

# Approximating Markov chains and $V$ -geometric ergodicity via weak perturbation theory

Loïc HERVÉ and James LEDOUX \*

September 20, 2013

## Abstract

Let  $P$  be a Markov kernel on a measurable space  $\mathbb{X}$  and let  $V : \mathbb{X} \rightarrow [1, +\infty)$ . This paper provides explicit connections between the  $V$ -geometric ergodicity of  $P$  and that of finite-rank nonnegative sub-Markov kernels  $\widehat{P}_k$  approximating  $P$ . A special attention is paid to obtain an efficient way to specify the convergence rate for  $P$  from that of  $\widehat{P}_k$  and conversely. Furthermore, explicit bounds are obtained for the total variation distance between the  $P$ -invariant probability measure and the  $\widehat{P}_k$ -invariant positive measure. The proofs are based on the Keller-Liverani perturbation theorem which requires an accurate control of the essential spectral radius of  $P$  on usual weighted supremum spaces. Such computable bounds are derived in terms of standard drift conditions. Our spectral procedure to estimate both the convergence rate and the invariant probability measure of  $P$  is applied to truncation of discrete Markov kernels on  $\mathbb{X} := \mathbb{N}$ .

AMS subject classification : 60J10; 47B07

Keywords : rate of convergence, essential spectral radius, drift condition, quasi-compactness, truncation of discrete kernels.

## 1 Introduction

Throughout the paper  $P$  is a Markov kernel on a measurable space  $(\mathbb{X}, \mathcal{X})$ , and  $\{\widehat{P}_k\}_{k \geq 1}$  is a sequence of nonnegative sub-Markov kernels on  $(\mathbb{X}, \mathcal{X})$ . For any positive measure  $\mu$  on  $\mathbb{X}$  and any  $\mu$ -integrable function  $f : \mathbb{X} \rightarrow \mathbb{C}$ ,  $\mu(f)$  denotes the integral  $\int f d\mu$ . Let  $V : \mathbb{X} \rightarrow [1, +\infty)$  be a measurable function such that  $V(x_0) = 1$  for some  $x_0 \in \mathbb{X}$ . Let  $(\mathcal{B}_1, \|\cdot\|_1)$  denote the weighted-supremum Banach space

$$\mathcal{B}_1 := \left\{ f : \mathbb{X} \rightarrow \mathbb{C}, \text{ measurable} : \|f\|_1 := \sup_{x \in \mathbb{X}} |f(x)|V(x)^{-1} < \infty \right\}.$$

---

\*INSA de Rennes, F-35708, France & IRMAR CNRS-UMR 6625, Rennes, F-35000, France; Université Européenne de Bretagne, France. Loïc.Herve@insa-rennes.fr, James.Ledoux@insa-rennes.fr

In the sequel,  $PV/V$  and each  $\widehat{P}_k V/V$  are assumed to be bounded on  $\mathbb{X}$ , so that both  $P$  and  $\widehat{P}_k$  are bounded linear operators on  $\mathcal{B}_1$ . Moreover  $P$  (resp. every  $\widehat{P}_k$ ) is assumed to have an invariant probability measure  $\pi$  (resp. an invariant bounded positive measure  $\widehat{\pi}_k$ ) on  $(\mathbb{X}, \mathcal{X})$  such that  $\pi(V) < \infty$  (resp.  $\widehat{\pi}_k(V) < \infty$ ). We suppose that there exists a bounded measurable function  $\phi_k : \mathbb{X} \rightarrow [0, +\infty)$  such that  $\widehat{P}_k \phi_k = \phi_k$  and  $\widehat{\pi}_k(\phi_k) = 1$ . Throughout the paper, every  $\widehat{P}_k$  is of finite rank on  $\mathcal{B}_1$  and the sequence  $\{\widehat{P}_k\}_{k \geq 1}$  converges to  $P$  in the weak sense ( $C_{0,1}$ ) below. Finally we assume that

$$\lim_{k \rightarrow +\infty} \pi(\phi_k) = 1 \quad \text{and} \quad \lim_{k \rightarrow +\infty} \widehat{\pi}_k(1_{\mathbb{X}}) = 1. \quad (1)$$

These two conditions are necessary for the sequence  $\{\widehat{\pi}_k\}_{k \geq 1}$  to converge to  $\pi$  in total variation distance since  $\pi(\phi_k) - 1 = (\pi - \widehat{\pi}_k)(\phi_k)$  and  $\widehat{\pi}_k(1_{\mathbb{X}}) - 1 = (\widehat{\pi}_k - \pi)(1_{\mathbb{X}})$ .

When  $P$  is a Markov kernel on  $\mathbb{X} := \mathbb{N}$ , a typical instance of sequence  $\{\widehat{P}_k\}_{k \geq 1}$  is obtained by considering the extended sub-Markov kernel  $\widehat{P}_k$  derived from the linear augmentation (in the last column) of the  $(k+1) \times (k+1)$  northwest corner truncation  $P_k$  of  $P$  (e.g. see [33]). Then  $\widehat{\pi}_k$  is the (extended) probability measure on  $\mathbb{N}$  derived from the  $P_k$ -invariant probability measure and  $\phi_k = 1_{B_k}$  with  $B_k := \{0, \dots, k\}$ .

In this work, the connection between the  $V$ -geometrical ergodicity of  $P$  and that of  $\widehat{P}_k$  is investigated. These properties are defined as follows.

*$P$  (resp.  $\widehat{P}_k$ ) is said to be  $V$ -geometrically ergodic if there exist some rate  $\rho \in (0, 1)$  (resp.  $\rho_k \in (0, 1)$ ) and constant  $C > 0$  (resp.  $C_k > 0$ ) such that*

$$\forall n \geq 0, \quad \sup_{f \in \mathcal{B}_1, \|f\|_1 \leq 1} \|P^n f - \pi(f)1_{\mathbb{X}}\|_1 \leq C \rho^n \quad (V)$$

$$\text{respectively : } \forall n \geq 0, \quad \sup_{f \in \mathcal{B}_1, \|f\|_1 \leq 1} \|\widehat{P}_k^n f - \widehat{\pi}_k(f)\phi_k\|_1 \leq C_k \rho_k^n. \quad (V_k)$$

Specifically, the two following issues are studied.

(Q1) Suppose that  $P$  is  $V$ -geometrically ergodic.

- (a) Is  $\widehat{P}_k$  a  $V$ -geometrically ergodic kernel for  $k$  large enough?
- (b) When  $(\rho, C)$  is known in (V), can we deduce explicit  $(\rho_k, C_k)$  in  $(V_k)$  from  $(\rho, C)$ ?
- (c) Does the total variation distance  $\|\widehat{\pi}_k - \pi\|_{TV}$  go to 0 when  $k \rightarrow +\infty$ , and can we obtain an explicit bound for  $\|\widehat{\pi}_k - \pi\|_{TV}$  when  $(\rho, C)$  is known in (V)?

(Q2) Suppose that  $\widehat{P}_k$  is  $V$ -geometrically ergodic for every  $k$ .

- (a) Is  $P$  a  $V$ -geometrically ergodic kernel?
- (b) When  $(\rho_k, C_k)$  is known in  $(V_k)$ , can we deduce explicit  $(\rho, C)$  in (V) from  $(\rho_k, C_k)$ , and consequently obtain a bound for  $\|\widehat{\pi}_k - \pi\|_{TV}$  using the last part of (Q1(c))?

A natural way to solve (Q1) is to see  $\widehat{P}_k$  as a perturbed operator of  $P$ , and vice versa for (Q2). The standard perturbation theory requires that  $\{\widehat{P}_k\}_{k \geq 1}$  converges to  $P$  in operator

norm on  $\mathcal{B}_1$ . Unfortunately this condition may be very restrictive even if  $\mathbb{X}$  is discrete (e.g. see [32, 10]): for instance, in our application to truncation of discrete kernels in Section 6, this condition never holds (see Remark 6.2). Here we use the weak perturbation theory due to Keller and Liverani [16, 21] (see also [6]) which invokes the weakened convergence property

$$\|\widehat{P}_k - P\|_{0,1} := \sup_{f \in \mathcal{B}_0, \|f\|_0 \leq 1} \|\widehat{P}_k f - P f\|_1 \xrightarrow[k \rightarrow +\infty]{} 0, \quad (C_{0,1})$$

where  $\mathcal{B}_0$  is the Banach space of bounded measurable  $\mathbb{C}$ -valued functions on  $\mathbb{X}$  equipped with its usual norm  $\|f\|_0 := \sup_{x \in \mathbb{X}} |f(x)|$ . In the truncation context of Section 6 where  $\mathbb{X} := \mathbb{N}$ , Condition  $(C_{0,1})$  holds provided that  $\lim_k V(k) = +\infty$ . The price to pay for using  $(C_{0,1})$  is that two functional assumptions are needed. The first one involves the Doeblin-Fortet inequalities: such dual inequalities can be derived for  $P$  and every  $\widehat{P}_k$  from Condition (UWD) below. The second one requires to have an accurate bound of the essential spectral radius  $r_{ess}(P)$  of  $P$  acting on  $\mathcal{B}_1$ .

Issues (Q1) and (Q2) are solved in Sections 3 and 4. The question in (Q1)-(Q2) concerning  $\|\widehat{\pi}_k - \pi\|_{TV}$  is solved by arguments inspired by [21, Lem. 7.1] (see Proposition 2.1). The other questions in both (Q1)-(Q2) are addressed using [21]. Theorems 3.1 and 4.1 provide a positive and explicit answer to the issues (Q1) and (Q2) under Conditions  $(V)$ ,  $(V_k)$ ,  $(C_{0,1})$  and the following uniform weak drift condition:

$$\exists \delta \in (0, 1), \exists L > 0, \forall k \in \mathbb{N}^* \cup \{\infty\}, \quad \widehat{P}_k V \leq \delta V + L 1_{\mathbb{X}} \quad (\text{UWD})$$

where by convention  $\widehat{P}_\infty := P$ . The notion of essential spectral radius and the material on quasi-compactness used for bounding  $r_{ess}(P)$  are reported in Section 5. Our results are applied in Section 6 to truncation of discrete Markov kernels. Some useful theoretical complements on the weak perturbation theory are postponed to Section 7.

The estimates of Theorems 3.1 and 4.1 are all the more precise that the real number  $\delta$  in (UWD) and the bound for  $r_{ess}(P)$  are accurate. The weak drift inequality used in (UWD) is a simple and well-known condition introduced in [23] for studying  $V$ -geometric ergodicity. Managing  $r_{ess}(P)$  is more delicate, even in discrete case. To that effect, it is important not to confuse  $r_{ess}(P)$  with  $\rho$  in  $(V)$ . Actually Inequality  $(V)$  gives<sup>1</sup>  $r_{ess}(P) \leq \rho$ , but this bound may be very inaccurate since there exist Markov kernels  $P$  such that  $\rho - r_{ess}(P)$  is close to one (see Definition 5.1). Moreover note that no precise rate  $\rho$  in  $(V)$  is known a priori in Issue (Q2). Accordingly, our study requires to obtain a specific control of the essential spectral radius  $r_{ess}(P)$  of a general Markov kernel  $P$  acting on  $\mathcal{B}_1$ . This study, which has its own interest, is presented in Section 5 where the two following results are presented.

- (i) *If  $P$  satisfies some classical drift/minorization conditions, then  $r_{ess}(P)$  can be bounded in terms of the constants of these drift/minorization conditions (see Theorem 5.2).*
- (ii) *If  $PV \leq \delta V + L 1_{\mathbb{X}}$  for some  $\delta \in (0, 1)$  and  $L > 0$ , and if  $P : \mathcal{B}_0 \rightarrow \mathcal{B}_1$  is compact, then  $r_{ess}(P) \leq \delta$  (see Proposition 5.4).*

To the best of our knowledge, the general result (i) is not known in the literature. Result (ii) is a simplified version of [34, Th. 3.11]. For convenience we present a direct and short proof

<sup>1</sup>(See e.g. [17] or apply Definition 5.1 with  $H := \{f \in \mathcal{B}_1 : \pi(f) = 0\}$ .)

using [12, Cor. 1]. The bound  $r_{ess}(P) \leq \delta$  obtained in (ii) is more precise than that obtained in (i) (and sometimes is optimal). The compactness property in (ii) must not be confused with that of  $P : \mathcal{B}_1 \rightarrow \mathcal{B}_1$  or  $P : \mathcal{B}_0 \rightarrow \mathcal{B}_0$ , which are much stronger conditions. If  $P$  is an infinite matrix ( $\mathbb{X} := \mathbb{N}$ ), then  $P : \mathcal{B}_0 \rightarrow \mathcal{B}_1$  is compact when  $\lim_k V(k) = +\infty$ , while in general  $P$  is compact neither on  $\mathcal{B}_1$  nor on  $\mathcal{B}_0$ .

Let us give a brief review of previous works related to Issues (Q1) or (Q2). Various probabilistic methods have been developed to derive explicit rate and constant  $(\rho, C)$  in Inequality (V) from the constants of drift conditions (see [24, 22, 7] and the references therein). To the best of our knowledge, these methods, which are not concerned with approximation issues, provide a computable rate  $\rho$  which is often unsatisfactory, except for reversible or stochastically monotone  $P$ . Convergence of  $\{\hat{\pi}_k\}_{k \geq 1}$  to  $\pi$  has been studied in [33] for truncation approximations of  $V$ -geometrically ergodic Markov kernels with discrete  $\mathbb{X}$ . Specifically it is proved that  $\sup_{|f| \leq V} |\hat{\pi}_k(f) - \pi(f)|$  goes to 0 when  $k \rightarrow +\infty$ . In particular  $\|\hat{\pi}_k - \pi\|_{TV}$  goes to 0. In the special case when  $P$  is stochastically monotone, the following rate of convergence is obtained [33, Th. 4.2,(46)]

$$\forall k \in \mathbb{N}^*, \quad \|\hat{\pi}_k - \pi\|_{TV} = O\left(\frac{\ln V(k)}{V(k)}\right). \quad (2)$$

In this estimate, explicit constants only depending on constants involved in drift/minorization conditions are also provided. A similar result has been obtained for polynomially ergodic Markov chains in [20]. A bound of type (2) is provided by Theorems 3.1 and 4.1 under Conditions (V),  $(V_k)$ ,  $(C_{0,1})$  and (UWD). This result is new even for discrete truncation approximation since the stochastic monotonicity is not required in our work. Questions related to Issue (Q2) have been investigated in [29] for discrete reversible Markov kernels. The main result in [29, Cor. 3] is that  $\|\mu P^n - \pi\|_{TV} \leq c\beta^n$  for some explicit positive constant  $c$ , provided that  $\liminf \beta_k \leq \beta$ , where  $\beta_k$  denotes the so-called second eigenvalue of the  $k$ -th truncated kernels augmented on the diagonal. However, except in special cases (see [29]), finding a nontrivial upper bound  $\beta$  of  $\liminf \beta_k$  is a difficult problem. Note that our Theorem 4.1 only needs to compute the second eigenvalue of  $\hat{P}_k$  for some suitable  $k$ . Finally mention that perturbations bounds have been obtained in [25] by a different way for uniformly ergodic Markov chains (i.e.  $V \equiv 1$  in (V)).

The weak perturbation results of [16, 21] have been fully used in the framework of dynamical systems (e.g. see [3, 9, 4, 5, 36]). There, Markov kernels and their invariant probability measure are replaced by Perron-Frobenius operators and their so-called SRB measure. In the context of  $V$ -geometrically ergodic Markov kernels, the Keller-Liverani theorem has been already used in [10] to study general perturbation issues. When applied to our context, [10, Th. 1] gives a positive answer to (Q1)(a) and the convergence of  $\|\hat{\pi}_k - \pi\|_{TV}$  to 0. The others questions in (Q1)-(Q2) are not addressed in [10]. In the present work, the duality arguments introduced in [10] are applied together with the explicit bounds of [21].

## 2 Notations and preliminary results

Let us first introduce notations used in the paper. Let  $[\cdot]$  denote the integer part function. We denote by  $(\mathcal{L}(\mathcal{B}_0, \mathcal{B}_1), \|\cdot\|_{0,1})$  the space of all the bounded linear maps from  $\mathcal{B}_0$  to  $\mathcal{B}_1$ , equipped with its usual norm:

$$\|T\|_{0,1} = \sup \{ \|Tf\|_1, f \in \mathcal{B}_0, \|f\|_0 \leq 1 \}.$$

We write  $\mathcal{L}(\mathcal{B}_1)$  for  $\mathcal{L}(\mathcal{B}_1, \mathcal{B}_1)$  and  $\|T\|_1$  for  $\|T\|_{1,1}$ . Let  $(\mathcal{B}'_1, \|\cdot\|_1)$  be the dual space of  $\mathcal{B}_1$ . Note that we make a slight abuse of notation in writing again  $\|\cdot\|_1$  for the operator and dual norms on  $\mathcal{B}_1$ . Recall that the total variation distance between  $\hat{\pi}_k$  and  $\pi$  is defined by

$$\|\hat{\pi}_k - \pi\|_{TV} := \sup_{\|f\|_0 \leq 1} |\hat{\pi}_k(f) - \pi(f)|.$$

The next proposition is relevant to estimate  $\|\hat{\pi}_k - \pi\|_{TV}$ .

**Proposition 2.1** *Assume that Condition (UWD) holds. Set  $\Delta_k := \|\hat{P}_k - P\|_{0,1}$  and*

$$A := 1 + \frac{L}{1 - \delta}. \quad (3)$$

(a) *If, for some  $k \geq 1$ ,  $\hat{P}_k$  is  $V$ -geometrically ergodic with rate and constant  $(\rho_k, C_k)$  in  $(V_k)$ , then*

$$\|\hat{\pi}_k - \pi\|_{TV} \leq \hat{\pi}_k(1_{\mathbb{X}}) |1 - \pi(\phi_k)| + \frac{L}{1 - \delta} \left( \frac{2C_k}{\rho_k} + \frac{A}{\ln(\rho_k^{-1})} |\ln \Delta_k| \right) \Delta_k. \quad (4)$$

(b) *If  $P$  is  $V$ -geometrically ergodic with rate and constant  $(\rho, C)$  in  $(V)$ , then*

$$\forall n \geq 1, \quad \|\hat{\pi}_n - \pi\|_{TV} \leq |1 - \hat{\pi}_n(1_{\mathbb{X}})| + \frac{L}{1 - \delta} \left( \frac{2C}{\rho} + \frac{A}{\ln \rho^{-1}} |\ln \Delta_n| \right) \Delta_n. \quad (5)$$

From Assumption (1), we know that the first term in the right-hand side of both (4) and (5) converges to 0. When applied to truncation of discrete kernels in Section 6, the first term in the right-hand side of (4) (resp. of (5)) is  $O(\Delta_k)$  (resp. zero).

Inequality (4) is interesting since rate and constant  $(\rho_k, C_k)$  in  $(V_k)$  are expected to be computable. However further properties on the  $\rho_k$ 's and the  $C_k$ 's are required to deduce  $\lim_k \|\hat{\pi}_k - \pi\|_{TV} = 0$  from (4). They are derived from  $(V)$  in Theorem 3.1. By contrast, Inequality (5) gives  $\lim_n \|\hat{\pi}_n - \pi\|_{TV} = 0$  when  $\|\hat{P}_n - P\|_{0,1} \rightarrow 0$ . But the bound  $\|\hat{\pi}_n - \pi\|_{TV} = O(|\ln \Delta_n| \Delta_n)$  provided by (5) is only computable when some explicit rate and constant  $(\rho, C)$  are known in  $(V)$ . Such explicit  $(\rho, C)$  is derived from  $(V_k)$  in Theorem 4.1.

*Proof of Proposition 2.1.* In a first step, we prove that we have with  $A_n := \max_{0 \leq j \leq n-1} \|P^j\|_1$

$$\|\hat{\pi}_k - \pi\|_{TV} \leq \hat{\pi}_k(1_{\mathbb{X}}) |1 - \pi(\phi_k)| + \frac{L}{1 - \delta} (2C_k \rho_k^n + n A_n \Delta_k). \quad (6)$$

Observe that, for any  $n \geq 0$ , we can write

$$\|P^n - \widehat{P}_k^n\|_{0,1} \leq nA_n \|P - \widehat{P}_k\|_{0,1}.$$

This inequality follows from  $\|\widehat{P}_k\|_0 \leq 1$  and an easy induction based on

$$P^n - \widehat{P}_k^n = P^{n-1}(P - \widehat{P}_k) + (P^{n-1} - \widehat{P}_k^{n-1})\widehat{P}_k.$$

Using the triangle inequality, we obtain for any  $f \in \mathcal{B}_0$  such that  $\|f\|_0 \leq 1$

$$\begin{aligned} |\widehat{\pi}_k(f) - \pi(f)| &= |\widehat{\pi}_k(\widehat{P}_k^n f) - \pi(P^n f)| \\ &\leq |(\widehat{\pi}_k - \pi)(\widehat{P}_k^n f)| + |\pi(\widehat{P}_k^n f - P^n f)| \\ &\leq \widehat{\pi}_k(|f|) |(\widehat{\pi}_k - \pi)(\phi_k)| + |(\widehat{\pi}_k - \pi)(\widehat{P}_k^n f - \widehat{\pi}_k(f)\phi_k)| + \|\pi\|_1 \|\widehat{P}_k^n - P^n\|_{0,1} \\ &\leq \widehat{\pi}_k(1_{\mathbb{X}}) |\pi(\phi_k) - 1| + \|\widehat{\pi}_k - \pi\|_1 \|\widehat{P}_k^n f - \widehat{\pi}_k(f)\phi_k\|_1 + \pi(V) nA_n \Delta_k \\ &\leq \widehat{\pi}_k(1_{\mathbb{X}}) |\pi(\phi_k) - 1| + (\widehat{\pi}_k(V) + \pi(V)) C_k \rho_k^n \|f\|_1 + \pi(V) nA_n \Delta_k. \end{aligned}$$

Note that

$$\max(\pi(V), \widehat{\pi}_k(V)) \leq L/(1 - \delta) \quad (7)$$

from (UWD) since  $\pi$  (resp.  $\widehat{\pi}_k$ ) is  $P$ -invariant (resp.  $\widehat{P}_k$ -invariant). Inequality (6) then follows from  $\|f\|_1 \leq \|f\|_0$ . Now observe that  $A_n \leq A$  from  $P^j V \leq \delta^j V + L \sum_{i=0}^{j-1} \delta^i \leq AV$ . Setting  $n := \max(0, \lfloor (\ln \rho_k)^{-1} \ln \Delta_k \rfloor)$  in (6) allows us to derive Inequality (4).

The proof of (5) is similar by exchanging the role of  $P$  and  $\widehat{P}_k$ .  $\square$

Now we prove that Condition (UWD) of Introduction provides some uniform dual Doebelin-Fortet inequalities for the family  $\{P, \widehat{P}_k, k \geq 1\}$ . For any  $Q \in \mathcal{L}(\mathcal{B}_1)$ , we denote by  $Q'$  its adjoint operator. Define the following auxiliary semi-norm on  $\mathcal{B}'_1$ :

$$\forall f' \in \mathcal{B}'_1, \quad \|f'\|_0 := \sup \{|f'(f)|, f \in \mathcal{B}_0, \|f\|_0 \leq 1\}. \quad (8)$$

**Lemma 2.2** *Let  $Q$  be any non-negative linear operator on the space  $\mathcal{B}_1$  satisfying*

$$\exists \delta \in (0, 1), \exists L > 0, \quad QV \leq \delta V + L 1_{\mathbb{X}}.$$

*Then:  $\forall f' \in \mathcal{B}'_1, \|Q'f'\|_1 \leq \delta \|f'\|_1 + L \|f'\|_0$ .*

For a Markov kernel  $Q$ , the proof of this lemma is given in [10]. Since only the non-negativity of the operator  $Q$  plays a role in this proof, the details are omitted. From Lemma 2.2 we obtain the following statement.

**Lemma 2.3** *If Condition (UWD) holds with parameters  $(\delta, L)$ , then the kernels  $Q := P$  and  $Q := \widehat{P}_k$  for  $k \geq 1$  satisfy the following uniform Doebelin-Fortet inequality on  $\mathcal{B}'_1$ :*

$$\forall f' \in \mathcal{B}'_1, \quad \|Q'f'\|_1 \leq \delta \|f'\|_1 + L \|f'\|_0.$$

### 3 From $P$ to $\widehat{P}_k$ : solution to (Q1)

For any  $(a, \theta) \in \mathbb{C} \times (0, +\infty)$ , let us define  $D(a, \theta) := \{z \in \mathbb{C} : |z - a| < \theta\}$  and  $\overline{D}(a, \theta) := \{z \in \mathbb{C} : |z - a| \leq \theta\}$ . For any  $T \in \mathcal{L}(\mathcal{B}_1)$ , the spectrum of  $T$  is denoted by  $\sigma(T)$ , and for any  $(r, \vartheta) \in (0, 1)^2$ , we introduce the following subsets of the complex plan

$$\mathcal{V}(r, \vartheta, T) := \{z \in \mathbb{C} : |z| \leq r \text{ or } d(z, \sigma(T)) \leq \vartheta\} \quad \text{and} \quad \mathcal{V}(r, \vartheta, T)^c := \mathbb{C} \setminus \mathcal{V}(r, \vartheta, T),$$

where  $d(z, \sigma(T)) := \inf\{|z - \lambda|, \lambda \in \sigma(T)\}$ . Below  $P$  is assumed to be  $V$ -geometrically ergodic. In particular  $P$  is quasi-compact on  $\mathcal{B}_1$ , that is:  $r_{ess}(P) < 1$  (see Section 5). Under the additional Condition (UWD), we set

$$\hat{\alpha} := \max(r_{ess}(P), \delta). \quad (9)$$

Lemma 2.3 ensures that  $Q := P$  and  $Q := \widehat{P}_k$  for  $k \geq 1$  satisfy the following Doeblin-Fortet inequalities on  $\mathcal{B}'_1$  (use the fact that  $P$  and  $\widehat{P}_k$  are contraction on  $\mathcal{B}_0$ ):

$$\forall n \in \mathbb{N}^*, \forall f' \in \mathcal{B}'_1, \quad \|Q^n f'\|_1 \leq \hat{\alpha}^n \|f'\|_1 + B \|f'\|_0 \quad \text{with} \quad B := \frac{L}{1 - \hat{\alpha}}. \quad (10)$$

For any  $r > \hat{\alpha}$  and  $\vartheta > 0$ , we set:

$$H \equiv H(r, \vartheta, P) := \sup_{z \in \mathcal{V}(r, \vartheta, P)^c} \|(zI - P)^{-1}\|_1 \quad (11a)$$

$$n_1 \equiv n_1(r) := \left\lfloor \frac{\ln 2}{\ln(r/\hat{\alpha})} \right\rfloor + 1 \quad n_2 \equiv n_2(r, \vartheta, P) := \left\lfloor \frac{\ln(8B(B+3)r^{-n_1}H)}{\ln(r/\hat{\alpha})} \right\rfloor + 1$$

$$\varepsilon_1 \equiv \varepsilon_1(r, \vartheta, P) := \frac{r^{n_1+n_2}}{8B(HB + (1-r)^{-1})}. \quad (11b)$$

**Theorem 3.1** *Let  $P$  be a  $V$ -geometrically ergodic Markov kernel with rate  $\rho \in (0, 1)$  in  $(V)$ . Assume that Conditions  $(C_{0,1})$  and (UWD) hold. Let  $(r, \vartheta) \in (0, 1)^2$  be such that*

$$\max(\hat{\alpha}, \rho) + \vartheta < r < 1 - \vartheta. \quad (12)$$

*Then, for any  $k \in \mathbb{N}^*$  such that the convergence condition  $(C_{0,1})$  holds at rate  $\varepsilon_1$  that is*

$$\|\widehat{P}_k - P\|_{0,1} \leq \varepsilon_1, \quad (\mathcal{E}_{0,1})$$

*and such that  $(V_k)$  holds with some rate  $\rho_k$  satisfying  $\rho_k < 1 - \vartheta$ , the following inequalities are valid.*

1. *Inequality  $(V_k)$  holds with rate  $r$ , more precisely*

$$\forall n \geq 0, \quad \|\widehat{P}_k^n - \widehat{\pi}_k(\cdot)\phi_k\|_1 \leq c r^{n+1} \quad \text{with} \quad c \equiv c(r, \vartheta, P) := \frac{4(B+1)}{r^{n_1}(1-r)} + \frac{1}{2\varepsilon_1}. \quad (13)$$

2. *We have, with  $A$  defined in (3) and  $\Delta_k := \|\widehat{P}_k - P\|_{0,1}$ ,*

$$\|\widehat{\pi}_k - \pi\|_{TV} \leq \widehat{\pi}_k(1_{\mathbb{X}}) |1 - \pi(\phi_k)| + \frac{L}{1-\delta} \left( 2c + \frac{A}{\ln r^{-1}} |\ln \Delta_k| \right) \Delta_k. \quad (14)$$

Under Conditions (V), (C<sub>0,1</sub>) and (UWD), the conclusions (13)-(14) hold true for  $k$  large enough, so that Issue (Q1) is solved. Indeed, the fact that Condition (E<sub>0,1</sub>) is fulfilled for  $k$  large enough follows from (C<sub>0,1</sub>). That (V<sub>k</sub>) holds with some rate  $\rho_k < 1 - \vartheta$  for  $k$  large enough is more difficult to establish. This follows from Proposition 7.1 which ensures that a sufficient condition for such a property to hold is that  $k \geq k_0$ , with  $k_0$  defined in (43). However in practice we do not need to use  $k_0$  since Property (V<sub>k</sub>) with  $\rho_k < 1 - \vartheta$  may be fulfilled for  $k \ll k_0$ . This explains why  $k_0$  is not introduced in Theorem 3.1. Actually, under Conditions (V), (C<sub>0,1</sub>) and (UWD), the results [21, Prop. 3.1] and the spectral rank-stability property [21, Cor. 3.1] directly provide Property (13) when Condition (E<sub>0,1</sub>) is replaced with  $\|\widehat{P}_k - P\|_{0,1} \leq \varepsilon_0$  with  $\varepsilon_0 \equiv \varepsilon_0(r, \vartheta, P)$  defined in (42). In this case the assumption (V<sub>k</sub>) with  $\rho_k < 1 - \vartheta$  can be dropped in Theorem 4.1. When  $\varepsilon_0$  is strictly less than  $\varepsilon_1$  given in (11b), using Conditions (E<sub>0,1</sub>) and  $\rho_k < 1 - \vartheta$  invokes smaller  $k$ .

In fact Theorem 3.1 is only relevant when the rate  $\rho$  and the bound  $C$  in (V) are known. That  $\rho$  must be known is necessary to choose  $(r, \vartheta)$  in (12). Constant  $C$  must be known for an effective computation of an upper bound of  $H \equiv H(r, \vartheta, P)$  in (11a) (see Remark 3.2). Note that  $H$  is involved in the definition of  $\varepsilon_1(r, \vartheta, P)$ , thus in the computation of the bound  $c(r, \vartheta, P)$  in (13). Furthermore note that, although Inequality (14) is interesting from a theoretical point of view since it provides  $\lim_k \|\widehat{\pi}_k - \pi\|_{TV} = 0$ , the bound in (14) is only computable when some explicit rate and constant in (V) are known. Another way for obtaining a computable bound for  $\|\widehat{\pi}_k - \pi\|_{TV}$  is presented in Theorem 4.1

**Remark 3.2** *When  $P$  is  $V$ -geometrically ergodic with some rate and constant  $(\rho, C)$  in (V), we have for any  $(r, \vartheta) \in (0, 1)^2$  such that  $\rho + \vartheta < r < 1 - \vartheta$ :*

$$H(r, \vartheta, P) = \sup_{z \in \overline{D}(0, r)^c \cap \overline{D}(1, \vartheta)^c} \|(zI - P)^{-1}\|_1 \leq \frac{\pi(V)}{\vartheta} + \frac{C}{r - \rho} \leq \frac{\pi(V) + C}{\vartheta}.$$

*These properties are derived by writing  $g \in \mathcal{B}_1$  as  $g = \pi(g)1_{\mathbb{X}} + h$  with  $h := g - \pi(g)1_{\mathbb{X}}$ . Using (V) first implies that  $(zI - P)^{-1}$  is well-defined in  $\mathcal{L}(\mathcal{B}_1)$  for any  $z \in \mathbb{C}$  such that  $|z| > \rho$  and  $z \neq 1$ , with:  $(zI - P)^{-1}(g) = \pi(g)(zI - P)^{-1}1_{\mathbb{X}} + (zI - P)^{-1}(h)$ . The first term is  $(\pi(g)/(z - 1))1_{\mathbb{X}}$ . The second one can be bounded in norm  $\|\cdot\|_1$  by using Neumann series. Since  $\pi$  may be unknown, note that  $\pi(V) \leq L/(1 - \delta)$  under (UWD) (see (7)).*

Using Conditions (E<sub>0,1</sub>) and (10) together with the condition  $\rho_k < 1 - \vartheta$  allow us to derive Theorem 3.1 from the first part of [21, Prop. 3.1] as follows.

*Proof.* For any  $T \in \mathcal{L}(\mathcal{B}_1)$ , recall that  $\|T'\|_1 = \|T\|_1$  and  $\sigma(T) = \sigma(T')$ . Given any real numbers  $r > \hat{\alpha}$  and  $\vartheta > 0$ , we use the quantities introduced in (11a)-(11b). Note that  $H$  in (11a) may be equivalently defined with  $P'$  in place of  $P$  since the resolvents of  $P$  and  $P'$  have the same norm in  $\mathcal{L}(\mathcal{B}_1)$  and  $\mathcal{L}(\mathcal{B}'_1)$  respectively and  $\mathcal{V}(r, \vartheta, P) = \mathcal{V}(r, \vartheta, P')$ .

The first assertion of Theorem 3.1 is proved in two steps. In a first step, using (V), (10) and (E<sub>0,1</sub>) for some  $k \in \mathbb{N}^*$ , we show that we have with  $c(r, \vartheta, P)$  given in (13)

$$\sigma(\widehat{P}_k) \subset \mathcal{V}(r, \vartheta, P) \quad \text{and} \quad \sup_{z \in \mathcal{V}(r, \vartheta, P)^c} \|(zI - \widehat{P}_k)^{-1}\|_1 \leq c(r, \vartheta, P). \quad (15)$$



Clearly (15) will hold true if we establish that the following properties are valid:

$$\sigma(\widehat{P}_k') \subset \mathcal{V}(r, \vartheta, P) \quad \text{and} \quad \sup_{z \in \mathcal{V}(r, \vartheta, P)^c} \|(zI - \widehat{P}_k')^{-1}\|_1 \leq c(r, \vartheta, P). \quad (16)$$

In fact we prove that (16), so (15), hold for any  $\vartheta > 0$  and  $r > \hat{\alpha}$ . To that effect, the Keller-Liverani perturbation theorem [16] is applied to the adjoint operators of  $P$  and  $\widehat{P}_k$  acting on  $\mathcal{B}'_1$ . The auxiliary semi-norm  $\|\cdot\|_0$  on  $\mathcal{B}'_1$  has been introduced in (8). We have that  $\forall f' \in \mathcal{B}'_1$ ,  $\|f'\|_0 \leq \|f'\|_1$ . Moreover we have for any  $f' \in \mathcal{B}'_1$

$$\|P'f'\|_0 \leq \|f'\|_0 \quad \text{and} \quad \forall k \geq 1, \|\widehat{P}_k'f'\|_0 \leq \|f'\|_0.$$

Indeed, for  $K := P$  and  $K := \widehat{P}_k$ , we obtain  $\|Kf\|_0 \leq \|f\|_0$  from the positivity of  $K$  and  $K1_{\mathbb{X}} \leq 1_{\mathbb{X}}$ , so that we have for any  $f' \in \mathcal{B}'_1$  and  $f \in \mathcal{B}_0$ ,  $\|f\|_0 \leq 1$ :

$$|(K'f')(f)| = |f'(Kf)| \leq \|f'\|_0 \|Kf\|_0 \leq \|f'\|_0.$$

Next, using (UWD), we know from Lemma 2.3 that  $P'$  and  $\widehat{P}_k'$  (for every  $k \geq 1$ ) satisfy the uniform Doeblin-Fortet inequalities (10) on  $\mathcal{B}'_1$ . Moreover we obtain by duality and from Inequality ( $\mathcal{E}_{0,1}$ )

$$\|\widehat{P}_k' - P'\|_{1,0} := \sup_{f' \in \mathcal{B}'_1, \|f'\|_{\mathcal{B}'_1} \leq 1} \|\widehat{P}_k'f' - P'f'\|_0 = \|\widehat{P}_k - P\|_{0,1} \leq \varepsilon_1.$$

Finally we know that  $P'$  is quasi-compact on  $\mathcal{B}'_1$  with essential spectral radius less than  $\hat{\alpha}$ , and that the essential spectral radius of  $\widehat{P}_k'$  is zero since  $\widehat{P}_k$  is of finite rank. The previous facts and the first part of [21, Prop. 3.1] then give (16).

In a second step, we prove that Inequality (13) holds under the assumptions of Theorem 3.1. Since  $P$  is  $V$ -geometrically ergodic, we have  $\sigma(P) \subset D(0, \rho) \cup \{1\}$ . Thus, it follows from (12) and the first inclusion in (15) that  $\sigma(\widehat{P}_k) \subset \overline{D}(0, r) \cup \overline{D}(1, \vartheta)$ . Moreover, since  $\widehat{P}_k$  is assumed to be  $V$ -geometrically ergodic with rate  $\rho_k < 1 - \vartheta$ , we obtain that

$$\sigma(\widehat{P}_k) \subset \overline{D}(0, r) \cup \{1\}$$

and that  $(1/2i\pi) \oint_{|z-1|=\vartheta} (zI - \widehat{P}_k)^{-1} dz = \widehat{\pi}_k(\cdot) \phi_k$  from standard spectral theory. Thus we have for any  $\kappa \in (r, 1 - \vartheta)$

$$\begin{aligned} \forall n \geq 1, \quad \widehat{P}_k^n &= \frac{1}{2i\pi} \oint_{|z-1|=\vartheta} (zI - \widehat{P}_k)^{-1} dz + \frac{1}{2i\pi} \oint_{|z|=\kappa} z^n (zI - \widehat{P}_k)^{-1} dz \\ &= \widehat{\pi}_k(\cdot) \phi_k + \frac{1}{2i\pi} \oint_{|z|=\kappa} z^n (zI - \widehat{P}_k)^{-1} dz. \end{aligned}$$

Finally it follows from (15) that  $\|\widehat{P}_k^n - \widehat{\pi}_k(\cdot) \phi_k\|_1 \leq c(r, \vartheta, P) \kappa^{n+1}$ . Since  $\kappa \in (r, 1 - \vartheta)$  is arbitrary, this inequality holds with  $r$  in place of  $\kappa$ , and it gives (13).

Using (13), the second assertion of Theorem 3.1 follows from Inequality (4) of Proposition 2.1. The proof of Theorem 3.1 is complete.  $\square$

## 4 From $\widehat{P}_k$ to $P$ : solution to (Q2)

Let  $\rho_V(P)$  denote the spectral gap of  $P$ , that is the infimum bound of the rates  $\rho$  in Inequality (V). Recall that  $\rho_V(P)$  is unknown in general and that, even when it is known, finding an explicit constant  $C$  associated with  $\rho > \rho_V(P)$  in (V) is a difficult question. As mentioned in Introduction, there exist various methods [24, 22, 7] providing rates and constants  $(\rho, C)$  in Inequality (V), but they yield a rate  $\rho$  which is much far from  $\rho_V(P)$  except in specific cases. The first purpose of Theorem 4.1 below is to derive a rate  $r_k$  in (V) from  $\rho_k$  in  $(V_k)$ , which is all the more close to  $\rho_V(P)$  that  $k$  is large and  $\rho_k$  in  $(V_k)$  is close to the so-called second eigenvalue of the finite-rank operator  $\widehat{P}_k$ . The second purpose of Theorem 4.1 is to provide an explicit constant  $c_k \equiv c(r_k)$  associated with rate  $r_k$  in (V). Then Proposition 2.1 can be applied to derive a bound for  $\|\widehat{\pi}_k - \pi\|_{TV}$ .

Let us briefly explain why the passage from the  $V$ -geometric ergodicity of  $\widehat{P}_k$  to that of  $P$  is theoretically more difficult than the converse one studied in Section 3. Assume that Conditions (V),  $(C_{0,1})$  and (UWD) hold. Let  $r > \hat{\alpha}$  with  $\hat{\alpha}$  given in (9) and let  $\vartheta > 0$ . With  $\varepsilon_1(r, \vartheta, P)$  defined in (11b), Property (15) reads as follows

$$\|\widehat{P}_k - P\|_{0,1} \leq \varepsilon_1(r, \vartheta, P) \implies \sigma(\widehat{P}_k) \subset \mathcal{V}(r, \vartheta, P), \quad \sup_{z \in \mathcal{V}(r, \vartheta, P)^c} \|(zI - \widehat{P}_k)^{-1}\|_1 < \infty. \quad (17)$$

Next, for any fixed  $k \in \mathbb{N}^*$ , exchanging the role of  $P$  and  $\widehat{P}_k$  in (15) gives the following implication:

$$\|P - \widehat{P}_k\|_{0,1} \leq \varepsilon_1(r, \vartheta, \widehat{P}_k) \implies \sigma(P) \subset \mathcal{V}(r, \vartheta, \widehat{P}_k), \quad \sup_{z \in \mathcal{V}(r, \vartheta, \widehat{P}_k)^c} \|(zI - P)^{-1}\|_1 < \infty. \quad (18)$$

There is a significant difference between the implications (17) and (18) since the inequality  $\|\widehat{P}_k - P\|_{0,1} \leq \varepsilon_1(r, \vartheta, P)$  in (17) is satisfied for  $k$  large enough from Condition  $(C_{0,1})$ , while the inequality  $\|P - \widehat{P}_k\|_{0,1} \leq \varepsilon_1(r, \vartheta, \widehat{P}_k)$  in (18) could fail for every  $k$ . Fortunately, such a failure cannot occur when Conditions (V),  $(C_{0,1})$  and (UWD) hold (see Proposition 7.4).

Now we introduce the material used in Theorem 4.1. Under Condition (UWD), we define the following constants for any  $\vartheta > 0$ ,  $k \in \mathbb{N}^*$  and  $r_k \in (0, 1)$ , with  $B$  defined in (10):

$$\varepsilon_1(k) \equiv \varepsilon_1(r_k, \vartheta, \widehat{P}_k) := \frac{r_k^{n_1(k) + n_2(k)}}{8B(H_k B + (1 - r_k)^{-1})} \quad (19a)$$

$$\text{with } H_k \equiv H(r_k, \vartheta, \widehat{P}_k) := \sup_{z \in \mathcal{V}(r_k, \vartheta, \widehat{P}_k)^c} \|(zI - \widehat{P}_k)^{-1}\|_1 \quad (19b)$$

$$n_1(k) := \left\lceil \frac{\ln 2}{\ln(r_k/\hat{\alpha})} \right\rceil + 1, \quad n_2(k) \equiv n_2(r_k, \vartheta, P_k) := \left\lceil \frac{\ln(8B(B+3)r_k^{-n_1(k)}H_k)}{\ln(r_k/\hat{\alpha})} \right\rceil + 1. \quad (19c)$$

To apply Theorem 4.1, some preliminary (possibly poor) rate  $\rho$  in (V) is assumed to be available. But it is worth noticing that no constant associated with  $\rho$  in (V) is required. Inequality (23) then provides a new rate  $r_k$  in (V) with an explicit associated constant  $c_k$ .

**Theorem 4.1** *Assume that  $P$  is  $V$ -geometrically ergodic with some rate  $\rho \in (0, 1)$  and that Conditions  $(C_{0,1})$  and (UWD) hold. Let  $\vartheta$  be such that*

$$0 < \vartheta < \min\left(\frac{1 - \hat{\alpha}}{3}, \frac{1 - \rho}{3}\right) \quad (20)$$

and let  $\mathcal{I}_\vartheta$  denote the following subset of integers:

$$\mathcal{I}_\vartheta := \left\{ k \in \mathbb{N}^* : \widehat{P}_k \text{ satisfies } (V_k) \text{ with some rate } \rho_k < 1 - 2\vartheta \right\}. \quad (21)$$

Then, for any  $k \in \mathcal{I}_\vartheta$  and for any  $r_k$  such that

$$\max(\hat{\alpha}, \rho_k) + \vartheta < r_k < 1 - \vartheta, \quad (22)$$

the estimates in (a) and (b) below are valid provided that the convergence condition  $(C_{0,1})$  holds at rate  $\varepsilon_1(k)$ , that is

$$\|P - \widehat{P}_k\|_{0,1} \leq \varepsilon_1(k). \quad (\mathcal{E}_{0,1}(k))$$

(a) *The iterates of  $P$  converge to  $\pi(\cdot)1_{\mathbb{X}}$  with the following explicit rate of convergence:*

$$\forall n \geq 0, \quad \|P^n - \pi(\cdot)1_{\mathbb{X}}\|_1 \leq c_k r_k^{n+1} \text{ with } c_k \equiv c_k(r_k, \vartheta) := \frac{4(B+1)}{r_k^{n_1(k)}(1-r_k)} + \frac{1}{2\varepsilon_1(k)}. \quad (23)$$

(b) *We have, with  $A$  defined in (3) and  $\Delta_n := \|\widehat{P}_n - P\|_{0,1}$ ,*

$$\forall n \geq 1, \quad \|\widehat{\pi}_n - \pi\|_{TV} \leq |1 - \widehat{\pi}_n(1_{\mathbb{X}})| + \frac{L}{1 - \delta} \left( 2c_k + \frac{A}{\ln r_k^{-1}} |\ln \Delta_n| \right) \Delta_n. \quad (24)$$

Assertion (a) of Theorem 4.1 can be established as in Theorem 3.1 by exchanging the role of  $P$  and  $\widehat{P}_k$ . Assertion (b) follows from (23) and Inequality (5) of Proposition 2.1.

Under the assumptions of Theorem 4.1, the conclusions (23) and (24) are valid for  $k$  large enough, so that Issue (Q2) is solved. Indeed, first Proposition 7.1 ensures that a sufficient condition for  $k$  to belong to  $\mathcal{I}_\vartheta$  is that  $\|\widehat{P}_k - P\|_{0,1} \leq \varepsilon_0$  with  $\varepsilon_0 \equiv \varepsilon_0(r, \vartheta, P)$  defined in (42) (use that  $\max(\hat{\alpha}, \rho) + \vartheta < 1 - 2\vartheta$  from (20)). Second Proposition 7.4 ensures that a sufficient condition for the convergence property  $(\mathcal{E}_{0,1}(k))$  to hold is that  $k \in [\tilde{k}, +\infty) \cap \mathcal{I}_\vartheta$  for some integer  $\tilde{k}$  (greater than  $k_0$ ). Since integer  $k \in \mathcal{I}_\vartheta$  satisfying  $(\mathcal{E}_{0,1}(k))$  may exist for  $k < \max(\tilde{k}, k_0)$ , an effective application of Theorem 4.1 needs to find such a  $k$  with a value as small as possible. This can be done in testing the validity of both properties  $k \in \mathcal{I}_\vartheta$  and  $(\mathcal{E}_{0,1}(k))$  for increasing values of  $k$ . Finally note that the bound  $H_k$  in (19b) can be bounded under Condition  $(V_k)$  following the lines of Remark 3.2 with  $\widehat{P}_k$  in place of  $P$  (see Lemma 6.3 for discrete Markov kernels). An algorithm based on Theorem 4.1 is proposed in Subsection 6.2 for truncating discrete Markov kernels. Numerical results for random walks on  $\mathbb{X} := \mathbb{N}$  are reported in Subsection 6.3.

The whole results of [21], applied with  $\widehat{P}_k$  and  $P$  viewed as unperturbed and perturbed operators respectively, directly provide Property (23) when Condition  $(\mathcal{E}_{0,1}(k))$  is replaced with  $\|\widehat{P}_k - P\|_{0,1} \leq \varepsilon_0(k)$ , where  $\varepsilon_0(k) \equiv \varepsilon_0(r_k, \vartheta, \widehat{P}_k)$  is defined as in (42) with  $r_k, H_k$  in

place of  $r, H$ . In this case no preliminary bound  $\rho$  in  $(V)$  is required in Theorem 4.1, and Condition (20) is :  $0 < \vartheta < (1 - \hat{\alpha})/3$  (in fact the  $V$ -geometric ergodicity assumption in Theorem 4.1 may be replaced by the quasi-compactness of  $P$  on  $\mathcal{B}_1$ ). When  $\varepsilon_0(k)$  is strictly less than  $\varepsilon_1(k)$  given in (19a) and a preliminary bound  $\rho$  in  $(V)$  is known, the alternative Assertion (a) in Theorem 4.1 is of interest since it involves smaller integers  $k$ . Note that the second eigenvalue of  $\hat{P}_k$  will be all the more tractable that  $k$  is small since the rank of  $\hat{P}_k$  is expected to increase with respect to  $k$ .

When  $r_k$  gets closer to  $\rho_V(P)$ , the constant  $c_k$  associated with  $r_k$  in (23) goes to  $+\infty$ . Numerical evidence of this fact is provided by Table 1. More generally, an efficient use of Theorem 4.1 needs to make a trade-off between the rate  $r_k$  and the constant  $c_k$  in (23) since the choice of  $(\vartheta, r_k)$  greatly affects the bound  $H_k$  in (19b), thus the constant  $c_k$ .

**Remark 4.2** *Property (24) provides a bound  $\|\hat{\pi}_n - \pi\|_{TV} = O(\Delta_n |\ln \Delta_n|)$  with computable constants since  $r_k$  and  $c_k$  are deduced from computations involving the finite-rank operator  $\hat{P}_k$  for some  $k$ . Consequently, given some error term  $\varepsilon > 0$ , this bound can be used to find an integer  $n \equiv n(\varepsilon) \geq 1$  such that  $\|\hat{\pi}_n - \pi\|_{TV} \leq \varepsilon$ . However, as illustrated in Section 6, applying Inequality (4) for larger and larger  $k$  to test the previous property is more efficient since the constant  $C_k$  in  $(V_k)$  is much smaller than the constant  $c_k$  in (23), which in practice is derived from  $C_k$  (see the above comments on  $H_k$ ).*

## 5 Bounds on the essential spectral radius of quasi-compact kernels on $\mathcal{B}_1$

From the definition of  $\hat{\alpha}$  in (9), the new rates in  $(V_k)$  provided by Theorem 3.1, or in  $(V)$  by Theorem 4.1, are all the more interesting that an accurate bound of  $r_{ess}(P)$  is known (e.g. see Condition (22) for the new rate in  $(V)$ ). Estimates of  $r_{ess}(P)$  are provided under two different kind of assumptions on the Markov kernel  $P$ . The first bound for  $r_{ess}(P)$  (Theorem 5.2) is derived from a result on positive operators [30]. The second one (Proposition 5.4) is obtained by duality from the quasi-compactness criterion of [12].

First we recall some basic facts on quasi-compactness and on the essential spectral radius of a linear bounded operator  $T$  with positive spectral radius on a complex Banach space  $(\mathcal{B}, \|\cdot\|)$ . Its spectral radius  $r(T) := \lim_n \|T^n\|^{1/n}$ , where  $\|\cdot\|$  also stands for the operator norm on  $\mathcal{B}$ , is assumed to be 1 (if not, replace  $T$  with  $r(T)^{-1}T$ ). We denote by  $I$  the identity operator on  $\mathcal{B}$ . The next definitions of quasi-compactness and essential spectral radius are from [12] (to be compared with the reduction of matrices or compact operators). Additional materials are developed in [26, 19, 1].

**Definition 5.1**  *$T$  is quasi-compact on  $\mathcal{B}$  if there exist  $r_0 \in (0, 1)$ ,  $m \in \mathbb{N}^*$  and  $(\lambda_i, p_i) \in \mathbb{C} \times \mathbb{N}^*$  for  $i = 1, \dots, m$  such that:*

$$\mathcal{B} = \bigoplus_{i=1}^m \text{Ker}(T - \lambda_i I)^{p_i} \oplus H, \quad (25a)$$

where the  $\lambda_i$ 's are such that

$$|\lambda_i| \geq r_0 \quad \text{and} \quad 1 \leq \dim \text{Ker}(T - \lambda_i I)^{p_i} < \infty, \quad (25b)$$

and  $H$  is a closed  $T$ -invariant subspace such that

$$\inf_{n \geq 1} \left( \sup_{h \in H, \|h\| \leq 1} \|T^n h\| \right)^{1/n} < r_0. \quad (25c)$$

Then the essential spectral radius of  $T$ , denoted by  $r_{ess}(T)$ , is given by

$$r_{ess}(T) = \inf \left\{ r_0 \in (0, 1) \text{ such that we have (25a), (25b), (25c)} \right\}. \quad (26)$$

Note that the infimum bound in the left-hand side of (25c) is nothing else but the spectral radius of the restriction of  $T$  to  $H$ . It is well-known that the essential spectral radius of  $T$  is usually defined by  $r_{ess}(T) := \lim_n (\inf \|T^n - K\|)^{1/n}$ , where the infimum is taken over the ideal of compact operators  $K$  on  $\mathcal{B}$ . Consequently,  $T$  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  $r(T^{n_0} - K_0) < 1$ . Under the previous condition we have

$$r_{ess}(T) \leq (r(T^{n_0} - K_0))^{1/n_0}. \quad (27)$$

Finally recall that  $r_{ess}(T) = r_{ess}(T^p)^{1/p}$  for every  $p \geq 1$  since  $\lim_n (\inf \|T^n - K\|)^{1/n} = \lim_k (\inf \|T^{pk} - K\|)^{1/(pk)}$ .

## 5.1 Bounds on $r_{ess}(P)$ under drift/minorization conditions

In the next theorem, a simple bound of  $r_{ess}(P)$  is given in terms of the parameters of the drift/minorization conditions (28a)-(28b) below (note that no irreducibility/apperiodicity condition is assumed).

**Theorem 5.2** *Assume that  $P$  satisfies the following drift/minorization conditions: there exist a bounded set  $S \in \mathcal{X}$  and a positive measure  $\nu$  on  $(\mathbb{X}, \mathcal{X})$  such that*

$$\exists \delta \in (0, 1), \exists L > 0, \quad PV \leq \delta V + L 1_S, \quad (28a)$$

$$\forall x \in \mathbb{X}, \forall A \in \mathcal{X}, \quad P(x, A) \geq \nu(1_A) 1_S(x). \quad (28b)$$

Then  $P$  is a power-bounded quasi-compact operator on  $\mathcal{B}_1$  with

$$r_{ess}(P) \leq \frac{\delta \nu(1_{\mathbb{X}}) + \tau}{\nu(1_{\mathbb{X}}) + \tau} \quad \text{where } \tau := \max(0, L - \nu(V)). \quad (29)$$

In the long version [13] of [14], the quasi-compactness of  $P$  on  $\mathcal{B}_1$  is proved under the assumptions of Theorem 5.2. The bound obtained for  $r_{ess}(P)$  in [13, Th. IV.2] is less tractable than (29) since it is expressed in terms of the hitting time for  $S$ . Also mention that, in the unpublished paper [8], the quasi-compactness of Markov kernels is obtained on the subspace

of continuous functions of  $\mathcal{B}_1$  under some drift/minorization conditions. No bound on the essential spectral radius is presented in [8].

The short proof of Theorem 5.2 illuminates the role of the drift and minorization conditions to obtain good spectral properties of  $P$  on  $\mathcal{B}_1$ . In particular, using [13, Cor. IV.3], this provides a simple proof of the fact that under the assumptions of Theorem 5.2, together with irreducibility and aperiodicity assumptions,  $P$  is  $V$ -geometrically ergodic. This fact is well-known (e.g. see [23]).

*Proof.* Condition (28a) implies that  $PV \leq \delta V + L 1_{\mathbb{X}}$ , thus

$$\sup_{n \geq 0} \|P^n\|_1 \leq \frac{1 - \delta + L}{1 - \delta} \quad (30)$$

so that  $P$  is power-bounded on  $\mathcal{B}_1$ . Then, from  $P1_{\mathbb{X}} = 1_{\mathbb{X}}$  and  $1_{\mathbb{X}} \in \mathcal{B}_1$ , we have  $r(P) = 1$ . Moreover, since  $\|PV\|_1 < \infty$ , we deduce from (28b) that  $\nu(V) < \infty$ . Thus we can define the following rank-one operator on  $\mathcal{B}_1$ :  $Tf := \nu(f) 1_S$ . Set  $R := P - T$ . From  $T \geq 0$  and from (28b), it follows that  $0 \leq R \leq P$ , so  $r(R) \leq 1$ . Set  $r := r(R)$ . If  $r = 0$ , then  $P$  is quasi-compact with  $r_{ess}(P) = 0$  from (27). Now assume that  $r \in (0, 1]$ . Then, we know from [30, Appendix, Cor.2.6] that there exists a nontrivial non-negative  $\eta \in \mathcal{B}'_1$  such that  $\eta \circ R = r \eta$ . From  $P = T + R$ , we have  $\eta \circ P = \eta \circ T + r \eta$ , thus  $\eta(P1_{\mathbb{X}}) = \eta(1_{\mathbb{X}}) = \eta(T1_{\mathbb{X}}) + r \eta(1_{\mathbb{X}})$ . Hence  $\eta(T1_{\mathbb{X}}) = (1 - r)\eta(1_{\mathbb{X}})$ , from which we deduce that

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

Next, we have  $RV = PV - TV = PV - \nu(V)1_S \leq \delta V + (L - \nu(V))1_S$ . Hence, setting  $\tau := \max(0, L - \nu(V)) \geq 0$ , we obtain

$$r \eta(V) = \eta(RV) \leq \delta \eta(V) + \tau \eta(1_S) \leq \delta \eta(V) + \tau \frac{(1 - r)\eta(V)}{\nu(1_{\mathbb{X}})}. \quad (31)$$

Since  $\eta \neq 0$ , we have  $\eta(V) > 0$ , and since  $\delta \in (0, 1)$ , we cannot have  $r = 1$ . Thus  $r \in (0, 1)$ , and  $P$  is quasi-compact from (27) with  $r_{ess}(P) \leq r(P - T) = r$ . Then (31) gives (29).  $\square$

**Remark 5.3** Recall that  $A \in \mathcal{X}$  is said to be an atom for  $P$  if  $P(a, \cdot) = P(a', \cdot)$  for any  $(a, a') \in A^2$ . Any Markov model having a regenerative structure is concerned with such a property (e.g. see [27, 2]). If  $P$  satisfies (28a) with  $S := A$ , then  $r_{ess}(P) \leq \delta$ . Indeed, note that  $P$  satisfies the minorization condition (28b) with  $A$  and  $\nu(1_{\cdot}) := P(a_0, \cdot)$  for any  $a_0 \in A$ . Choose  $L := \sup_{x \in A} (PV)(x)$  in (28a). Since  $A$  is an atom, we have  $L = (PV)(a_0) = \nu(V)$  so that  $\tau = 0$  in (29).

## 5.2 Bound on $r_{ess}(P)$ under a weak drift condition

The key idea to obtain quasi-compactness in Proposition 5.4 below is to use the dual Doeblin-Fortet inequality obtained in Lemma 2.2. Despite its great simplicity, this duality approach seems to be unknown in the literature. In particular it allows us to greatly simplify the

arguments used in [34] since the well known statement [12, Cor. 1] gives the bound  $r_{ess}(P) \leq \delta$  provided that  $P^\ell$  is compact from  $\mathcal{B}_0$  to  $\mathcal{B}_1$  for some  $\ell \geq 1$ . This compactness condition is much simpler than the assumptions of [34] based on sophisticated parameters  $\beta_w(P)$  and  $\beta_\tau(P)$  as measure of non-compactness of  $P$ . Precise comparisons with [34] and complements are presented in [11, Sect. 2.3]. Simple sufficient conditions for this compactness property are presented in [11]. For instance, this holds for any discrete Markov chains or for functional autoregressive models on  $\mathbb{X} := \mathbb{R}^q$  with absolutely continuous noise with respect to the Lebesgue measure. Finally note that Equality  $r_{ess}(P) = \delta$  holds in several cases (see [11]).

**Proposition 5.4** *Assume that  $P$  satisfies the following condition*

$$\exists \delta \in (0, 1), \exists L > 0, \quad PV \leq \delta V + L 1_{\mathbb{X}} \quad (\text{WD})$$

and that  $P^\ell : \mathcal{B}_0 \rightarrow \mathcal{B}_1$  is compact for some  $\ell \geq 1$ . Then  $P$  is a power-bounded quasi-compact operator on  $\mathcal{B}_1$ , with  $r_{ess}(P) \leq \delta$ .

*Proof.* Iterating (WD) ensures that  $P$  is power-bounded on  $\mathcal{B}_1$  (see (30)). Since we have  $r_{ess}(P) = (r_{ess}(P^\ell))^{1/\ell}$ , we only consider the case  $\ell := 1$ , that is  $P : \mathcal{B}_0 \rightarrow \mathcal{B}_1$  is compact. Let  $\mathcal{B}'_1$  and  $\mathcal{B}'_0$  denote the dual spaces of  $\mathcal{B}_1$  and  $\mathcal{B}_0$  respectively. Let  $P'$  denote the adjoint operator of  $P$  on  $\mathcal{B}'_1$ . In fact, we prove that  $P'$  is a quasi-compact operator on  $\mathcal{B}'_1$  with  $r_{ess}(P') \leq \delta$ , so that  $P$  satisfies the same properties on  $\mathcal{B}_1$ . Since the operator  $P : \mathcal{B}_0 \rightarrow \mathcal{B}_1$  is assumed to be compact, so is  $P' : \mathcal{B}'_1 \rightarrow \mathcal{B}'_0$ . Then we deduce from the Doeblin-Fortet inequality of Lemma 2.2 and from [12, Cor. 1] that  $P'$  is a quasi-compact operator on  $\mathcal{B}'_1$  with  $r_{ess}(P') \leq \delta$ .  $\square$

**Remark 5.5** *If Conditions (28a)-(28b) are fulfilled for some iterate  $P^N$  in place of  $P$  (with parameters  $\delta_N < 1$ ,  $L_N > 0$  and positive measure  $\nu_N(\cdot)$ ), then the conclusions of Theorem 5.2 are valid with (29) replaced by*

$$r_{ess}(P) = r_{ess}(P^N)^{1/N} \leq \left( \frac{\delta_N \nu_N(1_{\mathbb{X}}) + \tau_N}{\nu_N(1_{\mathbb{X}}) + \tau_N} \right)^{1/N} \quad \text{where } \tau_N := \max(0, L_N - \nu_N(V)).$$

Similarly Proposition 5.4 still holds when (WD) is replaced by the following condition

$$\exists \delta \in (0, 1), \exists N \in \mathbb{N}^*, \exists L > 0, \quad P^N V \leq \delta^N V + L 1_{\mathbb{X}}.$$

Indeed the dual Doeblin-Fortet inequality of Lemma 2.2 extends by replacing  $P', \delta$  by  $P'^N$  and  $\delta^N$  respectively.

## 6 Application to truncation of discrete Markov kernels

In this section we assume that  $P := (P(i, j))_{(i, j) \in \mathbb{N}^2}$  is a Markov kernel on  $\mathbb{X} := \mathbb{N}$ . Let  $B_k := \{0, \dots, k\}$  for any  $k \geq 1$ . We consider the  $k$ -th truncated (and augmented) matrix  $P_k$ :

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

Such a matrix is generally called a linear augmentation (in the last column here) of the  $(k+1) \times (k+1)$  northwest corner truncation of  $P$ . Other kinds of augmentation, as the censored Markov chain [35], could be considered. Truncation approximation of an infinite stochastic matrix has a long story (e.g. see [31, 33, 20] and the references therein).

The associated (extended) sub-Markov kernel  $\widehat{P}_k$  on  $\mathbb{N}$  is defined by:

$$\forall (i, j) \in \mathbb{N}^2, \quad \widehat{P}_k(i, j) := \begin{cases} P_k(i, j) & \text{if } (i, j) \in B_k^2 \\ 0 & \text{if } (i, j) \notin B_k^2. \end{cases}$$

### 6.1 Theorem 4.1 for truncation of discrete Markov kernels

Consider any unbounded increasing sequence  $V := \{V(j)\}_{j \in \mathbb{N}} \in [1, +\infty)^{\mathbb{N}}$  with  $V(0) = 1$ , and the associated weighted space  $(\mathcal{B}_1, \|\cdot\|_1)$  given by

$$\mathcal{B}_1 := \left\{ f \in \mathbb{C}^{\mathbb{N}} : \|f\|_1 = \sup_{j \in \mathbb{N}} |f(j)| V(j)^{-1} < \infty \right\}.$$

The following lemma helps us to check the assumptions of Theorems 3.1 and 4.1.

**Lemma 6.1** *Assume that  $P$  satisfies Condition (WD) with parameters  $(\delta, L)$ . Then the following assertions hold.*

- (i)  $P$  is a power-bounded quasi-compact operator on  $\mathcal{B}_1$  with  $r_{ess}(P) \leq \delta$ .
- (ii) Condition (UWD) is fulfilled, that is:  $\forall k \in \mathbb{N}^* \cup \{\infty\}$ ,  $\widehat{P}_k V \leq \delta V + L 1_{\mathbb{N}}$ .
- (iii) Property  $(C_{0,1})$ , that is  $\lim_k \|\widehat{P}_k - P\|_{0,1} = 0$ , holds true since

$$\forall k \in \mathbb{N}^*, \quad \|\widehat{P}_k - P\|_{0,1} \leq \frac{K}{V(k)} \quad \text{with } K := \max(2(\delta + L), 1). \quad (32)$$

If the stochastic matrix  $P_k$  has an invariant probability measure  $\pi_k$ , then  $\widehat{\pi}_k$  is the probability measure on  $\mathbb{N}$  defined by

$$\forall j \in \mathbb{N}, \quad \widehat{\pi}_k(\{j\}) := \begin{cases} \pi_k(\{j\}) & \text{if } j \in B_k \\ 0 & \text{if } j \notin B_k. \end{cases}$$

**Remark 6.2** *Assume that  $\limsup_i (PV)(i)/V(i) > 0$  and that  $P$  satisfies (WD), so that the infimum of  $\delta$  such that (WD) holds is non zero (these assumptions are satisfied in almost all  $V$ -geometrically ergodicity models). Then the strong convergence property  $\lim_k \|P - \widehat{P}_k\|_1 = 0$  required in the standard perturbation theory does not hold. Indeed, using  $\widehat{P}_k V(i) = 0$  when  $i \notin B_k$ , we obtain*

$$\sup_{i \notin B_k} \frac{(PV)(i)}{V(i)} = \sup_{i \notin B_k} \frac{|(PV)(i) - (\widehat{P}_k V)(i)|}{V(i)} \leq \|PV - \widehat{P}_k V\|_1 \leq \|P - \widehat{P}_k\|_1.$$



If  $\lim_k \|P - \widehat{P}_k\|_1 = 0$ , then  $\lim_k \sup_{i \notin B_k} (PV)(i)/V(i) = 0$  which cannot hold from the hypothesis. Actually using this strong convergence condition leads to difficulties or restrictions in other approximation questions. For instance in [18], some iterate  $P^N$  of the Markov kernel  $P$  is approached by finite rank kernels in operator norm on  $\mathcal{B}_1$  (for other purpose than truncation issue). Thus  $P^N$  is compact on  $\mathcal{B}_1$ . Consequently this property in operator norm  $\|\cdot\|_1$  leads to suppose that  $r_{\text{ess}}(P) = 0$  which is restrictive for  $V$ -geometrically ergodic kernel  $P$  since  $r_{\text{ess}}(P) < 1$  but  $r_{\text{ess}}(P) \neq 0$  in general.

*Proof of Lemma 6.1.* Assertion (i) follows from Proposition 5.4 since the injection from  $\mathcal{B}_0$  into  $\mathcal{B}_1$  is compact from Cantor's diagonal argument and  $\lim_j V(j) = +\infty$ . Next let  $i \in B_k$ . Then we obtain using  $V(j) \geq V(k)$  for  $j \geq k$

$$(\widehat{P}_k V)(i) = \sum_{j=0}^{k-1} P(i, j)V(j) + V(k) \sum_{j \geq k} P(i, j) \leq PV(i)$$

so  $(\widehat{P}_k V)(i) \leq \delta V(i) + L$  from (WD). If  $i \notin B_k$ , then  $(\widehat{P}_k V)(i) = 0$ . This proves (ii). To derive (iii), consider any  $f \in \mathcal{B}_0$  such that  $\|f\|_0 \leq 1$ . First, we obtain for every  $i \in B_k$ :

$$\begin{aligned} |(Pf)(i) - (\widehat{P}_k f)(i)| &= \left| \sum_{j \geq k} P(i, j)(f(j) - f(k)) \right| \leq 2 \sum_{j \geq k} P(i, j) = 2 \sum_{j \geq k} P(i, j)V(j) \frac{1}{V(j)} \\ &\leq \frac{2(PV)(i)}{V(k)} \leq \frac{2(\delta + L)}{V(k)} V(i) \quad (\text{since } PV \leq (\delta + L)V \text{ from (WD)}). \end{aligned}$$

Second let  $i \notin B_k$ . Then  $(\widehat{P}_k f)(i) = 0$ , so that  $|(Pf)(i) - (\widehat{P}_k f)(i)| \leq \sum_{j \in \mathbb{N}} P(i, j)|f(j)| \leq 1$ . This gives (32) in (iii).  $\square$

To conclude this subsection we present two lemmas. The first one provides a useful bound for the constants  $C_k$  in  $(V_k)$  and  $H_k$  in (19b). It is convenient to consider  $P_k$  as an endomorphism on the finite dimensional space of functions  $h : B_k \rightarrow \mathbb{C}$  equipped with the following norm (still denoted by  $\|\cdot\|_1$  for the sake of simplicity) defined by:  $\|h\|_1 := \sup_{i \in B_k} |h(i)|/V(i)$ . The associated operator norm is also denoted by  $\|\cdot\|_1$ .

**Lemma 6.3** *The matrix  $\widehat{P}_k$  is assumed to be  $V$ -geometrically ergodic, that is 1 is a simple eigenvalue of the stochastic matrix  $P_k$  and is the unique eigenvalue of modulus one. Suppose that an explicit upper bound  $\tilde{\rho}_k \in (0, 1)$  of the second eigenvalue of  $P_k$  is known. Let any  $\rho_k$  be such that*

$$\tilde{\rho}_k < \rho_k < 1 \tag{33}$$

and define  $s \equiv s(\rho_k) \in \mathbb{N}^*$  as the smallest integer such that

$$\|P_k^s - \pi_k(\cdot)1_{B_k}\|_1 \leq \rho_k^s. \tag{34}$$

Then, we obtain the following estimate

$$\begin{aligned} \forall n \geq 0, \quad \|\widehat{P}_k^n - \widehat{\pi}_k(\cdot)1_{B_k}\|_1 &\leq C_k \rho_k^n \leq \overline{C}_k \rho_k^n \\ \text{with } C_k &:= \frac{\max_{0 \leq r \leq s-1} \|P_k^r - \pi_k(\cdot)1_{B_k}\|_1}{\rho_k^{s-1}} \quad \overline{C}_k := \frac{1 - \delta + 2L}{(1 - \delta)\rho_k^{s-1}}. \end{aligned} \tag{35}$$

Moreover, for any  $(r_k, \vartheta) \in (0, 1)^2$  such that  $\rho_k + \vartheta < r_k < 1 - \vartheta$ , the following bounds hold for  $H_k$  defined in (19b):

$$H_k \leq \overline{H}_k := \max\left(\frac{L + C_k(1 - \delta)}{\vartheta(1 - \delta)}, \frac{1}{\vartheta}\right) \leq \overline{H}_k' := \max\left(\frac{L + \overline{C}_k(1 - \delta)}{\vartheta(1 - \delta)}, \frac{1}{\vartheta}\right). \quad (36)$$

*Proof.* Let  $n \geq 0$ . Writing  $n = sm + r$  with  $r \in \{0, \dots, s - 1\}$  we deduce from (34) that

$$\begin{aligned} \|P_k^n - \pi_k(\cdot)1_{B_k}\|_1 &= \|(P_k^r - \pi_k(\cdot)1_{B_k}) \circ (P_k^{sm} - \pi_k(\cdot)1_{B_k})\|_1 \\ &\leq \max_{0 \leq r \leq s-1} \|P_k^r - \pi_k(\cdot)1_{B_k}\|_1 \|P_k^s - \pi_k(\cdot)1_{B_k}\|_1^m \\ &\leq C_k \rho_k^n. \end{aligned}$$

with  $C_k$  given in (35). Inequality  $C_k \leq \overline{C}_k$  follows from

$$\forall r = 0, \dots, s - 1, \quad \|P_k^r - \pi_k(\cdot)1_{B_k}\|_1 \leq \|P_k^r\|_1 + \|\pi_k(\cdot)1_{B_k}\|_1 \leq 1 + \frac{2L}{1 - \delta}$$

since Lemma 6.1(ii) gives: first  $P_k^r V_{|B_k} \leq \delta^r V_{|B_k} + L(1 - \delta^r)/(1 - \delta)$ , thus  $\|P_k^r\|_1 \leq 1 + L/(1 - \delta)$ ; second  $\pi_k(V_{|B_k}) \leq L/(1 - \delta)$  (see (7)). The first assertion of Lemma 6.3 is proved.

Now let  $(r, \vartheta) \in (0, 1)^2$  be such that  $\rho_k + \vartheta < r_k < 1 - \vartheta$ . Since  $r_k > \rho_k$ , we obtain

$$\mathcal{V}(r_k, \vartheta, P_k)^c = \overline{D}(0, r_k)^c \cap \overline{D}(1, \vartheta)^c.$$

Then Remark 3.2 applied to  $P_k$  (with rate  $\rho_k$  and bound  $C_k$ ) gives

$$\sup_{z \in \mathcal{V}(r_k, \vartheta, P_k)^c} \|(zI - P_k)^{-1}\|_1 = \sup_{z \in \overline{D}(0, r_k)^c \cap \overline{D}(1, \vartheta)^c} \|(zI - P_k)^{-1}\|_1 \leq \frac{\pi_k(V_{|B_k}) + C_k}{\vartheta}. \quad (37)$$

Moreover observe that we have for any  $z \in \mathbb{C}^* \setminus \sigma(P_k)$ ,

$$\forall g \in \mathcal{B}_1, \forall i \in \mathbb{N}, \quad ((zI - \widehat{P}_k)^{-1}g)(i) := \begin{cases} ((zI_k - P_k)^{-1}g_{B_k})(i) & \text{if } i \in B_k \\ g(i)/z & \text{if } i \notin B_k. \end{cases} \quad (38)$$

If  $z \in \mathcal{V}(r_k, \vartheta, \widehat{P}_k)^c$ , then  $|z| > r_k > \vartheta$ , so that (37) and (38) give

$$H_k = \sup_{z \in \mathcal{V}(r_k, \vartheta, \widehat{P}_k)^c} \|(zI - \widehat{P}_k)^{-1}\|_1 \leq \frac{\max(\pi_k(V_{B_k}) + C_k, 1)}{\vartheta}.$$

Then the bounds on  $H_k$  in (36) are deduced from (7) and  $C_k \leq \overline{C}_k$ .  $\square$

The results for truncation of discrete Markov kernels on  $\mathbb{X} := \mathbb{N}$  deduced from Theorem 4.1 are gathered in the following lemma. This is the basic material for the algorithm proposed in the next subsection.

**Lemma 6.4** *Let  $P$  be a  $V$ -geometrically ergodic Markov kernel on  $\mathbb{N}$  with some rate  $\rho \in (0, 1)$  and satisfying Condition (WD) with parameters  $(\delta, L)$ . The  $k$ -th truncated matrix  $P_k$  is supposed to satisfy the assumptions of Lemma 6.3. Pick  $\rho_k$  as in (33). Let  $\vartheta$  be such that*

$$0 < \vartheta < \min\left(\frac{1 - \delta}{3}, \frac{1 - \rho}{3}, \frac{1 - \rho_k}{2}\right). \quad (39)$$

Let  $r_k \in (0, 1)$  be such that  $\max(\delta, \rho_k) + \vartheta < r_k < 1 - \vartheta$ . If  $k$  is such that  $V(k) \geq K/\varepsilon_1(k)$  with  $\varepsilon_1(k), K$  defined in (19a) and (32) respectively, then we have the following estimates

$$\forall n \geq 0, \quad \|P^n - \pi(\cdot)1_{\mathbb{N}}\|_1 \leq c_k r_k^{n+1} \quad (40a)$$

$$\forall n \geq n_K, \quad \|\widehat{\pi}_n - \pi\|_{TV} \leq \frac{LK}{1-\delta} \left( 2c_k + \frac{A \ln V(n)}{\ln(r_k^{-1})} \right) \frac{1}{V(n)} \quad (40b)$$

where  $n_K := \min\{n \in \mathbb{N}^*, V(n) \geq K\}$ ,  $A$  is introduced in (3) and  $c_k$  in (23).

To derive the final form (40b) of (24), note that, for  $n \geq n_K$ , we have  $1 \geq \Delta_n \geq \|P1_{\mathbb{N}} - \widehat{P}_n 1_{\mathbb{N}}\|_1 = \|1_{\mathbb{N} \setminus B_n}\|_1 = 1/V(n)$  so that  $|\ln \Delta_n| \leq \ln V(n)$ .

Recall that, if no rate  $\rho$  in  $(V)$  is known, then Lemma 6.4 applies provided that  $k \geq k_0$ , with  $k_0$  defined as in (43) where  $\varepsilon_0$  is replaced by  $\varepsilon_0(k)$  given as in (42) using  $(r_k, H_k)$  in place of  $(r, H)$  (see the comments after Theorem 4.1). In this case, delete  $(1 - \rho)/3$  in (39).

Property (40b) provides a bound  $\|\widehat{\pi}_n - \pi\|_{TV} = O(\ln(V(n))/V(n))$  with computable constants since  $r_k$  and  $c_k$  are deduced from calculations involving the matrix  $\widehat{P}_k$  for some  $k$ . Consequently, given some error term  $\varepsilon > 0$ , this bound can be used to find an integer  $n \equiv n(\varepsilon) \geq 1$  such that  $\|\widehat{\pi}_n - \pi\|_{TV} \leq \varepsilon$  (see Table 3-(b) for an illustration). However, as mentioned in Remark 4.2, using the direct Inequality (4) is the best way to estimate the total variation distance between  $\pi$  and  $\widehat{\pi}_k$ . This point is illustrated in Subsection 6.3 (see Table 2). To specify (4), observe that the first term of the bounds satisfies from  $\phi_k := 1_{B_k}$  and (7)

$$\widehat{\pi}_k(1_{\mathbb{N}}) |1 - \pi(\phi_k)| \leq \sum_{j \notin B_k} \pi(j) \leq \frac{1}{V(k)} \sum_{j \notin B_k} \pi(j)V(j) \leq \frac{\pi(V)}{V(k)} \leq \frac{L}{(1-\delta)V(k)}$$

so that Inequality (4) reads as follows

$$\|\widehat{\pi}_k - \pi\|_{TV} \leq \frac{L}{1-\delta} \left( 1 + \frac{2KC_k}{\rho_k} + \frac{AK}{\ln(\rho_k^{-1})} \ln V(k) \right) \frac{1}{V(k)}. \quad (41)$$

## 6.2 Algorithm for assessing the bounds

Let  $P$  be a  $V$ -geometrically ergodic Markov kernel on  $\mathbb{X} := \mathbb{N}$  with some rate  $\rho \in (0, 1)$  and satisfying Condition (WD) with parameters  $(\delta, L)$ . Assume that, for any  $k \geq 1$ , 1 is a simple eigenvalue of the stochastic matrix  $P_k$  and is the unique eigenvalue of modulus one. Lemma 6.4 is used to propose the following generic algorithm which allows us to assess the bounds (23) and (24) provided by Theorem 4.1 for  $\|P^n - \pi(\cdot)1_{\mathbb{N}}\|_1$  and  $\|\pi - \widehat{\pi}_k\|_{TV}$  respectively. Numerical illustrations are proposed in the next subsection.

First compute the constants in (3), (40b), (9) and (10):

$$A := 1 + \frac{L}{1-\delta} \quad K := \max(2(\delta + L), 1) \quad \hat{\alpha} = \max(\delta, r_{ess}(P)) = \delta, \quad B := \frac{L}{1-\hat{\alpha}}.$$

1. Pick a (small)  $k \in \mathbb{N}^*$ .

2. Compute the second eigenvalue  $\tilde{\rho}_k$  of  $P_k$ .
3. Pick  $\rho_k \in (\max(\hat{\alpha}, \tilde{\rho}_k), 1)$ .
4. Compute the invariant probability measure  $\pi_k$  of  $P_k$  and the matrix  $G_k := P_k - \pi_k(\cdot)1_{B_k}$ .
5. Compute  $s := \inf\{n \in \mathbb{N}^* : \|G_k^n\|_1 \leq \rho_k^n\}$  (see (34)).
6. Compute  $C_k := \max\{\|G_k^r\|_1/\rho_k^{s-1} : 0 \leq r \leq s-1\}$ .
7. **Compute the bound for  $\|\pi - \hat{\pi}_k\|_{TV}$  in (41).**
8. Pick  $\vartheta$  and  $r_k$  such that  $0 < \vartheta < \min\left(\frac{1-\rho_k}{2}, \frac{1-\rho}{3}, \frac{1-\delta}{3}\right)$  and  $\rho_k + \vartheta < r_k < 1 - \vartheta$ .
9. Compute  $n_1(k) := \left\lfloor \frac{\ln 2}{\ln(r_k/\delta)} \right\rfloor + 1$ ,  $n_2(k) := \left\lfloor \frac{\ln(8B(B+3)r_k^{-n_1(k)}H_k)}{\ln(r_k/\delta)} \right\rfloor + 1$ ,  $H_k := \frac{L+C_k(1-\delta)}{\vartheta(1-\delta)}$   
and

$$\varepsilon_1(k) := \frac{r_k^{n_1(k)+n_2(k)}}{8B(H_k B + (1-r_k)^{-1})} \quad k_1 \equiv k_1(k) := \min\{n : V(n) \geq K/\varepsilon_1(k)\}.$$

10. Check Condition  $(\mathcal{E}_{0,1}(k))$ , that is:

if  $k < k_1$  then go to Step 1 setting  $k := k + 1$ ;

**else compute constant  $c_k$  in (23) and use bounds in (40a)-(40b) with  $(c_k, r_k)$ .**

You can use  $\bar{C}_k$  in (35) in place of  $C_k$  in Step 6, so that  $C_k, H_k$  are replaced by  $\bar{C}_k, \bar{H}_k'$  in the sequel of the algorithm.

### 6.3 A numerical example

Let us consider an instance of random walk on  $\mathbb{X} := \mathbb{N}$  with identically distributed bounded increments, that is a Markov chain with transition kernel  $P$  defined, for some  $c, g, d \in \mathbb{N}^*$ , by

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

$$\forall i \geq g, \forall j \in \mathbb{N}, \quad P(i, j) = \begin{cases} a_{j-i} & \text{if } i-g \leq j \leq i+d \\ 0 & \text{if not} \end{cases}$$

where  $(a_{-g}, \dots, a_d) \in [0, 1]^{g+d+1}$  satisfies  $\sum_{k=-g}^d a_k = 1$  for all  $i \geq g$ . This kind of kernels arises, for instance, from time-discretization of Markovian queuing models. Under the negative mean increment condition

$$\sum_{k=-g}^d k a_k < 0$$

the following properties are known:

1. there exists  $\hat{\gamma} > 1$  such that  $\phi(\hat{\gamma}) := \sum_{k=-g}^d a_k \hat{\gamma}^k = \min_{\gamma > 1} \phi(\gamma) < 1$ . Let  $V \equiv \{\hat{\gamma}^n\}_{n \in \mathbb{N}}$ .
2.  $P$  satisfies (WD) with  $\delta = \phi(\hat{\gamma})$  and  $L = \max(0, \sum_{j=0}^c P(i, j) \hat{\gamma}^j - \phi(\hat{\gamma}) : i = 1, \dots, g-1)$ . Moreover,  $r_{ess}(P) = \delta$  (see [15]).
3.  $P$  is  $V$ -geometrically ergodic with an invariant probability measure  $\pi$  such that  $\pi(V) < \infty$ .

A general procedure to estimate the infimum bound  $\rho_V(P)$  of the rates  $\rho$  in Inequality (V) for such models is given in [15]. But the assessment of the constant is not addressed. The next tables give such constants as well as explicit bounds for the total variation distance  $\|\pi - \hat{\pi}_k\|_{TV}$ . To the best of our knowledge, such results are not known for this kind of models.

In the present context, the integer  $k_1$  of Step 9 has the following form

$$k_1 := \left\lceil \frac{1}{\ln \hat{\gamma}} \left( \ln(8BK) + \ln(H_k B + (1 - r_k)^{-1}) + (n_1(k) + n_2(k)) \ln(r_k^{-1}) \right) \right\rceil + 1.$$

As a matter of example, we take  $c := 2, g := 2, d := 1$  with  $a_{-2}, a_1 > 0$  and boundary probabilities

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

It is easily seen that  $P$  and  $\{P_k\}_{k \geq 1}$  are irreducible and aperiodic. Since  $P(2, 0) = a_{-2} > 0$  and  $P(0, 2) = 0$ ,  $P$  is not reversible. The form of boundary probabilities is chosen for convenience. Other (finitely many) boundary probabilities could be considered provided that  $P$  and  $P_k$   $k \geq 1$  are irreducible and aperiodic. In order to provide numerical evidence for the effectiveness of the algorithm, we also specify the values

$$a_{-2} := 1/2, \quad a_{-1} := 1/3, \quad a_0 := 0, \quad a_1 := 1/6.$$

We obtain  $\hat{\gamma} \approx 2.18$  and  $\delta \approx 0.621$ . Moreover, using the procedure in [15], we can obtain a numerical approximation of  $\rho_V(P)$ , so that any bound on this estimate gives the required initial value of  $\rho$  in (V) to use Theorem 4.1. Note that  $P$  is stochastically monotone iff  $a \geq b \geq 5/6$ .

We report in Table 2 and Table 3-(a) the value of the different bounds (40a)-(40b) for  $(a, b) := (1/2, 1/2)$  which gives a kernel  $P$  which has no eigenvalue in the annulus  $\{z \in \mathbb{C} : \delta < |z| < 1\}$  so that  $\rho_V(P) = \delta \approx 0.621$  (see [15] for details).  $P$  is not reversible and not stochastically monotone. Therefore the results of [33, 28, 7] are not relevant since the bound for  $\|\hat{\pi}_k - \pi\|_{TV}$  in [33] is only obtained for stochastically monotone Markov kernels, and the rate  $\rho$  in (V) obtained in [28, 7] is very close to 1 (and far from the best rate) in the non reversible case. We also report in Table 2 the results obtained from the direct Inequality (41). These results are in agreement with the expected fact that Inequality (41) gives better bounds for  $\|\hat{\pi}_k - \pi\|_{TV}$  than (40b). We also report in Table 3-(b) the value of  $n(\varepsilon)$  such that  $\|\hat{\pi}_n - \pi\|_{TV} \leq \varepsilon$  using (40b).

**Remark 6.5** *Note that the assumptions for using the algorithm (see Lemma 6.4) are not sensitive to approximate values of  $\hat{\gamma}$  and  $\delta$  since properties as  $r_{ess}(P) = \delta = \phi(\gamma) < 1$  and  $V$ -geometric ergodicity are valid for any  $\gamma \in (1, \gamma_0)$  with  $\gamma_0 = 5.541$ .*

$(a, b) = (1/2, 1/2)$			
$r_{45}$	0.87	0.78	0.76
$c_{45}$	$1.924 \times 10^7$	$4.610 \times 10^{11}$	$1.348 \times 10^{14}$

Table 1: Improvement of the rate vs degradation of the constant using Theorem 4.1

$(a, b) = (1/2, 1/2)$										
$k$	$ \tilde{\rho}_k $	$\rho_k$	$s$	$C_k$	(41)	$\vartheta$	$r_k$	$k_1$	$c_k$	(40b) for $n := k$
15	0.6018	0.75	4	4.1539	$8.44 \times 10^{-2}$	0.09	0.9	20		
25	0.6142	0.75	4	4.1540	$5.712 \times 10^{-5}$	0.09	0.9	20	$4.715 \times 10^5$	$1.112 \times 10^{-1}$
35	0.6177	0.75	4	4.1540	$3.277 \times 10^{-8}$	0.09	0.9	20	$4.715 \times 10^5$	$4.616 \times 10^{-5}$
45	0.6192	0.75	4	4.3736	$1.733 \times 10^{-11}$	0.09	0.9	20	$4.816 \times 10^5$	$1.946 \times 10^{-8}$

Table 2: Bounds for  $\|\pi - \pi_k\|_{TV}$  using (40b) / (41)

## 7 Complements on Theorems 3.1 and 4.1

In this section, we assume that  $P$  is quasi-compact on  $\mathcal{B}_1$  and that Conditions  $(C_{0,1})$  and (UWD) hold. We use the second part of [21, Prop. 3.1] and the spectral rank-stability property [21, Cor. 3.1] to prove the two following properties:

- if  $P$  satisfies (V), then  $\widehat{P}_k$  is  $V$ -geometrically ergodic with explicit rates for  $k$  large enough. See Proposition 7.1.
- There exists an integer  $\tilde{k}$  such that, for every  $k \geq \tilde{k}$ , the condition  $\|P - \widehat{P}_k\|_{0,1} \leq \varepsilon_1(k)$  holds, where  $\varepsilon_1(k)$  is defined in (19a). See Proposition 7.4.

As discussed in Sections 3 and 4, the previous properties may be valid for  $k$  much smaller than the integers  $k_0$  and  $\tilde{k}$  provided in this section. It is the reason why these two integers are not introduced in Theorems 3.1 and 4.1. However, since the  $V_k$ -geometric ergodicity of  $P_k$  is required for some  $k$  in these two theorems, as well as the condition  $\|P - \widehat{P}_k\|_{0,1} \leq \varepsilon_1(k)$  in the second theorem, obtaining the theoretical existence of  $k_0$  and  $\tilde{k}$  is important.

### 7.1 Complements on Theorem 3.1

We use the notations introduced for Theorem 3.1. For any  $r \in (0, 1)$ , set  $\eta \equiv \eta(r) := 1 - \ln r / \ln \hat{\alpha} \in (0, 1)$ , and with  $\varepsilon_1 \equiv \varepsilon_1(r, \vartheta, P)$  defined in (11b)

$$\varepsilon_2 \equiv \varepsilon_2(r, \vartheta, P) := \left\{ \frac{r^{n_1}}{4B(H(2B+3) + 2(1+B) + (1-r)^{-1})} \right\}^{1/\eta},$$

and  $\varepsilon_0 \equiv \varepsilon_0(r, \vartheta, P) := \min(\varepsilon_1, \varepsilon_2)$ . (42)

From  $(C_{0,1})$  there exists a positive integer  $k_0 \equiv k_0(r, \vartheta, P)$  such that

$$\forall k \geq k_0, \quad \|\widehat{P}_k - P\|_{0,1} \leq \varepsilon_0. \quad (43)$$

$(a, b) = (1/2, 1/2)$ with $r_k = 0.925$		
$k$	$c_k$	Bound (40a)
30	$1.675 \times 10^6$	$1.497 \times 10^{-7}$
50	$6.533 \times 10^6$	$5.839 \times 10^{-7}$

(a)

$(a, b) = (1/2, 1/2)$ with $r_{35} := 0.9$			
$\varepsilon$	$10^{-2}$	$10^{-4}$	$10^{-6}$
$n(\varepsilon)$	28	34	40

(b)

Table 3: (a): bound (40a) for  $\|P^{300} - \pi(\cdot)1_{\mathbb{N}}\|_1$  and (b):  $\|\pi_{n(\varepsilon)} - \pi\|_{TV} \leq \varepsilon$  from (40b)

**Proposition 7.1** *Assume that  $P$  is  $V$ -geometrically ergodic with rate  $\rho$  and that Conditions  $(C_{0,1})$  and (UWD) hold. Let  $\vartheta > 0$  be such that  $\max(\hat{\alpha}, \rho) < 1 - 2\vartheta$  and  $r$  such that  $\max(\hat{\alpha}, \rho) + \vartheta < r < 1 - \vartheta$ . Then, for every  $k \geq k_0$ , Inequality  $(V_k)$  holds with  $\rho_k := r$ .*

Using duality as in the proof of Theorems 3.1, Proposition 7.1 is based on the next lemma which follows from the rank-stability property of spectral projections [21, cor. 3.1].

**Lemma 7.2** *Assume that  $P$  is quasi-compact on  $\mathcal{B}_1$  and that Conditions  $(C_{0,1})$  and (UWD) hold. Let  $r > \hat{\alpha}$  and  $\vartheta > 0$ . For every  $k \geq k_0$  the spectral projections of  $P'$  and  $\hat{P}'_k$  associated with any connected component of  $\mathcal{V}(r, \vartheta, P)$  (not containing 0) have the same rank.*

*Proof of Proposition 7.1.* From the  $V$ -geometrical ergodicity of  $P$ , we have  $\sigma(P') = \sigma(P) \subset D(0, \rho) \cup \{1\}$ . Thus it follows from (16) that  $\sigma(\hat{P}'_k) \subset \overline{D}(0, r) \cup \overline{D}(1, \vartheta)$ . Next, from  $\|\hat{P}'_k - P\|_{0,1} \leq \varepsilon_0$  and Lemma 7.2, we obtain that  $\sigma(\hat{P}'_k) \cap \overline{D}(1, \vartheta) = \{\lambda\}$  for some eigenvalue  $\lambda \in \mathbb{C}$  of  $\hat{P}'_k$  and that there exists an associated rank-one projection  $\hat{\Pi}'_{k,\lambda}$  on  $\mathcal{B}'_1$  such that we have for any  $\kappa \in (r, 1 - \vartheta)$ :

$$\forall n \geq 1, \quad \hat{P}'_k{}^n - \lambda^n \hat{\Pi}'_{k,\lambda} = \frac{1}{2i\pi} \oint_{|z|=\kappa} z^n (zI - \hat{P}'_k)^{-1} dz.$$

It follows from (16) that

$$\forall n \geq 1, \quad \|\hat{P}'_k{}^n - \lambda^n \hat{\Pi}'_{k,\lambda}\|_1 \leq c(r, \vartheta, P) \kappa^{n+1}. \quad (44)$$

Since  $\kappa$  is arbitrarily close to  $r$ , this gives the expected conclusion in Proposition 7.1 using duality and the next lemma.  $\square$

**Lemma 7.3** *The eigenvalue  $\lambda$  in (44) is equal to 1. Moreover there exists a  $\hat{P}'_k$ -invariant positive measure  $\hat{\pi}_k$  on  $(\mathbb{X}, \mathcal{X})$  such that  $\hat{\pi}_k(V) < \infty$ , and the rank-one projection  $\hat{\Pi}'_{k,\lambda}$  is the adjoint of the following rank-one projection  $\hat{\Pi}_k$  on  $\mathcal{B}_1$ :*

$$\forall f \in \mathcal{B}_1, \quad \hat{\Pi}_k f := \hat{\pi}_k(f) \phi_k.$$

*Proof.* Since  $|\lambda| > \kappa$ , we deduce from  $\hat{P}'_k \phi_k = \phi_k$  and (44) that, for any  $f' \in \mathcal{B}'_1$ , we have  $\lim_n \lambda^{-n} f'(\phi_k) = \lim_n \lambda^{-n} (\hat{P}'_k{}^n f')(\phi_k) = (\hat{\Pi}'_{k,\lambda} f')(\phi_k)$ . Since there exists  $f' \in \mathcal{B}'_1$  such that  $f'(\phi_k) \neq 0$ ,  $\{\lambda^{-n}\}_{n \in \mathbb{N}}$  converges in  $\mathbb{C}$ . Thus either  $\lambda = 1$ , or  $|\lambda| > 1$ . Moreover the sequence  $\{\hat{P}'_k{}^n\}_{n \in \mathbb{N}}$  is bounded in  $\mathcal{L}(\mathcal{B}_1)$  (proceed as in (30)). Thus  $\{\hat{P}'_k{}^n\}_{n \in \mathbb{N}}$  is bounded in  $\mathcal{L}(\mathcal{B}'_1)$ , so that Inequality (44) implies that  $\{\lambda^n\}_{n \in \mathbb{N}}$  is bounded in  $\mathbb{C}$ . Therefore  $\lambda = 1$ .

Now we omit  $\lambda$  in  $\widehat{\Pi}'_{k,\lambda}$ . We can prove as in [10, proof of Th. 1] that  $\widehat{\Pi}'_k$  is the adjoint of the rank-one projection  $\widehat{\Pi}_k$  defined on  $\mathcal{B}_1$  by

$$\forall f \in \mathcal{B}_1, \quad \widehat{\Pi}_k f := \lim_n \widehat{P}_k^n f \quad \text{in } \mathcal{B}_1,$$

and that  $\widehat{\Pi}_k$  is of the form  $\widehat{\Pi}_k f = \widehat{e}_k'(f) \phi_k$  ( $f \in \mathcal{B}_1$ ) for some non-negative element  $\widehat{e}_k' \in \mathcal{B}'_1$  and the function  $\phi_k$  of the Introduction. Let  $x_0 \in \mathbb{X}$  such that  $\phi_k(x_0) \neq 0$ . Then we obtain from the previous facts:  $\forall A \in \mathcal{X}, \lim_n \widehat{P}_k^n(x_0, A) = \widehat{e}_k'(1_A) \phi_k(x_0)$ . Since  $\widehat{P}_k^n(x_0, \cdot)$  is a positive measure on  $(\mathbb{X}, \mathcal{X})$ , it follows from the Vitali-Hahn-Saks theorem that, for all  $A \in \mathcal{X}$   $\widehat{e}_k'(1_A) = \widehat{\pi}_k(1_A)$ , where  $\widehat{\pi}_k$  is the positive measure on  $(\mathbb{X}, \mathcal{X})$  of Introduction. Proceeding as in [10], we can prove that  $\widehat{\pi}_k(V) < \infty$  and that  $\widehat{e}_k'$  coincide with  $\widehat{\pi}_k$  on  $\mathcal{B}_1$ .  $\square$

## 7.2 Complements on Theorem 4.1

We use the notations introduced before Theorem 4.1.

**Proposition 7.4** *Assume that  $P$  is quasi-compact on  $\mathcal{B}_1$  and that Conditions  $(C_{0,1})$  and (UWD) hold. Let  $\vartheta \in (0, (1 - \hat{\alpha})/2)$ . There exists  $\tilde{k} \equiv \tilde{k}(\vartheta) \in \mathbb{N}^*$  such that, for every  $k \in [\tilde{k}, +\infty) \cap \mathcal{I}_\vartheta$ , the property  $(\mathcal{E}_{0,1}(k))$  holds.*

This proposition is inspired from [21, Lem. 4.2]. Since the proof in [21] is only sketched and the choice of  $r_k$  (involved in (19a)) must be carefully examined, the derivation of Proposition 7.4 is detailed in this subsection. First note that, from Definition 5.1 and quasi-compactness of  $P$ , all the spectral values of  $P$  strictly larger than  $\hat{\alpha}$  are eigenvalues since  $\hat{\alpha} \geq r_{ess}(P)$ . More precisely, for any  $R > \hat{\alpha}$ , the operator  $P$  has a finite number of eigenvalues of modulus larger than  $R$ . The same property holds for every  $\widehat{P}_k$  since  $\widehat{P}_k$  is of finite rank (thus  $r_{ess}(\widehat{P}_k) = 0$ ).

**Lemma 7.5** *Let  $R > \hat{\alpha}$  and  $\theta > 0$ . Then there exists  $\tilde{k}_0 \equiv \tilde{k}_0(R, \theta) \in \mathbb{N}^*$  such that, for any eigenvalue of  $P$  satisfying  $|\lambda| > R$  and for every  $k \geq \tilde{k}_0$ , the open disk  $D(\lambda, \theta)$  contains at least an eigenvalue of  $\widehat{P}_k$ .*

*Proof.* Let  $\mathcal{D}_R$  denote the set of the eigenvalues  $z$  of  $P$  such that  $|z| > R$ . Define  $a := R - \hat{\alpha}$ ,  $b := \min\{|z - z'|, z, z' \in \mathcal{D}_R, z \neq z'\}$  and  $c := \min\{|z| - R, z \in \mathcal{D}_R\}$ . Without loss of generality we can suppose that  $\theta < \min(a, b/2, c/2)$ . Let  $\vartheta \in (0, \theta)$  and  $r := \hat{\alpha} + \theta$ . Let  $\tilde{k}_0$  be the smallest integer such that:  $\forall k \geq \tilde{k}_0, \|\widehat{P}_k - P\|_{0,1} \leq \varepsilon_0$ , with  $\varepsilon_0 \equiv \varepsilon_0(r, \vartheta, P)$  given in (42). Note that assumptions of Lemma 7.2 are satisfied.

Next, consider any  $k \geq \tilde{k}_0$  and any  $\lambda \in \mathcal{D}_R$ . Let  $\theta' \in (\vartheta, \theta)$  be such that  $(zI - \widehat{P}_k)^{-1}$  is well-defined in  $\mathcal{L}(\mathcal{B}_1)$  for every  $z \in C(\lambda, \theta') := \{z \in \mathbb{C} : |z - \lambda| = \theta'\}$ . Such  $\theta'$  exists from the quasi-compactness of  $\widehat{P}_k$ . Note that  $(zI - P)^{-1}$  is also well-defined in  $\mathcal{L}(\mathcal{B}_1)$  for every  $z \in C(\lambda, \theta')$ , more precisely:  $C(\lambda, \theta') \subset \mathcal{V}(r, \vartheta, P)^c$ . Indeed let  $z \in C(\lambda, \theta')$ . Then

$$|z| \geq |\lambda| - \theta' \geq R + c - \theta' > R + 2\theta - \theta' > R + \vartheta,$$



thus  $d(z, \overline{D}(0, R)) > \vartheta$ . Next, let  $z' \in \sigma(P)$  be such that  $|z'| > R$ . Then  $z' \in \mathcal{D}_R$ . If  $z' = \lambda$  then  $|z' - z| = |\lambda - z| = \theta' > \vartheta$ . If  $z' \neq \lambda$  then, using the triangle inequality  $|z' - \lambda| \leq |z' - z| + |z - \lambda|$ , we obtain  $|z - z'| \geq b - \theta' > b - \theta > \theta > \vartheta$ . We have proved that  $d(z, \sigma(P)) > \vartheta$ . Finally we have  $|z| > r$  since  $R > r$  (use  $r = \hat{\alpha} + \theta < \hat{\alpha} + a = \hat{\alpha} + R - \hat{\alpha} = R$ ). Thus  $z \in \mathcal{V}(r, \vartheta, P)^c$ .

Now, the spectral projections

$$\Pi'_\lambda := \frac{1}{2i\pi} \oint_{C(\lambda, \theta')} (zI - P')^{-1} dz \quad \text{and} \quad \Pi'_{k, \lambda} := \frac{1}{2i\pi} \oint_{C(\lambda, \theta')} (zI - \widehat{P}'_k)^{-1} dz$$

have the same rank from  $k \geq \widetilde{k}_0$  and Lemma 7.2. Since  $\Pi'_\lambda$  has a nonzero rank from  $\lambda \in \sigma(P')$ , so is  $\Pi'_{k, \lambda}$ . Thus we have  $D(\lambda, \theta') \cap \sigma(\widehat{P}'_k) \neq \emptyset$ .  $\square$

Now, to prove Proposition 7.4, we consider any  $\vartheta \in (0, (1 - \hat{\alpha})/2)$  and we set  $\tilde{r} := \hat{\alpha} + \vartheta/2$ .

**Lemma 7.6** *There exists  $\widetilde{k}_0 \equiv \widetilde{k}_0(\tilde{r}, \vartheta) \in \mathbb{N}^*$  such that  $\forall k \geq \widetilde{k}_0$ ,  $\mathcal{V}(\tilde{r}, \vartheta/4, P) \subset \mathcal{V}(\tilde{r}, \vartheta, \widehat{P}_k)$ .*

*Proof.* Let  $u \in \mathcal{V}(\tilde{r}, \vartheta/4, P)$ . Thus  $|u| \leq \tilde{r}$  or  $d(u, \sigma(P)) \leq \vartheta/4$ . If  $|u| \leq \tilde{r}$ , then  $u \in \mathcal{V}(\tilde{r}, \vartheta, \widehat{P}_k)$ . Now assume that  $|u| > \tilde{r}$  and  $d(u, \sigma(P)) \leq \vartheta/4$ . Since  $\sigma(P)$  is compact, there exists  $\lambda \in \sigma(P)$  such that  $|u - \lambda| \leq \vartheta/4$ . We have  $|\lambda| > \hat{\alpha} + \vartheta/4$  from

$$|\lambda| \geq |u| - \frac{\vartheta}{4} > \tilde{r} - \frac{\vartheta}{4} = \hat{\alpha} + \frac{\vartheta}{4}.$$

Then it follows from Lemma 7.5 with  $R := \hat{\alpha} + \vartheta/4$  and  $\theta := \vartheta/4$  that there exists  $\widetilde{k}_0 \equiv \widetilde{k}_0(R, \theta) \in \mathbb{N}^*$  such that, for every  $k \geq \widetilde{k}_0$ , the disk  $D(\lambda, \vartheta/4)$  contains an eigenvalue of  $\widehat{P}_k$ , say  $\lambda_k$ . We obtain  $d(u, \sigma(\widehat{P}_k)) \leq \vartheta$  since  $|u - \lambda_k| \leq |u - \lambda| + |\lambda - \lambda_k| \leq \vartheta/2$ . Thus  $u \in \mathcal{V}(\tilde{r}, \vartheta, \widehat{P}_k)$ .  $\square$

From the definition of  $\mathcal{V}(\tilde{r}, \vartheta, \widehat{P}_k)$ , we have:  $z \in \mathcal{V}(\tilde{r}, \vartheta, \widehat{P}_k)^c \Rightarrow d(z, \sigma(\widehat{P}_k)) > \vartheta$ . Thus, the following constant is well-defined for every  $k \geq 1$ :

$$\widetilde{H}_k := \sup_{z \in \mathcal{V}(\tilde{r}, \vartheta, \widehat{P}_k)^c} \|(zI - \widehat{P}_k)^{-1}\|_1.$$

**Lemma 7.7** *The sequence  $\{\widetilde{H}_k\}_{k \geq 1}$  is bounded.*

*Proof.* Let  $\varepsilon_1(\tilde{r}, \vartheta/4, P)$  be defined as in (11b). From  $(C_{0,1})$  there exists  $k_1 \equiv k_1(\tilde{r}, \vartheta) \in \mathbb{N}^*$  such that

$$\forall k \geq k_1, \quad \|\widehat{P}_k - P\|_{0,1} \leq \varepsilon_1(\tilde{r}, \vartheta/4, P).$$

It follows from Lemma 7.6 that

$$\forall k \geq \widetilde{k}_0, \quad \mathcal{V}(\tilde{r}, \vartheta, \widehat{P}_k)^c \subset \mathcal{V}(\tilde{r}, \vartheta/4, P)^c$$

so that, for every  $k \geq \max(\widetilde{k}_0, k_1)$  we have from (15) with  $r := \tilde{r}$  and  $\vartheta/4$  in place of  $\vartheta$  (note that, from [16], Property (15) holds for any  $r > \hat{\alpha}$  and  $\vartheta > 0$  under Conditions  $(C_{0,1})$ , (UWD) and  $r_{\text{ess}}(P) < 1$ ):

$$\sup_{z \in \mathcal{V}(\tilde{r}, \vartheta, \widehat{P}_k)^c} \|(zI - \widehat{P}_k)^{-1}\|_1 \leq \sup_{z \in \mathcal{V}(\tilde{r}, \vartheta/4, P)^c} \|(zI - \widehat{P}_k)^{-1}\|_1 \leq c(\tilde{r}, \vartheta/4, P) < \infty.$$

This gives the expected assertion.  $\square$

*Proof of Proposition 7.4.* Let  $k \in \mathcal{I}_\vartheta$  and  $H_k \equiv H(r_k, \vartheta, \widehat{P}_k)$  be defined by (21) and (19b) respectively. Then  $H_k \leq \widetilde{H}_k$ . Indeed we have  $\mathcal{V}(r_k, \vartheta, \widehat{P}_k)^c \subset \mathcal{V}(\tilde{r}, \vartheta, \widehat{P}_k)^c$  since (use (22))

$$\tilde{r} = \hat{\alpha} + \frac{\vartheta}{2} \leq \max(\hat{\alpha}, \rho_k) + \vartheta < r_k.$$

It follows from Lemma 7.7 that  $\{H_k\}_{k \geq 1}$  is bounded. Then, the sequences  $\{n_1(k)\}_{k \geq 1}$  and  $\{n_2(k)\}_{k \geq 1}$  given in (19c) are bounded since  $\hat{\alpha} + \vartheta < r_k < 1 - \vartheta$ . Therefore, the sequence  $\{\varepsilon_1(k)\}_{k \geq 1}$  in (19a) is uniformly bounded away from zero, that is  $\alpha_1 := \inf_{k \geq 1} \varepsilon_1(k) > 0$ . Finally, from  $(C_{0,1})$  there exists  $\tilde{k} \in \mathbb{N}$  such that:  $\forall k \geq \tilde{k}$ ,  $\|\widehat{P}_k - P\|_{0,1} \leq \alpha_1$ . Thus, for every  $k \geq \tilde{k}$ ,  $\|\widehat{P}_k - P\|_{0,1} \leq \varepsilon_1(k)$ .  $\square$

## 8 Conclusion

In all the cases where the probabilistic works cited in Introduction do not provide a satisfactory rate  $\rho$  in (V), Theorem 4.1 can be applied to obtain a new rate  $r_k$  and constant  $c_k$  in (V) derived from  $(\rho_k, C_k)$  in  $(V_k)$  for some  $k$  large enough (see (23)). This new pair  $(r_k, c_k)$  will be all the more interesting that

1. the pair  $(\rho_k, C_k)$  in  $(V_k)$  is precise;
2. the bound on the essential spectral radius  $r_{ess}(P)$  is accurate.

The estimate  $\|\widehat{\pi}_n - \pi\|_{TV} = O(\Delta_n |\ln \Delta_n|)$  provided by (24) then involves a computable constant depending on  $(r_k, c_k)$ . As mentioned in Remark 4.2 and illustrated in Section 6, the direct Inequality (4) provides the best estimate of the total variation distance between  $\pi$  and  $\widehat{\pi}_k$  when the  $C_k$ 's are accurately computed.

The point 1. above is of computational nature (see Section 6 for discrete  $\mathbb{X}$ ). The point 2. is addressed in Theorem 5.2 and Proposition 5.4. Note that  $r_{ess}(P)$  is less than the contractive coefficient  $\delta$  of (WD) in Proposition 5.4 and in atomic case (see Remark 5.3). It would be of interest to know whether inequality  $r_{ess}(P) \leq \delta$  extends to other situations, and more generally whether the general bound in Theorem 5.2 can be improved.

**Acknowledgment** We wish to thank the referee for her/his useful advices and encouragements which allowed us to improve our article. We also thank Bernard Delyon for stimulating discussions concerning the results of Section 5.

## References

- [1] W. Arendt, A. Grabosch, G. Greiner, U. Groh, H. P. Lotz, U. Moustakas, R. Nagel, F. Neubrander, and U. Schlotterbeck. *One-parameter semigroups of positive operators*, volume 1184 of *Lecture Notes in Math.* Springer-Verlag, Berlin, 1986.

- [2] S. Asmussen. *Applied probability and queues*. Springer-Verlag, New York, 2003.
- [3] W. Bahsoun. Rigorous numerical approximation of escape rates. *Nonlinearity*, 19(11):2529–2542, 2006.
- [4] W. Bahsoun and C. Bose. Quasi-invariant measures, escape rates and the effect of the hole. *Discrete Contin. Dyn. Syst.*, 27(3):1107–1121, 2010.
- [5] W. Bahsoun and C. Bose. Invariant densities and escape rates: rigorous and computable approximations in the  $L^\infty$ -norm. *Nonlinear Anal.*, 74(13):4481–4495, 2011.
- [6] V. Baladi. *Positive transfer operators and decay of correlations*. World Scientific Publishing Co. Inc., River Edge, NJ, 2000.
- [7] P. H. Baxendale. Renewal theory and computable convergence rates for geometrically ergodic Markov chains. *Ann. Appl. Probab.*, 15(1B):700–738, 2005.
- [8] D. Delyon. Geometric convergence of Markov chains through a local Doeblin condition. Technical Report 1115, IRISA Rennes, 1997.
- [9] M. F. Demers and C. Liverani. Stability of statistical properties in two-dimensional piecewise hyperbolic maps. *Trans. Amer. Math. Soc.*, 360(9):4777–4814, 2008.
- [10] D. Ferré, L. Hervé, and J. Ledoux. Regular perturbation of  $V$ -geometrically ergodic Markov chains. *J. Appl. Probab.*, 50:184–194, 2013.
- [11] D. Guibourg, L. Hervé, and J. Ledoux. Quasi-compactness of Markov kernels on weighted-supremum spaces and geometrical ergodicity. arXiv : 1110.3240, 2011.
- [12] H. Hennion. Sur un théorème spectral et son application aux noyaux lipchitziens. *Proc. Amer. Math. Soc.*, 118:627–634, 1993.
- [13] H. Hennion. Quasi-compactness and absolutely continuous kernels, applications to Markov chains. arXiv:math/0606680, June 2006.
- [14] H. Hennion. Quasi-compactness and absolutely continuous kernels. *Probab. Theory Related Fields*, 139:451–471, 2007.
- [15] L. Hervé and J. Ledoux. Spectral analysis of Markov kernels and application to the convergence rate of discrete random walks. Revised for publication in *Applied Probab. J.*, 2013.
- [16] G. Keller and C. Liverani. Stability of the spectrum for transfer operators. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze Sér. 4*, 28:141–152, 1999.
- [17] I. Kontoyiannis and S. P. Meyn. Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab.*, 13(1):304–362, 2003.
- [18] I. Kontoyiannis and S. P. Meyn. Large deviations asymptotics and the spectral theory of multiplicatively regular Markov processes. *Electron. J. Probab.*, 10:61–123, 2005.
- [19] U. Krengel. *Ergodic Theorems*. de Gruyter, Berlin-New York, 1985.

- [20] Y. Liu. Augmented truncation approximations of discrete-time Markov chains. *Oper. Res. Lett.*, 38(3):218–222, 2010.
- [21] C. Liverani. Rigorous numerical investigation of the statistical properties of piecewise expanding maps. A feasibility study. *Nonlinearity*, 14(3):463–490, 2001.
- [22] R. B. Lund and R. L. Tweedie. Geometric convergence rates for stochastically ordered Markov chains. *Math. Oper. Res.*, 21(1):182–194, 1996.
- [23] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Springer-Verlag London Ltd., London, 1993.
- [24] S. P. Meyn and R. L. Tweedie. Computable bounds for geometric convergence rates of Markov chains. *Ann. Probab.*, 4:981–1011, 1994.
- [25] A. Y. Mitrophanov. Sensitivity and convergence of uniformly ergodic Markov chains. *J. Appl. Probab.*, 42(4):1003–1014, 2005.
- [26] J. Neveu. *Bases mathématiques du calcul des probabilités*. Masson, Paris, 1964.
- [27] E. Nummelin. *General irreducible Markov chains and nonnegative operators*. Cambridge University Press, Cambridge, 1984.
- [28] 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.
- [29] J. S. Rosenthal. Markov chain convergence: from finite to infinite. *Stochastic Process. Appl.*, 62(1):55–72, 1996.
- [30] H. H. Schaefer. *Topological vector spaces*. Springer, New York, 1971.
- [31] E. Seneta. *Nonnegative matrices and Markov chains*. Springer, New York, 1981.
- [32] T. Shardlow and A. M. Stuart. A perturbation theory for ergodic Markov chains and application to numerical approximations. *SIAM J. Numer. Anal.*, 37:1120–1137, 2000.
- [33] R. L. Tweedie. Truncation approximations of invariant measures for Markov chains. *J. Appl. Probab.*, 35(3):517–536, 1998.
- [34] L. Wu. Essential spectral radius for Markov semigroups. I. Discrete time case. *Probab. Theory Related Fields*, 128(2):255–321, 2004.
- [35] Y. Q. Zhao and D. Liu. The censored Markov chain and the best augmentation. *J. Appl. Probab.*, 33(3):623–629, 1996.
- [36] Y. Zheng. Stochastic perturbations of fat solenoidal attractors. *Acta Math. Sin.*, 26(5):923–936, 2010.