Quantitative approximation of the invariant distribution of a Markov chain. A new approach

Loïc Hervé, James Ledoux

Quantitative approximation of the invariant distribution of a Markov chain. A new approach

Loïc HERVÉ, and James LEDOUX *

version: Friday 11th March, 2022 – 09:53

Abstract

In this paper, we deal with a Markov chain on a measurable state space $(\mathbb{X}, \mathcal{X})$ with transition kernel P admitting some small-set $S \in \mathcal{X}$, that is such that $P(x, A) \geq \nu(1_A)1_S(x)$ for any $x \in \mathbb{X}$, $A \in \mathcal{X}$ and for some positive measure ν. Under this condition, we propose a constructive characterisation of the existence of an P-invariant probability measure π on $(\mathbb{X}, \mathcal{X})$ such that $\pi(1_S) > 0$. When such an π exists, it is approximated in total variation norm by a finite linear combination of non-negative measures only depending on ν, P and S. Next, using standard drift-type conditions, we provide geometric/subgeometric convergence bounds of the approximation. Theses bounds are fully explicit and as simple as possible. Anyway the rates of convergence are accurate, and they are optimal in the atomic case. Note that the rate of convergence for approximating the iterates of P by the finite-rank submarkovian kernels introduced in [HL20b] is also discussed. This is a new approach for approximating π in the sense that it is not based on the convergence of the iterates of P to π. Thus we need no aperiodicity condition. Moreover, the proofs are direct and simple. They use neither the split chain in the non-atomic case, nor the renewal theory, nor the coupling method, nor the spectral theory. In some sense, this approach for Markov chains with a small-set is self-contained.

AMS subject classification : 60J05

Keywords : Small set, Invariant probability, Finite-rank approximating, Rate of convergence, Drift conditions

1 Introduction

Throughout the paper P is a Markov kernel on a measurable space $(\mathbb{X}, \mathcal{X})$. Let $(X_n)_{n\geq 0}$ be a Markov chain with state space $\mathbb{X}$ and transition kernel P. If $(X_n)_{n\geq 0}$ admits an invariant distribution denoted by π, the two following issues are of interest for any $A \in \mathcal{X}$.

(Q1) How to approximate the value of $\pi(1_A)$ and to control the error?

(Q2) How to approximate the value of $\mathbb{P}(X_n \in A)$ and to control the error?

*Univ Rennes, INSA Rennes, CNRS, IRMAR-UMR 6625, F-35000, France. Loic.Herve@insa-rennes.fr, James.Ledoux@insa-rennes.fr

The standard way is to use, when n is large enough, $\mathbb{P}(X_n \in A)$ to approximate $\pi(1_A)$, and $\pi(1_A)$ to approximate $\mathbb{P}(X_n \in A)$. Of course this approach is supported by all the classical results related to the convergence in distribution of $(X_n)_{n\geq 0}$ to π, or in other words by all the results of convergence of the iterates P^n to the rank-one operator $\pi(\cdot)1_{\mathbb{X}}$. It is worth noticing that the approximation of $\mathbb{P}(X_n \in A)$ by $\pi(1_A)$ is effective only when π is known. Unfortunately, in practice π is often unknown, in which case (Q1) becomes a central issue. Concerning (Q1), observe that π may be approximated by something other than the iterates of P, provided that the approximation procedure is effective and that the error is well controlled. Then, once the stationary distribution π is well estimated, then the classical bounds on $|P^n(x, A) - \pi(1_A)|$ can be used to solve (Q2) thanks to the triangular inequality.

The main objective of this work is to propose a new approach to address (Q1), which is not directly based on the convergence of P^n to π. Specifically, when P has a small-set S and has an invariant probability distribution π such that $\pi(1_S) > 0$, we present a general and effective procedure for approximating π. The central point here is that all the convergence bounds are fully explicit and as simple as possible. Anyway the rates of convergence are accurate, and they are optimal in the atomic case.

Let $\mathcal{M}^+$ (resp. $\mathcal{M}_*^+$) denote the set of finite non-negative (resp. positive) measures on $(\mathbb{X}, \mathcal{X})$. For any $\mu \in \mathcal{M}^+$ and any μ-integrable function $f : \mathbb{X} \to \mathbb{C}$, $\mu(f)$ denotes the integral $\int f d\mu$. Throughout the paper, the existence of a small-set S for P is assumed, that is

$$\exists S \in \mathcal{X},\ \exists \nu \in \mathcal{M}_*^+, \quad \forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad P(x, A) \geq \nu(1_A)\, 1_S(x). \tag{\textbf{S}}$$

Under Assumption (**S**), we can use the following sequence $(\beta_k)_{k\geq 1} \in (\mathcal{M}^+)^{\mathbb{N}}$ introduced in [HL20b], which is recursively defined by

$$\beta_1(\cdot) := \nu(\cdot) \quad \text{and} \quad \forall n \geq 2, \quad \beta_n(\cdot) := \nu\big(P^{n-1}\,\cdot\,\big) - \sum_{k=1}^{n-1} \nu\big(P^{n-k-1} 1_S\big)\, \beta_k(\cdot). \tag{1}$$

Note that no spectral theory is used here in contrast to [HL20b]. Under Assumption (**S**), the following results are obtained.

- In Section 2 (Theorem 2.1), we prove that there exists a $P-$invariant probability measure π on $(\mathbb{X}, \mathcal{X})$ such that $\pi(1_S) > 0$ if, and only if,

$$\sum_{k=1}^{+\infty} \beta_k(1_{\mathbb{X}}) < \infty. \tag{2}$$

  Actually, under this condition, set $\mu := \sum_{k=1}^{+\infty} \beta_k \in \mathcal{M}_*^+$. Then $\mu(1_S) = 1$ and

$$\pi := \frac{\mu}{\mu(1_{\mathbb{X}})} \tag{3}$$

  is an $P-$invariant probability measure on $(\mathbb{X}, \mathcal{X})$ such that $\pi(1_S) = 1/\mu(1_{\mathbb{X}}) > 0$.

In the next items Condition (2) is assumed, and for every $n \geq 1$ we consider $\mu_n \in \mathcal{M}_*^+$ and the probability measure $\widetilde{\mu}_n$ on $(\mathbb{X}, \mathcal{X})$ defined by:

$$\mu_n := \sum_{k=1}^{n} \beta_k \quad \text{and} \quad \widetilde{\mu}_n := \frac{1}{\mu_n(1_{\mathbb{X}})}\, \mu_n.$$

- In Section 3 (Theorem 3.1), we prove that the invariant distribution π given by Formula (3) can be approximated in total variation norm by either $(\mu_n/\mu(1_{\mathbb{X}}))_n$ or $(\widetilde{\mu}_n)_n$ with the following error estimates

$$\|\pi - \mu(1_{\mathbb{X}})^{-1}\mu_n\|_{TV} = \mu(1_{\mathbb{X}})^{-1}\,\varepsilon_n \quad \text{and} \quad \|\pi - \widetilde{\mu}_n\|_{TV} \leq \left(\mu(1_{\mathbb{X}})^{-1} + \nu(1_{\mathbb{X}})^{-1}\right)\varepsilon_n \quad (4)$$
$$\text{with} \quad \varepsilon_n := \sum_{k=n+1}^{+\infty} \beta_k(1_{\mathbb{X}}).$$

  Note that $\lim_n \varepsilon_n = 0$ from (2).

- In Section 4, geometric drift conditions are used to obtain geometric rates of convergence for the above sequence $(\varepsilon_n)_{n\geq 0}$. First, under the usual drift condition $PV \leq \delta\, V + L\, 1_S$ for some constants $\delta \in (0,1)$, $L > 0$ and some measurable function $V : \mathbb{X} \to [1, +\infty)$, we establish that $\theta_0 := \limsup_n [\beta_n(1_{\mathbb{X}})]^{1/n} < 1$ and that: $\forall \theta \in (\theta_0, 1)$, $\varepsilon_n = \mathrm{O}(\theta^n)$ (see Theorem 4.1). Second, in order to obtain computable rates of convergence for ε_n, the following condition is introduced

$$\exists \delta \in (0,1), \quad PV \leq \delta\, V + \nu(V)\, 1_S. \quad (5)$$

  Under Condition (5), we prove that (Theorem 4.2)

$$\varepsilon_n \leq \frac{\nu(V)}{1-\delta}\,\delta^n. \quad (6)$$

  Condition (5), which always holds when S is an atom (see Corollary 4.1), may fail in the non atomic case. It is shown in Corollary 4.2 that, if P satisfies $PV \leq \delta\, V$ on $S^c := \mathbb{X}\backslash S$ and if PV is bounded on S, then the bound (6) still holds, but with δ^{α_0} and V^{α_0} in place of δ and V for some easily computable $\alpha_0 \in (0,1)$. Finally some properties involved in the proof of Theorem 4.2 (resp. of Corollary 4.2) are used in Theorem 4.3 to obtain a rate of convergence for $P^n - T_n$, where T_n is the submarkovian finite-rank kernel defined in (12). This rate of convergence, which is expressed in $V-$weighted (resp. $V^{\alpha_0}-$weighted) total variation norm, enable us to specify the error bound obtained in [HL20b, HL20a] for the $V-$geometrical ergodicity of P. Using the triangular inequality, any such error bounds can be combined with (4) to solve (Q2), see Theorem 4.4.

- In Section 5 the following subgeometric drift-type conditions are introduced to study the rate of convergence of $(\varepsilon)_{n\geq 0}$: for $m \geq 1$ there exists $m+1$ measurable functions $V_i : \mathbb{X} \to [1, +\infty)$, $i = 0, \ldots, m$, such that

$$\forall i \in \{0, \ldots, m-1\}, \quad (P - T)V_i \leq V_i - V_{i+1} \quad \text{with} \quad T\cdot := \nu(\cdot)1_S. \quad (7)$$

  Under this assumption, $(\varepsilon)_{n\geq 0}$ is proved to satisfy $\lim_n n^{m-1}\varepsilon_n = 0$ (Theorem 5.1). Moreover, if $m \geq 2$ and if the sequence $(\beta_k(V_m))_{k\geq 1}$ with $\beta_k(\cdot)$ defined in (1) is decreasing, then (Theorem 5.2)

$$\forall n \geq 1, \quad \varepsilon_n \leq \frac{C_m\, \nu(V_0)}{(m-1)\, n^{m-1}} \quad \text{with} \quad C_m := 2^{\frac{m(m+1)}{2}-1}. \quad (8)$$

  If the decrease of $(\beta_k(V_m))_{k\geq 1}$ is not (or cannot be) checked and if $m \geq 3$, then (8) holds with $m-1$ in place of m. Next it is shown in Corollary 5.1 that the subgeometric

drift conditions (7) are fulfilled under the more explicit following ones:

$$\forall i \in \{0, \dots, m-1\}, \quad \begin{cases} V_{i+1} \leq V_i \\ PV_i \leq \big(V_i - V_{i+1}\big) + \nu(V_i)\, 1_S. \end{cases} \tag{9}$$

When S is an atom, the first condition in (9) implies the second one. Then, using an iterative procedure based on [JR02, Lem. 3.5], we prove that, if PV is bounded on S and if P satisfies the condition $PV \leq V - c_1\, V^\alpha$ on S^c for some constants $\alpha \in [0,1)$, $c_1 > 0$, and some measurable function $V : \mathbb{X} \to [1, +\infty)$, then the bound (8) holds with $m := \lfloor (1-\alpha)^{-1} \rfloor$, where $\lfloor \cdot \rfloor$ denotes the integer part function on $\mathbb{R}$ (see Corollary 5.3). Of course, in the non atomic case the second condition in (9) does not hold automatically with $V_0 = V$ and $V_1 = c_1\, V^\alpha$, but the iterative procedure of the atomic case still applies, provided that $PV \leq V - c_1\, V^\alpha$ is replaced by $P\widehat{V} \leq \widehat{V} - \widehat{c}_1\, \widehat{V}^{\widehat{\alpha}}$ with $\widehat{V} = V^{\eta_0}$ for some explicit $\eta_0 \in (0,1]$, to initialize the procedure. Then, if $\eta_0 \geq 1 - \alpha$ and V, PV are bounded on S, the bound (8) holds with $m := \lfloor \eta_0 (1-\alpha)^{-1} \rfloor$ (see Corollary 5.4). Finally in Theorem 5.3 the rate of convergence for $P^n - T_n$ with T_n given in (12) is specified under the subgeometric drift conditions (7).

We recall that this work is not directly based on the convergence in distribution of the Markov chain $(X_n)_n$ of transition kernel P. In particular no aperiodicity condition is introduced. We use neither renewal theory, nor coupling method, nor even spectral theory. Actually our main statements are concerned with the rate of convergence in (4), in which the positive measure μ_n and the probability measure $\widetilde{\mu}_n$ write as a linear combination of the non-negative measures $\nu, \nu \circ P, \dots, \nu \circ P^{n-1}$ with explicit coefficients only depending on ν, P and S. Therefore, precise qualitative or quantitative comparisons with the classical works recalled below are difficult to address.

The basic fact is that our assumptions are quite close to usual ones. Indeed, the central assumption (**S**) is the existence of a small-set S. But we do not introduce the strong aperiodicity condition $\nu(1_S) > 0$ in order to get a minorization condition as in [MT09, p 98] or in [DMPS18, Chap. 11]. Thus no use of the split chain is needed for proving our results in the non-atomic case. Next, Condition (2) is proved to be equivalent to $\sum_{k=1}^{+\infty} \beta_k(1_S) = 1$ and $\sum_{k=1}^{+\infty} k\, \beta_k(1_S) < \infty$ in Theorem 2.1. When S is an atom, this last condition is nothing else but the usual condition of finite expectation of the first return time in S, see (22). Formula (3), which has been obtained in the V-geometric ergodicity context [HL20b], extends a well-known formula when P satisfies the Doeblin condition ($\mathbb{X}$ is a small-set), see [LC14], or when P is irreducible and recurrent positive according to [Num84, p 74]. Next, the use of geometric or subgeometric drift conditions is standard for investigating the rate of convergence of the iterates P^n of the transition kernel P to π. Recall that in these works, the error term is usually computed in some weighted-type total variation. In property (4), π is approximated by μ_n or $\widetilde{\mu}_n$ in total variation. Under irreducibility and aperiodicity conditions, if P satisfies Assumption (**S**) and the geometric drift condition $PV \leq \delta\, V + L\, 1_S$ for some constants $\delta \in (0,1)$, $L > 0$, and some measurable function $V : \mathbb{X} \to [1, +\infty)$, then P is V-geometrically ergodic, e.g. see [RR04, Bax05, MT09, DMPS18] (see also [Hen06, Hen07, HM11, Del17, HL20b] for alternative approaches). Moreover the previous drift condition bas been proved to be useful to derive computable rates of convergence in the V-geometric ergodicity property, e.g. see [MT94, LT96, RT99, RT00, Ros02, Bax05]. However recall that deriving effective and accurate bounds in the V-geometric ergodicity property is a difficult issue. Similarly non-geometric

(for instance polynomial) rates of convergence can be derived under subgeometric drift conditions, see [DMPS18, and the references therein] and [Del17] for an operator-type approach. The second condition in (9) is an instance of such drift conditions. More general subgeometric drift conditions can be found in [DFMS04, DMPS18] and quantitative bounds of polynomial rates for the convergence of P^n to π are obtained in [AF05, AFV15].

The estimates in (4) do not give directly any information on the convergence of the iterates of P to π, but they do provide an approximation of $\pi(1_A)$ for all $A \in \mathcal{X}$. Anyway the error bounds obtained in both geometrical case (Section 4) and subgeometrical case (Section 5) are simple and explicit. The proofs in this paper are quite simple. It appears that the initial idea of approximating π by μ_n or $\widetilde{\mu}_n$ rather than with the iterates of P simplifies the error computations.

2 Existence of π under Assumption (S)

We denote by $\mathcal{B}$ the space of real-valued bounded measurable functions on $(\mathbb{X}, \mathcal{X})$, equipped with its usual supremum norm: $\forall f \in \mathcal{B},\ \|f\| := \sup_{x\in\mathbb{X}} |f(x)|$. If Q_1 and Q_2 are bounded linear operators on $\mathcal{B}$, we write $Q_1 \leq Q_2$ when the following property holds: $\forall f \in \mathcal{B},\ f \geq 0,\ Q_1 f \leq Q_2 f$.

Let P be a Markov kernel satisfying Condition (**S**). Note that P is a bounded linear operator on $\mathcal{B}$ since P is a Markov kernel, and that $f \mapsto \nu(f)$ is a continuous linear form on $\mathcal{B}$, with $\nu \in \mathcal{M}_*^+$ given in (**S**). We set $\beta_1(\cdot) := \nu(\cdot)$, and for every $n \geq 2$

$$\forall f \in \mathcal{B}, \quad \beta_n(f) := \nu\big(P^{n-1} f\big) - \sum_{k=1}^{n-1} \nu\big(P^{n-k-1} 1_S\big)\, \beta_k(f). \tag{10}$$

Moreover let T be the rank-one operator on $\mathcal{B}$ defined by :

$$\forall f \in \mathcal{B}, \quad Tf := \nu(f)\, 1_S = \beta_1(f)\, 1_S. \tag{11}$$

It follows from the positivity of ν and from (**S**) that $0 \leq T \leq P$.

Proposition 2.1 *Assume that P satisfies Assumption* (**S**). *Then we have for every $n \geq 1$*

$$T_n := P^n - (P-T)^n = \sum_{k=1}^{n} \beta_k(\cdot) P^{n-k} 1_S \qquad \text{and} \qquad 0 \leq T_n \leq P^n \tag{12}$$

$$T_n - T_{n-1} P = (P^{n-1} - T_{n-1}) T \tag{13}$$

with the convention $T_0 = 0$. Moreover, for every $n \geq 1$, $\beta_n \in \mathcal{M}^+$, that is: there exists a non-negative measure on $(\mathbb{X}, \mathcal{X})$ (still denoted by β_n) such that $\int_{\mathbb{X}} d\beta_n < \infty$ and, such that, for every $f \in \mathcal{B}$, we have $\beta_n(f) = \int_{\mathbb{X}} f\, d\beta_n$. Finally we have

$$\forall n \geq 1, \quad \beta_n = \nu \circ \big(P^{n-1} - T_{n-1}\big) = \nu \circ (P-T)^{n-1} \quad \text{and} \quad \beta_{n+1} = \beta_n \circ (P-T) \tag{14}$$

with the convention that P^0 and $(P-T)^0$ stand for the identity map on $\mathcal{B}$.

Proof. The first equality in (12) is just the definition of T_n. That $0 \leq T_n \leq P^n$ follows from $0 \leq T \leq P$. The second equality in (12) for $n = 1$ is obvious from the definition of T. Now assume that this second equality holds true for some $n \geq 1$. Then

$$P^{n+1} - T_{n+1} := (P-T)^{n+1} \quad = \quad (P-T)(P^n - T_n) = P^{n+1} - PT_n - TP^n + TT_n$$

from which we deduce that, for every $f \in \mathcal{B}$

$$\begin{aligned} T_{n+1}f &= PT_nf + TP^nf - TT_nf \qquad (15) \\ &= \sum_{k=1}^{n} \beta_k(f) P^{n-k+1}1_S + \left(\beta_1(P^nf) - \sum_{k=1}^{n} \beta_k(f)\nu(P^{n-k}1_S)\right)1_S \\ &= \sum_{k=1}^{n} \beta_k(f) P^{n+1-k}1_S + \beta_{n+1}(f)1_S \end{aligned}$$

with $\beta_{n+1}(\cdot)$ defined in (10). This provides the second equality in (12) by induction. Next we obtain that for every $n \geq 1$

$$P^n - T_n := (P-T)^n = (P^{n-1} - T_{n-1})(P-T) = P^n - P^{n-1}T - T_{n-1}P + T_{n-1}T$$

so that

$$T_n - T_{n-1}P = P^{n-1}T - T_{n-1}T = (P^{n-1} - T_{n-1})T.$$

Formula (13) is proved. Now we deduce from (10) and then from (12) that

$$\beta_n(f) = \nu\big(P^{n-1}f\big) - \sum_{k=1}^{n-1} \beta_k(f)\, \nu\big(P^{n-k-1}1_S\big) = \nu\big(P^{n-1}f - T_{n-1}f\big).$$

This gives the first equality in (14), from which the second one is easily deduced. Finally note that $\beta_1(\cdot) = \nu(\cdot)$ is a positive measure on $(\mathbb{X}, \mathcal{X})$, so that for every $n \geq 1$ $\beta_n(\cdot)$ is a linear combination of non-negative measures on $(\mathbb{X}, \mathcal{X})$ (by induction). That β_n is a finite non-negative measure follows from (14) since $0 \leq P^{n-1} - T_{n-1} \leq P^{n-1}$ □

Now we can prove the main theorem of this section.

Theorem 2.1 *Assume that P satisfies Assumption* (**S**). *Then the four following assertions are equivalent.*

(i) There exists a $P-$invariant probability measure π on $(\mathbb{X}, \mathcal{X})$ such that $\pi(1_S) > 0$.

(ii) $\displaystyle\sum_{k=1}^{+\infty} \beta_k(1_{\mathbb{X}}) < \infty.$

(iii) $\displaystyle\sum_{k=1}^{+\infty} \beta_k(1_S) = 1$ *and* $\displaystyle\sum_{k=1}^{+\infty} k\, \beta_k(1_S) < \infty.$

(iv) $\displaystyle\lim_{n \to +\infty} \beta_k(1_{\mathbb{X}}) = 0$ *and* $\displaystyle\sum_{k=1}^{+\infty} k\, \beta_k(1_S) < \infty.$

Moreover, under any of these four conditions, we have

$$\sum_{k=1}^{+\infty} k\,\beta_k(1_S) = \frac{1}{\nu(1_{\mathbb{X}})} \sum_{k=1}^{+\infty} \beta_k(1_{\mathbb{X}}) \tag{16}$$

and

$$\pi := \frac{1}{\sum_{k=1}^{+\infty} \beta_k(1_{\mathbb{X}})} \sum_{k=1}^{+\infty} \beta_k$$

is an $P-$invariant probability measure on $(\mathbb{X}, \mathcal{X})$ with

$$\pi(1_S) = \frac{1}{\sum_{k=1}^{+\infty} \beta_k(1_{\mathbb{X}})} > 0. \tag{17}$$

Proof. Assume that Property (i) holds. We deduce from (12) that

$$0 \leq \pi\big((P^n - T_n)1_{\mathbb{X}}\big) = 1 - \pi(T_n 1_{\mathbb{X}}) = 1 - \pi(1_S) \sum_{k=1}^{n} \beta_k(1_{\mathbb{X}}),$$

from which it follows that $\sum_{k=1}^{+\infty} \beta_k(1_{\mathbb{X}}) \leq 1/\pi(1_S) < \infty$ since $\pi(1_S) > 0$ by hypothesis. This gives (ii). Conversely assume that Property (ii) holds. Then

$$\mu := \sum_{k=1}^{+\infty} \beta_k \ \in \mathcal{M}_*^+$$

since $\mu(1_{\mathbb{X}}) \geq \beta_1(1_{\mathbb{X}}) = \nu(1_{\mathbb{X}}) > 0$. Also note that, for every $f \in \mathcal{B}$, the series $\sum_{k=1}^{+\infty} \beta_k(f)$ absolutely converges in $\mathbb{C}$ since $|\beta_k(f)| \leq \|f\|\,\beta_k(1_{\mathbb{X}})$ (the series $\sum_{k=1}^{+\infty} \beta_k$ is absolutely convergent in the topological dual space of $\mathcal{B}$). We obtain that for every $f \in \mathcal{B}$ (by convention we set $T_0 := 0$)

$$\begin{aligned}
\mu(Pf) &= \sum_{k=1}^{+\infty} \nu\big(P^k f - T_{k-1} P f\big) \qquad \text{(from (14))} && (18)\\
&= \sum_{k=1}^{+\infty} \nu\big(P^k f - T_k f\big) + \sum_{k=1}^{+\infty} \nu\big(P^{k-1} T f - T_{k-1} T f\big) \qquad \text{(from (13))} && (19)\\
&= \mu(f) - \nu(f) + \mu(Tf) \qquad \text{(from (14) and } \beta_1(f) = \nu(f)\text{)}\\
&= \mu(f) - \nu(f) + \mu(1_S)\,\nu(f) \qquad \text{(from the definition of } T\text{)}\\
&= \mu(f) - \nu(f)\big(1 - \mu(1_S)\big).
\end{aligned}$$

Note that the equality from (18) to (19) holds since both series in the right-hand side of (19) are absolutely convergent: indeed they equal to $\sum_{k=1}^{+\infty} \beta_{k+1}(f)$ and $\sum_{k=1}^{+\infty} \beta_k(Tf)f$ respectively. With $f = 1_{\mathbb{X}}$ the previous equality gives $\mu(1_S) = 1$ since $P1_{\mathbb{X}} = 1_{\mathbb{X}}$ and $\nu(1_{\mathbb{X}}) > 0$. Thus μ is a $P-$invariant non-negative measure such that $\mu(1_{\mathbb{X}}) > 0$ and $\mu(1_S) = 1$, so that $\pi = \mu/\mu(1_{\mathbb{X}})$ is a $P-$invariant probability measure on $(\mathbb{X}, \mathcal{X})$ such that $\pi(1_S) = 1/\mu(1_{\mathbb{X}}) > 0$. We have proved that $(i) \Leftrightarrow (ii)$ and that Equality (17) holds true under any of these two conditions.

Next we have

$$\forall k \geq 1, \quad \beta_{k+1}(1_{\mathbb{X}}) = \beta_k \circ (P - T)(1_{\mathbb{X}}) = \beta_k(1_{\mathbb{X}}) - \nu(1_{\mathbb{X}})\,\beta_k(1_S) \tag{20}$$

from (14) and the definition of T. Set: $\forall k \geq 1,\ b_k := \beta_k(1_{\mathbb{X}})$ and $c_k := \nu(1_{\mathbb{X}})\,\beta_k(1_S)$. Note that $b_k, c_k \geq 0$ and that $(b_k)_k$ is decreasing. From (20) (i.e. $c_k = b_k - b_{k+1}$) we deduce that

$$\sum_{k=1}^{n} k\, c_k = \sum_{k=1}^{n} \big(b_k\ - b_{n+1}\big) = \sum_{k=1}^{+\infty} \phi^{(n)}(k) \qquad \text{with} \quad \phi^{(n)}(k) := \big(b_k\ - b_{n+1}\big)\, 1_{[1,n]}(k).$$

Note that $0 \leq \phi^{(n)} \leq \phi^{(n+1)}$. Moreover, if $\lim_n b_n = 0$, then we have $\forall k \geq 1,\ \lim_n \phi^{(n)}(k) = b_k$ and the following equalities hold in $[0, +\infty]$

$$\sum_{k=1}^{+\infty} k\, c_k = \lim_{n\,\to\,+\infty} \sum_{k=1}^{n} k\, c_k = \lim_{n\,\to\,+\infty} \sum_{k=1}^{+\infty} \phi^{(n)}(k) = \sum_{k=1}^{+\infty} b_k$$

from the monotone convergence theorem with respect to the counting measure. This gives the equivalence $(ii) \Leftrightarrow (iv)$ and Equality (16). Equivalence $(iii) \Leftrightarrow (iv)$ follows from

$$\forall n \geq 1, \quad \sum_{k=1}^{n-1} c_k = \sum_{k=1}^{n-1} \big(b_k - b_{k+1}\big) = \nu(1_{\mathbb{X}}) - b_n \tag{21}$$

due to (20) and $\beta_1(1_{\mathbb{X}}) = \nu(1_{\mathbb{X}})$.

□

Recall that a set $S \in \mathcal{X}$ is said to be an atom for P if: $\forall (a, a') \in S^2,\ P(a, \cdot) = P(a', \cdot)$. Note that Assumption (**S**) then holds with $\nu(\cdot) := P(a_0, \cdot)$ with some (any) $a_0 \in S$. In the atomic case, Conditions (ii) or (iii) of Theorem 2.1 corresponds to the well-known condition involving the first return time in an atom. More precisely, let us assume that S is an atom for P and define R_S as the first return time in S:

$$R_S := \inf\{n \geq 1 : X_n \in S\}.$$

Then, we have

$$\forall n \geq 1, \quad \beta_n(1_S) = \mathbb{P}_{a_0}\big(R_S = n\big) \qquad \text{and} \qquad \beta_n(1_{\mathbb{X}}) = \mathbb{P}_{a_0}\big(R_S \geq n\big) \tag{22}$$

with $\beta_k(\cdot)$ defined from S and $\nu(\cdot) := P(a_0, \cdot)$ with some $a_0 \in S$. Hence Conditions (ii) of Theorem 2.1 rewrites as $\sum_{k=1} \mathbb{P}(R_S \geq k) < \infty$ and Conditions (iii) as $\mathbb{P}_{a_0}(R_S < \infty) = 1$ and $\sum_{k=1} k\mathbb{P}(R_S = k) < \infty$. Both conditions read as the usual moment condition of the return time in S: $\mathbb{E}_{a_0}[R_S] < \infty$.

3 Approximation of π in total variation norm

Recall that, if λ_1 and λ_2 belong to $\mathcal{M}^+$, the total variation norm $\|\lambda_1 - \lambda_2\|_{TV}$ is defined by

$$\|\lambda_1 - \lambda_2\|_{TV} := \sup_{f \in \mathcal{B}, \|f\| \leq 1} \big|\lambda_1(f) - \lambda_2(f)\big|.$$

If λ_1 and λ_2 are probability measures on $(\mathbb{X}, \mathcal{X})$, then $\|\lambda_1 - \lambda_2\|_{TV}$ corresponds to their standard total variation distance.

Under Assumption (**S**), for every $n \geq 1$ let us define on $(\mathbb{X}, \mathcal{X})$ the following finite positive measure μ_n and probability measure $\widetilde{\mu}_n$:

$$\forall n \geq 1, \quad \mu_n := \sum_{k=1}^{n} \beta_k \quad \text{and} \quad \widetilde{\mu}_n = \frac{1}{\mu_n(1_{\mathbb{X}})}\, \mu_n.$$

Theorem 3.1 *Assume that P satisfies Assumption* (**S**) *and that one of the four conditions* (i), (ii), (iii) *or* (iv) *of Theorem 2.1 holds. Define*

$$\forall n \geq 1, \quad \varepsilon_n := \sum_{k=n+1}^{+\infty} \beta_k(1_{\mathbb{X}}). \tag{23}$$

Then the $P-$invariant probability measure $\pi := \mu/\mu(1_{\mathbb{X}})$ on $(\mathbb{X}, \mathcal{X})$, with $\mu := \sum_{k=1}^{+\infty} \beta_k$, satisfies the following properties:

$$\forall n \geq 1, \quad \left\|\pi - \mu(1_{\mathbb{X}})^{-1}\mu_n\right\|_{TV} = \mu(1_{\mathbb{X}})^{-1}\, \varepsilon_n \leq \nu(1_{\mathbb{X}})^{-1}\, \varepsilon_n \tag{24}$$

$$\forall n \geq 1, \quad \left\|\pi - \widetilde{\mu}_n\right\|_{TV} \leq \left(\mu(1_{\mathbb{X}})^{-1} + \nu(1_{\mathbb{X}})^{-1}\right) \varepsilon_n \leq 2\, \nu(1_{\mathbb{X}})^{-1}\, \varepsilon_n. \tag{25}$$

Proof. We have

$$\|\pi - \mu_n/\mu(1_{\mathbb{X}})\|_{TV} = (\mu - \mu_n)(1_{\mathbb{X}})/\mu(1_{\mathbb{X}}) = \varepsilon_n/\mu(1_{\mathbb{X}})$$

since $\pi = \mu/\mu(1_{\mathbb{X}})$ and $\mu - \mu_n$ is a non-negative measure on $(\mathbb{X}, \mathcal{X})$, so that $\|\mu - \mu_n\|_{TV} = (\mu - \mu_n)(1_{\mathbb{X}}) = \varepsilon_n$ from (23). The last inequality in (24) follows from $\mu(1_{\mathbb{X}}) \geq \beta_1(1_{\mathbb{X}}) = \nu(1_{\mathbb{X}})$. To prove (25) consider any $f \in \mathcal{B}$ such that $\|f\| \leq 1$. Then

$$\begin{aligned}
\left|\pi(f) - \widetilde{\mu}_n(f)\right| &= \left|\pi(f) - \frac{\mu_n(f)}{\mu_n(1_{\mathbb{X}})}\right| \\
&\leq \left|\pi(f) - \frac{\mu_n(f)}{\mu(1_{\mathbb{X}})}\right| + |\mu_n(f)| \times \left|\frac{1}{\mu(1_{\mathbb{X}})} - \frac{1}{\mu_n(1_{\mathbb{X}})}\right| \\
&\leq \frac{\varepsilon_n}{\mu(1_{\mathbb{X}})} + \mu(1_{\mathbb{X}})\left|\frac{\mu_n(1_{\mathbb{X}}) - \mu(1_{\mathbb{X}})}{\mu(1_{\mathbb{X}})\mu_n(1_{\mathbb{X}})}\right| = \frac{\varepsilon_n}{\mu(1_{\mathbb{X}})} + \frac{1}{\mu_n(1_{\mathbb{X}})}\left|\mu_n(1_{\mathbb{X}}) - \mu(1_{\mathbb{X}})\right| \\
&\leq \left(\frac{1}{\mu(1_{\mathbb{X}})} + \frac{1}{\nu(1_{\mathbb{X}})}\right) \varepsilon_n
\end{aligned}$$

by using (24) and the following facts: first $|\mu_n(f)| \leq \mu_n(1_{\mathbb{X}}) \leq \mu(1_{\mathbb{X}})$, second (23) to get $|\mu_n(1_{\mathbb{X}}) - \mu(1_{\mathbb{X}})| = \varepsilon_n$ and finally $\mu_n(1_{\mathbb{X}}) \geq \nu(1_{\mathbb{X}})$. □

Remark 3.1 *If P satisfies Assumption* (**S**) *with $S = \mathbb{X}$, then P is uniformly ergodic, and we have $\sup_{x \in \mathbb{X}} \|P^n(x, \cdot) - \pi\|_{TV} \leq (1 - \nu(1_{\mathbb{X}}))^n$, e.g. see [RR04]. Note that, in this case, we also have $\varepsilon_n = (1 - \nu(1_{\mathbb{X}}))^n$ in (24)-(25) since it follows from an easy induction that: $\forall k \geq 1,\ \beta_k(1_{\mathbb{X}}) = \nu(1_{\mathbb{X}})\,(1 - \nu(1_{\mathbb{X}}))^{k-1}$.*

4 Explicit bound under geometric drift conditions

Throughout the Sections 4 and 5, any measurable function $V : \mathbb{X} \to [1, +\infty)$ will be called a Lyapunov function. For the sake of simplicity, the Lyapunov function V considered in this section is assumed to satisfy: $\forall x \in \mathbb{X}$, $(PV)(x) < \infty$. Hence, under Assumption (**S**), we have

$$\nu(V) < \infty.$$

The following theoretical statement is derived from Theorem 3.1 and [HL20b, Th. 3.1].

Theorem 4.1 ([HL20b]) *Assume that P satisfies Assumption (**S**). Moreover assume that*

$$\exists \delta \in (0,1),\ \exists L > 0, \quad PV \leq \delta V + L\, 1_S, \tag{**D**}$$

with respect to some Lyapunov function V. Then we have

$$\theta_0 := \limsup_n [\beta_n(1_{\mathbb{X}})]^{\frac{1}{n}} < 1.$$

That is, for every $\theta \in (\theta_0, 1)$ there exists a positive constant C_θ such that

$$\forall k \geq 1, \quad \beta_k(1_{\mathbb{X}}) \leq C_\theta\, \theta^k. \tag{26}$$

Hence Properties (24) and (25) hold with

$$\varepsilon_n \leq \frac{C_\theta}{1-\theta}\, \theta^{n+1}. \tag{27}$$

Proof. We know from [HL20b, Th. 3.1] that $\limsup_n [\beta_n(V)]^{\frac{1}{n}} < 1$ under Assumptions (**S**) and (**D**). Thus $\theta_0 < 1$ since $1_{\mathbb{X}} \leq V$. Then (27) follows from (23).

□

As mentioned in [HL20b, Rem. 5.4], the real number $\limsup_n [\beta_n(V)]^{1/n}$ may be strictly less than the so-called spectral gap related to the V-geometric ergodicity of P. In this case the rate of convergence in (27) is better than that given by the V-geometric ergodicity. Moreover note that (27) concerns the total variation norm rather than the usual $V-$weighted norm involved in V-geometric ergodicity. Recall that finding explicit rate and bound in the V-geometric ergodicity property are difficult issues. Similarly, finding explicit bound for θ_0 and for the constant C_θ in (27) is difficult a priori, because the above mentioned property $\limsup_n [\beta_n(V)]^{1/n} < 1$ is obtained in [HL20b, Th. 3.1] thanks to spectral arguments.

Below various statements specify the explicit control of the error term ε_n in (24)-(25) under Assumption (**S**) and the following drift condition:

$$\exists \delta \in (0,1),\ \forall x \in S^c, \quad (PV)(x) \leq \delta\, V(x). \tag{$\boldsymbol{D_{S^c}}$}$$

Note that Condition ($\boldsymbol{D_{S^c}}$) is equivalent to (**D**) when PV is bounded on S.

Theorem 4.2 *Assume that P satisfies (**S**) for some $S \in \mathcal{X}$ and $\nu \in \mathcal{M}_*^+$. Moreover assume that there exists a Lyapunov function V such that P satisfies ($\boldsymbol{D_{S^c}}$) and the following condition on S*

$$\forall x \in S, \quad (PV)(x) \leq \delta\, V(x) + \nu(V). \tag{$\boldsymbol{D_S}$}$$

Then the estimates (24)-(25) hold true with

$$\varepsilon_n \leq \frac{\nu(V)}{1-\delta}\, \delta^n. \tag{28}$$

Note that Conditions $(\boldsymbol{D}_{\boldsymbol{S}^c})$-$(\boldsymbol{D}_{\boldsymbol{S}})$ rewrite as

$$PV \le \delta\, V + \nu(V)1_S.$$

However, in view of the proof below and of the next Corollary 4.2, it may be convenient to well separate the condition on S^c and that on S.

Proof. Recall that $T = \nu(\cdot)1_S$. Then

$$\begin{aligned}(P-T)V &= 1_{S^c}\big(PV - \nu(V)1_S\big) + 1_S\big(PV - \nu(V)1_S\big)\\ &= 1_{S^c}\, PV + 1_S\big(PV - \nu(V)1_S\big)\\ &\le \delta\, V\, 1_{S^c} + 1_S\big(PV - \nu(V)1_S\big) = \delta\, V + 1_S\big(PV - \delta\, V - \nu(V)1_S\big)\\ &\le \delta\, V\end{aligned}$$

from $(\boldsymbol{D}_{\boldsymbol{S}^c})$ and $(\boldsymbol{D}_{\boldsymbol{S}})$. Using $P - T \ge 0$ and iterating the previous inequality gives

$$0 \le (P-T)^n V \le \delta^n\, V. \tag{29}$$

Next, it follows from (14) that

$$\forall n \ge 1, \quad \beta_n(V) = \nu\big((P-T)^{n-1}V\big) \le \nu(V)\,\delta^{n-1}$$

from which we deduce that

$$\beta_n(1_{\mathbb{X}}) \le \beta_n(V) \le \nu(V)\,\delta^{n-1}. \tag{30}$$

from the positivity of β_n. Then Conditions (ii) of Theorem 2.1 is fulfilled, and (28) is due to the definition of ε_n in (23).

□

When Assumption $(\mathbf{S})$ holds for an atom S and for $\nu(\cdot) := P(a_0, \cdot)$ with some (any) $a_0 \in S$, then Condition $(\boldsymbol{D}_{\boldsymbol{S}})$ is fulfilled since

$$\forall x \in S, \quad PV(x) - \delta\, V(x) - \nu(V) = -\delta\, V(x) \le 0.$$

Consequently the next corollary follows from Theorem 4.2.

Corollary 4.1 (Atomic case) *Assume that P satisfies Assumption $(\mathbf{S})$ with an atom S and with $\nu(\cdot)$ defined by $\nu(\cdot) := P(a_0, \cdot)$ with some (any) $a_0 \in S$. Moreover assume that there exists a Lyapunov function V such that P satisfies the drift condition $(\boldsymbol{D}_{\boldsymbol{S}^c})$. Then the geometrical bound (28) for ε_n holds true.*

Remark 4.1 *In the atomic case, the bound $\beta_n(1_{\mathbb{X}}) \le \nu(V)\,\delta^{n-1}$ (see (30)) may be derived from well-known results under Assumption $(\boldsymbol{D}_{\boldsymbol{S}^c})$. Indeed we know from (22) that $\beta_n(1_{\mathbb{X}}) = \mathbb{P}_{a_0}\big(R_S \ge n\big)$, where R_S is the first return time in S. Moreover $(\boldsymbol{D}_{\boldsymbol{S}^c})$ gives*

$$PV \le \delta\, V + (c - \delta\vartheta)\, 1_S \qquad \text{with } c := \nu(V) \text{ and } \vartheta := \inf_{x\in S} V(x).$$

Then we deduce from [DMPS18, Prop.4.3.3(ii)] that

$$\forall x \in \mathbb{X}, \quad \mathbb{E}_x\big[\delta^{-R_S}\big] \le V(x) + (c - \delta\vartheta)\,\delta^{-1}. \tag{31}$$

hence

$$\mathbb{E}_{a_0}\big[\delta^{-R_S}\big] \leq \vartheta + (c - \delta\vartheta)\delta^{-1} = c\,\delta^{-1}.$$

Consequently we obtain the same estimate as in (30) since

$$\beta_n(1_{\mathbb{X}}) = \mathbb{P}_{a_0}\big(R_S \geq n\big) = \mathbb{P}_{a_0}\big(\delta^{-R_S} \geq \delta^{-n}\big) \leq c\,\delta^{n-1}$$

from Markov's inequality.

According to the previous discussion, the bound $\beta_n(1_{\mathbb{X}}) \leq \nu(V)\,\delta^{n-1}$ obtained in (30), and consequently the resulting bound (28) for ε_n, are not only simple and explicit but also quite optimal. Although Condition $(\boldsymbol{D_S})$ is automatically satisfied in the atomic case and may hold in the non atomic case too, this condition is nevertheless restrictive. In the next statement, Condition $(\boldsymbol{D_{S^c}})$ is preserved, but the function V is replaced with V^{α_0} for some suitable $\alpha_0 \in (0,1]$ in Condition $(\boldsymbol{D_S})$. The price to be paid is that the geometrical bound (28) for ε_n will hold with $\delta^{\alpha_0 n}$ in place of the expected rate δ^n. But the benefit will be that the bound for ε_n remains simple and explicit.

Let V be a Lyapunov function such that PV is bounded on S. Then

$$\exists \alpha_0 \in (0,1],\ \forall x \in S, \quad (PV^{\alpha_0})(x) \leq \delta^{\alpha_0}\,V(x)^{\alpha_0} + \nu(V^{\alpha_0}). \tag{32}$$

Indeed, set $M_S := \sup_S PV$. Then, for every $\alpha \in (0,1]$, we have $1 \leq \sup_S PV^\alpha \leq {M_S}^\alpha$ from $1_{\mathbb{X}} \leq V^\alpha$ and $PV^\alpha \leq (PV)^\alpha$ (Jensen's inequality). Moreover

$$\forall x \in S, \quad (PV^\alpha)(x) - \delta^\alpha\,V(x)^\alpha - \nu(V^\alpha) \leq {M_S}^\alpha - \delta^\alpha - \nu(1_{\mathbb{X}})$$

from $1_{\mathbb{X}} \leq V$. Passing to the limit when $\alpha \to 0$ gives (32) since $\nu(1_{\mathbb{X}}) > 0$.

Corollary 4.2 *Assume that P satisfies Assumption* (**S**) *and that there exists a Lyapunov function V such that P satisfies Condition $(\boldsymbol{D_{S^c}})$ and PV is bounded on S (so that the usual drift condition* (**D**) *holds). Let $\alpha_0 \in (0,1]$ be such that Condition* (32) *is fulfilled. Then the estimates (24)-(25) hold true with*

$$\varepsilon_n \leq \frac{\nu(V^{\alpha_0})}{1-\delta^{\alpha_0}}\,\delta^{\alpha_0 n}. \tag{33}$$

Proof. We have

$$\forall x \in S^c, \quad (PV^{\alpha_0})(x) \leq \delta^{\alpha_0}\,V(x)^{\alpha_0} \tag{34}$$

from $PV^{\alpha_0} \leq (PV)^{\alpha_0}$ (Jensen's inequality) and from $(\boldsymbol{D_{S^c}})$. Moreover (32) holds. Then Corollary 4.2 follows from Theorem 4.2 applied to V^{α_0} and δ^{α_0} in place of V and δ.

□

The previous statements concern the approximation of the stationary distribution π. To conclude this section recall that the usual V-weighted operator norm of $P^n - T_n$ with T_n given in (12) is proved in [HL20b, Cor. 2.3] to converge to zero with geometric rate of convergence under Assumptions (**S**) and (**D**). Using Inequality (29), we can here specify this rate of convergence under the assumptions of Theorem 4.2 or Corollary 4.2.

Theorem 4.3 *Assume that P satisfies the assumptions of Theorem 4.2. Then*

$$\sup_{|f|\leq V}\sup_{x\in\mathbb{X}}\frac{\big|(P^nf)(x)-(T_nf)(x)\big|}{V(x)}\leq\delta^n\quad\textit{with}\quad T_nf=\sum_{k=1}^n\beta_k(f)P^{n-k}1_S \tag{35}$$

where the functions f are assumed to be real-valued and measurable on $(\mathbb{X},\mathcal{X})$. Similarly, if P satisfies the assumptions of Corollary 4.2, then Property (35) holds with V^{α_0} and δ^{α_0} in place of V and δ.

Proof. If P satisfies the assumptions of Theorem 4.2, then it follows from (29) that for every measurable function $f:\mathbb{X}\to\mathbb{R}$ such that $|f|\leq V$

$$|P^nf-T_nf|=|(P-T)^nf|\leq(P-T)^n|f|\leq(P-T)^nV\leq\delta^n\,V, \tag{36}$$

from which we deduce (35). Under the assumptions of Corollary 4.2, we know that P satisfies the assumptions of Theorem 4.2 with V^{α_0} and δ^{α_0} in place of V and δ, so that (36) holds with V^{α_0} and δ^{α_0} in place of V and δ too.

□

If P satisfies the assumptions of Theorem 4.2 and the strong aperiodicity condition $\nu(1_S)>0$, then the bound (35) can be used to obtain a rate of convergence in the V-geometrical ergodicity property, which simply depends on $\delta\in(0,1)$ in $(\boldsymbol{D_S})$ and on the real number

$$\varrho_S:=\limsup_{n\to+\infty}\left(\sup_{x\in\mathbb{X}}\frac{\big|(P^n1_S)(x)-\pi(1_S)\big|}{V(x)}\right)^{\frac{1}{n}}$$

introduced in [HL20b]. More precisely we know from [HL20b] that $\varrho_S<1$. In Theorem 4.4 below we consider any $\varrho\in(\varrho_S,1)$ and we define

$$\alpha:=\max(\delta,\varrho)\qquad\text{and}\quad D_\varrho:=\sup_{n\geq0}\varrho^{-n}\sup_{x\in\mathbb{X}}\frac{\big|(P^n1_S)(x)-\pi(1_S)\big|}{V(x)}<\infty.$$

Theorem 4.4 *Assume that P satisfies the assumptions of Theorem 4.2 and that $\nu(1_S)>0$. Then we have*

$$\sup_{|f|\leq V}\sup_{x\in\mathbb{X}}\frac{\big|(P^nf)(x)-\pi(f)\big|}{V(x)}\leq\frac{\nu(V)+1-\delta}{1-\delta}\;\delta^n+\frac{\nu(V)\,D_\varrho}{\delta}\,n\,\alpha^n. \tag{37}$$

Moreover, setting $c:=\mu(1_{\mathbb{X}})^{-1}+\nu(1_{\mathbb{X}})^{-1}\leq2\nu(1_{\mathbb{X}})^{-1}$, the following inequality holds for every $n\geq1$ and for every $A\in\mathcal{X}$:

$$\left|P^n(x,A)-\frac{\mu_n(1_A)}{\mu_n(1_{\mathbb{X}})}\right|\leq\left(\frac{(1+c)\nu(V)+1-\delta}{1-\delta}\;\delta^n+\frac{\nu(V)\,D_\varrho}{\delta}\,n\,\alpha^n\right)V(x). \tag{38}$$

Similarly, if P satisfies the assumptions of Corollary 4.2 and if $\nu(1_S)>0$, then the bounds (37) and (38) hold true with V^{α_0} and δ^{α_0} in place of V and δ (the function V must be replaced by V^{α_0} in the definitions of ϱ_S and D_ϱ).

Proof. Property (37) can be easily obtained by using the bound (35) in the proof of [HL20b, Th. 5.3]. Then Inequality (38) follows from (37), (25) and (28) by using $V \geq 1_{\mathbb{X}}$ and the triangular inequality.

□

Similar inequality to (38) can be obtained with $\mu_n(1_A)/\mu(1_{\mathbb{X}})$ from (24). Of course any bound known for $|P^n(x,A)-\pi(1_A)|$ combined with (24) or (25) can be used to obtain an approximate value of $P^n(x,A)$.

Remark 4.2 *Let r be the spectral radius of the operator $P-T$ on the $V-$weighted supremum space $(\mathcal{B}_V, \|\cdot\|_V)$ composed of the complex-valued measurable functions $f : \mathbb{X} \to \mathbb{C}$ such that $\|f\|_V := \sup_{\mathbb{X}} |f|/V < \infty$. Then (35) gives $r \leq \delta$. Consequently the proofs of [HL20b, Th. 5.3] and [HL20a, Th. A.1] can be easily adapted to obtain the following alternative:*

- *either $\varrho_S \leq \delta$*
- *or $\varrho_S = \theta^{-1}$ with $\theta := \min\left\{|z| : z \in \mathbb{C},\ 1 < |z| < 1/\delta,\ B_{1_{\mathbb{X}}}(z) = 0\right\}$, where $B_{1_{\mathbb{X}}}$ is the power series defined by $B_{1_{\mathbb{X}}}(z) := \sum_{k=1}^{+\infty} \beta_k(1_{\mathbb{X}})\, z^k$.*

This alternative is due to the following fact: if $\lambda \in \mathbb{C}$ is such that $\delta < |\lambda| \leq 1$, then λ is an eigenvalue of P on $\mathcal{B}_V$ if, and only if, $B_{1_{\mathbb{X}}}(\lambda^{-1}) = 0$ (see [HL20a, Prop. A.2]). We can observe that the bound (28) in Theorem 4.2 or Corollary 4.1 does not take into account the possible eigenvalues λ of P such that $\delta < |\lambda| < 1$, whereas the bound (37), thus (38), depend on the real number ϱ_S which equals to θ^{-1}, thus is strictly greater than δ, when such eigenvalues exist. Recall that atomic instances with such eigenvalues exist, e.g. see [HL14, Prop. 4.1]. Consequently, at least in the geometric and atomic case, the use of the non-negative measures β_n seems to erase the effect of intermediate eigenvalues between δ and the eigenvalue 1. By contrast, the (necessary non-atomic) case when the rate of convergence in (4) is only $O(\delta^{\alpha_0 n})$ for some $\alpha_0 \in (0,1)$ rather than $O(\delta^n)$ (Corollary 4.2) could coincide with the existence of such intermediate eigenvalues, which therefore would slow down the rate of convergence in (4) as in (37).

5 Explicit bounds under subgeometric drift conditions

For the sake of simplicity, any Lyapunov function V in this section is assumed to satisfy: $\forall x \in \mathbb{X},\ (PV)(x) < \infty$.

5.1 Theoretical results

Let P be a Markov kernel on $(\mathbb{X}, \mathcal{X})$ satisfying Assumption (**S**). Let $T(\cdot) := \nu(\cdot)1_S$. For $m \geq 1$ let us introduce the following assumption: there exists a collection $\{V_i\}_{i=0}^m$ of Lyapunov functions such that

$$\forall i \in \{0, \ldots, m-1\}, \quad (P-T)V_i \leq V_i - V_{i+1}. \tag{39}$$

Note that the properties (39) and $P - T \geq 0$ from (**S**) give

$$V_m \leq V_{m-1} \leq \cdots \leq V_1 \leq V_0.$$

Since $(PV_0)(\cdot) < \infty$ by hypothesis, we have under Assumption $(\mathbf{S})$

$$\nu(V_0) < \infty. \tag{40}$$

In this section we first present a theoretical result which shows that the estimates (24)-(25) hold with a polynomial rate of convergence under Assumptions (39). Second we propose further statements in which an explicit polynomial rate of convergence is obtained. Denote by $(\vartheta_j)_{j\geq 0}$ the recurrent sequence of positive real numbers defined by

$$\vartheta_0 = 1 \quad \text{and} \quad \forall \ell \geq 1,\ \vartheta_\ell = \sum_{j=0}^{\ell-1} C_\ell^j \vartheta_j. \tag{41}$$

Theorem 5.1 *Let P be a Markov kernel on $(\mathbb{X}, \mathcal{X})$ satisfying Assumption $(\mathbf{S})$. Moreover assume that there exists a collection $\{V_i\}_{i=0}^m$ of Lyapunov functions satisfying Assumptions (39). Then we have*

$$\forall j \in \{0, \ldots, m-1\}, \quad \sum_{k=1}^{+\infty} k^j\, \beta_k(V_{j+1}) \leq \vartheta_j\, \nu(V_0). \tag{42}$$

Moreover the estimates (24)-(25) hold true with $(\varepsilon_n)_{n\geq 0}$ satisfying

$$\lim_{n\to+\infty} n^{m-1}\varepsilon_n = 0. \tag{43}$$

Proof. Assume that (39) holds with $m = 1$, that is $(P - T)V_0 \leq V_0 - V_1$, or equivalently: $V_1 \leq V_0 - (P - T)V_0$. Thus

$$\forall k \geq 0, \quad (P-T)^k V_1 \leq (P-T)^k V_0 - (P-T)^{k+1} V_0$$

from which we deduce that

$$\forall n \geq 1, \quad \sum_{k=0}^{n} (P-T)^k V_1 \leq \sum_{k=0}^{n} \left[(P-T)^k V_0 - (P-T)^{k+1} V_0\right] \leq V_0.$$

It follows from (14) that

$$\forall n \geq 1, \quad \sum_{k=1}^{n+1} \beta_k(V_1) \leq \nu(V_0).$$

This proves (42) when $m = 1$. Now assume that Inequalities (42) hold true for some $m \geq 1$. Assume that (39) holds at order $m + 1$. Then using $V_{m+1} \leq V_m - (P - T)V_m$, we get

$$\forall k \geq 0, \quad (P-T)^k V_{m+1} \leq (P-T)^k V_m - (P-T)^{k+1} V_m$$

hence we have for every $n \geq 1$

$$\begin{aligned}
\sum_{k=0}^{n}(k+1)^m\,(P-T)^k V_{m+1} &\leq \sum_{k=0}^{n}(k+1)^m\,(P-T)^k V_m - \sum_{k=0}^{n+1}k^m\,(P-T)^k V_m \\
&\leq \sum_{k=0}^{n}\left[(k+1)^m - k^m\right](P-T)^k V_m \\
&\leq \sum_{j=0}^{m-1} C_m^j \sum_{k=0}^{n} k^j\,(P-T)^k V_m \\
&\leq \sum_{j=0}^{m-1} C_m^j \sum_{k=0}^{n} k^j\,(P-T)^k V_{j+1}
\end{aligned}$$

by using $\forall j \in \{0,\ldots,m-1\},\ V_m \leq V_{j+1}$ for the last inequality. It follows from (14) that

$$\sum_{k=1}^{+\infty} k^m\,\beta_k(V_{m+1}) \leq \sum_{j=0}^{m-1} C_m^j \sum_{k=0}^{+\infty} k^j\,\beta_{k+1}(V_{j+1}) \leq \sum_{j=0}^{m-1} C_m^j \sum_{k=1}^{+\infty} k^j\,\beta_k(V_{j+1}) \leq \left(\sum_{j=0}^{m-1} C_m^j \vartheta_j\right)\nu(V_0)$$

from the induction hypothesis. This gives Inequalities (42) at order $m+1$. Finally, to prove (43), note that

$$\varepsilon_n = \sum_{k=n+1}^{+\infty}\beta_k(1_{\mathbb{X}}) \leq \sum_{k=n+1}^{+\infty}\beta_k(V_m) \leq \frac{1}{(n+1)^{m-1}}\sum_{k=n+1}^{+\infty} k^{m-1}\beta_k(V_m)$$

since $1_{\mathbb{X}} \leq V_m$. This gives (43) since $\sum_{k=1}^{+\infty} k^{m-1}\,\beta_k(V_m) < \infty$ from (42) for $j = m-1$.

□

Theorem 5.2 *Let P be a Markov kernel on $(\mathbb{X},\mathcal{X})$ satisfying Assumption (**S**). Moreover assume that there exists there exist a collection $\{V_i\}_{i=0}^m$ of Lyapunov functions satisfying Assumptions (39). Then the following assertions hold true.*

(i) $\forall i \in \{0,\ldots,m\},\ \forall k \geq 1,\ \beta_k(V_i) < \infty$.

(ii) If the sequence $(\beta_k(V_m))_{k\geq 1}$ is decreasing, then

$$\forall n \geq 1, \quad \beta_n(V_m) \leq \frac{C_m\,\nu(V_0)}{n^m} \quad \textit{with} \quad C_m := 2^{\frac{m(m+1)}{2}-1}. \tag{44}$$

Moreover, if $m \geq 2$, then the estimates (24)-(25) hold true with

$$\forall n \geq 1, \quad \varepsilon_n \leq \frac{C_m\,\nu(V_0)}{(m-1)}\,\frac{1}{n^{m-1}}. \tag{45}$$

(iii) If the sequence $\left(\beta_k(V_m)\right)_{k\geq 1}$ is decreasing and if $\mu(V_0) := \sum_{k=1}^{+\infty}\beta_k(V_0) < \infty$, then

$$\forall n \geq 1, \quad \beta_n(V_m) \leq \frac{D_m\,\mu(V_0)}{n^{m+1}} \quad \textit{with} \quad D_m := 2^{\frac{(m+1)(m+2)}{2}+1}. \tag{46}$$

Moreover the estimates (24)-(25) hold true with

$$\forall n \geq 1, \quad \varepsilon_n \leq \frac{D_m\,\mu(V_0)}{m}\,\frac{1}{n^m}. \tag{47}$$

Note that, using the triangular inequality, any quantitative error bounds on $|P^n(x,A)-\pi(1_A)|$ as in [AF05, AFV15] can be combined with (45) or (47) to solve (Q2).

Remark 5.1 *A sufficient condition for the sequence $(\beta_k(V_m))_{k\geq 1}$ to be decreasing is that there exists an additional Lyapunov function V_{m+1} satisfying $PV_m \leq V_m - V_{m+1}$ on S^c and $V_{m+1} \leq V_m$. This fact follows from Assertion (b) of Lemma 5.1 applied with $V = V_m$ and $W = V_{m+1}$. Consequently, in practice, if the decrease of the sequence $(\beta_k(V_m))_{k\geq 1}$ cannot be directly checked, then this last condition can be dropped in the assumptions of Theorem 5.2 provided that (39) is assumed at order $m+1$. But it is worth noticing that, in this case, the conclusions (44)-(45) and (46)-(47) are only guaranteed at the order m (as stated in Theorem 5.2). Indeed the conditions $PV_m \leq V_m - V_{m+1}$ on S^c and $V_{m+1} \leq V_m$ are then only used to prove that the sequence $(\beta_k(V_m))_{k\geq 1}$ is decreasing (the decrease of the sequence $(\beta_k(V_{m+1}))_{k\geq 1}$ is not guaranteed).*

Lemma 5.1 *Assume that P satisfies Assumption (**S**). Let V and W be two Lyapunov functions such that*

$$(P-T)V \leq V - W \quad \text{where} \quad T(\cdot) := \nu(\cdot)1_S. \tag{48}$$

Then the following statements hold true.

(a) $\forall k \geq 1,\ \beta_k(V) < \infty$

(b) The sequence $\big(\beta_k(V)\big)_{k\geq 1}$ *is decreasing.*

(c) If the sequence $\big(\beta_k(W)\big)_{k\geq 1}$ *is decreasing, then we have for every* $k \geq 1$ *and* $\varepsilon \in \{0,1\}$

$$\beta_k(W) \leq \frac{\nu(V)}{k} \quad \text{and} \quad \beta_{2k-\varepsilon}(W) \leq \frac{\beta_k(V)}{k}.$$

(d) If $\mu(V) := \sum_{k=1}^{+\infty} \beta_k(V) < \infty$ *and if the sequence* $\big(\beta_k(W)\big)_{k\geq 1}$ *is decreasing, then*

$$\forall n \geq 1, \quad \beta_n(W) \leq \frac{16\,\mu(V)}{n^2}.$$

Proof. Note that $W \leq V$ from (48) and $P - T \geq 0$. Next we deduce from (48) that $\forall j \geq 1,\ (P-T)^j V \leq (P-T)^{j-1}(V-W)$. Then (14) gives

$$\forall j \geq 1, \quad \beta_{j+1}(V) \leq \beta_j(V) - \beta_j(W) \leq \beta_j(V) \quad \text{in } [0,+\infty].$$

Using $\beta_1(V) = \nu(V) < \infty$, Assertion (a) is obtained by induction, and Assertion (b) is then obvious. Next rewrite the previous inequalities as

$$\forall j \geq 1, \quad 0 \leq \beta_j(W) \leq \beta_j(V) - \beta_{j+1}(V) \tag{49}$$

and assume that $(\beta_j(W))_{j\geq 1}$ is decreasing. Then it follows from (49) that

$$\forall k \geq 1, \quad k\,\beta_k(W) \leq \sum_{j=1}^{k} \beta_j(W) \leq \beta_1(V) - \beta_{k+1}(V) \leq \nu(V),$$

from which we deduce the first inequality in Assertion (c). Moreover (49) also gives

$$\forall k \geq 1,\ \forall \varepsilon \in \{0,1\} \quad k\,\beta_{2k-\varepsilon}(W) \leq \sum_{j=k}^{2k-\varepsilon} \beta_j(W) \leq \beta_k(V) - \beta_{2k-\varepsilon+1}(V) \leq \beta_k(V), \tag{50}$$

from which we deduce the second inequality in Assertion (c). Finally, to prove Assertion (d), note that for every $\ell \geq 1$ and every $\varepsilon \in \{0,1\}$

$$\ell\,\beta_{2\ell-\varepsilon}(V) \leq \sum_{j=\ell}^{2\ell-\varepsilon} \beta_j(V) \leq \mu(V) < \infty \tag{51}$$

since $(\beta_j(V))_{j\geq 1}$ is decreasing (Assertion (b)). Let $n \geq 1$ and write $n = 2(2\ell - \varepsilon_1) - \varepsilon_2$ with $\ell \geq 1$ and $(\varepsilon_1, \varepsilon_2) \in \{0,1\}^2$. Then it follows from (50) and (51) that

$$\beta_n(W) \leq \frac{\beta_{2\ell-\varepsilon_1}(V)}{2\ell - \varepsilon_1} \leq \frac{\mu(V)}{\ell(2\ell-1)} \leq \frac{\mu(V)}{\ell^2} \leq \frac{16\,\mu(V)}{(n+2\varepsilon_1+\varepsilon_2)^2} \leq \frac{16\,\mu(V)}{n^2}.$$

□

Proof of Theorem 5.2. Assertion (a) of Lemma 5.1 applied with $V = V_0$ and $W = V_1$ proves that: $\forall k \geq 1,\ \beta_k(V_0) < \infty$. Then statement (i) of Theorem 5.2 holds since $V_i \leq V_0$. Now let us prove by induction on the positive integer m that Property (44) holds. If $m = 1$, then the first inequality in Assertion (c) of Lemma 5.1 applied with $V = V_0$ and $W = V_1$ provides

$$\forall n \geq 1, \quad \beta_n(V_1) \leq \frac{\nu(V_0)}{n}.$$

Hence (44) holds with $C_1 = 1$ when $m = 1$. Now assume that (44) holds for some $m \geq 1$. Let $\{V_i\}_{i=0}^{m+1}$ be a collection of Lyapounov functions such that

$$\forall i \in \{0, \ldots, m\}, \quad (P - T)V_i \leq V_i - V_{i+1}$$

and finally such that the sequence $(\beta_k(V^{m+1}))_{k\geq 1}$ is decreasing. Note that Assertion (b) of Lemma 5.1 applied with $V = V_m$ and $W = V_{m+1}$ ensures that the sequence $(\beta_k(V_m))_{k\geq 1}$ is decreasing. Consequently we have

$$\forall k \geq 1, \quad \beta_k(V_m) \leq \frac{C_m\,\nu(V_0)}{k^m} \quad \text{with} \quad C_m := 2^{\frac{m(m+1)}{2}-1} \tag{52}$$

from the induction hypothesis. Next let $n \geq 1$ and write $n = 2k - \varepsilon$ with $k \geq 1$ and $\varepsilon \in \{0,1\}$. Then the second inequality in Assertion (c) of Lemma 5.1 applied with $V = V_m$ and $W = V_{m+1}$ gives

$$\beta_n(V_{m+1}) \leq \frac{\beta_k(V_m)}{k} \tag{53}$$

so that $\beta_n(V_{m+1}) \leq C_m\nu(V_0)/k^{m+1}$ from (52). Hence

$$\beta_n(V_{m+1}) \leq \frac{2^{m+1}\,C_m\,\nu(V_0)}{(n+\varepsilon)^{m+1}} \leq \frac{C_{m+1}\,\nu(V_0)}{n^{m+1}} \quad \text{with} \quad C_{m+1} = 2^{m+1}\,C_m = 2^{\frac{(m+1)(m+2)}{2}-1}.$$

We have proved (44) by induction.

Now, if $m \geq 2$, then Conditions (ii) of Theorem 2.1 is fulfilled due to (44) and $1_{\mathbb{X}} \leq V_m$. Moreover Property (45) is deduced from

$$\forall n \geq 1, \quad \varepsilon_n \leq \sum_{k=n+1}^{+\infty} \beta_k(V_m) \leq C_m \nu(V_0) \sum_{k=n+1}^{+\infty} \frac{1}{k^m} \leq C_m \nu(V_0) \int_n^{+\infty} \frac{dt}{t^m} = \frac{C_m\, \nu(V_0)}{(m-1)\, n^{m-1}}$$

since $\beta_k(1_{\mathbb{X}}) \leq \beta_k(V_m)$. Statement (ii) of Theorem 5.2 is established.

The proof of Statement (iii) of Theorem 5.2 follows the same induction procedure. Indeed, if $m = 1$, then Assertion (d) of Lemma 5.1 applied with $V = V_0$ and $W = V_1$ provides

$$\forall n \geq 1, \quad \beta_n(V_1) \leq \frac{16\, \mu(V_0)}{n^2}.$$

Hence (46) holds with $D_1 = 16$ when $m = 1$. Now, assume that (46) is true at order m for some $m \geq 1$, and consider a collection $\{V_i\}_{i=0}^{m+1}$ of Lyapunov functions as in the above induction proof. Then, writing $n \geq 1$ as $n = 2k - \varepsilon$ with $k \geq 1$ and $\varepsilon \in \{0, 1\}$, we deduce from (53) and from the induction hypothesis that

$$\beta_n(V_{m+1}) \leq \frac{\beta_k(V_m)}{k} \leq \frac{D_m\, \mu(V_0)}{k^{m+2}} \quad \text{with} \quad D_m := 2^{\frac{(m+1)(m+2)}{2}+1}.$$

Hence

$$\beta_n(V_{m+1}) \leq \frac{2^{m+2}\, D_m\, \mu(V_0)}{(n+\varepsilon)^{m+2}} \leq \frac{D_{m+1}\, \mu(V_0)}{n^{m+2}} \quad \text{with} \quad D_{m+1} = 2^{m+2}\, D_m.$$

This proves (46). Then Condition (ii) of Theorem 2.1 is fulfilled, and Property (47) can be derived as above.

□

5.2 Applications

Let P be a Markov kernel on $(\mathbb{X}, \mathcal{X})$ satisfying Assumption $(\mathbf{S})$ for some $S \in \mathcal{X}$ and $\nu \in \mathcal{M}_*^+$. For $m \geq 1$ let us introduce the following conditions: there exists a collection $\{V_i\}_{i=0}^m$ of Lyapunov functions such that

$$\forall i \in \{0, \ldots, m-1\}, \quad \begin{cases} \forall x \in \mathbb{X}, & V_{i+1}(x) \leq V_i(x) \\ \forall x \in S^c, & (PV_i)(x) \leq V_i(x) - V_{i+1}(x) \\ \forall x \in S, & PV_i(x) \leq \big(V_i(x) - V_{i+1}(x)\big) + \nu(V_i). \end{cases} \tag{54}$$

Note that the second condition in (54) implies that $V_{i+1} \leq V_i$ on S^c, so that the first condition in (54) may be replaced with $V_{i+1} \leq V_i$ on S. Also note that the term $(V_i(x) - V_{i+1}(x))$ in the third condition of (54) is non-negative. Finally observe that Assumption (54) rewrites in a more concise form as follows

$$\forall i \in \{0, \ldots, m-1\}, \quad \begin{cases} V_{i+1} \leq V_i \\ PV_i \leq \big(V_i - V_{i+1}\big) + \nu(V_i)\, 1_S. \end{cases}$$

However, as in the previous section, it could be convenient to well separate the conditions on S^c and S respectively.

Corollary 5.1 *Assume that P satisfies Assumption* (**S**) *and that there exists a collection $\{V_i\}_{i=0}^m$ of Lyapunov functions satisfying Assumptions (54). Then the statements (i)-(ii)-(iii) of Theorem 5.2 hold true.*

Proof. Prove that (54) imply (39). For $i = 0, \dots, m-1$ we have

$$\begin{aligned}(P-T)V_i &= 1_{S^c}\big(PV_i - \nu(V_i)1_S\big) + 1_S\big(PV_i - \nu(V_i)1_S\big)\\ &= 1_{S^c}\, PV_i + 1_S\big(PV_i - \nu(V_i)1_S\big)\\ &\le 1_{S^c}\big(V_i - V_{i+1}\big) + 1_S\big(PV_i - \nu(V_i)1_S\big)\\ &= V_i - V_{i+1} + 1_S\big(PV_i - V_i + V_{i+1} - \nu(V_i)1_S\big)\\ &\le V_i - V_{i+1}.\end{aligned}$$

This gives (39). □

Corollary 5.2 (Atomic case) *Let P be a Markov kernel on $(\mathbb{X}, \mathcal{X})$ satisfying Assumption* (**S**) *with an atom S and with $\nu(\cdot)$ defined by $\nu(\cdot) := P(a_0, \cdot)$ for $a_0 \in S$. Moreover assume that there exists a collection $\{V_i\}_{i=0}^m$ of Lyapunov functions such that*

$$\forall i \in \{0, \dots, m-1\}, \quad \begin{cases} \forall x \in S, & V_{i+1}(x) \le V_i(x) \\ \forall x \in S^c, & (PV_i)(x) \le V_i(x) - V_{i+1}(x). \end{cases} \tag{55}$$

Then the statements (i)-(ii)-(iii) of Theorem 5.2 hold true.

Proof. Prove that (55) implies (54). First note that the second condition of (55) ensures that $V_{i+1} \le V_i$ on S^c, so that the first condition in (54) holds. The second condition in (55) is the same as in (54). Finally, for every $i = 0, \dots, m-1$ and every $x \in S$ we have $PV_i(x) = \nu(V_i)$ since S in an atom, so that

$$\forall x \in S, \quad PV_i(x) - \big(V_i(x) - V_{i+1}(x)\big) - \nu(V_i) = V_{i+1}(x) - V_i(x) \le 0.$$

This proves the third condition of (54).

□

Now we apply the two previous corollaries under the following subgeometric drift condition

$$\exists \alpha \in [0,1), \exists c_1 > 0,\ \forall x \in S^c, \quad (PV)(x) \le V(x) - c_1\, V(x)^\alpha \tag{$\boldsymbol{Sub_{\alpha, S^c}}$}$$

where V is some Lyapunov function. We begin with the atomic case. For any $\alpha \in [0,1)$ define the integer $m \equiv m_\alpha \ge 1$ by

$$m := \big\lfloor (1-\alpha)^{-1} \big\rfloor \tag{56}$$

where $\lfloor \cdot \rfloor$ denotes the integer part function on $\mathbb{R}$.

Corollary 5.3 (Atomic case) *Let P be a Markov kernel on $(\mathbb{X}, \mathcal{X})$ satisfying Assumptions* (**S**) *and* ($\boldsymbol{Sub_{\alpha, S^c}}$) *with an atom S and with $\nu(\cdot)$ defined by $\nu(\cdot) := P(a_0, \cdot)$ for $a_0 \in S$. Assume that PV is bounded on S. Let $m \equiv m_\alpha$ given in (56). Then the estimates (24)-(25) hold true with*

$$\forall n \ge 1, \quad \varepsilon_n \le \frac{d\, C_m\, \nu(V)}{(m-1)}\, \frac{1}{n^{m-1}} \quad \text{with} \quad C_m = 2^{\frac{m(m+1)}{2}-1} \tag{57}$$

where d is a positive constant (specified in the proof). If moreover $\mu(V) := \sum_{k=1}^{+\infty} \beta_k(V) < \infty$, then the estimates (24)-(25) hold true with

$$\forall n \geq 1, \quad \varepsilon_n \leq \frac{d\, D_m\, \mu(V)}{m} \frac{1}{n^m} \quad \textit{with} \quad D_m = 2^{\frac{(m+1)(m+2)}{2}+1}. \tag{58}$$

To prove Corollary 5.3 we use the following lemma which is based on [JR02, Lem. 3.5].

Lemma 5.2 *Let $S \in \mathcal{X}$, and let W be a Lyapunov function such that PW is bounded on S. Let $0 < \theta_2 < \theta_1 < 1$ be such that*

$$\exists c > 0,\ \forall x \in S^c, \quad (PW^{\theta_1})(x) \leq W(x)^{\theta_1} - c\, W(x)^{\theta_2}.$$

Then

$$\exists c' > 0,\ \forall x \in S^c, \quad (PW^{\theta_2})(x) \leq W(x)^{\theta_2} - c'\, W(x)^{\theta_3} \qquad \textit{with} \quad \theta_3 = 2\theta_2 - \theta_1.$$

Proof. The hypothesis of Lemma 5.2 writes as $PW^{\theta_1} \leq W^{\theta_1} - c\left(W^{\theta_1}\right)^{\theta_2/\theta_1}$ on S^c. It follows from [JR02, Lem. 3.5] that

$$\forall \eta \in (0,1],\ \exists c' > 0, \quad PW^{\eta\theta_1} \leq W^{\eta\theta_1} - c'\left(W^{\theta_1}\right)^{\frac{\theta_2}{\theta_1}+\eta-1} \quad \text{on } S^c.$$

With $\eta := \theta_2/\theta_1$ this gives

$$PW^{\theta_2} \leq W^{\theta_2} - c'\, W^{2\theta_2-\theta_1} \quad \text{on } S^c.$$

□

Proof of Corollary 5.3. Note that the implication in Lemma 5.2 does not apply when $\theta_2 = 0$ since $P1_{\mathbb{X}} = 1_X$ (this would give $c' = 0$). Let $\alpha_1 := 1 - 1/m \in [0,1)$ with m given in (56). Note that $\alpha_1 \leq \alpha$. It then follows from $(\boldsymbol{Sub}_{\boldsymbol{\alpha},\boldsymbol{S^c}})$ that

$$PV \leq V - c_1\, V^{\alpha_1} \quad \text{on } S^c. \tag{59}$$

Note that we can choose $c_1 < 1$ in (59).

- If $\alpha_1 = 0$ (i.e. $\alpha \in [0, 1/2)$), then Assumptions (55) of Corollary 5.2 hold with $m = 1$ and with $V_0 = c_1^{-1}V$ and $V_1 = 1_{\mathbb{X}}$ (note that $1_{\mathbb{X}} = V_1 \leq V_0$).
- If $\alpha_1 = 1/2$ (i.e. $\alpha \in [1/2, 2/3)$), then we deduce from (59) and Lemma 5.2 that

  $$\exists c_2 > 0, \quad PV^{\alpha_1} \leq V^{\alpha_1} - c_2\, V^{\alpha_2} \quad \text{on } S^c \tag{60}$$

  with $\alpha_2 := 2\alpha_1 - 1 = 0$. Again note that we can choose $c_2 < 1$. Then the procedure stops, and Assumptions (55) of Corollary 5.2 hold with $m = 2$ and with $V_0 = c_1^{-1}c_2^{-1}V$, $V_1 = c_2^{-1}V^{\alpha_1}$, and $V_2 = 1_{\mathbb{X}}$ (note that $1_{\mathbb{X}} = V_2 \leq V_1 \leq V_0$).
- More generally, if $\alpha_1 > 1/2$, then Lemma 5.2 can be repeated recursively: this provides inequalities of the form $PV^{\alpha_{i-1}} \leq V^{\alpha_{i-1}} - c_i\, V^{\alpha_i}$ on S^c with $c_i < 1$ and with

  $$\alpha_i = 2\alpha_{i-1} - \alpha_{i-2} = (\alpha_1 - 1)\, i + 1.$$

  Actually Lemma 5.2 can only be repeated until the value $i = m$ since $\alpha_m = 0$ and $\alpha_i < 0$ for $i > m$. Then Assumptions (55) of Corollary 5.2 hold with

  $$V_0 = (c_1 c_2 \cdots c_m)^{-1}V,\ V_1 = (c_2 \cdots c_m)^{-1}\, V^{\alpha_1}, \ldots, V_{m-1} = c_m^{-1}V^{\alpha_{m-1}},\ V_m = 1_{\mathbb{X}}$$

  (note that $1_{\mathbb{X}} = V_m \leq \cdots \leq V_0$).

Finally, since $V_m = 1_{\mathbb{X}}$, the sequence $(\beta_k(V_m))_{k\geq 1}$ is decreasing from (20). Then the conclusions of Corollary 5.3 follows from Corollary 5.2.

□

Now we consider the general case when P satisfies Assumptions (**S**) and ($\boldsymbol{Sub_{\alpha,S^c}}$). Using Corollary 5.1, we prove below that the procedure in the atomic case (Corollary 5.3) based on the iteration of Lemma 5.2, extends to the non atomic case provided that Condition ($\boldsymbol{Sub_{\alpha,S^c}}$) can be modified thanks to Lemma 5.2 in order to fulfil the third condition in (54). To that effect, assume that PV is bounded on S. Then for every $\varepsilon \in (0, \nu(1_{\mathbb{X}}))$

$$\exists \eta_0 \equiv \eta_0(\varepsilon) \in (0,1],\ \forall \eta \in (0,\eta_0],\ \forall x \in S, \quad (PV^\eta)(x) \leq V(x)^\eta + \nu(V^\eta) - \varepsilon. \tag{61}$$

Indeed we have

$$\forall x \in S, \quad (PV^\eta)(x) - V(x)^\eta - \nu(V^\eta) \leq (\sup_S PV)^\eta - 1 - \nu(1_{\mathbb{X}})$$

from Jensen's inequality and $1_{\mathbb{X}} \leq V^\eta$. Then (61) follows from the following property

$$\exists \eta_0 \in (0,1],\ \forall \eta \in (0,\eta_0], \quad (\sup_S PV)^\eta - 1 \leq \nu(1_{\mathbb{X}}) - \varepsilon$$

which holds since $(\sup_S PV)^\eta \to 1$ when $\eta \to 0$. Next, if $\eta_0 \geq 1-\alpha$, define the positive integer $m \equiv m(\varepsilon, \alpha, \eta_0)$ as follows

$$m := \lfloor \frac{\eta_0}{1-\alpha} \rfloor. \tag{62}$$

Corollary 5.4 *Assume that P satisfies Assumptions (**S**) and ($\boldsymbol{Sub_{\alpha,S^c}}$) for some $S \in \mathcal{X}$, $\nu \in \mathcal{M}_*^+$, $\alpha \in [0,1)$ and for some Lyapunov function V. Moreover assume that V and PV are bounded on S. Let $\varepsilon \in (0, \nu(1_{\mathbb{X}}))$, and assume that the real number η_0 given in (61) is such that $\eta_0 \geq 1-\alpha$. Finally let m be the positive integer defined in (62). Then the estimates (24)-(25) hold true with*

$$\forall n \geq 1, \quad \varepsilon_n \leq \frac{d\, C_m\, \nu(V)}{(m-1)} \frac{1}{n^{m-1}} \quad \textit{with} \quad C_m = 2^{\frac{m(m+1)}{2}-1} \tag{63}$$

where d is a positive constant (specified in the proof). If moreover $\mu(V) := \sum_{k=1}^{+\infty} \beta_k(V) < \infty$, then the estimates (24)-(25) hold true with

$$\forall n \geq 1, \quad \varepsilon_n \leq \frac{d\, D_m\, \mu(V)}{m} \frac{1}{n^m} \quad \textit{with} \quad D_m = 2^{\frac{(m+1)(m+2)}{2}+1}. \tag{64}$$

Proof. Note that the third condition in (54) associated with Assumption ($\boldsymbol{Sub_{\alpha,S^c}}$) may fail, that is the inequality $PV \leq V - c_1 V^\alpha + \nu(V)$ on S may be false. To initialize the procedure, we apply [JR02, Lem. 3.5] from ($\boldsymbol{Sub_{\alpha,S^c}}$) with the exponent η_0 given in (61), that is:

$$\exists c_{\eta_0} > 0,\ \forall x \in S^c, \quad (PV^{\eta_0})(x) \leq V(x)^{\eta_0} - c_{\eta_0}\, V(x)^{\alpha+\eta_0-1}. \tag{65}$$

If $\alpha + \eta_0 - 1 < 0$, then Property (65) cannot be used to apply Corollary 5.1 since the function V_1 in Assumptions (54) must take its value in $[a, +\infty)$ for some $a > 0$. Now assume that $\alpha + \eta_0 - 1 \geq 0$ and prove that the third condition in (54) associated with (65) is satisfied. Let

$M_1 := \sup_S V$ and $M_2 := \sup_S PV$. Recall that $\varepsilon \in (0, \nu(1_{\mathbb{X}}))$ and note that c_{η_0} in (65) can be chosen such that $c_{\eta_0} M_1^{\alpha+\eta_0-1} \le \varepsilon$ (up to reduce the value of c_{η_0}). Then we have

$$\forall x \in S, \quad (PV^{\eta_0})(x) - V(x)^{\eta_0} + c_{\eta_0}\, V(x)^{\alpha+\eta_0-1} - \nu(V^{\eta_0}) \le 0 \tag{66}$$

from (61). Now, starting from (65)-(66), we iterate Lemma 5.2 as many times as possible. Namely, let

$$\widehat{V} := V^{\eta_0} \quad \text{and} \quad \widehat{\alpha}_1 := 1 - \frac{1}{m}$$

with m defined in (62). Note that $m = \lfloor (1-\widehat{\alpha})^{-1} \rfloor$ with $\widehat{\alpha} = 1-(1-\alpha)/\eta_0$, and that $\widehat{\alpha}_1 \le \widehat{\alpha}$. Also set $\widehat{c}_1 = c_{\eta_0}$. Then

$$\begin{cases} \forall x \in S^c, \quad (P\widehat{V})(x) \le \widehat{V}(x) - \widehat{c}_1\, \widehat{V}(x)^{\widehat{\alpha}_1} \\ \forall x \in S, \quad (P\widehat{V})(x) \le \widehat{V}(x) - \widehat{c}_1\, \widehat{V}(x)^{\widehat{\alpha}_1} + \nu(\widehat{V}) \end{cases} \tag{67}$$

from (65)-(66) and $\widehat{\alpha}_1 \le \widehat{\alpha}$. Then, starting from (67) and iterating Lemma 5.2, we can proceed exactly as in the proof of Corollary 5.3, provided that the third condition in (54) holds at each step (this was automatically fulfilled in the atomic case). More precisely, at each step, Lemma 5.2 provides an inequality of the form

$$P\widehat{V}^{\widehat{\alpha}_{i-1}} \le \widehat{V}^{\widehat{\alpha}_{i-1}} - \widehat{c}_i\, \widehat{V}^{\widehat{\alpha}_i} \quad \text{on } S^c \tag{68}$$

with some $\widehat{c}_i > 0$ and with

$$\widehat{\alpha}_i = 2\widehat{\alpha}_{i-1} - \widehat{\alpha}_{i-2} = (\widehat{\alpha}_1 - 1)\, i + 1.$$

This procedure can be repeated only until the value $i = m$ since $\widehat{\alpha}_m = 0$ and $\widehat{\alpha}_i < 0$ for $i > m$, but we have moreover to check that the third condition in (54) associated with (68) holds. To verify this last point, note that $\widehat{\alpha}_{i-1} \le 1$ and that

$$P\widehat{V}^{\widehat{\alpha}_{i-1}} - \widehat{V}^{\widehat{\alpha}_{i-1}} - \nu(\widehat{V}^{\widehat{\alpha}_{i-1}}) = PV^{\eta_i} - V^{\eta_i} - \nu(V^{\eta_i}) \quad \text{with} \quad \eta_i := \eta_0 \widehat{\alpha}_{i-1} \in (0, \eta_0]$$

from $\widehat{V} := V^{\eta_0}$. It then follows from (61) and $\widehat{V}^{\widehat{\alpha}_{i-1}} \ge 1_{\mathbb{X}}$ that

$$\forall x \in S, \quad (P\widehat{V}^{\widehat{\alpha}_{i-1}})(x) - \widehat{V}^{\widehat{\alpha}_{i-1}}(x) + \widehat{c}_i\, \widehat{V}^{\widehat{\alpha}_i}(x) - \nu(\widehat{V}^{\widehat{\alpha}_{i-1}}) \le \widehat{c}_i\, V^{\eta_0}(x) - \varepsilon \le 0$$

since $\widehat{c}_i$ in (68) can be chosen such that $\widehat{c}_i M_1^{\eta_0} \le \varepsilon$ (recall that $M_1 := \sup_S V$). Then Assumptions (54) of Corollary 5.1 hold with

$$V_0 = (\widehat{c}_1 \widehat{c}_2 \cdots \widehat{c}_m)^{-1} \widehat{V},\ V_1 = (\widehat{c}_2 \cdots \widehat{c}_m)^{-1}\, \widehat{V}^{\alpha_1}, \ldots, V_{m-1} = \widehat{c}_m^{-1} \widehat{V}^{\alpha_{m-1}},\ V_m = 1_{\mathbb{X}}$$

(note that $1_{\mathbb{X}} = V_m \le \cdots \le V_0$). Finally, since $V_m = 1_{\mathbb{X}}$, the sequence $(\beta_k(V_m))_{k \ge 1}$ is decreasing from (20), so that (63) and (64) follows from Corollary 5.1. □

Remark 5.2 *In practice, for the choice of $\varepsilon \in (0, \nu(1_{\mathbb{X}}))$ in Corollary 5.4, a trade-off must be made with respect to Condition (61) versus the resulting positive constant $c\nu(V)$ and $d\mu(V)$ in (63)-(64). Indeed, the smaller ε is, the larger η_0 in (61) will be, so the larger m in (62) will be. However, the smaller ε is, the larger $(\widehat{c}_1 \widehat{c}_2 \cdots \widehat{c}_m)^{-1}$ will be in the above definition of V_0, so that the larger constants $c\nu(V) = \nu(V_0)$ and $d\mu(V) = \mu(V_0)$ in (63)-(64) will be.*

We conclude this section by presenting a result on the approximation of P^n by the submarovian kernel T_n given in (12) under the subgeometric drift conditions (39).

Theorem 5.3 *Assume that P satisfies the assumptions of Theorem 5.1 with some $m \geq 1$. Then*

$$\forall x \in \mathbb{X}, \quad \sum_{k=0}^{+\infty} (k+1)^{m-1} \left| (P^k f)(x) - (T_k f)(x) \right| \leq \vartheta_{m-1} V_0(x) \tag{69}$$

with ϑ_{m-1} defined in (41).

Proof. If $m = 1$, then (12) and the positivity of $P - T$ give (see the proof of Theorem 5.1)

$$\sum_{k=0}^{+\infty} \left| (P^k f)(x) - (T_k f)(x) \right| \leq \sum_{k=0}^{+\infty} \big((P-T)^k |f| \big)(x) \leq \sum_{k=0}^{+\infty} \big((P-T)^k V_1 \big)(x) \leq V_0(x).$$

This proves (69) for $m = 1$. Inequality (69) for $m \geq 1$ easily follows by induction from the following fact: if $(P - T)V_m \leq V_m - V_{m+1}$, then we have

$$\forall k \geq 0, \quad (P-T)^k V_{m+1} \leq (P-T)^k V_m - (P-T)^{k+1} V_m,$$

from which we deduce that for every measurable function $f : \mathbb{X} \to \mathbb{R}$ such that $|f| \leq V_{m+1}$ (see the proof of Theorem 5.1)

$$\begin{aligned}
\sum_{k=0}^{+\infty} (k+1)^m \left| P^k f - T_k f \right| &\leq \sum_{k=0}^{+\infty} (k+1)^m (P-T)^k V_{m+1} \\
&\leq \sum_{j=0}^{m-1} C_m^j \sum_{k=0}^{+\infty} k^j (P-T)^k V_m \\
&\leq \sum_{j=0}^{m-1} C_m^j \sum_{k=0}^{+\infty} (k+1)^j (P-T)^k V_{j+1} \\
&\leq \left(\sum_{j=0}^{m-1} C_m^j \vartheta_j \right) V_0 \qquad \text{(from induction hypothesis)} \\
&\leq \vartheta_m V_0.
\end{aligned}$$

□

References

[AF05] C. Andrieu and G. Fort. Explicit control of subgeometric ergodicity. Technical Report 05:17, Univ. Bristol, 2005.

[AFV15] C. Andrieu, G. Fort, and M. Vihola. Quantitative convergence rates for subgeometric Markov chains. *J. Appl. Probab.*, 52(2):391–404, 2015.

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

[Del17] B. Delyon. Convergence rate of the powers of an operator. Applications to stochastic systems. *Bernoulli*, 23(4A):2129–2180, 2017.

[DFMS04] R. Douc, G. Fort, E. Moulines, and P. Soulier. Practical drift conditions for subgeometric rates of convergence. *Ann. Appl. Probab.*, 14(3):1353–1377, 2004.

[DMPS18] R. Douc, E. Moulines, P. Priouret, and P. Soulier. *Markov chains*. Springer Series in Operations Research and Financial Engineering. Springer, 2018.

[Hen06] H. Hennion. Quasi-compactness and absolutely continuous kernels, applications to Markov chains. arXiv:math/0606680, June 2006.

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

[HL14] L. Hervé and J. Ledoux. Spectral analysis of Markov kernels and aplication to the convergence rate of discrete random walks. *Adv. in Appl. Probab.*, 46(4):1036–1058, 2014.

[HL20a] L. Hervé and J. Ledoux. Additional material on V-geometrical ergodicity of Markov kernels via finite-rank approximations. *hal-02410491v4*, 2020.

[HL20b] L. Hervé and J. Ledoux. V-geometrical ergodicity of Markov kernels via finite-rank approximations. *Electronic Communications in Probab.*, 25(23):1–12, March 2020.

[HM11] M. Hairer and J. C. Mattingly. Yet another look at Harris' ergodic theorem for Markov chains. In R. Dalang, M. Dozzi, and F. Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications VI*, pages 109–117, Basel, 2011. Springer Basel.

[JR02] S. F. Jarner and G. O. Roberts. Polynomial convergence rates of Markov chains. *Ann. Appl. Probab.*, 12(1):224–247, 2002.

[LC14] M. E. Lladser and S. R. Chestnut. Approximation of sojourn-times via maximal couplings: motif frequency distributions. *J. Math. Biol.*, 69(1):147–182, 2014.

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

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

[MT09] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Cambridge Univ. Press, London, 2009.

[Num84] E. Nummelin. *General irreducible Markov chains and nonnegative operators*. Cambridge University Press, Cambridge, 1984.

[Ros02] J. S. Rosenthal. Quantitative convergence rates of Markov chains: a simple account. *Electron. Comm. Probab.*, 7:123–128, 2002.

[RR04] G. O. Roberts and J. S. Rosenthal. General state space Markov chains and MCMC algorithms. *Probab. Surv.*, 1:20–71 (electronic), 2004.

[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.

[RT00] G. O. Roberts and R. L. Tweedie. Rates of convergence of stochastically monotone and continuous time Markov models. *J. Appl. Probab.*, 37(2):359–373, 2000.