(x)} < \infty$$

where $V : \mathbb{X} \to [1, +\infty)$. Let $(\mathcal{B}_0, \|\cdot\|_0)$ be the usual Banach space composed of all the bounded measurable functions $f : \mathbb{X} \to \mathbb{C}$ equipped with the supremum norm $\|f\|_0 := \sup_{x \in \mathbb{X}} |f(x)|$.

The first purpose of the paper is to study the quasi-compactness of P on $\mathcal{B}_V$ with a control of its essential spectral radius $r_{ess}(P)$. Recall that $r_{ess}(P)$ is the infimum bound of the positive real numbers r_0 for which the following property holds: the spectral values of P of modulus greater than r_0 are finitely many eigenvalues having a finite-dimensional characteristic space. P is said to be quasi-compact on $\mathcal{B}_V$ if $r_{ess}(P)$ is strictly less than the spectral radius of P (see Section 2 for details). The second purpose of the paper is to specify the link between quasi-compactness and the so-called V-geometric ergodicity [MT93], namely with the convergence of P^n to π in operator norm on $\mathcal{B}_V$, where π denotes the P-invariant probability measure. In this case, we are interested in finding upper bounds for the convergence rate $\rho_V(P)$ defined by

$$\rho_V(P) := \inf \big\{ \rho \in (0,1), \sup_{\|f\|_V \le 1} \|P^n f - \pi(f)\|_V = O(\rho^n) \big\}. \tag{1}$$

Finally the third purpose of the paper is to derive the V-geometric ergodicity of P, with a control of $\rho_V(P)$, from some strong ergodicity property assumed to hold on a subspace of $\mathcal{B}_V$.

Note that this paper is not directly concerned with the convergence rate of Markov chains, either with respect to the Lebesgue space $\mathbb{L}^2(\pi)$ (in place of $\mathcal{B}_V$) as studied for instance in [AP07] for Hastings and Metropolis algorithms using operator methods (see [FHL10, Section 2] for an overview), or with respect to $\mathcal{B}_V$ with a bounded function V, that is for uniformly ergodic Markov chains.

Let us give an account of the main results of the paper in regards to the previous objectives. Under irreducibility and aperiodicity assumptions, it is well-known that the V-geometric ergodicity is equivalent to the following so-called drift condition:

$$\exists m_0 \in \mathbb{N}^*,\ \exists \varrho \in (0,1),\ \exists M \in (0, +\infty), \quad P^{m_0} V \le \varrho V + M\, 1_S, \tag{\textbf{D}}$$

where $S \subset \mathbb{X}$, called a small set, satisfies the minorization condition

$$\forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad P^{m_0}(x, A) \ge \nu(A)\, 1_S(x), \tag{\textbf{S}}$$

for some positive measure ν on $(\mathbb{X}, \mathcal{X})$ such that $\nu(S) > 0$ (see [MT93]). In Theorem 1 (Subsection 2.1), without assuming any irreducibility or aperiodicity conditions, the quasi-compactness of P on $\mathcal{B}_V$ is proved under Conditions (**D**)-(**S**). This is an expected result, already obtained in [Hen06, Hen07]. We provide a simple and short proof of Theorem 1 which enables to well understand why the drift condition implies good spectral properties of P on

$\mathcal{B}_V$. Furthermore we obtain the following upper bound for $r_{ess}(P)$ which is more explicit than in [Hen06]:

$$r_{ess}(P) \leq \left(\frac{\varrho\,\nu(1_{\mathbb{X}}) + M}{\nu(1_{\mathbb{X}}) + M}\right)^{\frac{1}{m_0}}. \tag{2}$$

In Theorem 2 (Subsection 2.2), assuming that P is a compact operator from $\mathcal{B}_0$ to $\mathcal{B}_V$, P is shown to be power-bounded and quasi-compact on $\mathcal{B}_V$ under the following weak drift condition

$$\exists N \in \mathbb{N}^*,\ \exists d \in (0, +\infty),\ \exists \delta \in (0,1), \quad P^N V \leq \delta^N\, V + d\, 1_{\mathbb{X}}. \tag{\textbf{WD}}$$

Such a condition with $N = 1$ has been introduced in [MT93, Lem. 15.2.8] as an alternative to the drift condition [MT93, (V4)] under suitable assumption on V. Under Condition (**WD**), let us define the real number $\delta_V(P) \in (0,1)$ as the infimum of the real numbers $\delta \in [0,1)$ such that we have (**WD**):

$$\delta_V(P) := \inf\big\{\delta \in [0,1) : \exists N \in \mathbb{N}^*, \exists d \in (0,+\infty),\ P^N V \leq \delta^N\, V + d\, 1_{\mathbb{X}}\big\}. \tag{3}$$

Then the upper bound obtained in Theorem 2 for $r_{ess}(P)$ is more precise than (2), that is:

$$r_{ess}(P) \leq \delta_V(P).$$

The key idea to prove Theorem 2 is that Condition (**WD**) yields a Doeblin-Fortet inequality on the dual of $\mathcal{B}_V$. This fact has been already used in [FHL] to study regular perturbations of V-geometrically ergodic Markov chains.

When the Markov kernel P has an invariant probability distribution, the connection between the V-geometric ergodicity and the quasi-compactness of P is recalled in Theorem 3 (Subsection 2.3). Namely, P is V-geometrically ergodic if and only if P is a power-bounded quasi-compact operator on $\mathcal{B}_V$ for which $\lambda = 1$ is a simple eigenvalue and the unique eigenvalue of modulus one. In this case, if $\mathcal{V}$ denotes the set of all the eigenvalues λ of P such that $r_{ess}(P) < |\lambda| < 1$, then the convergence rate $\rho_V(P)$ is given by:

$$\rho_V(P) = r_{ess}(P) \ \text{ if } \ \mathcal{V} = \emptyset \quad \text{ and } \quad \rho_V(P) = \max\{|\lambda|,\, \lambda \in \mathcal{V}\} \ \text{ if } \ \mathcal{V} \neq \emptyset.$$

This result is valid for any quasi-compact operator, however we have not found such an explicit result in the literature on V-geometric ergodicity.

Theorem 4 proved in Subsection 2.4 is of great interest to investigate the eigenvalues of modulus one and the above set $\mathcal{V}$ in order to obtain the V-geometric ergodicity of P and, more importantly, an upper bound for $\rho_V(P)$ from Theorem 3. Namely, under Condition (**WD**), for any $\lambda \in \mathbb{C}$ such that $\delta \leq |\lambda| \leq 1$ where δ is given in (**WD**), and for any $p \in \mathbb{N}^*$, we obtain with $\beta(\lambda) := \ln|\lambda| / \ln \delta$:

$$f \in \mathcal{B}_V \cap \operatorname{Ker}(P - \lambda I)^p \ \Rightarrow\ \exists c \in (0, +\infty),\ |f| \leq c\, (\ln V)^{p(p-1)/2}\, V^{\beta(\lambda)}. \tag{4}$$

In particular, if λ is an eigenvalue such that $|\lambda| = 1$, then any associated eigen-function f is bounded on $\mathbb{X}$. By contrast, if $|\lambda|$ is close to $\delta_V(P)$, then $|f| \leq c\, V^{\beta(\lambda)}$ with $\beta(\lambda)$ close to 1.

In Section 3, applications of Theorems 2-4 to discrete Markov chains are presented. When $X := \mathbb{N}$ and $\lim_n V(n) = +\infty$, any Markov kernel P is compact from $\mathcal{B}_0$ to $\mathcal{B}_V$ from Corollary 1, and Theorem 2-Theorem 3 are then specially relevant: if P satisfies Condition (**WD**), then P is power-bounded and quasi-compact on $\mathcal{B}_V$; if in addition P is irreducible and aperiodic, then P is V-geometrically ergodic. In Subsection 3.4, Property (4) is used to compute the convergence rate $\rho_V(P)$ for binary random walks.

In Section 4, we are interested in the following question. Assume that there exists a subspace $\mathcal{B} \subset \mathcal{B}_V$ and a constant $\kappa_{\mathcal{B}} \in (0,1)$ such that

$$\forall f \in \mathcal{B}, \quad \|P^n f - \pi(f)1_{\mathbb{X}}\|_V = O(\kappa_{\mathcal{B}}^n). \tag{5}$$

Under what additional conditions do we have the V-geometrical ergodicity of P? Even if (5) has its own interest, the previous issue is of value when, for instance, the indicator functions $1_A,\ A \in \mathcal{X}$, are not in $\mathcal{B}$, since in this case, (5) does not provide estimates for the probabilities $\mathbb{P}(X_n \in A \,|\, X_0 = x) = (P^n 1_A)(x)$. In Theorem 5, assuming that P is compact from $\mathcal{B}_0$ to $\mathcal{B}_V$ and satisfies (**WD**), we prove that a natural additional condition to derive the V-geometric ergodicity from (5) is the following one: $\exists \tau \in [0,1],\ P(\mathcal{B}_{V^\tau}) \subset \mathcal{B}$. Furthermore we obtain that

$$\rho_V(P) \le \max\left(\kappa_{\mathcal{B}}, \delta_V(P)^\tau\right).$$

Property (4) is useful for proving this result. The case when $\mathcal{B}$ is a weighted-Lipschitz space is specially relevant. Applications to iterated function systems (IFS) are presented in Section 5.

To the best of our knowledge, Theorems 1-2-4 of Section 2 are new, as well as Theorem 5 of Section 4. Furthermore we try to provide upper bounds of the essential spectral radius $r_{ess}(P)$ and of the convergence rate $\rho_V(P)$ as explicit as possible. More precisely, in this work, the estimates of $\rho_V(P)$ are derived from the upper bound $r_{ess}(P) \le \delta_V(P)$ (Theorem 2) and from the precise study of the eigen-functions of P belonging to the eigenvalues λ such that $\delta_V(P) < |\lambda| < 1$ (Theorem 4). This approach is original and often provides the exact value of $\rho_V(P)$. But it is worth noticing that this method can only be used for Markov kernels P assumed to be compact from $\mathcal{B}_0$ to $\mathcal{B}_V$. Classical instances of V-geometrically ergodic Markov kernels concern the discrete state-space, the autoregressive models on $\mathbb{X} = \mathbb{R}^q$ with absolutely continuous noise with respect to the Lebesgue measure, and finally the MCMC algorithms. The compactness assumption on $P : \mathcal{B}_0 \to \mathcal{B}_V$ is fulfilled for the two first instances, see Section 3 and Subsection 5.4. Unfortunately, this assumption does not hold in general for non-discrete Markov kernels arising from Hastings and Metropolis algorithms. Concerning the last issue, we refer to the works [Bax05, Lun97, LT96, MT96, MT94, RT99] where the convergence rate $\rho_V(P)$ is investigated by probabilistic methods. The best rates are obtained in [LT96] under the stochastic monotonicity assumption for P which cannot be compared with our compactness assumption.

Throughout the paper, when the function $V \equiv V_c$ depends on some parameter c, we use the notation $\mathcal{B}_c \equiv \mathcal{B}_{V_c}$ so that $\mathcal{B}_c$ may stand for different sets from section to section.

2 Quasi-compactness on $\mathcal{B}_V$ and V-geometric ergodicity

Let $(\mathcal{B}, \|\cdot\|)$ be a complex Banach space, and let L be a bounded linear operator on $\mathcal{B}$ with positive spectral radius $r(L) := \lim_n \|L^n\|^{1/n}$, where $\|\cdot\|$ also stands for the operator norm on $\mathcal{B}$. For the sake of simplicity, we assume that $r(L) := 1$ (if not, replace L with $r(L)^{-1}L$). The restriction of L to a L-invariant subspace H is denoted by $L_{|H}$, and I stands for the identity operator on $\mathcal{B}$.

The simplest definition of quasi-compactness is the following (compare the definition below with the reduction of matrices or compact operators).

Definition 2.1 *L is quasi-compact on $\mathcal{B}$ if there exist $r_0 \in (0,1)$ and $m \in \mathbb{N}^*$, $\lambda_i \in \mathbb{C}$, $p_i \in \mathbb{N}^*$ ($i = 1, \ldots, m$) such that:*

$$\mathcal{B} = \overset{m}{\underset{i=1}{\oplus}} \operatorname{Ker}(L - \lambda_i I)^{p_i} \oplus H, \tag{6a}$$

where the λ_i's are such that

$$|\lambda_i| \geq r_0 \quad \text{ and } \quad 1 \leq \dim \operatorname{Ker}(L - \lambda_i I)^{p_i} < \infty, \tag{6b}$$

and H is a closed L-invariant subspace such that

$$\sup_{h \in H,\, \|h\|_V \leq 1} \|P^n h\|_V = O({r_0}^n). \tag{6c}$$

Concerning the essential spectral radius of L, denoted by $r_{ess}(L)$, here it is enough to have in mind that, if L is quasi-compact on $\mathcal{B}$, then we have (see for instance [Hen93])

$$r_{ess}(L) := \inf \big\{ r_0 \in (0,1) \text{ s.t. we have (6a) (6b) (6c)} \big\}$$

It is also well-known (e.g. see [Nev64, Kre85]) that $r_{ess}(L)$ is defined by

$$r_{ess}(L) := \lim_n \big(\inf \|L^n - K\| \big)^{\frac{1}{n}} \tag{7}$$

where the infimum is taken over the ideal of compact operators K on $\mathcal{B}$. Consequently L is quasi-compact if and only if there exist some $n_0 \in \mathbb{N}^*$ and some compact operator K_0 on $\mathcal{B}$ such that $r(L^{n_0} - K_0) < 1$. Under the previous condition we have

$$r_{ess}(L) \leq (r(L^{n_0} - K_0))^{1/n_0}. \tag{8}$$

Indeed, for all $k \geq 1$ we have $\|(L^{n_0} - K_0)^k\|^{1/(n_0 k)} = \|L^{n_0 k} - K_k\|^{1/(n_0 k)}$ with some compact operator K_k on $\mathcal{B}$. Then (7) gives: $r_{ess}(L) \leq \lim_n \|(L^{n_0} - K_0)^k\|^{1/(n_0 k)} = (r(L^{n_0} - K_0))^{1/n_0}$.

Throughout the paper, we consider a function $V : \mathbb{X} \to [1, +\infty)$ and a Markov kernel P on $(\mathbb{X}, \mathcal{X})$ such that PV/V is bounded on $\mathbb{X}$ (i.e. $\|PV\|_V < \infty$). So P continuously acts on $\mathcal{B}_V$.

2.1 Quasi-compactness on $\mathcal{B}_V$ under the drift condition

Theorem 1 *Let us assume that the Conditions* (**D**)-(**S**) *in Introduction hold true. Then P is a power-bounded quasi-compact operator on $\mathcal{B}_V$ with*

$$r_{ess}(P) \leq \left(\frac{\varrho\, \nu(1_{\mathbb{X}}) + M}{\nu(1_{\mathbb{X}}) + M} \right)^{\frac{1}{m_0}}. \tag{9}$$

The proof of Theorem 1 is based on the next lemma.

Lemma 1 *Let $Q(x,dy)$ be a Markov kernel on $(\mathbb{X},\mathcal{X})$ having a continuous action on $\mathcal{B}_V$ (i.e. $\|QV\|_V < \infty$) such that $Q = A + B$ for some nonnegative bounded linear operators A and B on $\mathcal{B}_V$. Let $r(B)$ denote the spectral radius of B which is assumed to be positive.*
Then, there exists a nontrivial nonnegative continuous linear form η on $\mathcal{B}_V$ such that $\eta\circ B = r(B)\,\eta$ and $\eta(A1_{\mathbb{X}}) = (1-r(B))\eta(1_{\mathbb{X}})$.

Proof. Since $B \geq 0$ and $r := r(B) > 0$, we know from [Sch71, App., Cor.2.6] that there exists a nontrivial nonnegative continuous linear form η on $\mathcal{B}_V$ such that $\eta\circ B = r\,\eta$ (see also Remark 1 and Appendix A). From $Q = A + B$, we have $\eta\circ Q = \eta\circ A + r\,\eta$, thus $\eta(Q1_{\mathbb{X}}) = \eta(1_{\mathbb{X}}) = \eta(A1_{\mathbb{X}}) + r\,\eta(1_{\mathbb{X}})$. Hence $\eta(A1_{\mathbb{X}}) = (1-r)\eta(1_{\mathbb{X}})$. □

Proof of Theorem 1. Condition (**D**) implies that $P^{m_0}V \leq \varrho V + M\,1_{\mathbb{X}}$. Iterating this inequality easily ensures that $\sup_k \|P^{km_0}V\|_V < \infty$. Thus P is power-bounded on $\mathcal{B}_V$, namely: $\sup_n \|P^nV\|_V < \infty$ (use the Euclidean division of $n\in\mathbb{N}$ by m_0). Then, from $P1_{\mathbb{X}} = 1_{\mathbb{X}}$ and $1_{\mathbb{X}} \in \mathcal{B}_V$, we have $r(P) = 1$. Besides, since $\|PV\|_V < \infty$, we deduce from (**S**) that $\nu(V) < \infty$. Thus we can define the following rank-one operator on $\mathcal{B}_V$: $Tf := \nu(f)\,1_S$. Let $Q := P^{m_0}$ and $R := Q - T$. From $T \geq 0$ and from (**S**), it follows that $0 \leq R \leq Q$, so $r(R) \leq 1$. Let us set $r := r(R)$. If $r = 0$, then P is quasi-compact with $r_{ess}(P) = 0$ from (8). Now assume that $r \in (0,1]$. Then, from Lemma 1, there exists a nontrivial nonnegative continuous linear form η on $\mathcal{B}_V$ such that $\eta\circ R = r\,\eta$ and $\eta(T1_{\mathbb{X}}) = (1-r)\eta(1_{\mathbb{X}})$, from which we deduce that

$$\eta(1_S) = \frac{(1-r)\eta(1_{\mathbb{X}})}{\nu(1_{\mathbb{X}})} \leq \frac{(1-r)\eta(V)}{\nu(1_{\mathbb{X}})}.$$

Besides, we have $RV \leq QV \leq \varrho V + M\,1_S$, hence

$$r\,\eta(V) = \eta(RV) \leq \varrho\,\eta(V) + M\,\eta(1_S) \leq \varrho\,\eta(V) + \frac{M(1-r)\eta(V)}{\nu(1_{\mathbb{X}})}.$$

Since $\eta \neq 0$, we have $\eta(V) > 0$, and since $\varrho \in (0,1)$, we cannot have $r = 1$. Thus $r \in (0,1)$, and P is quasi-compact from (8) with

$$r_{ess}(P) \leq r^{1/m_0} \leq \left(\frac{\varrho\,\nu(1_{\mathbb{X}}) + M}{\nu(1_{\mathbb{X}}) + M}\right)^{1/m_0}.$$

□

Remark 1 *The proof of Lemma 1 is based on the following result [Sch71, App., Cor.2.6]: if L is a positive operator on a Banach lattice $\mathcal{B}$ whose positive cone is normal and has interior points, then there exists a nontrivial nonnegative continuous linear form e' on $\mathcal{B}$ such that $e'\circ L = r(L)\,e'$. In fact $\mathcal{B}_V$ is the simplest (and generic) example of Banach lattices satisfying the last conditions, and we give in Appendix A a proof of the previous statement in this special case. Mention that this result also provides that the quasi-compactness of P on $\mathcal{B}_V$ is equivalent to the mean ergodicity with finite rank limit projection (see [Her08]).*

2.2 Quasi-compactness on $\mathcal{B}_V$ under the weak drift condition (WD)

Recall that $(\mathcal{B}_0, \|\cdot\|_0)$ denotes the Banach space of all the bounded measurable functions $f : \mathbb{X} \to \mathbb{C}$, equipped with the supremum norm $\|f\|_0 := \sup_{x\in\mathbb{X}} |f(x)|$, and that $\delta_V(P)$ is the infimum of the real numbers $\delta \in [0,1)$ such that we have (**WD**) (see (3)).

Theorem 2 *If Condition* (**WD**) *holds true and if $P : \mathcal{B}_0 \to \mathcal{B}_V$ is compact, then P is a power-bounded quasi-compact operator on $\mathcal{B}_V$, and we have*

$$r_{ess}(P) \leq \delta_V(P).$$

Proof. Iterating (**WD**) shows that P is power-bounded on $\mathcal{B}_V$ (proceed as in the beginning of the proof of Theorem 1).

Now let $(\mathcal{B}'_V, \|\cdot\|_V)$ (resp. $(\mathcal{B}'_0, \|\cdot\|_0)$) denote the dual space of $\mathcal{B}_V$ (resp. of $\mathcal{B}_0$). Note that we make a slight abuse of notation in writing again $\|\cdot\|_V$ and $\|\cdot\|_0$ for the dual norms. Let P^* denote the adjoint operator of P on $\mathcal{B}'_V$. In fact, we prove that P^* is a quasi-compact operator on $\mathcal{B}'_V$ with $r_{ess}(P^*) \leq \delta_V(P)$, so that P satisfies the same properties on $\mathcal{B}_V$. Since $P : \mathcal{B}_0 \to \mathcal{B}_V$ is assumed to be compact, then so is $P^* : \mathcal{B}'_V \to \mathcal{B}'_0$. Moreover P^* satisfies a Doeblin-Fortet inequality from Lemma 2 below. Then we deduce from Lemma 2 and [Hen93] that P^* is a quasi-compact operator on $\mathcal{B}'_V$, with $r_{ess}(P^*) \leq \delta$ for any $\delta \in (\delta_V(P), 1)$, so that $r_{ess}(P^*) \leq \delta_V(P)$. □

For the sake of simplicity we consider the same usual bracket notation $\langle \cdot, \cdot \rangle$ in both $\mathcal{B}'_V \times \mathcal{B}_V$ and $\mathcal{B}'_0 \times \mathcal{B}_0$. Recall that $\mathcal{B}_V, \mathcal{B}_0$ are Banach lattices, so are $\mathcal{B}'_V$, $\mathcal{B}'_0$. For each $g' \in \mathcal{B}'_V$ (resp. $g' \in \mathcal{B}'_0$), one can define the modulus $|g'|$ of g' in $\mathcal{B}'_V$ (resp. in $\mathcal{B}'_0$), see [Sch71]. For the next arguments, it is enough to have in mind that g' and $|g'|$ have the same norm in $\mathcal{B}'_V$ (resp. in $\mathcal{B}'_0$), more precisely:

$$\forall g' \in \mathcal{B}'_0, \quad \|g'\|_0 = \langle |g'|, 1_{\mathbb{X}} \rangle \quad \text{and} \quad \forall g' \in \mathcal{B}'_V, \quad \|g'\|_V = \langle |g'|, V \rangle.$$

Lemma 2 *Let $\delta \in (\delta_V(P), 1)$. Then, there exist $N \in \mathbb{N}^*$ and $d \in (0, +\infty)$ such that for all $f' \in \mathcal{B}'_V$ we have: $\|P^{*N} f'\|_V \leq \delta^N \|f'\|_V + d\|f'\|_0$.*

Proof. Let $f' \in \mathcal{B}'_V$ and $n \geq 1$. Since P^n is a nonnegative operator on $\mathcal{B}_V$, so is its adjoint operator P^{*n} on $\mathcal{B}'_V$, and we have for all $f \in \mathcal{B}_V$ such that $\|f\|_V \leq 1$ (ie. $|f| \leq V$):

$$\big|\langle (P^*)^n f', f \rangle\big| \leq \big\langle (P^*)^n |f'|, |f| \big\rangle \leq \big\langle (P^*)^n |f'|, V \big\rangle = \langle |f'|, P^n V \rangle.$$

By definition of $\delta_V(P)$ and from $\delta \in (\delta_V(P), 1)$, there exist $N \in \mathbb{N}^*$ and $d \in (0, +\infty)$ such that $P^N V \leq \delta^N V + d\, 1_{\mathbb{X}}$. Thus

$$\begin{aligned}
\|(P^*)^N f'\|_V &:= \sup_{f\in\mathcal{B}_V, \|f\|_V \leq 1} \big|\langle (P^*)^N f', f \rangle\big| \\
&\leq \langle |f'|, P^N V \rangle \\
&\leq \delta^N \langle |f'|, V \rangle + d \langle |f'|, 1_{\mathbb{X}} \rangle = \delta^N \|f'\|_V + d\|f'\|_0.
\end{aligned}$$

□

The following lemma is useful to prove that $P : \mathcal{B}_0 \to \mathcal{B}_V$ is compact.

Lemma 3 *Assume that* $(\mathbb{X}, d)$ *is a metric space in which every closed ball is compact, that* $\lim_{d(x,x_0)\to+\infty} V(x) = +\infty$ *for some* $x_0 \in \mathbb{X}$*, and finally that* $\{Pf,\, \|f\|_0 \le 1\}$ *is equicontinuous. Then* P *is a compact operator from* $\mathcal{B}_0$ *into* $\mathcal{B}_V$*.*

Proof. Let $(f_n)_n \in \mathcal{B}_0^{\mathbb{N}}$ such that $\|f_n\|_0 \le 1$. Since $\mathbb{X}$ is separable, it follows from Ascoli's theorem that there exists a subsequence $(Pf_{n_k})_k$ converging pointwise to some continuous function $g \in \mathcal{B}_0$, $\|g\|_0 \le 1$, with uniform convergence on each compact set of $\mathbb{X}$. Let $\varepsilon > 0$. We know from the assumptions that there exists $A > 0$ such that: $d(x, x_0) > A \Rightarrow V(x)^{-1} < \varepsilon/2$. Next, let $n_0 \in \mathbb{N}$ be such that: $\forall k \ge n_0,\ \sup_{d(x,x_0)\le A} |(Pf_{n_k})(x) - g(x)| < \varepsilon$. Since $V \ge 1$ and P is a contraction on $\mathcal{B}_0$, we obtain:

$$\sup_{x\in\mathbb{X}} \frac{|(Pf_{n_k})(x) - g(x)|}{V(x)} \le \max\left(\sup_{d(x,x_0)\le A} |(Pf_{n_k})(x) - g(x)|,\ \sup_{d(x,x_0)>A} \frac{|(Pf_{n_k})(x) - g(x)|}{V(x)} \right) < \varepsilon.$$

Thus $\lim_k Pf_{n_k} = g$ in $\mathcal{B}_V$. We have proved that $P : \mathcal{B}_0 \to \mathcal{B}_V$ is compact. □

As shown in the next corollary, Theorem 2 is relevant in the case when $P(x, dy)$ is absolutely continuous with respect to some positive measure on a metric space $\mathbb{X}$ equipped with its Borel σ-algebra $\mathcal{X}$.

Corollary 1 *Assume that* $(\mathbb{X}, d)$ *is a metric space in which every closed ball is compact, and that* $V : \mathbb{X} \to [1, +\infty)$ *satisfies* $\lim_{d(x,x_0)\to+\infty} V(x) = +\infty$ *for some* $x_0 \in \mathbb{X}$*. Furthermore assume that Condition* (**WD**) *holds true and that there exist a positive measure* η *on* $(\mathbb{X}, \mathcal{X})$ *and a function* $K : \mathbb{X}^2 \to [0, +\infty)$*, supposed to be continuous in the first variable, such that*

$$\forall x \in \mathbb{X}, \quad P(x, dy) = K(x, y)\, d\eta(y). \tag{10}$$

Then P *is a power-bounded quasi-compact operator on* $\mathcal{B}_V$*, and we have* $r_{ess}(P) \le \delta_V(P)$*.*

Proof. Thanks to Theorem 2, we only have to prove that $P : \mathcal{B}_0 \to \mathcal{B}_V$ is compact. Let $f \in \mathcal{B}_0$ such that $\|f\|_0 \le 1$ (ie. $|f| \le 1$). Then we have for all $(x, x') \in \mathbb{X}^2$:

$$\big|(Pf)(x') - (Pf)(x)\big| \le \int_{\mathbb{X}} \big|K(x', y) - K(x, y)\big|\, d\eta(y).$$

Since we have $K(\cdot,\cdot) \ge 0$, $\int K(\cdot, y) d\eta(y) = 1$, and $\lim_{x'\to x} K(x', y) = K(x, y)$, we deduce from Scheffé's theorem that $\lim_{x'\to x} \int_{\mathbb{X}} |K(x', y) - K(x, y)|\, d\eta(y) = 0$. This proves that $\{Pf,\, \|f\|_0 \le 1\}$ is equicontinuous. Thus $P : \mathcal{B}_0 \to \mathcal{B}_V$ is compact from Lemma 3. □

2.3 From quasi-compactness on $\mathcal{B}_V$ to V-geometrical ergodicity

Recall that a Markov chain $(X_n)_{n\in\mathbb{N}}$ with transition kernel P is V-geometrically ergodic if P has an invariant probability measure π such that

(VG1) $\pi(V) < \infty$

(VG2) $\lim_{n\to\infty} \sup_{f\in\mathcal{B}_V, \|f\|_V\leq 1} \|P^n f - \pi(f)\|_V = 0$.

Let Π denotes the rank-one projection defined on $\mathcal{B}_V$ by: $\Pi f = \pi(f)1_{\mathbb{X}}$. Note that the condition (VG2) is equivalent to the convergence to 0 of $\|P^n - \Pi\|_V$, the operator norm associated with $\|\cdot\|_V$. Moreover, using $P^n - \Pi = (P - \Pi)^n$, it can be shown that the convergence is geometric, that is, there exists $\rho \in (0, 1)$ and $c_\rho \in (0, +\infty)$ such that

$$\|P^n - \Pi\|_V \leq c_\rho\, \rho^n. \tag{11}$$

The infimum bound of the positive real numbers ρ such that (11) holds is denoted by $\rho_V(P)$. It will be called *the convergence rate* of P on $\mathcal{B}_V$.

In this subsection we propose a result which makes explicit the relationship between the quasi-compactness of P and the V-geometric ergodicity of the Markov chain $(X_n)_{n\in\mathbb{N}}$ with transition kernel P. Moreover, we provide an explicit formula for $\rho_V(P)$ in terms of the spectral elements of P. A key element is the essential spectral radius $r_{ess}(P)$.

Theorem 3 *Let P be a transition kernel which has an invariant probability measure. The two following assertions are equivalent:*

(a) P is V-geometrically ergodic.

(b) P is a power-bounded quasi-compact operator on $\mathcal{B}_V$, for which $\lambda = 1$ is a simple eigenvalue (i.e. $\mathrm{Ker}(P - I) = \mathbb{C}\cdot 1_{\mathbb{X}}$) and the unique eigenvalue of modulus one.

Under any of these conditions, we have $r_{ess}(P) \leq \rho_V(P)$. In fact, for $r_0 \in (r_{ess}(P), 1)$, denoting the set of all the eigenvalues λ of P such that $r_0 \leq |\lambda| < 1$ by $\mathcal{V}_{r_0}$, we have:

- *either $\rho_V(P) \leq r_0$ when $\mathcal{V}_{r_0} = \emptyset$,*
- *or $\rho_V(P) = \max\{|\lambda|,\, \lambda \in \mathcal{V}_{r_0}\}$ when $\mathcal{V}_{r_0} \neq \emptyset$.*

Moreover, if $\mathcal{V}_{r_0} = \emptyset$ for all $r_0 \in (r_{ess}(P), 1)$, then $\rho_V(P) = r_{ess}(P)$.

From Definition 2.1, for any $r_0 \in (r_{ess}(P), 1)$, the set of all the eigenvalues of λ of P such that $r_0 \leq |\lambda| \leq 1$ is finite.

Remark 2 *The property that P admits a spectral gap on $\mathcal{B}_V$ in the recent paper [KM11] corresponds here to the quasi-compactness of P (which is a classical terminology in spectral theory). The spectral gap in [KM11] corresponds to the value $1 - \rho_V(P)$. Then, [KM11, Prop. 1.1]) is a reformulation of the equivalence of properties (a) and (b) in Theorem 3 under ψ-irreducibility and aperiodicity assumptions (see also [KM11, Lem. 2.1]). The last statements in Theorem 3 provide the value of the convergence rate $\rho_V(P)$ for V-geometrically ergodic Markov chains from the essential spectral radius $r_{ess}(P)$ and the (possible) eigenvalues λ such that $r_{ess}(P) < |\lambda| < 1$.*

Proof. Note that we have $\mathcal{B}_V = \mathbb{C}\,1_{\mathbb{X}} \oplus H_0$, with $H_0 := \{f \in \mathcal{B}_V : \pi(f) = 0\}$ (write $f = \pi(f)1_{\mathbb{X}} + (f - \pi(f)1_{\mathbb{X}}))$. Since $\pi(V) < \infty$, π defines a bounded linear form on $\mathcal{B}_V$, so that H_0 is a closed subspace of $\mathcal{B}_V$. From the invariance of π, we obtain that $P(H_0) \subset H_0$.

Now assume that (a) is fulfilled. Then for any $\rho \in (\rho_V(P), 1)$ we have from (11):

$$\sup_{h\in H_0,\, \|h\|_V \le 1} \|P^n h\|_V = O(\rho^n).$$

It follows from Definition 2.1 that P is quasi-compact on $\mathcal{B}_V$, with $r_{ess}(P) \le \rho_V(P)$. The fact that P is power-bounded on $\mathcal{B}_V$ easily follows from (a).

Conversely, assume that (b) holds and prove that Property (a), together with the claimed properties on $\rho_V(P)$, are fulfilled. Since P is Markov and power-bounded on $\mathcal{B}_V$, we have $r(P) = 1$. From Definition 2.1 and the assumption on the peripheral eigenvalues of P, we obtain for any $r_0 \in (r_{ess}(P), 1)$:

$$\mathcal{B}_V = \mathbb{C}\,1_{\mathbb{X}} \oplus \big(\oplus_{\lambda\in\mathcal{V}_{r_0}} \operatorname{Ker}(P - \lambda I)^{p_\lambda}\big) \oplus H, \tag{12}$$

where H is a closed P-invariant subspace of $\mathcal{B}_V$ such that $\sup_{h\in H,\, \|h\|_V\le 1} \|P^n h\|_V = O({r_0}^n)$. Let $f \in \mathcal{B}_V$. Then we have

$$f - \pi(f)1_{\mathbb{X}} = \sum_{\lambda\in\mathcal{V}_{r_0}} f_\lambda + h, \tag{13}$$

with $f_\lambda \in \operatorname{Ker}(P - \lambda I)^{p_\lambda}$ and $h \in H$, and there exist some constants c_λ and c_H (independent of f) such that $\|f_\lambda\|_V \le c_\lambda \|f\|_V$ and $\|h\|_V \le c_H \|f\|_V$ (since the projections associated with the decomposition (12) are continuous).

When $\mathcal{V}_{r_0} = \emptyset$, then (13) yields

$$\|P^n f - \pi(f)1_{\mathbb{X}}\|_V = \|P^n\big(f - \pi(f)1_{\mathbb{X}}\big)\|_V \le O({r_0}^n)\, \|f\|_V.$$

Thus Property (a) holds and $\rho_V(P) \le r_0$. If $\mathcal{V}_{r_0} = \emptyset$ for all $r_0 \in (r_{ess}(P), 1)$, then $\rho_V(P) \le r_{ess}(P)$, so that $\rho_V(P) = r_{ess}(P)$ from the proof of $(a) \Rightarrow (b)$.

When $\mathcal{V}_{r_0} \ne \emptyset$, define $\nu := \max\{|\lambda|,\, \lambda \in \mathcal{V}_{r_0}\}$. We have for $\lambda \in \mathcal{V}_{r_0}$ and $n \ge p_\lambda$

$$\begin{aligned} \|P^n f_\lambda\|_V = \|(P - \lambda I + \lambda I)^n f_\lambda\|_V &\le \sum_{k=0}^{p_\lambda - 1} \binom{n}{k} |\lambda|^{n-k} \|(P - \lambda I)^k\|_V\, \|f_\lambda\|_V \\ &\le c_\lambda \Big(\sum_{k=0}^{p_\lambda - 1} \binom{n}{k} |\lambda|^{n-k} \|(P - \lambda I)^k\|_V\Big) \|f\|_V. \end{aligned}$$

Then for each $k = 0, \ldots, p_\lambda - 1$, we have $\binom{n}{k}|\lambda|^{n-k} = O(n^k |\lambda|^n) \le O(n^k \nu^n)$. Thus $\|P^n f_\lambda\|_V = O(\rho^n)\, \|f\|_V$ for any $\rho \in (\nu, 1)$. From (13) and $r_0 \le \rho$, we obtain:

$$\|P^n f - \pi(f)1_{\mathbb{X}}\|_V \le \sum_{\lambda\in\mathcal{V}_{r_0}} \|P^n f_\lambda\| + \|P^n h\|_V \le \big(O(\rho^n) + O({r_0}^n)\big)\, \|f\|_V = O(\rho^n)\, \|f\|_V.$$

Since $\rho \in (\nu, 1)$ is arbitrary, this gives $\rho_V(P) \le \nu$. Conversely, given any $\lambda \in \mathcal{V}_{r_0}$ and $f \in \mathcal{B}_V$ such that $Pf = \lambda f$, we have $\pi(f) = 0$ (use the invariance of π), and from $P^n f = \lambda^n f$ and the definition of $\rho_V(P)$, we easily deduce that $|\lambda| \le \rho_V(P)$. Thus $\nu \le \rho_V(P)$. □

The next lemma provides the existence of the P-invariant probability measure under the weak drift condition Condition (**WD**). This well-known statement (see Appendix B) will be of interest in our examples for the use of Theorem 3.

Lemma 4 *Assume that* $(\mathbb{X}, d)$ *is a separable complete metric space and that* $V : \mathbb{X} \to [1, +\infty)$ *is continuous and such that the set* $\{V \leq \alpha\}$ *is compact for every* $\alpha \in (0, +\infty)$. *Under Condition* (**WD**), *there exists a* P*-invariant probability measure* π *such that* $\pi(V) < \infty$.

2.4 Study of characteristic functions

As explained in Introduction, the next theorem plays an important role in our work. For the sake of simplicity, we write δ_V for $\delta_V(P)$ in (3).

Theorem 4 *Assume that the weak drift condition* (**WD**) *holds true. If* $\lambda \in \mathbb{C}$ *is such that* $\delta \leq |\lambda| \leq 1$, *with* δ *given in* (**WD**), *and if* $f \in \mathcal{B}_V \cap \mathrm{Ker}(P - \lambda I)^p$ *for some* $p \in \mathbb{N}^*$, *then there exists* $c \in (0, +\infty)$ *such that*

$$|f| \leq c\, V^{\frac{\ln|\lambda|}{\ln \delta}} \, (\ln V)^{\frac{p(p-1)}{2}}.$$

The proof of Theorem 4 is based on the following lemma.

Lemma 5 *Let* $\lambda \in \mathbb{C}$ *be such that* $\delta \leq |\lambda| \leq 1$. *Then*

$$\forall f \in \mathcal{B}_V,\ \exists c \in (0, +\infty),\ \forall x \in \mathbb{X}, \quad |\lambda|^{-n(x)} \big| (P^{n(x)} f)(x) \big| \leq c\, V(x)^{\frac{\ln|\lambda|}{\ln\delta}} \tag{14}$$

with, for any $x \in \mathbb{X}$, $n(x) := \lfloor \frac{-\ln V(x)}{\ln \delta} \rfloor$, *where* $\lfloor \cdot \rfloor$ *denotes the integer part function.*

Proof. First note that the iteration of (**WD**) gives

$$\forall k \geq 1, \quad P^{kN} V \leq \delta^{kN}\, V + d\big(\sum_{j=0}^{k-1} \delta^{jN}\big)\, 1_{\mathbb{X}} \leq \delta^{kN}\, V + \frac{d}{1 - \delta^N}\, 1_{\mathbb{X}}. \tag{15}$$

Let $g \in \mathcal{B}_V$ and $x \in \mathbb{X}$. Using (15), the positivity of P and $|g| \leq \|g\|_V\, V$, we obtain with $b := d/(1 - \delta^N)$:

$$\forall k \geq 1, \quad |(P^{kN} g)(x)| \leq (P^{kN}|g|)(x) \leq \|g\|_V\, (P^{kN} V)(x) \leq \|g\|_V \big(\delta^{kN} V(x) + b\big). \tag{16}$$

The previous inequality is also fulfilled with $k = 0$. Next, let $f \in \mathcal{B}_V$ and $n \in \mathbb{N}$. Writing $n = kN + r$, with $k \in \mathbb{N}$ and $r \in \{0, 1, \ldots, N-1\}$, and applying (16) to $g := P^r f$, we obtain with $\xi := \max_{0 \leq \ell \leq N-1} \|P^\ell f\|_V$ (use $P^n f = P^{kN}(P^r f)$):

$$\big|(P^n f)(x)\big| \leq \xi \big[\delta^{kN} V(x) + b\big] \leq \xi \big[\delta^{-r}\big(\delta^n V(x) + b\big)\big] \leq \xi\, \delta^{-N}\big(\delta^n V(x) + b\big). \tag{17}$$

Using the inequality

$$-\frac{\ln V(x)}{\ln \delta} - 1 \leq n(x) \leq -\frac{\ln V(x)}{\ln \delta}$$

and the fact that $\ln\delta \le \ln|\lambda| < 0$, Inequality (17) with $n := n(x)$ gives:

$$\begin{aligned}
|\lambda|^{-n(x)}\big|(P^{n(x)}f)(x)\big| &\le \xi\,\delta^{-N}\Big(\big(\delta|\lambda|^{-1}\big)^{n(x)}\,V(x) + b\,|\lambda|^{-n(x)}\Big)\\
&= \xi\,\delta^{-N}\Big(e^{n(x)(\ln\delta-\ln|\lambda|)}\,e^{\ln V(x)} + b\,e^{-n(x)\ln|\lambda|}\Big)\\
&\le \xi\,\delta^{-N}\Big(e^{(\frac{\ln V(x)}{\ln\delta}+1)\,(\ln|\lambda|-\ln\delta)}\,e^{\ln V(x)} + b\,e^{\frac{\ln V(x)}{\ln\delta}\ln|\lambda|}\Big)\\
&= \xi\,\delta^{-N}\Big(e^{\frac{\ln|\lambda|}{\ln\delta}\ln V(x)}\,e^{\ln|\lambda|-\ln\delta} + b\,V(x)^{\frac{\ln|\lambda|}{\ln\delta}}\Big)\\
&= \xi\,\delta^{-N}\big(e^{\ln|\lambda|-\ln\delta} + b\big)\,V(x)^{\frac{\ln|\lambda|}{\ln\delta}}.
\end{aligned}$$

This gives the desired conclusion with $c = \xi\,\delta^{-N}(e^{\ln|\lambda|-\ln\delta} + b)$. $\square$

Proof of Theorem 4. If $f \in \mathcal{B}_V \cap \mathrm{Ker}(P-\lambda I)$, then $|\lambda|^{-n(x)}|(P^{n(x)}f)(x)| = |f(x)|$, so that (14) gives the expected conclusion when $p = 1$. Next, let us proceed by induction. Assume that the conclusion of Theorem 4 holds for some $p \ge 1$. Let $f \in \mathcal{B}_V \cap \mathrm{Ker}(P-\lambda I)^{p+1}$. We can write

$$P^n f = (P-\lambda I+\lambda I)^n f = \lambda^n\,f + \sum_{k=1}^{\min(n,p)}\binom{n}{k}\lambda^{n-k}\,(P-\lambda I)^k f. \tag{18}$$

For $k \in \{1,\ldots,p\}$, we have $f_k := (P-\lambda I)^k f \in \mathrm{Ker}(P-\lambda I)^{p+1-k} \subset \mathrm{Ker}(P-\lambda I)^p$, thus we have from the induction hypothesis :

$$\exists c' \in (0,+\infty),\ \forall k \in \{1,\ldots,p\},\ \forall x \in \mathbb{X},\quad |f_k(x)| \le c'\,V(x)^{\frac{\ln|\lambda|}{\ln\delta}}\,(\ln V(x))^{\frac{p(p-1)}{2}}. \tag{19}$$

Now, we obtain from (18) (with $n := n(x)$), (19) and Lemma 5 that for all $x \in \mathbb{X}$:

$$\begin{aligned}
|f(x)| &\le |\lambda|^{-n(x)}\big|(P^{n(x)}f)(x)\big| + c'\,V(x)^{\frac{\ln|\lambda|}{\ln\delta}}\,(\ln V(x))^{\frac{p(p-1)}{2}}\,|\lambda|^{-\min(n,p)}\sum_{k=1}^{\min(n,p)}\binom{n(x)}{k}\\
&\le c\,V(x)^{\frac{\ln|\lambda|}{\ln\delta}} + c_1\,V(x)^{\frac{\ln|\lambda|}{\ln\delta}}\,(\ln V(x))^{\frac{p(p-1)}{2}}\,n(x)^p\\
&\le c_2 V(x)^{\frac{\ln|\lambda|}{\ln\delta}}\,(\ln V(x))^{\frac{p(p-1)}{2}+p}
\end{aligned}$$

with some constants $c_1, c_2 \in (0,+\infty)$ independent of x. Since $p(p-1)/2 + p = p(p+1)/2$, this gives the expected result. $\square$

To conclude this section, notice that the V-geometrical ergodicity clearly implies Condition (**WD**). However Condition (**WD**) is not sufficient for P to be V-geometrically ergodic, even if P is assumed to be compact from $\mathcal{B}_0$ from $\mathcal{B}_V$. In fact, the previous statements provide the following procedure to check the V-geometric ergodicity of P and to compute an upper bound for its convergence rate $\rho_V(P)$. Let P be a transition kernel which has an invariant probability measure. Theorem 2 shows that, if P is compact from $\mathcal{B}_0$ into $\mathcal{B}_V$ and satisfies the weak drift condition (**WD**), then P is quasi-compact on $\mathcal{B}_V$ and $r_{ess}(P) \le \delta_V(P)$. Next Theorem 3 ensures that the V-geometric ergodicity of P can be deduced from quasi-compactness provided that the following properties are satisfied :

(i) $\lambda = 1$ is a simple eigenvalue of P on $\mathcal{B}_V$, namely $\mathrm{Ker}(P - I) = \mathbb{C} \cdot 1_{\mathbb{X}}$;

(ii) $\lambda = 1$ is the unique eigenvalue of P of modulus one on $\mathcal{B}_V$.

Finally Theorem 4 can be useful to check (i)-(ii), and in a more general way to investigate the sets $\mathcal{V}_{r_0}$ of eigenvalues of P given in Theorem 3 in order to obtain an upper bound for the convergence rate $\rho_V(P)$. This procedure is applied in the next section.

3 Applications to discrete Markov chains

In this section, we are concerned with discrete Markov chains. For the sake of simplicity, we assume that $\mathbb{X} := \mathbb{N}$ throughout the section. Let P be a Markov kernel on $\mathbb{N}$. The main focus is on the estimation of the essential spectral radius $r_{ess}(P)$ from Condition (**WD**): a general statement is derived from Corollary 1 in Subsection 3.1, and applications to random walks (RW) with bounded state-dependent increments are presented in Subsection 3.2.

For irreducible and aperiodic discrete Markov chains, criteria for the V-geometrical ergodicity are well-known from the literature using, either the equivalence between geometric ergodicity and V-geometric ergodicity of $\mathbb{N}$-valued Markov chains [HS92, Prop. 2.4], or the strong drift Conditions (**D**)-(**S**) of Introduction [MT93]. In Subsection 3.3, we just explain as an alternative way how the quasi-compactness combined with irreducibility and aperiodicity conditions provide the V-geometrical ergodicity. Finally the procedure mentioned at the end of the previous section (see (i))-(ii)) is applied to compute the convergence rate of some random walks (see Example 4 and Subsection 3.4).

3.1 Quasi-compactness of discrete Markov chains

Let $P = (P(i,j))_{i,j\in\mathbb{N}^2}$ be a Markov kernel on $\mathbb{N}$. The function $V : \mathbb{N} \to [1, +\infty)$ is assumed to satisfy

$$\lim_n V(n) = +\infty \quad \text{and} \quad \sup_{n\in\mathbb{N}} \frac{(PV)(n)}{V(n)} < \infty.$$

Corollary 2 *The two following conditions are equivalent:*

(a) Condition (**WD**) *holds with* V*;*

(b) $L := \inf_{N\geq 1} (\ell_N)^{\frac{1}{N}} < 1$ *where* $\ell_N := \limsup_{n\to+\infty} (P^N V)(n)/V(n)$.

In this case, P *is power-bounded and quasi-compact on* $\mathcal{B}_V$ *with*

$$r_{ess}(P) \leq \delta_V(P) = L.$$

Proof. The last assertion follows from Corollary 1[1]. Let us prove the equivalence *(a)⇔(b)*, as well as the equality $\delta_V(P) = L$. First, Condition (**WD**) clearly gives $\ell_N \le \delta^N < 1$ (with δ in (**WD**)), thus $L \le \delta_V(P)$ by definition of $\delta_V(P)$. Conversely, assume that $L < 1$: there exists $N \ge 1$ such that $\ell_N < 1$. Let δ be such that $\ell_N < \delta^N < 1$. Then there exists $n_0 \in \mathbb{N}$ such that: $\forall n > n_0,\ (P^N V)(n)/V(n) \le \delta^N$. Hence

$$P^N V \le \delta^N V + d, \quad \text{with } d := \max_{0\le i\le n_0} \frac{(P^N V)(i)}{V(i)}.$$

This proves (**WD**), and $\delta_V(P) \le (\ell_N)^{1/N}$ since δ is arbitrary close to $(\ell_N)^{1/N}$. In fact, the last argument shows that $\delta_V(P) \le (\ell_N)^{1/N}$ provided that $\ell_N < 1$. From definition of L, there exists a sequence $(N_k)_{k\ge 0}$ such that $L = \lim_k (\ell_{N_k})^{1/N_k}$. Thus we have $\delta_V(P) \le (\ell_{N_k})^{1/N_k}$ for k large enough. Thus $\delta_V(P) \le L$. □

In the next subsections, Corollary 2 is applied to random walks on $\mathbb{N}$ with the following special sequence $V_\gamma := (\gamma^n)_{n\in\mathbb{N}}$ for some $\gamma \in (1, +\infty)$. The associated weighted-supremum space $\mathcal{B}_\gamma \equiv \mathcal{B}_{V_\gamma}$ is defined by:

$$\mathcal{B}_\gamma := \big\{(f(n))_{n\in\mathbb{N}} \in \mathbb{C}^{\mathbb{N}} : \sup_{n\in\mathbb{N}} \gamma^{-n}|f(n)| < \infty\big\}. \tag{20}$$

3.2 Quasi-compactness of RW with bounded state-dependent increments

Let us fix $b \in \mathbb{N}^*$, and assume that the kernel P on $\mathbb{X} := \mathbb{N}$ satisfies the following conditions:

$$\forall i \in \{0,\dots,b-1\}, \quad \sum_{j\ge 0} P(i,j) = 1;$$
$$\forall i \ge b, \forall j \in \mathbb{N}, \quad P(i,j) = \begin{cases} 0 & \text{if } |i-j| > b \\ a_{j-i}(i) & \text{if } |i-j| \le b \end{cases} \tag{21}$$

where $(a_{-b}(i), \dots, a_b(i)) \in [0,1]^{2b+1}$ satisfies $\sum_{k=-b}^{b} a_k(i) = 1$ for all $i \ge b$. This kind of kernels arises, for instance, from time-discretization of Markovian queueing models (see a basic example in Remark 4).

Proposition 1 *Assume that for every $k \in \mathbb{Z}$ such that $|k| \le b$*

$$\lim_n a_k(n) = a_k \in [0,1], \tag{22a}$$

and that $\gamma \in (1, +\infty)$ is such that

$$\phi(\gamma) := \sum_{k=-b}^{b} a_k\, \gamma^k < 1 \tag{22b}$$

$$\forall i \in \{0,\dots,b-1\}, \quad \sum_{j\ge 0} P(i,j)\gamma^j < \infty. \tag{22c}$$

[1]Note that the compactness of $P : \mathcal{B}_0 \to \mathcal{B}_V$ can be directly deduced from Cantor's diagonal argument, which ensures that the canonical injection from $\mathcal{B}_0$ into $\mathcal{B}_V$ is compact (use $\lim_n V(n) = +\infty$).

Then P is power-bounded and quasi-compact on $\mathcal{B}_\gamma$ with

$$r_{ess}(P) \leq \phi(\gamma).$$

Proof. Set $\phi_n(\gamma) := \sum_{k=-b}^{b} a_k(n)\,\gamma^k$. We have $(PV_\gamma)(n) = \phi_n(\gamma)V_\gamma(n)$ for each $n \geq b$, hence we obtain from (22a)

$$\limsup_n \frac{(PV_\gamma)(n)}{V_\gamma(n)} \leq \phi(\gamma).$$

The conclusion of Proposition 1 then follows from Corollaries 2 using (22b)-(22c). $\square$

Assume that $a_0 \neq 1$. Let $\phi^{(k)}$ be the k-th derivative of ϕ. It is checked in Appendix C.2 that there exists $1 \leq \ell \leq 2b$ such that

$$\forall k \in \{1, \ldots, \ell - 1\}, \quad \phi^{(k)}(1) = 0 \quad \text{and} \quad \phi^{(\ell)}(1) \neq 0, \tag{23}$$

according that the first condition is removed when $\ell = 1$. Since $\phi(1) = 1$, a sufficient condition for (22b) to hold for some $\gamma \in (1, +\infty)$ is that

$$\phi^{(\ell)}(1) < 0. \tag{24}$$

Example 1 (State-dependent birth-and-death Markov chains) *When $b := 1$ in (21), we obtain the standard class of state-dependent birth-and-death Markov chains. Namely, the stochastic kernel P is defined by:*

$$\forall n \geq 1,\ \ P(n, n-1) := p_n, \quad P(n, n) := r_n, \quad P(n, n+1) := q_n,$$

where the real numbers $(p_n, r_n, q_n) \in [0, 1]^2$ and $p_n + r_n + q_n = 1$. Assume that the following limits exist:

$$\lim_n p_n := p \in (0, 1], \quad \lim_n r_n := r \in [0, 1), \quad \lim_n q_n := q.$$

If $\gamma \in (1, +\infty)$ is such that

$$\phi(\gamma) := \frac{p}{\gamma} + r + q\gamma < 1 \quad \textit{and} \quad \sum_{n\geq 0} P(0, n)\gamma^n < \infty,$$

then it follows from Proposition 1 that P is power-bounded and quasi-compact on $\mathcal{B}_\gamma$ with

$$r_{ess}(P) \leq \frac{p}{\gamma} + r + q\gamma.$$

The condition $p/\gamma + r + q\gamma < 1$ is equivalent to the following ones (use $r = 1 - p - q$ for (i)):

(i) either $p > q > 0$ and $1 < \gamma < p/q$; (ii) or $q = 0$ and $\gamma > 1$.

- *When $p > q > 0$ and $1 < \gamma < p/q$: if $\sum_{n\geq 0} P(0, n)\gamma^n < \infty$, then P is power-bounded and quasi-compact on $\mathcal{B}_\gamma$ with $r_{ess}(P) \leq \phi(\gamma)$. Set $\widehat{\gamma} := \sqrt{p/q}$. Then*

  $$\min_{\gamma>1} \phi(\gamma) = \phi(\widehat{\gamma}) = r + 2\sqrt{pq} \ \in (r^2, 1). \tag{25}$$

  Consequently, if $\sum_{n\geq 0} P(0, n)(\widehat{\gamma})^n < \infty$, then the previous conclusions holds for $\gamma := \widehat{\gamma}$, with essential spectral radius on $\mathcal{B}_{\widehat{\gamma}}$ satisfying

  $$r_{ess}(P) \leq r + 2\sqrt{pq}.$$

- *When $q := 0$ and $\gamma > 1$: if $\sum_{n\geq 0} P(0,n)\gamma^n < \infty$, then for P is power-bounded and quasi-compact on $\mathcal{B}_\gamma$ with*
$$r_{ess}(P) \leq \phi(\gamma) = p/\gamma + r.$$
Such a case is illustrated by the next example.

Example 2 (Simulation of a Poisson distribution with parameter one) *The Markov kernel P on $\mathbb{X} := \mathbb{N}$ defined by*

$$P(0,0) = P(0,1) = \frac{1}{2}$$
$$\forall n \geq 1,\ P(n,n-1) := \frac{1}{2}, \quad P(n,n) := \frac{n}{2(n+1)}, \quad P(n,n+1) := \frac{1}{2(n+1)}.$$

arises from a Hastings-Metropolis sampler of a Poisson distribution. We have $p = r = 1/2$ and $q = 0$ with the notations of Example 1. Hence, for each $\gamma \in (1,+\infty)$, P is power-bounded and quasi-compact on $\mathcal{B}_\gamma$ and
$$r_{ess}(P) \leq 1/2 + 1/(2\gamma).$$

Remark 3 (Random walks with i.d. bounded increments) *Consider the case when the increments $a_k(n)$ do not depend on the state n, that is when the kernel P is*

$$\forall i \in \{0,\ldots,b-1\}, \quad \sum_{j\geq 0} P(i,j) = 1; \quad \forall i \geq b, \forall j \in \mathbb{N}, \quad P(i,j) = \begin{cases} a_{j-i} & \text{if } \ |i-j| \leq b \\ 0 & \text{if } \ |i-j| > b \end{cases}$$

where $(a_{-b},\ldots,a_b) \in [0,1]^{2b+1}$ and $\sum_{k=-b}^{b} a_k = 1$. Obviously the statements of Example 1 apply but some additional facts can be deduced for such Markov chains. First note that

$$\forall \gamma \in (1,+\infty),\ \forall N \geq 1,\ \forall n \geq Nb, \quad (P^N V_\gamma)(n) = \phi(\gamma)^N\, V_\gamma(n). \tag{26}$$

*Consequently, under the assumptions $\phi(\gamma) < 1$ and (22c), we obtain from Corollary 2 that Condition (**WD**) is fulfilled with V_γ and*

$$\delta_{V_\gamma}(P) = \phi(\gamma). \tag{27}$$

*Besides, it is shown in Appendix C.2.2 that, under the assumptions $a_0 \neq 1$ and (22c), Condition (**WD**) holds true with V_{γ_0} for some $\gamma_0 \in (1,\gamma]$ if and only if $\phi^{(\ell)}(0) < 1$ (see (23)-(24)). Finally, for the birth-and-death Markov chains, that is when $b := 1$, the convergence rate can be computed (see Subsection 3.4).*

3.3 V-geometrical ergodicity for discrete Markov chains

Let $P = (P(i,j))_{i,j\in\mathbb{N}^2}$ be a Markov kernel on $\mathbb{N}$. The following irreducibility and aperiodicity conditions for discrete Markov chains are well-known. For any $(i,j) \in \mathbb{N}^2$, define

$$\mathcal{R}_{i,j} := \big\{n \geq 1 : P^n(i,j) > 0\big\},$$

where P^n denotes the n-step Markov transition kernel. The Markov kernel P is said to be irreducible if

$$\forall (i,j) \in \mathbb{N}, \quad \mathcal{R}_{i,j} \neq \emptyset, \tag{$\mathcal{I}$}$$

and to be aperiodic if

$$\exists\, i \in \mathbb{N}, \quad \mathcal{R}_{i,i} - \mathcal{R}_{i,i} := \{n - m,\, (m,n) \in \mathcal{R}_{i,i} \times \mathcal{R}_{i,i}\} = \mathbb{Z}. \tag{$\mathcal{A}$}$$

Since $\mathcal{R}_{i,i}$ is stable under addition from the Chapman-Kolmogorov equation, the subgroup of $\mathbb{Z}$ generated by $\mathcal{R}_{i,i}$ coincides with $\mathcal{R}_{i,i} - \mathcal{R}_{i,i}$. Hence the aperiodicity Condition ($\mathcal{A}$) is equivalent to the usual one: the largest element $d = d(i) \in \mathbb{N}^*$ such that $\mathcal{R}_{i,i} \subset d \cdot \mathbb{N}^*$ (i.e. the g.c.d. of $\mathcal{R}_{i,i}$), called the period of i, is equal to 1. If P is irreducible then each state $j \in \mathbb{N}$ has the same period.

Corollary 3 *Under Condition (b) of Corollary 2, P has an invariant probability measure π such that $\pi(V) < \infty$. If the additional Conditions ($\mathcal{I}$)-($\mathcal{A}$) hold true, then P is V-geometrically ergodic.*

As already mentioned, the previous statement is well-known. It can also be derived from quasi-compactness (note that the first assertion of Corollary 3 follows from Lemma 4): apply Corollary 2 and Theorem 3 with Conditions *($\mathcal{I}$)-($\mathcal{A}$)* (see Appendix C.1 for completeness).

Under the additional Conditions *($\mathcal{I}$)-($\mathcal{A}$)*, all the statements of Subsection 3.2 can be completed in order to find again the V_γ-geometric ergodicity. For instance, in Example 1, the V_γ-geometrical ergodicity holds when $P(0,0) \in (0,1)$ and $p_n, q_n > 0$ for all $n \geq 1$ [2]. In Example 2, Conditions *($\mathcal{I}$)-($\mathcal{A}$)* are automatically fulfilled so that P is V_γ-geometrically ergodic without additional assumptions.

Note that the irreducibility condition is not necessary for P to be V-geometrically ergodic: in this case the use of Theorem 3 (via Corollary 2) is of interest to obtain the V-geometric ergodicity as illustrated in the following simple example.

Example 3 (An instance of binary RW) *Assume that*

$$\forall n \geq 1,\ P(n, n-1) := p \in (0,1], \quad P(n,n) := r = 1 - p.$$

We know from Example 1 that, under the assumptions $\gamma > 1$ and $\sum_{n\geq 0} P(0,n)\gamma^n < \infty$, P is power-bounded and quasi-compact on $\mathcal{B}_\gamma$ with $r_{ess}(P) \leq p/\gamma + r$. Note that Condition ($\mathcal{I}$) is not automatically fulfilled in this instance. Anyway, without additional assumptions, P is V_γ-geometrically ergodic. Indeed Theorem 3 applies. First the equation $Pf = f$ leads to: $\forall n \geq 1, f(n) = f(n-1)$, so that f is constant. Hence 1 is a simple eigenvalue of P. Second, given $\lambda \in \mathbb{C}$, $|\lambda| = 1$, $\lambda \neq 1$, any solution of $Pf = \lambda f$ is of the form: $f = (f(0) {z_\lambda}^n)_{n\in\mathbb{N}}$ with $z_\lambda := p/(\lambda - r)$. From $|\lambda - r| > 1 - r = p$, we obtain $|z_\lambda| < 1$, so that the equality $(Pf)(0) = \lambda f(0)$, namely $\lambda f(0) = f(0) \sum_{n\geq 0} P(0,n) {z_\lambda}^n$ is only possible when $f(0) = 0$. Hence 1 is the only eigenvalue of modulus one.

[2] Note that, if $P(0,0) = 0$, then the period $d(0)$ of $i := 0$ may be equal to 2. For instance this fact holds when $P(0,1) = 1$, and $\lambda = -1$ is then an eigenvalue of P: P is quasi-compact on $\mathcal{B}_{\widehat{\gamma}}$, but it is not $V_{\widehat{\gamma}}$-geometrically ergodic.

Finally recall that, as it was outlined at the end of Section 2, quasi-compactness is especially of interest for bounding the convergence rate of P. Example 4 below is a first simple illustration of this fact. Other applications to birth-and-death Markov chains are proposed in the next Subsection.

Example 4 (An instance of RW with unbounded increments) *Let us point out that Corollary 2 and Theorem 3 may be also useful for random walks on $X := \mathbb{N}$ with unbounded increments. For instance, let P be defined by [MS95]*

$$\forall n \geq 1,\ P(0,n) := q_n, \quad \forall n \geq 1,\ P(n,0) := p,\ P(n,n+1) := q = 1-p,$$

with $p \in (0,1)$ and $q_n \in [0,1]$ such that $\sum_{n\geq 1} q_n = 1$. For $\gamma \in (1,+\infty)$ and $V_\gamma := (\gamma^n)_{n\in\mathbb{N}}$, we have: $\forall n \geq 1$, $(PV_\gamma)(n) = p + q\gamma^{n+1} = (p/\gamma^n + q\gamma)V_\gamma(n)$. Thus, if $\gamma \in (1, 1/q)$ and $\sum_{n\geq 1} q_n\gamma^n < \infty$, then Condition (**WD**) *holds with V_γ and we have $\delta_{V_\gamma}(P) \leq q\gamma$. Therefore, under the previous conditions, if follows from Corollary 2 that P is power-bounded, quasi-compact on $\mathcal{B}_\gamma$ and*

$$r_{ess}(P) \leq q\gamma.$$

No additional assumptions are required to obtain the V_γ-geometric ergodicity: P is V_γ-geometrically ergodic provided that $\gamma \in (1, 1/q)$ and $\sum_{n\geq 1} q_n\gamma^n < \infty$, and the convergence rate $\rho_{V_\gamma}(P)$ of P on $\mathcal{B}_\gamma$ satisfies:

$$\rho_{V_\gamma}(P) \leq \max(q\gamma, p). \tag{28}$$

Proof of (28). Theorem 3 is applied with any $r_0 > \max(q\gamma, p)$. Let $\lambda \in \mathbb{C}$ be such that $\max(q\gamma, p) < |\lambda| \leq 1$, and let $f \in \mathcal{B}_\gamma$, $f \neq 0$, be such that $Pf = \lambda f$. We obtain $f(n) = (\lambda/q)f(n-1) - pf(0)/q$ for any $n \geq 2$, so that

$$\begin{aligned}\forall n \geq 2, \quad f(n) &= f(1)\left(\frac{\lambda}{q}\right)^{n-1} - \frac{pf(0)}{q}\left(\frac{1-(\frac{\lambda}{q})^{n-1}}{1-\frac{\lambda}{q}}\right) \\ &= \left(\frac{\lambda}{q}\right)^{n-1}\left(f(1) - \frac{pf(0)}{\lambda - q}\right) + \frac{pf(0)}{\lambda - q}.\end{aligned}$$

Since $f \in \mathcal{B}_\gamma$ and $|\lambda|/q > \gamma$, we must have $f(1) = pf(0)/(\lambda - q)$, and consequently: $\forall n \geq 1$, $f(n) = pf(0)/(\lambda - q)$. Next the equality $\lambda f(0) = (Pf)(0) = \sum_{n\geq 1} q_n f(n)$ gives: $\lambda f(0) = pf(0)/(\lambda - q)$ since $\sum_{n\geq 1} q_n = 1$. We have $f(0) \neq 0$ since we look for a solution $f \neq 0$. Thus λ satisfies $\lambda^2 - q\lambda - p = 0$, namely: $\lambda = 1$ or $\lambda = -p$. The case $\lambda = -p$ has not to be considered from assumption. If $\lambda = 1$, we have $f(n) = f(0)$ for each $n \in \mathbb{N}$, so that 1 is a simple eigenvalue. We have proved that 1 is a simple eigenvalue of P on $\mathcal{B}_\gamma$ and that $\lambda = 1$ is the only eigenvalue of P on $\mathcal{B}_\gamma$ such that $\max(q\gamma, p) < |\lambda| \leq 1$. Then Theorem 3 gives the estimate (28) of the convergence rate. Note that p cannot be dropped in (28) since $\lambda = -p$ is an eigenvalue of P on $\mathcal{B}_\gamma$ with corresponding eigenvector (up to a multiplicative constant) $f_p := (1, -p, -p, \dots)$.

3.4 Study of the convergence rate for the birth-and-death Markov chains

We consider real numbers $p, q, r \in [0,1]$ such that $p+r+q=1$, $p>q>0$, and we assume that P is defined on $\mathbb{X} := \mathbb{N}$ by

$$\begin{aligned}
&\forall n \geq 1,\ P(n,n-1) := p, \quad P(n,n) := r \in [0,1), \quad P(n,n+1) := q, \\
&P(0,0) \in (0,1), \quad \sum_{n\geq 0} P(0,n)\,(\widehat{\gamma})^n < \infty \text{ where } \widehat{\gamma} := \sqrt{\frac{p}{q}} \in (1,+\infty).
\end{aligned} \tag{29}$$

Let $V_{\widehat{\gamma}} := (\widehat{\gamma}^n)_{n\in\mathbb{N}}$. The weighted-supremum space $\mathcal{B}_{\widehat{\gamma}} := \mathcal{B}_{V_{\widehat{\gamma}}}$ associated to $V_{\widehat{\gamma}}$ is defined in (20). Note that Conditions *$(\mathcal{I})$-$(\mathcal{A})$* hold true. We know from Example 1 and Corollary 3 that P is $V_{\widehat{\gamma}}$-geometrically ergodic and from (25) (27) that

$$r_{ess}(P) \leq \delta_{V_{\widehat{\gamma}}}(P) = r + 2\sqrt{pq}.$$

As illustrated afterwards, thanks to Theorem 3 and Lemmas 6-7 below, the last bound on the essential spectral radius $r_{ess}(P)$ is relevant to compute the convergence rate $\rho_{V_{\widehat{\gamma}}}(P)$ (see (11)).

Lemma 6 *Assume that Conditions (29) hold true. If f is a nontrivial eigenvector in $\mathcal{B}_{\widehat{\gamma}}$ associated with a complex eigenvalue λ of P such that $r + 2\sqrt{pq} < |\lambda| \leq 1$ then*

$$\exists \alpha_1 \in \mathbb{C}\setminus\{0\},\ \forall n \geq 0,\ f(n) = \alpha_1\, {z_\lambda}^n, \tag{30}$$

with z_λ satisfying the following conditions:

$$|z_\lambda| < \widehat{\gamma}, \tag{31a}$$

$$q{z_\lambda}^2 + (r-\lambda)z_\lambda + p = 0, \tag{31b}$$

$$\sum_{n\geq 0} P(0,n){z_\lambda}^n = \lambda. \tag{31c}$$

Proof. Let $\lambda \in \mathbb{C}$ be such that $r + 2\sqrt{pq} < |\lambda| \leq 1$. Let $f \in \mathcal{B}_{\widehat{\gamma}}$, $f \neq 0$ such that $Pf = \lambda f$ so that

$$\forall n \geq 1, \quad \lambda f(n) = pf(n-1) + rf(n) + qf(n+1). \tag{32}$$

Let us denote by z_λ, z'_λ the two complex solutions of the characteristic equation

$$qz^2 + (r-\lambda)z + p = 0.$$

Observe that $z_\lambda z'_\lambda = p/q = \widehat{\gamma}^2$. Recall that the solutions of (32) are of the form, either $f(n) = \alpha_1 {z_\lambda}^n + \alpha_2 {z'_\lambda}^n$ if $z_\lambda \neq z'_\lambda$, or $f(n) = \alpha_1 {z_\lambda}^n + \alpha_2 n {z_\lambda}^n$ if $z_\lambda = z'_\lambda$, with $\alpha_1, \alpha_2 \in \mathbb{C}$.

We have $|z_\lambda| \neq |z'_\lambda|$. Indeed, Theorem 4 with $p := 1$ and $\delta := r + 2\sqrt{pq}$ implies that $|f| \leq cV^\tau$ with $\tau := \ln|\lambda|/\ln\delta \in (0,1)$ and some constant c. Consequently we have $|\alpha_1 {z_\lambda}^n + \alpha_2 {z'_\lambda}^n| \leq c\widehat{\gamma}^{\tau n}$ in case $z_\lambda \neq z'_\lambda$, and $|\alpha_1 {z_\lambda}^n + \alpha_2 n {z_\lambda}^n| \leq c\widehat{\gamma}^{\tau n}$ in case $z_\lambda = z'_\lambda$. If $|z_\lambda| = |z'_\lambda|$, then we would have $|z_\lambda| = |z'_\lambda| = \widehat{\gamma}$, but the two previous inequalities then easily imply that $\alpha_1 = \alpha_2 = 0$, that is $f = 0$.

From $|z_\lambda| \neq |z'_\lambda|$, we can suppose that (for instance) $|z_\lambda| < \widehat{\gamma}$ and $|z'_\lambda| > \widehat{\gamma}$. Since f, $({z_\lambda}^n)_n$ are in $\mathcal{B}_{\widehat{\gamma}}$ and $({z'_\lambda}^n)_n$ is not in $\mathcal{B}_{\widehat{\gamma}}$, we obtain: $\forall n \geq 0,\ f(n) = \alpha_1 {z_\lambda}^n$. Since $f \neq 0$ (i.e. $\alpha_1 \neq 0$), the equation $(Pf)(0) = \lambda f(0)$ implies that z_λ must satisfy (31c). $\square$

Lemma 7 *Under assumptions (29), we have:* $\rho_{V_{\widehat{\gamma}}}(P) \geq r + 2\sqrt{pq}$.

Proof. From (26), we have: $\forall i \geq n, (P^n V_{\widehat{\gamma}})(i) = (r+2\sqrt{pq})^n\, V_{\widehat{\gamma}}(i)$. Also recall that $\widehat{\gamma} := \sqrt{p/q}$ and $\phi(\widehat{\gamma}) = r + 2\sqrt{pq}$ (see (25)), where $\phi(\gamma) := \frac{p}{\gamma} + r + q\gamma$. We have

$$\forall \gamma \in (1, +\infty), \quad \gamma\, \phi(\gamma) - 1 = (\gamma - 1)(q\gamma + 1 - p) > 0,$$

thus $\widehat{\gamma}\, \phi(\widehat{\gamma}) = (r\sqrt{p} + 2p\sqrt{q})/\sqrt{q} > 1$. Now, the definition of $\rho_{V_{\widehat{\gamma}}}(P)$ ensures that, for each $\rho \in (\rho_{V_{\widehat{\gamma}}}(P), 1)$, we can write $|(P^n V_{\widehat{\gamma}})(n) - \pi(V_{\widehat{\gamma}})|/V_{\widehat{\gamma}}(n) \leq c\rho^n$ for some constant $c \equiv c(\rho)$. Thus we have

$$\begin{aligned}
\frac{|(P^n V_{\widehat{\gamma}})(n) - \pi(V_{\widehat{\gamma}})|}{V_{\widehat{\gamma}}(n)} &= \big|(r + 2\sqrt{pq})^n - \pi(V_{\widehat{\gamma}})\big(\frac{\sqrt{q}}{\sqrt{p}}\big)^n\big| \\
&= (r + 2\sqrt{pq})^n \left| 1 - \pi(V_{\widehat{\gamma}}) \left(\frac{\sqrt{q}}{r\sqrt{p} + 2p\sqrt{q}}\right)^n \right| \\
&\leq c\rho^n,
\end{aligned}$$

Hence we have $r + 2\sqrt{pq} \leq \rho$, thus $r + 2\sqrt{pq} \leq \rho_{V_{\widehat{\gamma}}}(P)$. $\square$

Proposition 2 *In addition to Conditions (29), the boundary transition probabilities are assumed to satisfy, for some* $a \in (0,1)$*:*

$$P(0,0) := a, \quad P(0,1) := 1 - a.$$

Then P *is* $V_{\widehat{\gamma}}$*-geometrically ergodic. Furthermore, defining* $a_0 := 1 - q - \sqrt{pq}$*, the convergence rate* $\rho_{V_{\widehat{\gamma}}}(P)$ *of* P *is given by:*

- *when* $a \in [a_0, 1)$*,*
$$\rho_{V_{\widehat{\gamma}}}(P) = r + 2\sqrt{pq}\,; \tag{33}$$
- *when* $a \in (0, a_0]$*, we have*
  - *(a) in case* $2p \leq \big(1 - q + \sqrt{pq}\big)^2$*:*
$$\rho_{V_{\widehat{\gamma}}}(P) = r + 2\sqrt{pq}\,; \tag{34}$$
  - *(b) in case* $2p > \big(1 - q + \sqrt{pq}\big)^2$*, setting* $a_1 := p - \sqrt{pq} - \sqrt{r\big(r + 2\sqrt{pq}\big)}$*:*
$$\rho_{V_{\widehat{\gamma}}}(P) = \left| a + \frac{p(1-a)}{a - 1 + q} \right| \quad \text{when } a \in (0, a_1] \tag{35a}$$
$$\rho_{V_{\widehat{\gamma}}}(P) = r + 2\sqrt{pq} \quad \text{when } a \in [a_1, a_0). \tag{35b}$$

When $r := 0$ in the previous lemma, we have $a_0 = a_1 = p - \sqrt{pq} = (p-q)/(1+\sqrt{q/p})$, and it can be easily checked that $2p > (1-q+\sqrt{pq})^2$. The properties (33) (35a) (35b) then rewrite as:

$$\rho_{V_{\widehat{\gamma}}}(P) = \frac{pq+(a-p)^2}{|a-p|} \text{ when } a \in (0, a_0] \tag{36a}$$

$$\rho_{V_{\widehat{\gamma}}}(P) = 2\sqrt{pq} \qquad \text{when } a \in (a_0, 1). \tag{36b}$$

Using Kendall's theorem, the properties (36a)-(36b) have been proved for $a < p$ in [RT99] and [Bax05, Ex. 8.4]. For $a \geq p$, (36b) can be derived from [LT96] using the fact that P is stochastically monotone. Our method gives a unified and simple proof of (36a)-(36b), and encompasses the case $r \neq 0$.

Proof of Proposition 2. By elimination, given some $\lambda \in \mathbb{C}$, a necessary and sufficient condition for the two following equations:

$$qz^2 + (r-\lambda)z + p = 0, \tag{37a}$$

$$a + (1-a)z = \lambda. \tag{37b}$$

to have a common solution $z \in \mathbb{C}$ is that

$$\begin{aligned} 0 = \begin{vmatrix} 1-a & a-\lambda & 0 \\ 0 & 1-a & a-\lambda \\ q & r-\lambda & p \end{vmatrix} &= (1-\lambda)\begin{vmatrix} 1-a & 1 & 0 \\ 0 & 1 & a-\lambda \\ q & 1 & p \end{vmatrix} \\ &= (1-\lambda)\big[(\lambda-a)(1-a-q) + p(1-a)\big]. \end{aligned} \tag{38}$$

Assume that $a \neq 1-q$. Then $\lambda = 1$ is a solution of (38), and the other solution of (38), say $\lambda(a)$, and the associated complex number in (37b), say $z(a)$, are given by:

$$\lambda(a) := a + \frac{p(1-a)}{a-1+q} \in \mathbb{R} \quad \text{and} \quad z(a) := \frac{p}{a+q-1} \in \mathbb{R}. \tag{39}$$

Now, let $\lambda \in \mathbb{C}$ be such that $r + 2\sqrt{pq} < |\lambda| < 1$, and assume that there exists $f \in \mathcal{B}_{\widehat{\gamma}}$, $f \neq 0$, such that $Pf = \lambda f$. Then Lemma 6 gives $f := ({z_\lambda}^n)_{n\geq 0}$ (up to a multiplicative constant), with $z_\lambda \in \mathbb{C}$ satisfying $|z_\lambda| < \widehat{\gamma}$ and Equations (37a)-(37b). Thus we have $\lambda = \lambda(a)$ and $z_\lambda = z(a)$, with $\lambda(a)$ and $z(a)$ given by (39). Conversely, we have $Pf_a = \lambda(a)f_a$ with $f_a = (z(a)^n)_{n\geq 0}$ since, by definition, $z(a)$ satisfies the equations (37a)-(37b) associated with $\lambda = \lambda(a)$. Now we must find the values $a \in (0,1)$ for which we have $r + 2\sqrt{pq} < |\lambda(a)| < 1$ and $|z(a)| \leq \widehat{\gamma}$. This is the relevant question since Theorem 3 gives the following properties:

(i) if $r + 2\sqrt{pq} < |\lambda(a)| < 1$ and $|z(a)| < \widehat{\gamma}$, then we have $\rho_{V_{\widehat{\gamma}}}(P) = |\lambda(a)|$ since $\lambda(a)$ is the only eigenvalue λ of P on $\mathcal{B}_{\widehat{\gamma}}$ such that $r + 2\sqrt{pq} < |\lambda| < 1$ (apply Theorem 3 with any r_0 such that $r + 2\sqrt{pq} < r_0 < |\lambda(a)|$),

(ii) if $\lambda(a)$ or $z(a)$ do not satisfy the previous conditions, then we have $\rho_{V_{\widehat{\gamma}}}(P) = r + 2\sqrt{pq}$ since there is no eigenvalue λ of P on $\mathcal{B}_{\widehat{\gamma}}$ such that $r+2\sqrt{pq} < |\lambda| < 1$ (apply Theorem 3 with any r_0 such that $r + 2\sqrt{pq} < r_0 < 1$).

First, observe that

$$|z(a)| \leq \widehat{\gamma} \;\Leftrightarrow\; |a-1+q| \geq \sqrt{pq}. \tag{40}$$

Hence, if $a \in (a_0, 1)$ (recall that $a_0 := 1-q-\sqrt{pq}$), then $|z(a)| > \widehat{\gamma}$. Then (ii) gives (33).

Second consider the case $a \in (0, a_0]$. Then we have $|z(a)| \leq \widehat{\gamma}$, and we have to study $\lambda(a)$. Note that $\lambda'(a) = 1 - pq/(a-1+q)^2$, so that the function $a \mapsto \lambda(a)$ is increasing on $(-\infty, a_0]$ from $-\infty$ to $\lambda(a_0) = r - 2\sqrt{pq}$. Thus

$$\forall a \in (0, a_0], \quad \lambda(a) \leq r - 2\sqrt{pq} < r + 2\sqrt{pq}.$$

and the equation $\lambda(a) = -(r+2\sqrt{pq})$ has a unique solution $a_1 \in (-\infty, a_0)$. For the continuation, it suffices to have in mind that $a_1 < a_0$ and $\lambda(a_1) = -(r+2\sqrt{pq})$, that $\lambda(0) = p/(q-1) \in [-1, 0)$ and finally that

$$\lambda(0) - \lambda(a_1) = p/(q-1) + r + 2\sqrt{pq} = \frac{(q-\sqrt{pq}-1)^2 - 2p}{1-q}.$$

When $2p \leq (1-q+\sqrt{pq})^2$, (34) follows from (ii). Indeed $|\lambda(a)| < r + 2\sqrt{pq}$ since

$$\forall a \in (0, a_0], \quad -(r+2\sqrt{pq}) = \lambda(a_1) \leq \lambda(0) < \lambda(a) < r + 2\sqrt{pq}.$$

When $2p > (1-q+\sqrt{pq})^2$, we have $a_1 \in (0, a_0]$ and:

- if $a \in (0, a_1)$, then (35a) follows from (i). Indeed $r + 2\sqrt{pq} < |\lambda(a)| < 1$ since
$$\forall a \in (0, a_1], \quad -1 \leq \lambda(0) < \lambda(a) < \lambda(a_1) = -(r+2\sqrt{pq})\,;$$
- if $a \in [a_1, a_0]$, then (35b) follows (ii). Indeed $|\lambda(a)| < r + 2\sqrt{pq}$ since
$$-(r+2\sqrt{pq}) = \lambda(a_1) \leq \lambda(a) < r + 2\sqrt{pq}.$$

It remains to study the special case $a = 1-q$. Then $\lambda = 1$ is the only solution of (38). Again let $\lambda \in \mathbb{C}$ be such that $r + 2\sqrt{pq} < |\lambda| < 1$, and let $f \in \mathcal{B}_{\widehat{\gamma}}$, $f \neq 0$, such that $Pf = \lambda f$. Then Lemma 6 gives $f := ({z_\lambda}^n)_{n\geq 0}$, with $z_\lambda \in \mathbb{C}$ satisfying Equations (37a)-(37b), thus Equation (38). Consequently there is no eigenvalue of P such that $r + 2\sqrt{pq} < |\lambda| < 1$. Theorem 3 applied with any $r_0 \in (r+2\sqrt{pq}, 1)$ then gives $\rho_{V_{\widehat{\gamma}}}(P) = r + 2\sqrt{pq}$. $\square$

Remark 4 *Let us consider the time-discretised M/M/1 queue obtained using the uniformization technique [HS92, Section 4.1]. The arrival and service rates are denoted by $\beta > 0$ and $\mu > 0$ respectively. For $0 < h < 1/(\beta+\mu)$, the kernel P_h is defined by $P_h = I + hQ$ where Q is the generator of the continuous time birth-and-death process $(X_t)_{t\geq 0}$ of the number of customers in a M/M/1 queue, so that*

$$P_h(0,0) := 1 - \beta h, \quad P_h(0,1) := \beta h;$$
$$\forall n \geq 1, \quad P_h(n, n-1) = \mu h, \quad P_h(n,n) := 1 - h(\beta+\mu), \quad P_h(n, n+1) = \beta h.$$

Assume that $\beta/\mu < 1$ which is the ergodicity condition of the M/M/1 queue. Note that $\mu h, \beta h, 1-h(\beta+\mu), 1-\beta h$ stand for p, q, r, a with the notations of Proposition 2 and that

$p > q$, $a = P_h(0,0) \in (0,1)$ and $a + q - 1 = 0$. Therefore, for any $0 < h < 1/(\beta+\mu)$, P_h is $V_{\widehat{\gamma}}$-geometrically ergodic with $\widehat{\gamma} = \sqrt{\mu/\beta}$ and

$$\rho_{V_{\widehat{\gamma}}}(P_h) = 1 - h(\sqrt{\mu} - \sqrt{\beta})^2.$$

Note that $\rho_{V_{\widehat{\gamma}}}(P_h)$ is decreasing as h growth to $1/(\beta+\mu)$. The minimum (not attained) $2\sqrt{pq} = 2h\sqrt{\mu\beta} = 2\sqrt{\mu\beta}/(\mu+\beta)$ would be obtained as $h := 1/(\beta+\mu)$ but in this case $r := 0$ and we retrieve a binary random walk (and its convergence rate $\rho_{V_{\widehat{\gamma}}} = 2\sqrt{\mu\beta}$) corresponding to the embedded Markov chain associated with the birth-and-death Markov process. This last value was found to be the $\mathbb{L}^1$-convergence rate in [LT96] using the monotone structure of the Markov chain and to be the $\mathbb{L}^2$-convergence rate in [RT01] from the equality of the two rates using the reversibility of the model. Finally, note that exponential bounds for $(X_t)_{t\geq 0}$ are easily derived using that its semi-group $(P_t)_{t\geq 0}$ satisfies $P_t = \exp(Qt) = \sum_{k=0}^{\infty} {P_h}^k \exp(-t/h)(t/h)^k/k!$ and that P_h has the same invariant probability measure than $(X_t)_{t\geq 0}$ for any $h < 1/(\beta+\mu)$. Indeed, we obtain that for any $0 < \rho < (\sqrt{\mu} - \sqrt{\beta})^2$:

$$\forall f \in \mathcal{B}_{\widehat{\gamma}}, \quad \|P_t f - \pi(f)\|_{V_{\widehat{\gamma}}} = \sup_{n\in\mathbb{N}} \frac{|(P_t f)(n) - \pi(f)|}{V_{\widehat{\gamma}}(n)} = O\big(\exp(-t\rho)\big).$$

Note that $(\sqrt{\mu} - \sqrt{\beta})^2$ is the $\mathbb{L}^2(\pi)$-spectral gap for such a process (see e.g. [Kar00]).

4 From convergence on a subspace of $\mathcal{B}_V$ to V-geometrical ergodicity

In this subsection, P is assumed to have an invariant probability measure π on $(\mathbb{X}, \mathcal{X})$ such that $\pi(V) < \infty$. We give another procedure to get an upper bound for the convergence rate $\rho_V(P)$. This method is based on Theorem 2 which involves the compactness of P from $\mathcal{B}_0$ to $\mathcal{B}_V$ and Condition (**WD**), and on the new following conditions:

Conditions (B). *There exists a subspace $\mathcal{B}$ of $\mathcal{B}_V$ such that we have:*

(a) $\exists \kappa_{\mathcal{B}} \in (0,1)$ *such that* $\forall f \in \mathcal{B}$, $\exists\, c_f \in [0,+\infty)$, $\quad \|P^n f - \pi(f)1_{\mathbb{X}}\|_V \leq c_f\, {\kappa_{\mathcal{B}}}^n$,

(b) $\exists \tau \in [0,1]$ *such that* $P(\mathcal{B}_{V^\tau}) \subset \mathcal{B}$, *where*

$$\mathcal{B}_{V^\tau} := \big\{f :\ \mathbb{X} \to \mathbb{C} \text{ measurable such that } \sup_{x\in\mathbb{X}} |f(x)|/V(x)^\tau < \infty.\big\}$$

Note that, for $\tau = 0$, the space $\mathcal{B}_{V^\tau}$ corresponds to the space $\mathcal{B}_0$ of all the bounded measurable $\mathbb{C}$-valued functions on $\mathbb{X}$.

Theorem 5 *The transition kernel P is assumed to satisfy Conditions* (**WD**) *and* (B) *and to be compact from $\mathcal{B}_0$ to $\mathcal{B}_V$. Then P is V-geometrically ergodic, and we have, with $\delta_V(P)$ defined in (3):*

$$\rho_V(P) \leq \max\big(\kappa_{\mathcal{B}}, \delta_V(P)^\tau\big). \tag{41}$$

Proof. From Theorem 2, P is a power-bounded quasi-compact operator on $\mathcal{B}_V$, with in addition $r_{ess}(P) \leq \delta_V(P)$. It follows from Theorem 3 that 1) P is V-geometrically ergodic; 2) Property (41) is fulfilled, provided that

(i) $\lambda = 1$ is a simple eigenvalue of P on $\mathcal{B}_V$;

(ii) $\lambda = 1$ is the unique eigenvalue of P on $\mathcal{B}_V$ in $\{z \in \mathbb{C} : \max(\kappa_{\mathcal{B}}, (\delta_V(P))^\tau) < |z| \leq 1\}$ for $\tau \neq 0$ and in $\{z \in \mathbb{C} : |z| = 1\}$ for $\tau = 0$.

For (ii), if $\tau \neq 0$, note that $r_{ess}(P) \leq \delta_V(P) \leq (\delta_V(P))^\tau \leq \max(\kappa_{\mathcal{B}}, (\delta_V(P))^\tau)$, so that Theorem 3 can be applied with any r_0 such that $\max(\kappa_{\mathcal{B}}, (\delta_V(P))^\tau) < r_0 < 1$. If $\tau := 0$, then only the equivalence $(a) \Leftrightarrow (b)$ of Theorem 3 is needed.

Let us prove (i)-(ii). Let $f \in \mathcal{B}_V$ such that $Pf = \lambda f$ with $r_0 \leq |\lambda| \leq 1$ in case $\tau \neq 0$, and with $|\lambda| = 1$ in case $\tau = 0$. Let $\delta \in (\delta_V(P), 1)$ be such that $\delta^\tau \leq |\lambda|$. From Theorem 4 with $p := 1$, there exists $c \in (0, +\infty)$ such that

$$|f| \leq c\, V^{\frac{\ln|\lambda|}{\ln\delta}} \leq c\, V^\tau.$$

Thus $f \in \mathcal{B}_{V^\tau}$. The equality $Pf = \lambda f$ and Condition (B)(b) then give $f \in \mathcal{B}$. Besides, since $|\lambda| > \kappa_{\mathcal{B}}$, we deduce from Condition (B)(a) the following properties: first, if $\lambda = 1$, then f is constant on $\mathbb{X}$; second, if $\lambda \neq 1$, then $f = 0$ (use Condition (B)(a) and the fact $\pi(f) = 0$ easily derived from $Pf = \lambda f$). This proves (i) and (ii). $\square$

Remark 5 *The real number c_f in Condition* (B)(a) *is not necessarily of the form $c\|f\|_V$ (with fixed $c > 0$). In other words Condition* (B)(a) *does not involve (a priori) the operator norm on $\mathcal{B}_V$. This fact occurs for instance in the following case. Suppose that there exists a space $(\mathcal{B}, \|\cdot\|)$ such that $\mathcal{B} \subset \mathcal{B}_V$ with continuous inclusion (i.e. $\exists d > 0$, $\forall f \in \mathcal{B}$, $\|f\|_V \leq d\,\|f\|$) and such that P continuously acts and is strongly ergodic on $(\mathcal{B}, \|\cdot\|)$, namely there exist $\kappa \in (0,1)$ and $c > 0$ such that*

$$\forall f \in \mathcal{B}, \quad \|P^n f - \pi(f)1_{\mathbb{X}}\| \leq c\,\kappa^n\, \|f\|.$$

Then Condition (B)(a) *is fulfilled with $c_f = c\,d\|f\|$.*

Remark 6 *If Condition* (B)(a) *holds and $V \in \mathcal{B}$, then Condition* (**WD**) *is fulfilled and*

$$\delta_V(P) \leq \kappa_{\mathcal{B}}.$$

Indeed let $\delta \in (\kappa_{\mathcal{B}}, 1)$ and let $N \in \mathbb{N}^*$ be such that ${c_V}^{1/N}\kappa_{\mathcal{B}} \leq \delta$. Then Condition (B)(a) *applied to $f := V$ gives: $P^N V \leq c_V\, \kappa_{\mathcal{B}}^N\, V + \pi(V) \leq \delta^N V + \pi(V)$.*

Remark 7 *Assume that Condition* (B)(a) *holds true and that $f : \mathbb{X} \to \mathbb{R}$ is such that $P^\ell f \in \mathcal{B}$ for some $\ell \in \mathbb{N}^*$. Then we have*

$$\forall n \geq \ell, \quad \|P^n f - \pi(f)1_{\mathbb{X}}\|_V \leq c_{P^\ell f}\, {\kappa_{\mathcal{B}}}^{n-\ell}.$$

Although this is not directly connected with the V-geometric ergodicity, such estimate may be of interest when the constant $c_{P^\ell f}$ is explicit. This is illustrated in the next section (see Remark 9 and Corollary 7).

5 V-geometrical ergodicity of iterated function systems

In this section we assume that $(\mathbb{X}, d)$ is a metric space equipped with its borel σ-algebra $\mathcal{X}$. Let $(\mathbb{V}, \mathcal{V})$ be a measurable space. Let us first recall the definition of an iterated function system (IFS) of Lipschitz maps (see [DF99, Duf97]).

Definition 1 (IFS of Lipschitz maps) *Let $(\vartheta_n)_{n\geq 1}$ be a sequence of $\mathbb{V}$-valued i.i.d. random variables, with common distribution denoted by ν. Let X_0 be a $\mathbb{X}$-valued r.v. which is assumed to be independent of the sequence $(\vartheta_n)_{n\geq 1}$. Finally, let $F : (\mathbb{V} \times \mathbb{X}, \mathcal{X} \otimes \mathcal{V}) \to (\mathbb{X}, \mathcal{X})$ be jointly measurable and Lipschitz continuous with respect to the second variable. The associated iterated function system (IFS) is the sequence of random variables $(X_n)_{n\in\mathbb{N}}$ which, given X_0, is recursively defined by:*

$$\forall n \geq 1, \quad X_n := F(\vartheta_n, X_{n-1}). \tag{42}$$

Clearly $(X_n)_{n\in\mathbb{N}}$ is a Markov chain, with transition kernel P:

$$\forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad P(x, A) = \mathbb{E}[1_A\big(F(\vartheta_1, x)\big)] = \int_{\mathbb{V}} 1_A\big(F(v, x)\big)\, d\nu(v). \tag{43}$$

Let $x_0 \in \mathbb{X}$ be fixed. For any $b \in [0, +\infty)$, we set

$$\forall x \in \mathbb{X},\ p(x) := 1 + d(x, x_0) \quad \text{and} \quad V_b(x) := p(x)^b.$$

We simply denote by $(\mathcal{B}_b, |\cdot|_b)$ the weighted-supremum Banach space $\mathcal{B}_{V_b}$ associated with $V_b(\cdot)$, that is

$$\mathcal{B}_b := \left\{ f : \mathbb{X} \to \mathbb{C} \text{ measurable such that } |f|_b := \sup_{x\in\mathbb{X}} \frac{|f(x)|}{p(x)^b} < \infty \right\}. \tag{44}$$

If $\psi : (\mathbb{X}, d) \to (\mathbb{X}, d)$ is a Lipschitz continuous function, we define

$$L(\psi) := \sup \left\{ \frac{d\big(\psi(x), \psi(y)\big)}{d(x, y)},\ (x, y) \in \mathbb{X}^2,\ x \neq y \right\}. \tag{45}$$

Let $a \in [1, +\infty)$. We denote by $\mathcal{L}_a$ the following space:

$$\mathcal{L}_a := \left\{ f : \mathbb{X} \to \mathbb{C}\ :\ m_a(f) := \sup \left\{ \frac{|f(x) - f(y)|}{d(x, y)\,(p(x) + p(y))^{a-1}},\ (x, y) \in \mathbb{X}^2,\ x \neq y \right\} < \infty \right\}. \tag{46}$$

Such Lipschitz-weighted spaces have been introduced in [LP83] to obtain quasi-compactness of Lipschitz kernels, see also [MR89, Duf97, Ben98, HH01].

Note that, for $f \in \mathcal{L}_a$, we have for all $x \in \mathbb{X}$: $|f(x)| \leq |f(x_0)| + 2^{a-1}\, m_a(f)\, p(x)^a$. Thus:

$$\forall f \in \mathcal{L}_a, \quad |f|_a := \sup_{x\in\mathbb{X}} \frac{|f(x)|}{p(x)^a} < \infty \quad \text{and} \quad \mathcal{L}_a \subset \mathcal{B}_a.$$

In this section we apply the results of Section 4. More specifically, in Subsection 5.1 we give standard contraction/moment conditions for P to satisfy Assertion (a) of Conditions (B) with $V := V_a$ and $\mathcal{B} := \mathcal{L}_a$ for some $a \in [1, +\infty)$. In Subsection 5.3 the passage to the V_a-geometric ergodicity is investigated under Conditions (B).

5.1 Basic inequalities for IFS

For all $x \in \mathbb{X}$, $v \in \mathbb{V}$ and $(v_1, \ldots, v_n) \in \mathbb{V}^n$ ($n \in \mathbb{N}^*$), define:

$$F_v x := F(v, x) \quad \text{and} \quad L(v) := L(F_v) \tag{47a}$$
$$F_{v_n:v_1} := F_{v_n} \circ \cdots \circ F_{v_1} \quad \text{and} \quad L(v_n : v_1) := L(F_{v_n:v_1}). \tag{47b}$$

By hypothesis we have $L(v) < \infty$, and so $L(v_n : v_1) < \infty$. Note that the limit

$$\hat{\kappa}_a := \lim_{n \to +\infty} \mathbb{E}\left[L(F_{\vartheta_n:\vartheta_1})^a\right]^{\frac{1}{na}}$$

exists in $[0, +\infty]$, since the sequence $(\mathbb{E}[L(\vartheta_n : \vartheta_1)^a])_{n \in \mathbb{N}^*}$ is submultiplicative. Let us consider the following classical moment/contraction conditions:

Conditions $(\mathcal{C}_a)$. *For some $a \in [1, +\infty)$:*

$$\mathbb{E}\left[d(F_{\vartheta_1}x_0, x_0)^a\right] < \infty \tag{48a}$$
$$\widehat{\kappa}_a < 1. \tag{48b}$$

Proposition 3 (see [Duf97, Ben98]) *Under Conditions $(\mathcal{C}_a)$, there exists a unique P-invariant distribution, denoted by π, on $(\mathbb{X}, \mathcal{X})$, and we have $\pi(d(x_0, \cdot)^a) < \infty$.*

Proposition 4 ([Duf97]) *Under Conditions $(\mathcal{C}_a)$, the transition kernel P continuously acts on $\mathcal{L}_a$, and for any $\kappa \in (\widehat{\kappa}_a, 1)$, there exists a constant $c \equiv c_\kappa$ such that we have:*

$$\forall n \geq 1,\ \forall f \in \mathcal{L}_a, \quad |P^n f - \pi(f)1_{\mathbb{X}}|_a \leq c\, \kappa^n\, m_a(f). \tag{49}$$

In particular, if $\kappa_1 := \mathbb{E}[L(\vartheta_1)^a]^{\frac{1}{a}} < 1$, then

$$\forall f \in \mathcal{L}_a,\ \forall n \geq 1, \quad |P^n f - \pi(f)1_{\mathbb{X}}|_a \leq c_1\, \kappa_1^n\, m_a(f), \tag{50}$$

where the constant c_1 is defined by $c_1 := \xi^{(a-1)/a}\, \|\pi\|_1 \big(1 + \|\pi\|_a\big)^{a-1}$, with

$$\xi := \sup_{n \geq 1} \sup_{x \in \mathbb{X}} \frac{(P^n V_a)(x)}{V_a(x)} < \infty \qquad \textit{and} \qquad \|\pi\|_b := \left(\int_{\mathbb{X}} p(y)^b\, d\pi(y)\right)^{1/b} \qquad \textit{for } b := 1, a.$$

Property (49) can be derived from the results of [Duf97, Chapter 6]. For convenience we give the proof of (50) in Appendix D. The proof of (49) is similar (replace P by P^N with N such that $\mathbb{E}[L(F_{\vartheta_N:\vartheta_1})^a] < 1$).

Note that the properties (49) and (50) do not provide the V_a-geometric ergodicity since they are only established for $f \in \mathcal{L}_a$. Indeed, in general the spaces $\mathcal{L}_a$ and $\mathcal{B}_a$ do not coincide, even for countable Markov chains.

Remark 8 *Under the conditions $(\mathcal{C}_a)$ and $\mathbb{E}[L(\vartheta_1)^a]^{\frac{1}{a}} < 1$, the proof of* (50) *(see Appendix D) gives the following bound for the constant ξ of Proposition 4. Given any real number δ satisfying $\mathbb{E}[L(\vartheta_1)^a] < \delta < 1$, choose r such that*

$$d(x, x_0) > r \ \Rightarrow\ \mathbb{E}\left[\left(\frac{1 + L(\vartheta_1)\, d(x, x_0) + d(F_{\vartheta_1}x_0, x_0)}{1 + d(x, x_0)}\right)^a\right] \leq \delta.$$

Then we have, with $\xi_1 := \mathbb{E}\big[\big(\max(1, L(\vartheta_1)) + d(F_{\vartheta_1}x_0, x_0)\big)^a\big]$,

$$\xi \leq 1 + \frac{\xi_1(1+r)^a}{1-\delta}.$$

Remark 9 *From Proposition 4 and Remark 7 we deduce the following fact. Assume that Conditions* $(\mathcal{C}_a)$ *holds and that* $f : \mathbb{X} \to \mathbb{R}$ *is such that* $P^\ell f \in \mathcal{L}_a$ *for some* $\ell \in \mathbb{N}^*$*. Then, for any* $\kappa \in (\widehat{\kappa}_a, 1)$*, we have*

$$\forall n \geq \ell, \quad |P^n f - \pi(f)1_{\mathbb{X}}|_a \leq c\,\kappa^{n-\ell}\, m_a(P^\ell f), \tag{51}$$

where $c \equiv c_\kappa$ *is the constant of Proposition 4. If* $\kappa = \kappa_1 := \mathbb{E}[L(\vartheta_1)^a]^{1/a} < 1$*, then* $c = c_1$*.*

Example 5 (A simple example) *Let* $(X_n)_{n\in\mathbb{N}}$ *be the real-valued IFS*

$$X_0 \in \mathbb{R}, \qquad \forall n \geq 1, \quad X_n := \vartheta_n X_{n-1},$$

associated with $F_v x := vx$ *and with a sequence* $(\vartheta_n)_{n\geq 1}$ *of i.i.d. random variables assumed to be independent of* X_0*. This kind of multiplicative Markov models are popular in finance. Let us assume that* ϑ_n*'s have a uniform probability distribution on* $[0,1]$*. The transition kernel* $P(x, dy)$ *of* $(X_n)_{n\geq 0}$ *is the uniform distribution on* $[0,x]$ *if* $x > 0$ *(resp. on* $[x, 0]$ *if* $x < 0$*). The Dirac distribution* δ_0 *at* 0 *is clearly* P*-invariant. Finally, setting* $x_0 := 0$ *and* $d(x,y) := |x-y|$*, we have:* $\forall a \in [1, +\infty)$

$$\mathbb{E}[d(F_{\vartheta_1}0, 0)^a] = 0 \quad \text{and} \quad \kappa_1 := \mathbb{E}[L(\vartheta_1)^a]^{\frac{1}{a}} = \mathbb{E}[\vartheta_1{}^a]^{\frac{1}{a}} = \left(\frac{1}{a+1}\right)^{\frac{1}{a}}.$$

Consequently Inequality (50) is valid. If $a := 1$*, then the constant* c_1 *in (50) is equal to* 1 *since we have* $p(x) = 1 + |x|$ *and* $\pi = \delta_0$*.*

Example 6 (Autoregressive model) *In this example, we prove that Inequality (50) is fulfilled with the (optimal) value* $\kappa_1 := \widehat{\kappa}_a$*. Let* $(X_n)_{n\in\mathbb{N}}$ *be the real-valued IFS*

$$X_0 \in \mathbb{R}, \qquad \forall n \geq 1, \quad X_n := \alpha\, X_{n-1} + \vartheta_n,$$

associated with $F_v x := \alpha x + v$ *where* $\alpha \in (-1, 1)$ *is fixed and with a sequence* $(\vartheta_n)_{n\geq 1}$ *of centered random variables. This is the so-called autoregressive model of order 1 with an arbitrary centered noise. We take* $d(x,y) := |x-y|$ *and* $x_0 := 0$*, so that* $L(v_n : v_1) = |\alpha|^n$ *and* $|F_v 0| = |v|$*. Then, we have* $\widehat{\kappa}_a = |\alpha|$ *for all* $a \in (1, +\infty)$*. In particular we have* $\widehat{\kappa}_a = \mathbb{E}[L(\vartheta_1)^a]^{1/a} < 1$*. Therefore, Conditions* $(\mathcal{C}_a)$ *reduce to the moment condition* $\mathbb{E}[|\vartheta_1|^a] < \infty$*, and under this condition, Inequality (50) holds for* $\kappa_1 := |\alpha|$*.*

Next, let us check that $|\alpha|$ *is the minimal value of* κ *(and* κ_1*) such that Inequality (49) (and Inequality (50)) is valid. Since* ϑ_1 *is centered, we have:* $\forall x \in \mathbb{X}$, $\mathbb{E}[X_1 \mid X_0 = x] = \mathbb{E}[\alpha x + \vartheta_1] = \alpha x$*. In other words, we have* $P\phi = \alpha\phi$ *where* ϕ *is the identity function on* $\mathbb{R}$ *(i.e.* $\phi(x) := x$*). Note that* ϕ *is in* $\mathcal{L}_a$ *for every* $a \geq 1$*, and that* $\pi(\phi) = 0$ *using* $P\phi = \alpha\phi$*. Consequently, under the condition* $\mathbb{E}[|\vartheta_1|^a] < \infty$*, Inequality (49) holds for* $f := \phi$*, and we can deduce from* $P^n\phi = \alpha^n\phi$ *that Inequality (49) cannot be valid with some* $\kappa < |\alpha|$*.*

5.2 Application to discrete Lindley's random walk

Recall that a Lindley random walk is defined on $\mathbb{X} := [0, +\infty)$ by $X_n := \max(0, X_{n-1} + \vartheta_n)$, where $(\vartheta_n)_{n\geq 1}$ is a sequence of $\mathbb{R}$-valued i.i.d. random variables independent of X_0. The rate of convergence of Lindley's random walks with respect to $\|\cdot\|_V$ is investigated in [Lun97, LT96]. More specifically, under the assumptions $\mathbb{E}[\gamma_0^{\vartheta_1}] < \infty$ for some $\gamma_0 \in (1, +\infty)$ and $\mathbb{E}[\vartheta_1] < 0$, it is proved that there exists $\gamma \in (1, \gamma_0]$ such that $\mathbb{E}[\gamma^{\vartheta_1}] < 1$, that P is V-geometrically ergodic with $V(x) = \gamma^x$, and that $\rho_V(P) = \mathbb{E}[\gamma^{\vartheta_1}]$. The constant c_ρ of (11) associated with any $\rho \in (\mathbb{E}[\gamma^{\vartheta_1}], 1)$ is not computed in [Lun97, LT96].

This subsection is devoted to the special case of discrete Lindley's random walks. More specifically, under the above assumption, setting $V_\gamma := (\gamma^n)_{n\in\mathbb{N}}$, we prove that the V_γ-geometric ergodicity property (11) holds true with the optimal rate $\rho := \mathbb{E}[\gamma^{\vartheta_1}]$ and with an explicit (and simple) constant c_ρ. This result is based on Proposition 3 and Proposition 4 using the distance $d(i,j) := |\gamma^i - \gamma^j|,\ (i,j) \in \mathbb{N}^2$.

Let X_0 be a $\mathbb{N}$-valued r.v. and $(\vartheta_n)_{n\geq 1}$ be a sequence of i.i.d. $\mathbb{Z}$-valued r.v., independent of X_0. Let us introduce the sequence of $\mathbb{N}$-valued r.v. $(X_n)_{n\in\mathbb{N}}$ defined by

$$\forall n \geq 1, \quad X_n := \max(0, X_{n-1} + \vartheta_n).$$

The common distribution $\nu := (\nu_j)_{j\in\mathbb{N}}$ of the ϑ_n's is assumed to be such that

$$\exists \gamma_0 \in (1, +\infty), \quad \mathbb{E}[\gamma_0^{\vartheta_1}] = \sum_{j\geq 0} \nu_j\, \gamma_0^j < \infty \quad \text{and} \quad \mathbb{E}[\vartheta_1] = \sum_{j\in\mathbb{Z}} j\, \nu_j < 0. \tag{52}$$

Proposition 5 *Under the assumptions (52), there exists $\gamma \in (1, \gamma_0]$ such that*

$$\kappa_1 := \mathbb{E}[\gamma^{\vartheta_1}] < 1,$$

and $(X_n)_{n\geq 0}$ is V_γ-geometrically ergodic with $V_\gamma := (\gamma^n)_{n\in\mathbb{N}}$. More precisely, we have the following properties:

$$\forall f \in \mathcal{B}_{V_\gamma},\ \forall n \geq 1,\ \forall i \in \mathbb{N}, \quad |(P^n f)(i) - \pi(f)| \leq c_1\, {\kappa_1}^n\, m_1(f)\, \gamma^i \tag{53a}$$

$$\forall (i,j) \in \mathbb{N}^2,\ \forall n \geq 1, \quad \big|\mathbb{P}[X_n = j \,|\, X_0 = i] - \pi(\{j\})\big| \leq \frac{c_1\, \gamma^{i+1}}{(\gamma - 1)\, \gamma^j}\ {\kappa_1}^n \tag{53b}$$

with $c_1 := \pi(V)$. In particular we have $\rho_V(P) \leq \kappa_1$. More precisely, Inequality (11) is fulfilled with $\rho := \kappa_1$ and $c_\rho := \pi(V)(\gamma+1)/(\gamma-1)$.

Proof. The first assertion holds since $G(\gamma) := \mathbb{E}[\gamma^{\vartheta_1}]$ satisfies $G(1) = 1$ and $G'(1) = \mathbb{E}[\vartheta_1] < 0$. To prove (53a)-(53b), we apply Proposition 4 with the distance[3]

$$\forall (i,j) \in \mathbb{N}^2, \quad d(i,j) := |\gamma^i - \gamma^j|. \tag{54}$$

[3] The fact that, in Proposition 5, the geometrical ergodicity is directly deduced from Proposition 4 is very particular. This is due to the choice of the distance in (54).

Note that we have with $x_0 = 0$: $\forall i \in \mathbb{N},\ p(i) := 1 + d(i,0) = \gamma^i$. Thus the space $\mathcal{B}_{V_\gamma}$ corresponds to $\mathcal{B}_1$ in (44). Next observe that the spaces $\mathcal{B}_1$ and $\mathcal{L}_1$ coincide. Indeed, for all $f = (f(n))_{n\in\mathbb{N}}$ such that $|f|_1 := \sup_{n\in\mathbb{N}} |f(n)|/\gamma^n < \infty$, we have (use $\sup_{n\in\mathbb{N}^*}(\gamma^n+1)/(\gamma^n-1) = (\gamma+1)/(\gamma-1)$)

$$m_1(f) := \sup\left\{\frac{|f(i)-f(j)|}{|\gamma^i-\gamma^j|},\ (i,j)\in\mathbb{N}^2, i\neq j\right\} \le \frac{\gamma+1}{\gamma-1}\,|f|_1. \tag{55}$$

Next we have: $\forall (v,i) \in \mathbb{Z}\times\mathbb{N},\ F_v i := \max(0, i+v)$. Besides we compute the Lipschitz (random) coefficient $L(\vartheta_1)$ with respect to the distance $d(i,j) := |\gamma^i - \gamma^j|$. We obtain for $(i,j)\in\mathbb{N}^2$ such that $i<j$ and for all $v\in\mathbb{Z}$:

(a) $d\big(F_v i, F_v j\big) = \gamma^v\,\big|\gamma^i - \gamma^j\big|$ when $i+v\ge 0$ and $j+v\ge 0$;

(b) $d\big(F_v i, F_v j\big) = \big|1-\gamma^{j+v}\big| = \gamma^v\,\big|\gamma^{-v}-\gamma^j\big|$ when $i+v<0$ and $j+v\ge 0$;

(c) $d\big(F_v i, F_v j\big) = 0$ when $i+v<0$ and $j+v<0$.

In Case (b), we have $i < -v \le j$, thus $|\gamma^{-v}-\gamma^j| \le |\gamma^i-\gamma^j|$. Thus

$$L(v) := \sup_{(i,j)\in\mathbb{N}^2, i\neq j} \frac{d\big(F_v i, F_v j\big)}{|\gamma^i-\gamma^j|} = \gamma^v.$$

Finally, we obtain $\mathbb{E}[d(F_{\vartheta_1}0,0)] = \mathbb{E}[|\gamma^{\max(0,\vartheta_1)}-1|] \le \mathbb{E}[\gamma^{\vartheta_1}]$. Thus Conditions (52) implies that Conditions $(\mathcal{C}_1)$ holds with $\mathbb{E}[L(\vartheta_1)] = \mathbb{E}[\gamma^{\vartheta_1}] < 1$. Consequently, P has an invariant distribution π such that $\pi(V)<\infty$ from Proposition 3. Then, Property (53a) follows from (50) with $a := 1$ (note that $c_1 = \int_{\mathbb{X}} p(y)\,d\pi(y)$). To obtain (53b), use the fact that $m_1(1_{\{j\}}) = (\gamma-1)^{-1}\gamma^{1-j}$. □

Example 7 (Simulation of a geometric distribution) *The Markov kernel P on $\mathbb{X} := \mathbb{N}$ defined for $p\in(0,1)$ by*

$$P(0,0) := 1-p/2, \quad P(0,1) := p/2$$
$$\forall i\in\mathbb{N}^*, \quad P(i,i-1) := 1/2, \quad P(i,i) := (1-p)/2, \quad P(i,i+1) := p/2.$$

arises from a Hastings-Metropolis sampler of a geometric distribution with parameter p. In [MT96, Example 2], P is shown to be V-geometrically ergodic with $V = (p^{-n/2})_{n\ge 0}$ and to have a convergence rate satisfying $\rho_V(P) \le \rho_0 := \sqrt{p} + (1-p)/2$. More specifically: Property (11) holds for any $\rho\in(\rho_0,1)$ with some constant c_ρ such that $\lim_{\rho\to\rho_0} c_\rho = +\infty$. Proposition 5 allows us to improve this result.

Indeed, P can also be viewed as the Markov kernel of the discrete Lindley random walk where ϑ_1 is distributed as: $\mathbb{P}(\vartheta_1=-1)=1/2$, $\mathbb{P}(\vartheta_1=0)=(1-p)/2$ and $\mathbb{P}(\vartheta_1=1)=p/2$. Consequently the estimates (53a)-(53b) are valid for this kernel, with here $\gamma := p^{-1/2}$ and $\kappa_1 := \mathbb{E}[\gamma^{\vartheta_1}] = \rho_0$. In particular, Property (11) holds true with $\rho := \rho_0$ and $c_{\rho_0} := (1+\sqrt{p})^2/(1-\sqrt{p})$, namely we have for all $f=(f(n))_{n\in\mathbb{N}}$ such that $|f|_1 := \sup_n |f(n)|\,p^{n/2} < \infty$:

$$\forall n\ge 1,\ \forall i\in\mathbb{N}, \quad p^{i/2}\,\big|P^n f(i) - \pi(f)\big| \le \frac{(1+\sqrt{p})^2}{1-\sqrt{p}}\,|f|_1\left(\sqrt{p}+\frac{1-p}{2}\right)^n.$$

5.3 V_a-Geometrical ergodicity of IFS

Recall that we have set: $\forall x \in \mathbb{X},\ V_a(x) := p(x)^a$. Let $(X_n)_{n\in\mathbb{N}}$ be an IFS. Under Conditions $(\mathcal{C}_a)$, Alsmeyer proved that, when $(X_n)_{n\in\mathbb{N}}$ is Harris recurrent and the support of π has a non-empty interior, $(X_n)_{n\in\mathbb{N}}$ is V_a-geometrically ergodic, see [Als03, Prop. 5.2][4]. In this subsection, we provide additional conditions to $(\mathcal{C}_a)$ for $(X_n)_{n\in\mathbb{N}}$, first to have a quasi-compact kernel P on $\mathcal{B}_a$ with an essential spectral radius such that $r_{ess}(P) \le \widehat{\kappa}_a$, second to be V_a-geometrically ergodic with a convergence rate such that $\rho_V(P) \le (\widehat{\kappa}_a)^\tau$ for some $\tau \in (0,1]$. Our additional conditions are stronger and quite different from those of [Als03]. In particular P is here assumed to be compact from $\mathcal{B}_0$ into $\mathcal{B}_a$.

If Conditions $(\mathcal{C}_a)$ are fulfilled, then Property (49) with $f := V_a$ and $n := 1$ gives $PV_a \le \xi_1 V_a$ for some $\xi_1 \in (0,+\infty)$, and so P continuously acts on $\mathcal{B}_a$. Besides Inequality (49) is nothing else but Condition (B)(a) of Section 4 with $\mathcal{B} := \mathcal{L}_a$ and any $\kappa_{\mathcal{B}} \in (\widehat{\kappa}_a, 1)$. Moreover, P satisfies Condition (**WD**) with the function V_a since $V_a \in \mathcal{L}_a$ (see Remark 6) and the real number $\delta_{V_a}(P)$ associated with V_a via the definition (3) satisfies:

$$\delta_{V_a}(P) \le \widehat{\kappa}_a. \tag{57}$$

We deduce the following statement from Theorem 2.

Corollary 4 *Assume that Conditions $(\mathcal{C}_a)$ hold true for some $a \in [1,+\infty)$ and that P is compact from $\mathcal{B}_0$ to $\mathcal{B}_a$. Then P is a power-bounded quasi-compact operator on $\mathcal{B}_a$, and*

$$r_{ess}(P) \le \delta_{V_a}(P) \le \widehat{\kappa}_a.$$

Next Theorem 5 gives the following.

Corollary 5 *Assume that Conditions $(\mathcal{C}_a)$ hold true for some $a \in [1,+\infty)$, that P is compact from $\mathcal{B}_0$ to $\mathcal{B}_a$ and satisfies the following condition:*
(b)′ $\exists \tau \in [0,1]$ such that $P(\mathcal{B}_{\tau a}) \subset \mathcal{L}_a$.
Then P is V_a-geometrically ergodic and its convergence rate on $\mathcal{B}_a$ verifies

$$\rho_{V_a}(P) \le \max\left(\widehat{\kappa}_a, \delta_{V_a}(P)^\tau\right) \le (\widehat{\kappa}_a)^\tau. \tag{58}$$

Remark 10 *In simple examples (as in Example 6), the optimal rate in (49) is equal to $\widehat{\kappa}_a$. In this case, we have $\rho_{V_a}(P) \ge \widehat{\kappa}_a$ since $\mathcal{L}_a \subset \mathcal{B}_a$.*

Remark 11 *Inequality (58) means that, for any real number κ such that $\max(\widehat{\kappa}_a, (\delta_{V_a}(P))^\tau) < \kappa < 1$, there exists a constant $e \equiv e_\kappa$ such that*

$$\forall n \ge 1,\ \forall f \in \mathcal{B}_a,\quad |P^n f - \pi(f)\, 1_{\mathbb{X}}|_a \le e\, \kappa^n\, |f|_a. \tag{59}$$

It is worth noticing that Theorem 5 does not give any information on the constant e of (59). Inequality (51) is more precise but in general is only valid for a smaller class of functions f.

[4] Under Conditions $(\mathcal{C}_a)$, we have $P^{m_0} V_a \le \varrho V_a + M\, 1_{B_r}$ where $m_0 \in \mathbb{N}^*$, $\varrho \in (0,1)$, $M \in (0,+\infty)$ and B_r is the ball in $\mathbb{X}$ centered at x_0 with some suitable radius $r > 0$ (see Appendix D). Under Alsmeyer's conditions, B_r is a small set (see [MT93]), so that the usual drift condition holds.

5.4 Applications to autoregressive models

Assume that $\mathbb{X} := \mathbb{R}^q$ and denote the Lebesgue measure on $\mathbb{R}^q$ by dy. Let $\|\cdot\|$ denote any norm of $\mathbb{R}^q$, and define $d(x,y) := \|x-y\|$ the associated distance on $\mathbb{R}^q$. Set $p(x) := 1+\|x\|$ ($x_0 := 0$) and let us consider $V_a(x) := (1+\|x\|)^a$ with $a \in (1,+\infty)$. We have $\lim_{\|x\|\to\infty} V_a(x) = \infty$. We know from the proof of Corollary 1 that any Markov kernel $P(x,dy) = K(x,y)\,dy$, associated with a continuous (in the first variable) function $K : \mathbb{R}^q \times \mathbb{R}^q \to [0,+\infty)$, is compact from $\mathcal{B}_0$ to $\mathcal{B}_a$. This fact allows us to apply Corollaries 4-5 to some classical IFSs. As an illustration, Properties (51) and (58) are detailed below for affine autoregressive (AR) models. Such applications can be easily extended for others IFSs, as for instance for functional autoregressive models and AR processes with ARCH errors (see Examples 9-10).

Let $(X_n)_{n\in\mathbb{N}}$ be the IFS

$$X_0 \in \mathbb{R}^q, \qquad \forall n \geq 1, \quad X_n := AX_{n-1} + \vartheta_n, \tag{60}$$

associated with $F(v,x) := Ax + v$ where $A = (a_{ij})$ is a fixed real $q \times q$-matrix. We have $L(v) = \|A\|$ where $\|A\|$ denotes the matrix norm of A associated with $\|\cdot\|$, and $d(F_v 0, 0) = \|v\|$. Consequently, Conditions $(\mathcal{C}_a)$ hold for $a \in [1,+\infty)$ provided that we have:

$$\|A\| < 1 \qquad \text{and} \qquad \mathbb{E}\big[\|\vartheta_1\|^a\big] < \infty. \tag{61}$$

Under these conditions, P has an invariant probability measure from Proposition 3, and we can easily prove that $\limsup_{\|x\|\to+\infty} PV_a(x)/V_a(x) \leq \|A\|^a$. Thus

$$\widehat{\kappa}_a = \|A\| = \mathbb{E}[L(\vartheta_1)^a]^{\frac{1}{a}} \qquad \text{and} \qquad \delta_{V_a}(P) \leq \|A\|^a.$$

Corollary 6 *Assume that Conditions (61) hold true for some $a \in [1,+\infty)$ and that the common distribution of $(\vartheta_n)_{n\geq 1}$ has a density with respect to the Lebesgue measure on $\mathbb{R}^q$. Then P is a power-bounded quasi-compact operator on $\mathcal{B}_a$, and we have*

$$r_{ess}(P) \leq \|A\|^a.$$

Proof. From Corollary 4, it remains to prove that P is compact from $\mathcal{B}_0$ to $\mathcal{B}_a$. Let $\nu(\cdot)$ denote the density of ϑ_1. Note that P has the form (10) with the Lebesgue measure on $\mathbb{R}^q$ and $K(x,y) := \nu\big(y - Ax\big)$ for $(x,y) \in \mathbb{R}^q \times \mathbb{R}^q$. If $\nu(\cdot)$ is continuous, then the desired property follows from the proof of Corollary 1. In the general case we can proceed as follows. Let $f \in \mathcal{B}_0$ such that $\|f\|_0 \leq 1$. Then we have

$$\forall (x,x') \in \mathbb{R}^q \times \mathbb{R}^q, \qquad \big|(Pf)(x') - (Pf)(x)\big| \leq \int_{\mathbb{R}^q} \big|\nu(y - A(x'-x)) - \nu(y)\big|\, dy.$$

Since $t \mapsto \nu(\cdot - t)$ is continuous from $\mathbb{R}^q$ to the Lebesgue space $\mathbb{L}^1(\mathbb{R}^q)$, it follows that $\{Pf,\, \|f\|_0 \leq 1\}$ is equicontinuous. Lemma 3 gives the desired property. □

That the model is V_a-geometrically ergodic under Conditions (61) is well-known. However, to the best of our knowledge, the rates of convergence obtained in the next statement are new. Recall that the total variation distance between two probability measures μ_1 and μ_2 on $\mathbb{X}$ is defined by $\|\mu_1 - \mu_2\|_{TV} = \sup_{B\in\mathcal{X}} |\mu_1(B) - \mu_2(B)|$. The gradient is denoted by ∇.

Corollary 7 *Assume that the assumptions of Corollary 6 are fulfilled, that the density* $\nu(\cdot)$ *of* ϑ_1 *is continuously differentiable on* $\mathbb{R}^q$*, and that there exist some positive constants* β *and* b *such that*

$$\forall v \in \mathbb{R}^q, \qquad \|\nabla\nu(v)\| \leq \frac{b}{(1+\|v\|)^\beta}. \tag{62}$$

Then the following assertions hold true:

(i) If $\beta > q+\gamma$ *for some* $\gamma \in [0, a-1]$*, then for each* $f \in \mathcal{B}_\gamma$*, we have* $Pf \in \mathcal{L}_a$ *and*

$$\forall n \geq 2,\ \forall x \in \mathbb{R}^q, \quad \big|\mathbb{E}_x[f(X_n)] - \pi(f)\big| \leq c_1\, d_f\, \|A\|^{n-1}\,(1+\|x\|)^a, \tag{63}$$

where $d_f := m_a(Pf)$ *and* c_1 *is the constant of Proposition 4.*

(ii) If $\beta > q$ *and if the initial distribution* μ *is such that* $I_\mu := \int (1+\|x\|)^a d\mu(x) < \infty$*, then:*

$$\forall n \geq 2, \quad \|\mathbb{P}_\mu(X_n \in \cdot) - \pi(\cdot)\|_{TV} \leq c_1\, d_0\, I_\mu\, \|A\|^{n-1} \tag{64}$$

where c_1 *is the constant of Proposition 4, and the constant* d_0 *can be easily expressed in function of the matrix* A *and the derivative of* ν *(in link with the norm* $\|\cdot\|$*). For instance, if* $\|\cdot\|$ *is the supremum norm on* $\mathbb{R}^q$ *then:* $d_0 := q\big(\max_k \sum_{i=1}^q |a_{ik}|\big) \int_{\mathbb{R}^q} \|\nabla\nu(y)\|\, dy$*.*

(iii) If $\beta > q+\gamma$ *for some* $\gamma \in [0, a-1]$*, then* P *is* V_a*-geometrically ergodic and*

$$\rho_{V_a}(P) \leq \max(\|A\|, \|A\|^\gamma).$$

Proof. Recall that $P(x,dy) = K(x,y)\,dy$ with $K(x,y) := \nu\big(y - Ax\big)$ so that the partial derivative of K in the direction x satisfies: $\partial_x K(x,y) = -A^*\nabla\nu\big(y-Ax\big)$ where A^* is the adjoint matrix of A.

Assertion *(i)* holds from Remark 9, if we prove that $P(\mathcal{B}_\gamma) \subset \mathcal{L}_a$. This is deduced from Proposition E.1 if we check Conditions (75)-(76) for $\tau := \gamma/a$. From (62) it can be easily seen that, for any $r > 0$, there exists a constant a_r such that we have for all $x \in \mathbb{R}^q$ satisfying $\|x\| \leq r$: $\forall y \in \mathbb{R}^q,\ \ \|\partial_x K(x,y)\| \leq a_r\,(1+\|y\|)^{-\beta}$. Since $\beta - \gamma > q$, Condition (75) holds with $\tau := \gamma/a$. Next, set $J(x,y) := (1+\|y\|)^\gamma\, \|\partial_x K(x,y)\|$ for $y \in \mathbb{R}^q$. We have

$$\begin{aligned}
\int_{\mathbb{R}^q} J(x,y)\,dy &= \int_{\mathbb{R}^q} \big(1+\|Ax+v\|\big)^\gamma\, \|A^*\nabla\nu(v)\|\,dv \\
&\leq C\,\Big(\,(1+\|x\|)^\gamma \int_{\mathbb{R}^q} \|\nabla\nu(v)\|\,dv + \int_{\mathbb{R}^q} \|v\|^\gamma\, \|\nabla\nu(v)\|\,dv\Big) \\
&\leq C'\,(1+\|x\|)^\gamma,
\end{aligned} \tag{65}$$

which proves Condition (76) (with $\tau := \gamma/a$). Thus Proposition E.1 gives $P(\mathcal{B}_\gamma) \subset \mathcal{L}_a$.

Under the assumptions of *(ii)*, setting $d_B := d_{1_B}$, we deduce from (i) (case $\gamma = 0$) that:

$$\forall B \in \mathcal{X},\ \forall n \geq 2,\ \forall x \in \mathbb{R}^q, \quad \big|\mathbb{P}_x(X_n \in B) - \pi(B)\big| \leq c_1\, d_B\, \|A\|^{n-1}\,(1+\|x\|)^a. \tag{66}$$

Assuming (for simplicity) that $\|\cdot\|$ is the supremum norm on $\mathbb{R}^q$, it follows from (77) that we have for all $B \in \mathcal{X}$:

$$\forall x \in \mathbb{R}^q, \quad \big\|\nabla(P1_B)(x)\big\| \leq \big(\max_k \sum_{i=1}^q |a_{ik}|\big) \int_{\mathbb{R}^q} \|\nabla\nu(y-Ax)\|\,dy.$$

Thus $d_B := m_a(P1_B) \leq d_0$ with d_0 given in Corollary 7, and (64) easily follows from (66).

Finally, we know from the proof of Corollary 6 that P is compact from $\mathcal{B}_0$ to $\mathcal{B}_a$, and under the condition of *(iii)* we have $P(\mathcal{B}_\gamma) \subset \mathcal{L}_a$ from the proof of Assertion (i). Then the desired statement follows from Corollary 5 (with $\tau := \gamma/a$). $\square$

Remark 12 *Under the conditions of Assertion* (i)*, an upper bound of the constant* $d_f := m_a(Pf)$ *for* $f \in \mathcal{B}_\gamma$ *can be easily derived (see Appendix E for details):*

$$d_f := m_a(Pf) \;\leq\; q|f|_\gamma\, b\, C_{\gamma,\beta}\,\big(\max_k \sum_{i=1}^{q} |a_{ik}|\big)$$

$$\textit{with}\;\; C_{\gamma,\beta} := \sup_{x\in\mathbb{R}^q} (1+\|x\|)^{-\gamma} \int_{\mathbb{R}^q} \frac{(1+\|y+Ax\|)^\gamma}{(1+\|y\|)^\beta}\, dy < \infty.$$

Remark 13 *The constant* c_1 *in (63)-(64) is that of Proposition 4. Let us give an upper bound of* c_1 *under Conditions* (61)*. Set*

$$M := \mathbb{E}[\|\vartheta_1\|^a]^{1/a}, \quad \varepsilon_0 := \frac{1-\|A\|}{2} \quad \textit{and} \quad r := \max\left(0, \frac{1+M-\varepsilon_0}{\varepsilon_0}\right).$$

Recall that $x_0 := 0$ *here. Then we have for any* $x \in \mathbb{R}^q$ *such that* $\|x\| \geq r$

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{1+L(\vartheta_1)\,d(x,x_0)+d(F_{\vartheta_1}x_0,x_0)}{1+d(x,x_0)}\right)^a\right]^{\frac{1}{a}} &= \mathbb{E}\left[\left(\frac{1+\|A\|\,\|x\|+\|\vartheta_1\|}{1+\|x\|}\right)^a\right]^{\frac{1}{a}} \\
&\leq \frac{1+\|A\|\,\|x\|}{1+\|x\|} + \frac{\mathbb{E}\big[\|\vartheta_1\|^a\big]^{1/a}}{1+r} \\
&\leq \|A\| + \frac{1+M}{1+r} \leq \frac{1+\|A\|}{2}.
\end{aligned}$$

Set $\rho := \big((1+\|A\|)/2\big)^a$*,* $\xi_1 := \mathbb{E}\big[(\|A\|+\|\vartheta_1\|)^a\big]$ *and* $\xi := 1+\xi_1(1+r)^a/(1-\rho)$*. Recall that* $\|\pi\|_b := \big(\int_{\mathbb{X}}(1+\|x\|)^b\, d\pi(y)\big)^{1/b}$ *for* $b := 1, a$*. Then we have from Proposition 4 and Remark 8*

$$c_1 \leq \xi^{\frac{a-1}{a}}\, \|\pi\|_1(1+\|\pi\|_a)^{a-1}.$$

Example 8 (Contracting normals) *Assume that* $\mathbb{X} := \mathbb{R}$ *and that* $P(x,dy)$ *is the Gaussian distribution* $\mathcal{N}(\theta x, 1-\theta^2)$ *for* $\theta \in (-1,1)$*. This kernel was studied in [Bax05, Example 8.3], in the specific case* $\theta := 1/2$ *in [Kol00] and the convergence of its ergodic averages in [RT99, Example 4]. Note that* P *is the transition kernel of the IFS defined by* $\forall n \geq 1,\ X_n := \theta X_{n-1} + \vartheta_n$*, where* $(\vartheta_n)_{n\geq 1}$ *is a sequence of* $\mathbb{R}$*-valued i.i.d. random variables, with common distribution* $\mathcal{N}(0, 1-\theta^2)$*. It can be easily checked that* P *has the standard Gaussian distribution* $\mathcal{N}(0,1)$ *as invariant probability measure* π*. Here* $\mathcal{B}_a := \{f : \mathbb{R}\to\mathbb{C},\ \sup_{x\in\mathbb{R}} |f(x)|/(1+|x|)^a < \infty\}$*.*

Let $a \in [1,+\infty)$*. Since Conditions (61) hold, we have* $r_{ess}(P) \leq |\theta|^a$ *with* P *considered as an operator on* $\mathcal{B}_a$*. Concerning the geometric ergodicity, since Condition (62) holds for any*

$\beta > 0$, *Assertion (iii) of Corollary 7 ensures that the convergence rate of* P *on* $\mathcal{B}_a$ *satisfies* $\rho_{V_a}(P) \leq |\theta|^{a-1}$ *when* $a \in [1,2]$, $\rho_{V_a}(P) = |\theta|$ *when* $a \in [2,+\infty)$ *(also use Remark 6 to obtain the last equality). This improves all the earlier bounds obtained for* $\rho_{V_a}(P)$ *in this example (compare with [Bax05] in case* $a := 2$*).*

Furthermore, for this example, Property (64) enables us to improve and simplify the results of [RT99, Section 5] concerning the total variation convergence bounds. In fact, for any initial distribution μ *such that* $I_\mu := \int (1+|x|)d\mu(x) < \infty$, *Inequality (64) with* $a := 1$ *gives (use* $c_1 := 1 + \sqrt{2/\pi}$, $d_0 := 2|\theta|/\sqrt{2\pi(1-\theta^2)}$*):*

$$\forall n \geq 2, \quad \|\mu P^n - \pi\|_{TV} \leq \frac{\sqrt{2\pi}+2}{\pi\sqrt{1-\theta^2}}\, I_\mu\, |\theta|^n.$$

Finally Assertion (i) of Corollary 7 provides an interesting alternative result between the last one and the V_a*-geometrical ergodicity. For instance, in case* $a := 2$, *Property (63) ensures that, for all* $f \in \mathcal{B}_1$, *we have* $Pf \in \mathcal{L}_2$ *and:*

$$\forall n \geq 2,\ \forall x \in \mathbb{R}^q, \quad \big|\mathbb{E}_x[f(X_n)] - \pi(f)\big| \leq c\, d_f\, |\theta|^n\, (1+|x|)^2, \tag{67}$$

with $c := 2\left(1+\sqrt{2}/\sqrt{\pi}\right)\left(1+\sqrt{2+2\sqrt{2}/\sqrt{\pi}}\ \right)/|\theta|$ *(first use that* $P^nW_2(x) = 2(1-\theta^{2n}) + \theta^{2n}W_2(x)$ *for* $n \geq 1$ *with* $W_2(x) = 1+x^2$, *so that* $\sup_{n\geq 1}\sup_{x\in\mathbb{R}} P^nW_2(x)/W_2(x) = 2$*; second deduce from* $V_2/2 \leq W_2 \leq V_2$ *that* $\xi \leq 4$*; third check that* $\|\pi\|_1 = 1 + \sqrt{2/\pi}$, $\|\pi\|_2{}^2 := 2(1+\sqrt{2/\pi})$*). Note that (67) does not involve the* V_a*-geometrical ergodicity (either on* $\mathcal{B}_1$*, or on* $\mathcal{B}_2$*). However the rate of convergence in (67) is optimal and the associated constant* C *is explicit. The weighted-Lipschitz constant* $d_f := m_2(Pf)$ *can be easily computed thanks to Remark 12.*

The two last examples are classical extensions of the affine ARs.

Example 9 (The functional autoregressive process) *Let* $(X_n)_{n\in\mathbb{N}}$ *be the IFS*

$$X_0 \in \mathbb{R}^q, \qquad \forall n \geq 1, \quad X_n := \psi(X_{n-1}) + \vartheta_n, \tag{68}$$

associated with $F(v,x) := \psi(x) + v$ *where* $\psi : \mathbb{R}^q \to \mathbb{R}^q$ *is a fixed differentiable function and* ϑ_1 *has a density* $\nu(\cdot)$*. Assume that*

$$\alpha := \sup_{x\in\mathbb{R}^q} \|\psi'(x)\| < 1 \quad \text{and} \quad \exists a \in (1,+\infty),\ \mathbb{E}\big[\|\vartheta_1\|^a\big] < \infty. \tag{69}$$

Then Corollary 6 extends to the IFS model (68) with α *in place of* $\|A\|$*, and a direct adaptation of the above arguments allows us to prove that, if the density* $\nu(\cdot)$ *of* ϑ_1 *satisfies* (62), *then all the conclusions of Corollary 7 hold true with again* α *in place of* $\|A\|$*.*

Example 10 (Autoregressive process with ARCH(1) error) *Let* $(X_n)_{n\in\mathbb{N}}$ *be the real-valued IFS*

$$X_0 \in \mathbb{R}, \qquad \forall n \geq 1, \quad X_n := aX_{n-1} + \sigma(X_{n-1})\vartheta_n, \tag{70}$$

where $\sigma : \mathbb{R} \to \mathbb{R}$ *is defined by* $\sigma(x) := \sqrt{b+cx^2}$ *with fixed* $a, b > 0$*. It can be easily seen that the associated Markov kernel is of the form* $P(x,dy) = K(x,y)dy$*, provided that the probability distribution of* ϑ_1 *has a density* $\nu(\cdot)$*. The conclusions of Corollaries 6-7 are still valid under suitable assumptions on* $\nu(\cdot)$*.*

A Positive eigenvectors of the adjoint of a nonnegative operator on $\mathcal{B}_V$

Proposition A.1 *If L is a positive bounded linear operator on $\mathcal{B}_V$ such that $r(L) = 1$, then there exists a nontrivial nonnegative continuous linear form η on $\mathcal{B}$ such that $\eta \circ L = \eta$.*

Proof. Since $r(L) = 1$ and the spectrum $\sigma(L)$ of L is closed, there exists $\lambda \in \sigma(L)$ such that $|\lambda| = 1$. Set $\lambda_n := \lambda(1 + 1/n)$. From the Banach-Steinhaus theorem, there exists $f_0 \in \mathcal{B}_V$, $f_0 \geq 0$, such that $\|(\lambda_n I - L)^{-1} f_0\|_V \to +\infty$ when $n \to +\infty$. Using the Neumann series, namely

$$|z| > 1 \ \Rightarrow \ (zI - L)^{-1} = \sum_{k\geq 0} z^{-(k+1)} L^k,$$

the positivity of L (which gives $|L^k f_0| \leq L^k |f_0|$), and finally the fact that $\mathcal{B}_V$ is a Banach lattice ($\forall (f,g) \in \mathcal{B}_V^2 : |f| \leq |g| \Rightarrow \|f\|_V = \| \, |f| \, \|_V \leq \| \, |g| \, \|_V = \|g\|_V$), we obtain:

$$\|(\lambda_n I - L)^{-1} f_0\|_V \leq \|(|\lambda_n| I - L)^{-1} |f_0| \, \|_V \to +\infty \qquad \text{as } n \to +\infty.$$

This implies that $1 \in \sigma(L)$. Now, let $\mathcal{B}'_V$ denote the dual space of $\mathcal{B}_V$ with the associated norm also denoted by $\| \cdot \|_V$, and let L^* be the adjoint of L. Since $\sigma(L^*) = \sigma(L)$, we have $1 \in \sigma(L^*)$.

Let us set $\beta_n := 1 + 1/n$. We deduce from the Banach-Steinhaus theorem that there exists $f'_0 \in \mathcal{B}'_V$, $f'_0 \geq 0$, (i.e. $\forall f \in \mathcal{B}_V : f \geq 0 \ \Rightarrow \ f'_0(f) \geq 0$) such that

$$b_n := \|(\beta_n I - L^*)^{-1} f'_0\|_V \to +\infty \quad \text{as } n \to +\infty,$$

where I denotes here the identity map on $\mathcal{B}'_V$. Let us define the following positive elements in the unit ball of $\mathcal{B}'_V$:

$$f'_n := \frac{1}{b_n} (\beta_n I - L^*)^{-1} f'_0 = \frac{1}{b_n} \sum_{k \geq 0} \beta_n^{-(k+1)} (L^*)^k f'_0 \qquad (n \in \mathbb{N}^*).$$

We have $f'_n \geq 0$, thus $\|f'_n\|_V = f'_n(V) = 1$. Thanks to the Banach-Alaoglu theorem, the sequence $(f'_n)_n$ has a limit point, say η, in the unit ball of $\mathcal{B}'_V$ for the weak topology in $\mathcal{B}'_V$, that is: for all neighborhood W of η and for all $N \geq 1$, there exists $n > N$ such that $f'_n \in W$. Now, given $f \in \mathcal{B}_V$, consider the following special neighborhoods of η with respect to the weak topology of $\mathcal{B}'_V$:

$$W(f,p) = \left\{ f' \in \mathcal{B}'_V : \left| f'(f) - \eta(f) \right| < \frac{1}{p}, \ \left| f'(Lf) - \eta(Lf) \right| < \frac{1}{p} \right\}.$$

Let us denote by $(n_p)_p \equiv (n_p(f))_p$ any increasing sequence of integer numbers such that we have $f'_{n_p} \in W(f,p)$ for all $p \geq 1$. First, pick any $f \in \mathcal{B}_V$, $f \geq 0$: then it follows from $f'_{n_p}(f) \geq 0$ and $|f'_{n_p}(f) - \eta(f)| < 1/p$ that $\eta(f) \geq 0$. So $\eta \geq 0$. Second, consider $f := V$: then we obtain $\eta(V) = 1$ from $f'_{n_p}(V) = 1$ and $|f'_{n_p}(V) - \eta(V)| < 1/p$, so $\eta \neq 0$. Finally, let us fix any $f \in \mathcal{B}_V$. We have

$$\begin{aligned} \beta_n (f'_n - \eta)(f) - (f'_n - \eta)(Lf) &= (\beta_n I - L^*)(f'_n)(f) - (\beta_n I - L^*)(\eta)(f) \\ &= \frac{1}{b_n} f'_0(f) - (\beta_n I - L^*)(\eta)(f). \end{aligned}$$

Replacing n with $n_p \equiv n_p(f)$ gives $(I - L^*)(\eta)(f) = 0$ as $p \to +\infty$. Namely: $\eta \circ L = \eta$. $\square$

B Proof of lemma 4

We know from (**WD**) that P is power-bounded on $\mathcal{B}_V$. Let $x_0 \in \mathbb{X}$. Then we have $K := \sup_n (P^n V)(x_0) < \infty$. Let π_n, $n \geq 1$, be the probability measure on $(\mathbb{X}, \mathcal{X})$ defined by: $\forall B \in \mathcal{X},\ \pi_n(1_B) = \frac{1}{n} \sum_{k=0}^{n-1} (P^k 1_B)(x_0)$. Then Markov's inequality gives

$$\forall n \geq 1,\ \forall \alpha \in (0, +\infty), \quad \pi_n\big(1_{\{V > \alpha\}}\big) \leq \frac{\pi_n(V)}{\alpha} \leq \frac{K}{\alpha}.$$

Thus the sequence $(\pi_n)_n$ is tight, and we can select a subsequence $(\pi_{n_k})_k$ weakly converging to a probability measure π, which is clearly P-invariant. Next, for $p \in \mathbb{N}^*$, define $V_p(\cdot) = \min(V(\cdot), p)$. Then $\forall k \geq 0,\ \forall p \geq 0,\ \pi_{n_k}(V_p) \leq \pi_{n_k}(V) \leq K$. Since V_p is continuous and bounded on $\mathbb{X}$, we obtain: $\forall p \geq 0,\ \lim_k \pi_{n_k}(V_p) = \pi(V_p) \leq K$. The monotone convergence theorem then gives $\pi(V) < \infty$.

C Additional material for discrete Markov chains

C.1 Complements for the proof of Corollary 3

Lemma 8 *Under Conditions $(\mathcal{I})$-$(\mathcal{A})$, 1 is a simple eigenvalue and the unique eigenvalue of modulus one of P on $\mathcal{B}_V$.*

Proof. First prove that the support of π coincides with $\mathbb{N}$. We have

$$\forall j \in \mathbb{N},\ \forall n \geq 1, \quad \pi(j) := \pi(P^n 1_{\{j\}}) = \sum_{i \geq 0} \pi(i)\, P^n(i,j).$$

Hence, if $\pi(j) = 0$ for some $j \in \mathbb{N}$, then we would obtain $\pi(i) = 0$ whenever $P^n(i,j) > 0$, thus $\pi(i) = 0$ for all $i \in \mathbb{N}$ from Condition $(\mathcal{I})$, which is impossible.

Second, we have the following implication: $\forall \lambda \in \mathbb{C},\ |\lambda| = 1,\ \forall f \in \mathcal{B}_V$,

$$Pf = \lambda f \ \Rightarrow\ P|f| = |f|. \tag{71}$$

Indeed, we deduce that $|f| \leq P|f|$ from $Pf = \lambda f$ and the positivity of P. Then it follows from $\pi(P|f| - |f|) = 0$ that $P|f| = |f|$ π-a.s. Thus $P|f| = |f|$ since the support of π is $\mathbb{N}$.

Third, we prove that 1 is a simple eigenvalue of P on $\mathcal{B}_V$. Let $g \in \mathcal{B}_V$ such that $Pg = g$, and set $f := g - g(0)1_{\mathbb{N}}$. Then $Pf = f$ so that $P|f| = |f|$. We have: $\forall n \geq 1,\ 0 = |f(0)| = \sum_{j \geq 0} P^n(0,j)\, |f(j)|$. Condition $(\mathcal{I})$ then yields $f \equiv 0$, namely g is constant.

Finally, let $\lambda \in \mathbb{C}$, $|\lambda| = 1$, and let $f \in \mathcal{B}_V$, $f \neq 0$, be such that $Pf = \lambda f$. It follows from (71) and the last statement that, $\forall n \in \mathbb{N},\ |f(n)| = 1$ (up to a multiplicative constant). From $|\lambda| = 1$, $|f| \equiv 1$, and $\forall n \geq 1,\ \lambda^n f(0) = \sum_{j \geq 0} P^n(0,j)\, f(j)$, we obtain: $P^n(0,j) > 0 \Rightarrow \lambda^n f(0) = f(j)$. In particular: $n \in \mathcal{R}_{0,0} \Rightarrow \lambda^n f(0) = f(0)$. This gives: $\forall (m,n) \in \mathcal{R}_{0,0} \times \mathcal{R}_{0,0},\ \lambda^{n-m} = 1$, hence $\lambda = 1$ by Condition $(\mathcal{A})$. $\square$

C.2 Random walks with bounded increments

C.2.1 Proof that the integer ℓ in (23) is well-defined

Set $A(\gamma) := \phi(\gamma)\gamma^b = \sum_{k=0}^{2b} a_{-b+k}\gamma^k$ where $\phi(\gamma)$ is defined in (22b) and $a_0 \neq 1$. The integer ℓ in (23), if well-defined, can be equivalently characterized from Leibniz's formula by

$$\forall k \in \{1,\ldots,\ell-1\}, \quad \phi^{(k)}(1) = A^{(k)}(1) - \prod_{j=0}^{k}(b-j) = 0 \quad \text{and} \quad \phi^{(\ell)}(1) = A^{(\ell)}(1) - \prod_{j=0}^{\ell-1}(b-j) \neq 0,$$

according that the first condition is removed when $\ell = 1$. To prove the existence of such an integer ℓ, observe that, if $A^{(k)}(1) = \prod_{j=0}^{k-1}(b-j)$ for $k = 1,\ldots,2b$, then Taylor's formula would give

$$A(\gamma) = A(1) + \sum_{k=1}^{2b}\left(\prod_{j=0}^{k-1}(b-j)\right)\frac{(\gamma-1)^k}{k!} = 1 + \sum_{k=1}^{b}\binom{b}{k}(\gamma-1)^k = \gamma^b,$$

which is impossible since $a_0 \neq 1$.

C.2.2 Condition (WD) for random walks with i.d. bounded increments

Let P be defined on $\mathbb{X} = \mathbb{N}$ by

$$\forall i \in \{0,\ldots,b-1\}, \quad \sum_{j\geq 0} P(i,j) = 1; \quad \forall i \geq b, \forall j \in \mathbb{N}, \quad P(i,j) = \begin{cases} a_{j-i} & \text{if } \ |i-j| \leq b \\ 0 & \text{if } \ |i-j| > b \end{cases}$$

where $b \in \mathbb{N}^*$, $(a_{-b},\ldots,a_b) \in [0,1]^{2b+1}$ and $\sum_{k=-b}^{b} a_k = 1$. Assume that $a_0 < 1$ and that there exists $\gamma \in (1,+\infty)$ such that

$$\forall i \in \{0,\ldots,b-1\}, \quad \sum_{j\geq 0} P(i,j)\gamma^j < \infty.$$

The integer ℓ in the next proposition is defined by (23), it is well-defined from Subsection C.2.1.

Proposition 6 *The following conditions are equivalent:*

(a) There exists $\gamma_0 \in (1,\gamma]$ such that P satisfies Condition (**WD**) *with $V_{\gamma_0} := ({\gamma_0}^n)_{n\in\mathbb{N}}$, and we have*

$$\delta_{V_{\gamma_0}}(P) = \frac{A(\gamma_0)}{\gamma_0^g} = \phi(\gamma_0);$$

(b) $A^{(\ell)}(1) < \prod_{j=0}^{\ell-1}(b-j)$, *i.e.* $\phi^{(\ell)}(1) < 0$.

Proof. Let us prove the equivalence $(a) \Leftrightarrow (b)$. Assume that $A^{(\ell)}(1) > \prod_{j=0}^{\ell-1}(b-j)$, i.e. $\phi^{(\ell)}(1) > 0$, and prove that, for all $\gamma_0 \in (1, \gamma]$, P does not satisfy (**WD**) with $V := V_{\gamma_0}$. From the definition of ℓ and from $\phi^{(\ell)}(1) > 0$, there exists $\gamma_2 \in (1, \gamma]$ such that $\phi(\gamma') > \phi(1) = 1$ for all $\gamma' \in (1, \gamma_2)$, so that (26) gives

$$\forall \gamma' \in (1, \gamma_2),\ \forall N \geq 1, \quad \limsup_{n\to+\infty} \frac{(P^N V_{\gamma'})(n)}{V_{\gamma'}(n)} = \phi(\gamma')^N > 1.$$

Hence, from Corollary 2, for all $\gamma' \in (1, \gamma_2)$, P does not satisfy Condition (**WD**) with $V := V_{\gamma'}$. This proves the desired result. Indeed, if (**WD**) holds with $V := V_{\gamma_0}$ for some $\gamma_0 \in (1, \gamma]$, then (**WD**) would be fulfilled for all $\gamma' \in (1, \gamma_0]$ from Jensen's inequality, which contradicts the last conclusion.

Conversely, assume that $\phi^{(\ell)}(1) < 0$. Then there exists $\gamma_1 \in (1, \gamma]$ such that $\phi(\gamma_0) < 1$ for all $\gamma_0 \in (1, \gamma_1)$. Let $\gamma_0 \in (1, \gamma_1)$. From (26) with $k = 1$, we obtain $(PV_{\gamma_0})(n) = \phi(\gamma_0)\, V_{\gamma_0}(n)$ for all $n \geq b$. Since by assumption we have $\sum_{j\geq 0} P(i,j)\gamma_0^j < \infty$ for each $0 \leq i \leq b-1$, it follows from Corollary 2 that P satisfies (**WD**) with $V := V_{\gamma_0}$ and that $\delta_{V_{\gamma_0}}(P) \leq \phi(\gamma_0)$. The converse inequality follows from (26) and Corollary 2. □

D Proof of Formula (50)

First, we prove that the constant ξ in Proposition 4 is well defined. Second, we obtain a basic estimate (74) of the distance between functional of the states occupied at time n of the IFS from two different initial probability distributions. Then, we complete the proof of Formula (50).

First, we have for any $x \in \mathbb{X}$

$$\begin{aligned}\frac{(PV_a)(x)}{V_a(x)} = \mathbb{E}\left[\left(\frac{1 + d(F_{\vartheta_1}x, x_0)}{1 + d(x, x_0)}\right)^a\right] &\leq \mathbb{E}\left[\left(\frac{1 + d(F_{\vartheta_1}x, F_{\vartheta_1}x_0) + d(F_{\vartheta_1}x_0, x_0)}{1 + d(x, x_0)}\right)^a\right] \\ &\leq \mathbb{E}\left[\left(\frac{1 + L(\vartheta_1)\, d(x, x_0)}{1 + d(x, x_0)} + \frac{d(F_{\vartheta_1}x_0, x_0)}{1 + d(x, x_0)}\right)^a\right].\end{aligned}$$

Since $\mathbb{E}[(\max(1, L(\vartheta_1)) + d(F_{\vartheta_1}x_0, x_0))^a] < \infty$, we obtain $\xi_1 := \sup_{x\in\mathbb{X}}(PV_a)(x)/V_a(x) < \infty$. Next Lebesgue's theorem ensures that $\limsup (PV_a)(x)/V_a(x)$ converges to $\mathbb{E}[L(\vartheta_1)^a]$ when $d(x, x_0) \to +\infty$. Now let δ be such that $\mathbb{E}[L(\vartheta_1)^a] < \delta < 1$. Then there exists $r > 0$ such that we have for all $x \in \mathbb{X}$ satisfying $d(x, x_0) > r$: $PV_a(x) \leq \delta V_a(x)$. Besides, if $d(x, x_0) \leq r$, then we obtain $(PV_a)(x) \leq \xi_1 V_a(x) \leq \xi_1(1 + r)^a$. Thus: $PV_a \leq \delta V_a + \xi_1(1 + r)^a 1_{\mathbb{X}}$. Therefore

$$P^n V_a \leq \delta^n\, V_a + \frac{\xi_1(1 + r)^a}{1 - \delta} 1_{\mathbb{X}} \leq \big(1 + \frac{\xi_1(1 + r)^a}{1 - \delta}\big)\, V_a. \tag{72}$$

This prove that the bound ξ given in Proposition 4 is finite.

Second, let us introduce some additional notations. If μ is a probability measure on $\mathbb{X}$ and $X_0 \sim \mu$, we make a slight abuse of notation in writing $(X_n^\mu)_{n\in\mathbb{N}}$ for the associated IFS. We simply write $(X_n^x)_{n\in\mathbb{N}}$ when $\mu := \delta_x$ is the Dirac mass at some $x \in \mathbb{X}$. We denote by $\mathcal{M}_a$

the set of all the probability measures μ on $\mathbb{X}$ such that $\|\mu\|_a := (\int_{\mathbb{X}} V_a(y)\, d\mu(y))^{1/a} < \infty$. Finally, for $n \in \mathbb{N}$ and for any probability measures μ_1 and μ_2 on $\mathbb{X}$, define:

$$\Delta_n(\mu_1, \mu_2) := d\big(X_n^{\mu_1}, X_n^{\mu_2}\big)\, \big(p(X_n^{\mu_1}) + p(X_n^{\mu_2})\big)^{a-1}.$$

Lemma D.1 *We have:* $\forall n \geq 1,\ \forall (\mu_1, \mu_2) \in \mathcal{M}_a \times \mathcal{M}_a$

$$\mathbb{E}\big[\Delta_n(\mu_1, \mu_2)\big] \leq \xi^{\frac{a-1}{a}}\, \kappa_1^n\, \mathbb{E}[d(X_0^{\mu_1}, X_0^{\mu_2})]\, \big(\|\mu_1\|_a + \|\mu_2\|_a\big)^{a-1}. \tag{73}$$

Furthermore we have for all $f \in \mathcal{L}_a$*:*

$$\mathbb{E}\big[|f(X_n^{\mu_1}) - f(X_n^{\mu_2})|\big] \leq \xi^{\frac{a-1}{a}}\, m_a(f)\, \kappa_1^n\, \mathbb{E}[d(X_0^{\mu_1}, X_0^{\mu_2})]\, \big(\|\mu_1\|_a + \|\mu_2\|_a\big)^{a-1}. \tag{74}$$

Proof. If $a = 1$, then (73) follows from the independence of the ϑ_n's and from the definition of $L(v)$ and κ_1. Now assume that $a \in (1, +\infty)$. Without loss of generality, one can suppose that the sequence $(\vartheta_n)_{n\geq 1}$ is independent from $(X_0^{\mu_1}, X_0^{\mu_2})$. Also note that, if $\mu \in \mathcal{M}_a$, then we have

$$\mathbb{E}\big[p(X_n^{\mu})^a\big] = \int_{\mathbb{X}} (P^n V_a)(x) d\mu(x) \leq \xi\, \|\mu\|_a^a.$$

From Holder's inequality (use $1 = 1/a + (a-1)/a$), we obtain

$$\begin{aligned}
\mathbb{E}\big[\Delta_n(\mu_1, \mu_2)\big] &= \mathbb{E}\left[d\big(F_{\vartheta_n:\vartheta_1} X_0^{\mu_1}, F_{\vartheta_n:\vartheta_1} X_0^{\mu_2}\big)\, \big(p(X_n^{\mu_1}) + p(X_n^{\mu_2})\big)^{a-1}\right] \\
&\leq \mathbb{E}[d(X_0^{\mu_1}, X_0^{\mu_2})]\, \mathbb{E}\left[L(\vartheta_n : \vartheta_1)\, \big(p(X_n^{\mu_1}) + p(X_n^{\mu_2})\big)^{a-1}\right] \\
&\leq \mathbb{E}[d(X_0^{\mu_1}, X_0^{\mu_2})]\, \mathbb{E}\big[L(\vartheta_n : \vartheta_1)^a\big]^{\frac{1}{a}}\, \mathbb{E}\big[\big(p(X_n^{\mu_1}) + p(X_n^{\mu_2})\big)^a\big]^{\frac{a-1}{a}} \\
&\leq \mathbb{E}[d(X_0^{\mu_1}, X_0^{\mu_2})]\, \mathbb{E}\big[L(\vartheta_1)^a\big]^{\frac{n}{a}}\, \xi^{\frac{a-1}{a}}\, (\|\mu_1\|_a + \|\mu_2\|_a)^{a-1}.
\end{aligned}$$

This proves (73). Property (74) follows from (73) and the definition of $m_a(f)$. □

Finally, we can prove (50). Property (74), applied to $\mu_1 := \delta_x$ and $\mu_2 := \pi$ gives

$$\begin{aligned}
\big|(P^n f)(x) - \pi(f)\big| &= \big|\mathbb{E}[f(X_n^x)] - \mathbb{E}[f(X_n^{\pi})]\big| \\
&\leq \mathbb{E}\big[\,|f(X_n^x) - f(X_n^{\pi})|\,\big] \\
&\leq \xi^{\frac{a-1}{a}}\, m_a(f)\, \kappa_1^n\, \mathbb{E}[d(x, X_0^{\pi})]\, \big(\|\delta_x\|_a + \|\pi\|_a\big)^{a-1}.
\end{aligned}$$

Finally observe that $\|\delta_x\|_a = p(x)$ and

$$\mathbb{E}[d(x, X_0^{\pi})] \leq \mathbb{E}\big[d(x, x_0) + d(x_0, X_0^{\pi})\big] \leq p(x) + \pi(d(x_0, \cdot)) \leq p(x)\, \|\pi\|_1.$$

Hence $\mathbb{E}[d(x, X_0^{\pi})]\, (\|\delta_x\|_a + \|\pi\|_a)^{a-1} \leq p(x)^a \|\pi\|_1\, (1 + \|\pi\|_a)^{a-1}$. The proof of (50) is complete.

E Additional material for P defined by a kernel K

Here $(X_n)_{n\in\mathbb{N}}$ is a Markov chain with state space $\mathbb{X} = \mathbb{R}^q$ $(q \in \mathbb{N}^*)$ equipped with any norm $\|\cdot\|$, and we assume that there exists $K : \mathbb{R}^q \times \mathbb{R}^q \to [0, +\infty)$ measurable such that, for all

$x \in \mathbb{X}$, $P(x, dy)$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^q$, namely: $P(x, dy) = K(x, y)\, dy$.

Let $a \in (1, +\infty)$ and $\tau \in [0, \frac{a-1}{a}]$. The next result is useful to obtain the set inclusion $P(\mathcal{B}_{\tau a}) \subset \mathcal{L}_a$, where $\mathcal{B}_{\tau a}$, $\mathcal{L}_a$ are defined by (44) and (46) page 26.

Proposition E.1 *Assume that, for all $x_0 \in \mathbb{R}^q$, there exist a Lebesgue-integrable function $g_{x_0} : \mathbb{R}^q \to [0, +\infty)$ and an open neighborhood $\mathcal{U}_{x_0}$ of x_0 in $\mathbb{R}^q$ such that:*

$$\forall x \in \mathcal{U}_{x_0}, \quad k = 1, \ldots, q, \quad \big(1 + \|y\|\big)^{\tau a} \big|\frac{\partial K}{\partial x_k}(x, y)\big| \leq g_{x_0}(y) \ \text{ for a.e. } y \in \mathbb{R}^q \tag{75}$$

and assume in addition that there exists a constant d such that

$$k = 1, \ldots, q, \quad \forall x \in \mathbb{R}^q, \quad \int_{\mathbb{R}^q} (1 + \|y\|)^{\tau a} \big|\frac{\partial K}{\partial x_k}(x, y)\big|\, dy \leq d\, (1 + \|x\|)^{a-1}. \tag{76}$$

Then we have $P(\mathcal{B}_{\tau a}) \subset \mathcal{L}_a$.

Proof. Let $f \in \mathcal{B}_{\tau a}$. From Lebesgue's theorem, one can easily deduce that the function $Pf : \mathbb{R}^q \to \mathbb{C}$ is differentiable on $\mathbb{R}^q$, and that its derivative is given by:

$$\forall k = 0, \ldots, q, \ \forall x \in \mathbb{R}^q, \quad \frac{\partial (Pf)}{\partial x_k}(x) = \int_{\mathbb{R}^q} f(y)\, \frac{\partial K}{\partial x_k}(x, y)\, dy. \tag{77}$$

For the sake of simplicity assume that $\|\cdot\|$ is the euclidean norm on $\mathbb{R}^q$. By (77) and (76) we obtain: $\forall x \in \mathbb{R}^q,\ \|\nabla(Pf)(x)\| \leq d\sqrt{q}\, |f|_{\tau a} \big(1 + \|x\|\big)^{a-1}$, where ∇ stands for the gradient operator. Then Taylor's inequality gives for any $(x_1, x_2) \in \mathbb{R}^q \times \mathbb{R}^q$

$$\begin{aligned}
\big|Pf(x_1) - Pf(x_2)\big| &\leq \|x_1 - x_2\| \sup_{t \in [0,1]} \|\nabla(Pf)(tx_1 + (1-t)x_2)\| \\
&\leq d\sqrt{q}\, |f|_{\tau a}\, \|x_1 - x_2\| \sup_{t \in [0,1]} \big(1 + \|tx_1 + (1-t)x_2\|\big)^{a-1} \\
&\leq d\sqrt{q}\, |f|_{\tau a}\, \|x_1 - x_2\| \big(1 + \|x_1\| + \|x_2\|\big)^{a-1}.
\end{aligned}$$

It follows that $Pf \in \mathcal{L}_a$. □

The following statement gives a simple sufficient condition for (76) to hold true.

Proposition E.2 *Let $a \in [1, +\infty)$ and $V_{a-1}(\cdot) := (1 + \|\cdot\|)^{a-1}$. Assume that PV_{a-1}/V_{a-1} is bounded, and*

$$M := \sup_{(x,y) \in \mathbb{R}^q \times \mathbb{R}^q} \frac{\big|(\partial_x K)(x, y)\big|}{|K(x, y)|} < \infty. \tag{78}$$

Then, for each $\tau \in [0, (a-1)/a]$, Condition (76) is fulfilled.

Proof. Let $x \in \mathbb{R}^q$. We have

$$\int_{\mathbb{R}^q} (1 + \|y\|)^{\tau a} \big|\partial_x K(x, y)\big|\, dy \leq M \int_{\mathbb{R}^q} (1 + \|y\|)^{a-1} K(x, y)\, dy = (PV_{a-1})(x) \leq C\, V_{a-1}(x)$$

for some constant C. □

Proof of Remark 12. Assume that $\|\cdot\|$ is the supremum norm on $\mathbb{R}^q$. We obtain from (77) with $K(x,y) := \nu\big(y - Ax\big)$ and from (62): $\forall f \in \mathcal{B}_\gamma$, $\forall x \in \mathbb{R}^q$,

$$\begin{aligned}
\big\|\nabla(Pf)(x)\big\| &\leq |f|_\gamma \,\big(\max_k \sum_{i=1}^q |a_{ik}|\big) \int_{\mathbb{R}^q} (1+\|y\|)^\gamma \,\|\nabla\nu(y-Ax)\|\, dy \\
&\leq |f|_\gamma \,\big(\max_k \sum_{i=1}^q |a_{ik}|\big) \int_{\mathbb{R}^q} (1+\|y+Ax\|)^\gamma \,\frac{b}{(1+\|y\|)^\beta}\, dy \\
&\leq b\, C_{\gamma,\beta}\, |f|_\gamma \,\big(\max_k \sum_{i=1}^q |a_{ik}|\big)\, (1+\|x\|)^\gamma.
\end{aligned}$$

We easily deduce that $m_a(Pf) \leq q\, b\, C_{\gamma,\beta}\, |f|_\gamma\, (\max_k \sum_{i=1}^q |a_{ik}|)$. □

References

[Als03] G. Alsmeyer. On the Harris recurrence of iterated random Lipschitz functions and related convergence rate results. *J. Theoret. Probab.*, 16(1):217–247, 2003.

[AP07] Y. F. Atchadé and F. Perron. On the geometric ergodicity of Metropolis-Hastings algorithms. *Statistics*, 41(1):77–84, 2007.

[Bax05] P. H. Baxendale. Renewal theory and computable convergence rates for geometrically ergodic Markov chains. *Ann. Appl. Probab.*, 15(1B):700–738, 2005.

[Ben98] M. Benda. A central limit theorem for contractive stochastic dynamical systems. *J. Appl. Probab.*, 35:200–205, 1998.

[DF99] P. Diaconis and D. Freedman. Iterated random functions. *SIAM Rev.*, 41:45–76, 1999.

[Duf97] M. Duflo. *Random Iterative Models*. Applications of Mathematics. Springer-Verlag, Berlin Heidelberg, 1997.

[FHL] D. Ferré, L. Hervé, and J. Ledoux. Regular perturbation of V-geometrically ergodic Markov chains. Submitted.

[FHL10] D. Ferré, L. Hervé, and J. Ledoux. Limit theorems for stationary Markov processes with L^2-spectral gap. *Ann. Inst. Henri Poincaré Probab. Stat.*, 2010. To appear.

[Hen93] H. Hennion. Sur un théorème spectral et son application aux noyaux lipchitziens. *Proceeding of the A.M.S*, 118:627–634, 1993.

[Hen06] H. Hennion. Quasi-compactness and absolutely continuous kernels, applications to Markov chains. *ArXiv Mathematics e-prints*, (arXiv:math/0606680), June 2006.

[Hen07] H. Hennion. Quasi-compactness and absolutely continuous kernels. *Probab. Theory Relat. Fields.*, 139:451–471, 2007.

[Her08] L. Hervé. Quasi-compactness and mean ergodicity for Markov kernels acting on weighted supremum normed spaces. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(6):1090–1095, 2008.

[HH01] H. Hennion and L. Hervé. *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, volume 1766 of *Lecture Notes in Mathematics*. Springer, 2001.

[HS92] A. Hordijk and F. Spieksma. On ergodicity and recurrence properties of a Markov chain with an application to an open Jackson network. *Adv. in Appl. Probab.*, 24(2):343–376, 1992.

[Kar00] N. V. Kartashov. Determination of the spectral ergodicity exponent for the birth and death process. *Ukrain. Math. J.*, 52(7):1018–1028, 2000.

[KM11] I. Kontoyiannis and S. Meyn. Geometric ergodicity and the spectral gap of non-reversible Markov chains. *Probability Theory and Related Fields*, pages 1–13, 2011. To appear.

[Kol00] J. E. Kolassa. Explicit bounds for geometric convergence of Markov chains. *J. Appl. Probab.*, 37(3):642–651, 2000.

[Kre85] U. Krengel. *Ergodic Theorems.* de Gruyter Studies in Mathematics. de Gruyter, Berlin-New York, 1985.

[LP83] É. Le Page. Théorèmes de renouvellement pour les produits de matrices aléatoires. Équations aux différences aléatoires. In *Séminaires de probabilités Rennes 1983*, Publ. Sém. Math., page 116. Univ. Rennes I, Rennes, 1983.

[LT96] R. B. Lund and R. L. Tweedie. Geometric convergence rates for stochastically ordered Markov chains. *Math. Oper. Res.*, 21(1):182–194, 1996.

[Lun97] R. B. Lund. The geometric convergence rate of a Lindley random walk. *J. Appl. Probab.*, 34(3):806–811, 1997.

[MR89] X. Milhaud and A. Raugi. Étude de l'estimateur du maximum de vraisemblance dans le cas d'un processus autorégressif : convergence, normalité asymptotique, vitesse de convergence. *Ann. Inst. H. Poincaré Probab. Statist.*, 25:383–428, 1989.

[MS95] V. A. Malyshev and F. M. Spieksma. Intrinsic convergence rate of countable Markov chains. *Markov Process. Related Fields*, 1(2):203–266, 1995.

[MT93] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability.* Springer Verlag, 1993.

[MT94] S. P. Meyn and R. L. Tweedie. Computable bounds for geometric convergence rates of Markov chains. *Ann. Probab.*, 4:981–1011, 1994.

[MT96] K. L. Mengersen and R. L. Tweedie. Rates of convergence of the Hastings and Metropolis algorithms. *Ann. Statist.*, 24(1):101–121, 1996.

[Nev64] J. Neveu. *Bases mathématiques du calcul des probabilités.* Masson et Cie, Éditeurs, Paris, 1964.

[RT99] G. O. Roberts and R. L. Tweedie. Bounds on regeneration times and convergence rates for Markov chains. *Stochastic Process. Appl.*, 80(2):211–229, 1999.

[RT01] 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.

[Sch71] H. H. Schaefer. *Topological vector spaces.* Springer-Verlag, New York, 1971. Third printing corrected, Graduate Texts in Mathematics, Vol. 3.