Truncated Markov chains and V-geometric ergodicity via weak perturbation theory

Loïc Hervé, James Ledoux

Truncated Markov chains and V-geometric ergodicity via weak perturbation theory

Loïc HERVÉ and James LEDOUX *

June 6, 2012

Abstract

Let P be a Markov kernel on a metric space $\mathbb{X}$ and let $V : \mathbb{X} \to [1, +\infty)$. This paper provides explicit connections between the V-geometrical ergodicity of P and that of its truncated (and augmented) kernels P_k. More specifically, we provide conditions under which, the geometrical ergodicity of P can be derived from that of P_k. A special attention is paid to obtain an efficient way to specify the convergence rate for P from that of P_k. The converse problems are also discussed. Furthermore, an explicit bound on the total variation between the invariant probability measures of P and P_k is provided. The proofs are based on the Keller-Liverani perturbation theorem which requires an accurate control of the essential spectral radius of both P and the extended kernel of P_k as linear operators on the Banach space $\mathcal{B}_1$ of all the measurable $\mathbb{C}$-valued functions f such that $\sup_{x\in\mathbb{X}} |f(x)|/V(x) < \infty$. Consequently, a part of this paper is devoted to the derivation of computable bounds on the essential spectral radius on $\mathcal{B}_1$ of a general Markov kernel from standard drift conditions.

AMS subject classification : 60J10; 47B07

Keywords : quasi-compact operator, drift condition.

1 Introduction

Let $(\mathbb{X}, d)$ be a metric space equipped with its Borel σ-algebra $\mathcal{X}$. For $A \in \mathcal{X}$ and $f : \mathbb{X} \to \mathbb{C}$, we denote by f_A the restriction of f to A. Let P be a Markov kernel on $(\mathbb{X}, \mathcal{X})$. We define the following truncation approximation of P.

Definition 1 *Let $x_0 \in \mathbb{X}$. For every $k \in \mathbb{N}^*$, we denote by B_k the closed ball centered at x_0 with radius k and ${B_k}^c := \mathbb{X}\backslash B_k$. The σ-algebra induced by $\mathcal{X}$ on B_k is called $\mathcal{X}_k$. Given*

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

$x_k \in \mathbb{X}$ such that $d(x_0, x_k) = k$, the k-th truncated (and augmented) Markov kernel P_k is defined on $(B_k, \mathcal{X}_k)$ by

$$\forall x \in B_k,\ \forall F \in \mathcal{X}_k, \quad P_k(x, F) := P(x, F) + 1_F(x_k)\, P(x, {B_k}^c). \tag{1}$$

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

$$\forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad \widehat{P}_k(x, A) := \begin{cases} P_k(x, A \cap B_k) & \text{if } x \in B_k \\ 0 & \text{if } x \in {B_k}^c. \end{cases} \tag{2}$$

For instance, when $\mathbb{X} := \mathbb{N}$, the k-th truncated (and augmented) kernel P_k of $P := (P(i,j))_{(i,j)\in\mathbb{N}^2}$ associated with $B_k := \{0, \ldots, k\}$ and $x_k := k$ is the following $(k+1) \times (k+1)$ matrix:

$$P_k = \begin{pmatrix} P(0,0) & \cdots & P(0,k-1) & \sum_{j\geq k} P(0,j) \\ \vdots & & \vdots & \vdots \\ P(k,0) & \cdots & P(k,k-1) & \sum_{j\geq k} P(k,j) \end{pmatrix}.$$

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 so-called censored Markov chain [ZL96], could be considered here with suitable modifications of Definition 1. Truncation approximation of an infinite stochastic matrix has a long story (e.g. see [Sen81, Twe98, Liu10] and reference therein).

Here we are interested in investigating the connection between the geometrical ergodicity of P and that of its truncated kernels P_k when $\mathbb{X}$ is any metric space. Namely, given $V : \mathbb{X} \to [1, +\infty)$, we study the link between the two following conditions.

Condition (V). *P is V-geometrically ergodic, that is: P has an invariant probability measure π on $(\mathbb{X}, \mathcal{X})$ such that $\pi(V) := \int_{\mathbb{X}} V(x)\pi(dx) < \infty$ and there exist some rate $\rho \in (0,1)$ and constant $C > 0$ such that for every measurable function $f : \mathbb{X} \to \mathbb{C}$ satisfying $|f| \leq V$*

$$\forall n \geq 1, \qquad \sup_{x\in\mathbb{X}} \frac{|(P^n f)(x) - \pi(f)|}{V(x)} \leq C\, \rho^n. \tag{V}$$

Condition (V_k). *P_k is V_{B_k}-geometrically ergodic, that is: P_k has an invariant probability measure π_k on $(B_k, \mathcal{X}_k)$ such that $\pi_k(V_{B_k}) < \infty$, and there exist some rate $\rho_k \in (0,1)$ and constant $C_k > 0$ such that for every measurable function $h : B_k \to \mathbb{C}$ satisfying $|h| \leq V_{B_k}$*

$$\forall n \geq 1, \qquad \sup_{x\in B_k} \frac{|({P_k}^n h)(x) - \pi_k(h)|}{V(x)} \leq C_k\, {\rho_k}^n. \tag{V_k}$$

Specifically, the two next issues are studied.

(Q1) When (V) holds, does the truncated Markov kernel P_k satisfy (V_k) for k large enough? Moreover, when (ρ, C) is known in (V), can we obtain explicit (ρ_k, C_k) in (V_k)? Finally can we then obtain an explicit bound for the total variation distance $\|\widehat{\pi}_k - \pi\|_{TV}$, where $\widehat{\pi}_k$ is the probability measure on $(\mathbb{X}, \mathcal{X})$ defined by: $\forall A \in \mathcal{X},\ \widehat{\pi}_k(1_A) := \pi_k(1_{A\cap B_k})$?

(Q2) Conversely, when (V_k) holds for k large enough, does P satisfy (V)? Moreover, when (ρ_k, C_k) is known in (V_k), can we obtain explicit (ρ, C) in (V)? Finally can we then obtain an explicit bound for the total variation distance $\|\widehat{\pi}_k - \pi\|_{TV}$?

Although (Q1) is interesting from a theoretical point of view, it is worth noticing that the computational issues involve rather (Q2). Indeed, since the state space of P_k is a closed ball in $\mathbb{X}$, finding rates ρ_k and constants C_k in (V_k) is expected to be easier than in (V). For instance, when the state space $\mathbb{X}$ is discrete, P_k is a finite stochastic matrix, while P is infinite.

Let $(\mathcal{B}_1, \|\cdot\|_1)$ be the weighted-supremum Banach space composed of measurable functions $f : \mathbb{X} \to \mathbb{C}$ such that

$$\|f\|_1 := \sup_{x \in \mathbb{X}} \frac{|f(x)|}{V(x)} < \infty.$$

In Section 3, Theorems 3 and 4 provide a positive answer to the issues (Q1) and (Q2) under the following condition.

Condition (P). *The Markov kernel P satisfies anyone of the two following properties.*

(PWD) *P is compact from $\mathcal{B}_0$ into $\mathcal{B}_1$, and P satisfies the following weak drift condition:*

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

(PSD) *The following so-called strong drift condition holds:*

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

where $S \in \mathcal{X}$ is bounded and satisfies the minorization condition:

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

for some positive measure ν on $(\mathbb{X}, \mathcal{X})$.

Inequalities (**WD**) and (**SD**) are well-known drift conditions introduced in [MT93] for studying the V-geometric ergodicity of P. Condition (PWD) does not involve a minorization condition but requires the compactness of $P : \mathcal{B}_0 \to \mathcal{B}_1$. In fact Condition (PSD) is quite more general than (PWD), but the bounds obtained for (Q1)-(Q2) under (PSD) are worst than those obtained under (PWD), excepted in some specific cases as, for instance, when the set S is an atom. Under Condition (PWD), the rate of convergence ρ derived in (Q2) is asymptotically optimal when $k \to +\infty$ provided that optimal rates ρ_k in (V_k) are known. Simple sufficient conditions for $P : \mathcal{B}_0 \to \mathcal{B}_1$ to be compact are presented in [GHL11]. For instance, this condition holds for any discrete Markov chains, for autoregressive models with ARCH(1) errors or for functional autoregressive models on $\mathbb{X} := \mathbb{R}^q$ with absolutely continuous noise with respect to the Lebesgue measure.

Considering the action of P and $\widehat{P}_k$ on $\mathcal{B}_1$, a natural way to solve (Q1) is to see $\widehat{P}_k$ as a perturbed operator of P, and vice versa to solve (Q2). The standard perturbation theory requires the continuity property

$$\|\widehat{P}_k - P\|_1 := \sup_{f \in \mathcal{B}_1, \|f\|_1 \leq 1} \|\widehat{P}_k f - Pf\|_1 \xrightarrow[k \to \infty]{} 0. \tag{$C_{1,1}$}$$

Unfortunately the last condition does not hold for unbounded function V under (**WD**) or (**SD**) even if $\mathbb{X}$ is discrete (see Remark 3). In other words, the standard perturbation theory cannot be applied to solve (Q1)-(Q2). Here we use the weak perturbation theory due to Keller and Liverani [KL99, Liv01] which only invokes the weakened continuity property

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

where $\mathcal{B}_0$ is the Banach space of the bounded measurable $\mathbb{C}$-valued functions on $\mathbb{X}$ equipped with the supremum norm $\|f\|_0 := \sup_{x \in \mathbb{X}} |f(x)|$. Actually Condition $(C_{0,1})$ holds provided that $V(x) \to +\infty$ when $d(x_0, x) \to +\infty$. The price to pay for using the weakened continuity property $(C_{0,1})$ is that two functional assumptions are needed in the weak perturbation method. The first one involves the so-called Doeblin-Fortet inequalities: here such dual inequalities are derived for P and $\widehat{P}_k$ from (**WD**) or (**SD**) (see Lemmas 1 and 3). The second functional assumption requires that the essential spectral radii of both P and $\widehat{P}_k$ on $\mathcal{B}_1$ are strictly less than one in a uniform way in k. The next estimate (3) obtained for the essential spectral radius $r_{ess}(P)$ of P under Condition (PWD) extends to a uniform control of $r_{ess}(\widehat{P}_k)$ from Lemmas 3 and 4. Similarly, Inequality (4) below obtained under Condition (PSD) extends to $r_{ess}(\widehat{P}_k)$ from Lemma 5. The two first questions in both (Q1)-(Q2) are then addressed using the first part of [Liv01, Prop. 3.1]. The question in (Q1)-(Q2) concerning $\|\widehat{\pi}_k - \pi\|_{TV}$ is solved by elementary spectral arguments (Lemma 6).

The study of the essential spectral radius of a Markov kernel acting on $\mathcal{B}_1$, in connection with its quasi-compactness property, plays a crucial role in the previous discussion. This study, which has its own interest, is presented in Section 2 for a Markov kernel P defined on a general state space. Specifically two bounds for $r_{ess}(P)$ are provided. The first one (Theorem 1) is obtained under Condition (PWD):

$$r_{ess}(P) \leq \delta. \qquad (3)$$

The second bound (Theorem 2) is obtained under Condition (PSD):

$$r_{ess}(P) \leq \frac{\delta\, \nu(1_{\mathbb{X}}) + \tau}{\nu(1_{\mathbb{X}}) + \tau} \qquad \text{where } \tau := \max(0, L - \nu(V)). \qquad (4)$$

Inequality (3) could be derived from [Wu04, Th. 3.11], but Wu's assumptions are more complicated and are not directly connected to those of [KL99] (in particular to the Doeblin-Fortet inequalities). To the best of our knowledge, the general bound in (4) is not known in the literature. This bound is worst than that in (3) because of the constant τ. In particular, if τ is too large, then the bound in (4) is too close to 1 (i.e. to the spectral radius of P) to provide good bounds in (Q1)-(Q2). However, when $\tau = 0$, which occurs for instance when the set S is an atom (see Remark 2), Inequality (4) reduces to the expected one $r_{ess}(P) \leq \delta$, so that the resulting bounds in (Q1)-(Q2) are similar to those derived from (PWD).

Let us close this introduction with a brief review of previous related works. Recall that an irreducible and aperiodic Markov kernel P is V-geometrical ergodic if the strong drift condition (PSD) is satisfied (see e.g. [MT93]). Various probabilistic methods have been developed to derive explicit rate and constant (ρ, C) in (V) from the constants δ, L and the set S (see [MT94, LT96, Bax05] and the references therein). To the best of our knowledge, these

methods, which are not concerned with truncation approximation, provide a computable rate ρ which is often unsatisfactory except for stochastically monotone P.

For truncation approximation with a discrete state space $\mathbb{X}$, the difference $\widehat{\pi}_k - \pi$ has been investigated under (V) in [Twe98], where it is proved that each matrix P_k has an invariant probability measure π_k such that the V-norm $\sup_{|f|\leq V} |\widehat{\pi}_k(f) - \pi(f)|$ goes to 0 when $k \to +\infty$. In particular $\|\widehat{\pi}_k - \pi\|_{TV}$ converges to 0. Moreover, when P is stochastically monotone (and V is increasing), the following rate of convergence is obtained in [Twe98, Th. 4.2,(46)]

$$\forall m \in \mathbb{N}^*,\ \forall k \in \mathbb{N}^*, \quad \|\widehat{\pi}_k - \pi\|_{TV} \leq c\,\delta^m + \frac{d\,m}{V(k)},$$

with explicit constants c, d only depending on constants involved in the strong drift condition. A similar result has been obtained for polynomially ergodic Markov chains in [Liu10]. To the best of our knowledge, these results have not been extended to non-discrete $\mathbb{X}$. In particular, for general Markov kernels, there are no known conditions ensuring that $\lim_k \|\widehat{\pi}_k - \pi\|_{TV} = 0$, and no known rates in the previous convergence, excepted the above mentioned result concerning stochastically monotone kernel on discrete $\mathbb{X}$. The two first parts of both issues (Q1) and (Q2) are not addressed in [Twe98, Liu10].

A weak perturbation approach, based on the refinement [Liv04] of the Keller-Liverani theorem, has been used in [FHL11] to study general perturbed Markov kernels (not necessarily defined by truncation) of a V-geometrically ergodic Markov kernel. When applied to our context, under Condition (V), [FHL11, Th. 1] gives a positive answer to (Q1). The explicit connection between (ρ_k, C_k) and (ρ, C) is not addressed in [FHL11].

The weak perturbation results of [KL99, Liv01] have been fully used in the framework of dynamical systems (e.g. see [Bah06, DL08, BB10, BB11, Zhe10]). There, Markov kernels are replaced by the so-called Perron-Frobenius operators, and the investigation mainly concerns the perturbed SRB measures. We want to point out that, in contrast with the previously cited papers, the second part of [Liv01, Prop. 3.1] and the spectral rank-stability property [Liv01, Cor. 3.1] are not invoked in our work to solve (Q1)-(Q2). To understand why this difference is relevant, let us turn back to our truncation issues. The use of the second part of [Liv01, Prop. 3.1] and of [Liv01, Cor. 3.1] would require to deal with truncated kernels P_k for $k \geq k_0$, with some finite but very large integer k_0. In fact, the first part of [Liv01, Prop. 3.1] allows us to solve (Q1)-(Q2) for truncated kernels P_k such that $k \geq k_1$ with some integer $k_1 << k_0$ (in substance $k_1 \approx {k_0}^{1/3}$), provided that ρ and ρ_k satisfy some mild additional assumptions. For instance, these additional conditions in Theorem 4 to solve (Q2) are: $\rho \leq 1 - \vartheta$ and $\rho_k \leq 1 - 2\vartheta$ for some arbitrary small $\vartheta > 0$. Of course, such a condition on ρ_k is expected to be easily checked. When $\mathbb{X}$ is discrete, ρ_k is linked to the second eigenvalue of the finite matrix P_k. The condition on ρ is not contradictory to Issue (Q2) since, in most of examples, some preliminary bound of ρ is available. Even if this bound is close to 1, such information is enough to pick $\vartheta > 0$ as stated above, so that Theorem 4 can be applied to find a better rate in (V) (close to ρ_k). Finally mention that Lemma 6 is of interest in our approach since it provides explicit bounds for $\|\widehat{\pi}_k - \pi\|_{TV}$ as soon as (V) and (V_k) hold with known rates and constants. This new approach could be used in any application of the weak perturbation method. In particular, for issue (Q2) in the context of dynamical systems, the whole results of [Liv01] can be used in a first stage to obtain a preliminary (possibly poor) rate of convergence

for the iterates of the unperturbed operator. Then, applying the method of Theorem 4 in a second stage will provide a new (better in general) rate of convergence with an explicit associated constant. Finally applying the method of Lemma 6 will give an explicit bound for the difference between the perturbed and unperturbed invariant measures, which is better than the one derived from the spectral rank-stability property of [Liv01].

2 Quasi-compactness on $\mathcal{B}_1$ and drift conditions

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

The simplest definition of quasi-compactness is the following (to be compared with the reduction of matrices or compact operators).

Definition 2 *T is quasi-compact on $\mathcal{B}$ if there exist $r_0 \in (0,1)$ and $m \in \mathbb{N}^*$, $\lambda_i \in \mathbb{C}$, $p_i \in \mathbb{N}^*$ $(i = 1, \ldots, m)$ such that:*

$$\mathcal{B} = \bigoplus_{i=1}^{m} \mathrm{Ker}(T - \lambda_i I)^{p_i} \oplus H, \tag{5a}$$

where the λ_i's are such that

$$|\lambda_i| \geq r_0 \quad \text{and} \quad 1 \leq \dim \mathrm{Ker}(T - \lambda_i I)^{p_i} < \infty, \tag{5b}$$

and H is a closed T-invariant subspace such that

$$\sup_{h \in H,\, \|h\| \leq 1} \|T^n h\| = O({r_0}^n). \tag{5c}$$

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

$$r_{ess}(T) := \inf \big\{ r_0 \in (0,1) \text{ s.t. we have (5a), (5b), (5c)} \big\}.$$

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

$$r_{ess}(T) := \lim_n \big(\inf \|T^n - K\| \big)^{\frac{1}{n}} \tag{6}$$

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

$$r_{ess}(T) \leq (r(T^{n_0} - K_0))^{1/n_0}. \tag{7}$$

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

Throughout this section, we consider a general measurable space $(\mathbb{X}, \mathcal{X})$, a function $V : \mathbb{X} \to [1, +\infty)$ and a Markov kernel Q on $(\mathbb{X}, \mathcal{X})$ such that $\|QV\|_1 < \infty$. So Q continuously acts on $\mathcal{B}_1$. The first bound for $r_{ess}(Q)$ is derived by duality from the quasi-compactness criteria of [Hen93]. The second one is derived from a result on positive operators [Sch71].

2.1 Quasi-compactness on $\mathcal{B}_1$ under a weak drift condition

The key idea to obtain quasi-compactness in Theorem 1 below is that, under the following weak drift condition

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

the adjoint operator of the kernel Q satisfies a Doeblin-Fortet inequality on the dual space of $\mathcal{B}_1$ (Lemma 1). This remark allows us to greatly simplify the arguments used in [Wu04] since the well known statement [Hen93, Cor. 1] gives the bound $r_{ess}(Q) \leq \delta$ provided that Q^ℓ is compact from $\mathcal{B}_0$ to $\mathcal{B}_1$ for some $\ell \geq 1$. Furthermore, this last compactness condition is much simpler than the assumptions of [Wu04] based on sophisticated parameters $\beta_w(Q)$ and $\beta_\tau(Q)$ as measure of non-compactness of Q. More precise comparisons with [Wu04] and complements are presented in [GHL11, Sect. 2.3].

Under (WDN), the infimum $\delta_V(Q)$ of the real numbers $\delta \in [0,1)$ such that (WDN) holds is well defined:

$$\delta_V(Q) := \inf \left\{ \delta \in [0,1) : \exists N \in \mathbb{N}^*,\, \exists L > 0,\ Q^N V \leq \delta^N V + L\, 1_{\mathbb{X}} \right\}. \tag{8}$$

Theorem 1 *If Condition* (WDN) *holds true and if $Q^\ell : \mathcal{B}_0 \to \mathcal{B}_1$ is compact for some $\ell \geq 1$, then Q is a power-bounded quasi-compact operator on $\mathcal{B}_1$, and we have*

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

Proof. Iterating (WDN) easily ensures that $\sup_k \|Q^{kN} V\|_1 < \infty$. Thus Q is power-bounded on $\mathcal{B}_1$, that is $\sup_n \|Q^n V\|_1 < \infty$ (use the Euclidean division of $n \in \mathbb{N}$ by N). Since $\delta_V(Q) = (\delta_V(Q^\ell))^{1/\ell}$ and $r_{ess}(Q) = (r_{ess}(Q^\ell))^{1/\ell}$, we only consider the case $\ell := 1$, that is $Q : \mathcal{B}_0 \to \mathcal{B}_1$ is compact.

Now let $(\mathcal{B}_1', \|\cdot\|_1)$ (resp. $(\mathcal{B}_0', \|\cdot\|_0)$) denote the dual space of $\mathcal{B}_1$ (resp. of $\mathcal{B}_0$). Note that we make a slight abuse of notation in writing again $\|\cdot\|_1$ and $\|\cdot\|_0$ for the dual norms. Let Q' denote the adjoint operator of Q on $\mathcal{B}_1'$. In fact, we prove that Q' is a quasi-compact operator on $\mathcal{B}_1'$ with $r_{ess}(Q') \leq \delta_V(Q)$, so that Q satisfies the same properties on $\mathcal{B}_1$. The operator $Q : \mathcal{B}_0 \to \mathcal{B}_1$ is assumed to be compact, so is $Q' : \mathcal{B}_1' \to \mathcal{B}_0'$. Moreover Q' satisfies a Doeblin-Fortet inequality from Lemma 1 below. Then we deduce from Lemma 1 and [Hen93, Cor. 1] that Q' is a quasi-compact operator on $\mathcal{B}_1'$, with $r_{ess}(Q') \leq \delta$ for any $\delta \in (\delta_V(Q), 1)$, so that $r_{ess}(Q') \leq \delta_V(Q)$. □

Lemma 1 *Let $\delta \in (\delta_V(Q), 1)$. Then, there exist $N \in \mathbb{N}^*$ and $L > 0$ such that for all $f' \in \mathcal{B}_1'$ we have:*

$$\|{Q'}^N f'\|_1 \leq \delta^N \|f'\|_1 + L \|f'\|_0.$$

Proof. Recall that $\mathcal{B}_1, \mathcal{B}_0$ are Banach lattices, so are $\mathcal{B}_1'$, $\mathcal{B}_0'$. For each $g' \in \mathcal{B}_1'$ (resp. $g' \in \mathcal{B}_0'$), one can define the modulus $|g'|$ of g' in $\mathcal{B}_1'$ (resp. in $\mathcal{B}_0'$), see [Sch71]. For the next arguments,

it is enough to have in mind that g' and $|g'|$ have the same norm in $\mathcal{B}_1'$ (resp. in $\mathcal{B}_0'$), more precisely:

$$\forall g' \in \mathcal{B}_0', \quad \|g'\|_0 = |g'|(1_{\mathbb{X}}) \quad \text{and} \quad \forall g' \in \mathcal{B}_1', \quad \|g'\|_1 = |g'|(V).$$

Let $f' \in \mathcal{B}_1'$ and $n \geq 1$. Since Q^n is a nonnegative operator on $\mathcal{B}_1$, so is its adjoint operator Q'^n on $\mathcal{B}_1'$, and we have for all $f \in \mathcal{B}_1$ such that $\|f\|_1 \leq 1$ (i.e. $|f| \leq V$):

$$\left|(Q'^n f')(f)\right| \leq \left(Q'^n |f'|\right)(|f|) \leq \left(Q'^n |f'|\right)(V) = |f'|(Q^n V).$$

From the definition of $\delta_V(Q)$ and from $\delta \in (\delta_V(Q), 1)$, there exist $N \in \mathbb{N}^*$ and $L > 0$ such that $Q^N V \leq \delta^N V + L\, 1_{\mathbb{X}}$. Thus

$$\begin{array}{rcl} \|Q'^N f'\|_1 & := & \sup\limits_{f \in \mathcal{B}_1, \|f\|_1 \leq 1} \left|(Q'^N f')(f)\right| \\ & \leq & |f'|(Q^N V) \\ & \leq & \delta^N\, |f'|(V) + L\, |f'|(1_{\mathbb{X}}) = \delta^N \|f'\|_1 + L\|f'\|_0. \end{array}$$

□

2.2 Quasi-compactness on $\mathcal{B}_1$ under the strong drift condition

Theorem 2 below gives a simple bound of $r_{ess}(Q)$ in terms of the parameters involved in the strong drift condition.

Theorem 2 *Assume that*

$$\exists \delta \in (0,1),\ \exists L > 0, \quad QV \leq \delta\, V + L\, 1_S \tag{9}$$

with $S \in \mathcal{X}$ satisfying

$$\forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad Q(x, A) \geq \nu(1_A)\, 1_S(x) \tag{10}$$

for some positive measure ν on $(\mathbb{X}, \mathcal{X})$. Then Q is a power-bounded quasi-compact operator on $\mathcal{B}_1$ with

$$r_{ess}(Q) \leq \frac{\delta\, \nu(1_{\mathbb{X}}) + \tau}{\nu(1_{\mathbb{X}}) + \tau} \quad \textit{where } \tau := \max(0, L - \nu(V)). \tag{11}$$

It is well-known (see [MT93]) that, under usual irreducibility and aperiodicity assumptions, Conditions (9)-(10) ensure the V-geometrical ergodicity of Q (i.e. $\|Q^n - \pi(\cdot)1_{\mathbb{X}}\|_1 = O(\rho^n)$ for some $\rho < 1$ where π denotes the Q-invariant probability measure). The V-geometrical ergodicity of Q obviously implies that Q is quasi-compact on $\mathcal{B}_1$ with $r_{ess}(Q) \leq \rho$ (see e.g. [KM03]). This follows from Definition 2 with $H := \{f \in \mathcal{B}_1 : \pi(f) = 0\}$. However the bound $r_{ess}(Q) \leq \rho$ has no interest for our truncation purpose. In fact, some good estimations on the rate ρ in (V) is expected to be known for Issue (Q1), but such known value of ρ may be much greater than $r_{ess}(P)$ (even the best rate of convergence may be strictly greater than $r_{ess}(P)$). More importantly, Issue (Q2) is only interesting when no good estimation on ρ is known. Consequently the results of both Theorem 3 and 4 will be all the more precise that an accurate upper bound

of $r_{ess}(P)$ is known. In the unpublished paper [Hen06, Th. IV.2], the quasi-compactness of Q on $\mathcal{B}_1$ is directly derived from conditions (9)-(10), and the V-geometrical ergodicity of Q is then obtained as a corollary under the additional irreducibility and aperiodicity assumptions (see [Hen06, Cor. IV.3] and the related classical references therein). Unfortunately the bound obtained for $r_{ess}(Q)$ in [Hen06, Th. IV.2] is less tractable since it is expressed in terms of the hitting time in S.

Theorem 2 provides a much simpler bound of $r_{ess}(Q)$ in terms of the parameters of (9)-(10). Moreover the short proof of Theorem 2, based on Lemma 2 below, illuminates the role of the drift and minorization conditions to obtain good spectral properties of Q on $\mathcal{B}_1$. In particular, using [Hen06, Cor. IV.3], Theorem 2 provides a simple proof of the fact that Conditions (9)-(10), together with irreducibility and aperiodicity assumptions, imply the V-geometrical ergodicity of Q.

Lemma 2 *Assume that $Q = A + B$ for some nonnegative bounded linear operators A and B on $\mathcal{B}_1$. Let $r(B)$ denote the spectral radius of B which is assumed to be positive. Then, there exists a nontrivial nonnegative continuous linear form η on $\mathcal{B}_1$ such that $\eta \circ B = r(B)\,\eta$ and $\eta(A1_{\mathbb{X}}) = (1 - r(B))\eta(1_{\mathbb{X}})$.*

Proof. Since $B \geq 0$ and $r := r(B) > 0$, we know from [Sch71, App., Cor.2.6] that there exists a nontrivial nonnegative continuous linear form η on $\mathcal{B}_1$ such that $\eta \circ B = r\,\eta$. From $Q = A + B$, we have $\eta \circ Q = \eta \circ A + r\,\eta$, thus $\eta(Q1_{\mathbb{X}}) = \eta(1_{\mathbb{X}}) = \eta(A1_{\mathbb{X}}) + r\,\eta(1_{\mathbb{X}})$. Hence $\eta(A1_{\mathbb{X}}) = (1 - r)\eta(1_{\mathbb{X}})$. □

Proof of Theorem 2. Condition (9) implies that $QV \leq \delta\, V + L\, 1_{\mathbb{X}}$. Thus Q is power-bounded on $\mathcal{B}_1$. Then, from $Q1_{\mathbb{X}} = 1_{\mathbb{X}}$ and $1_{\mathbb{X}} \in \mathcal{B}_1$, we have $r(Q) = 1$. Moreover, since $\|QV\|_1 < \infty$, we deduce from (10) that $\nu(V) < \infty$. Thus we can define the following rank-one operator on $\mathcal{B}_1$: $Tf := \nu(f)\, 1_S$. Set $R := Q - T$. From $T \geq 0$ and from (10), it follows that $0 \leq R \leq Q$, so $r(R) \leq 1$. Set $r := r(R)$. If $r = 0$, then Q is quasi-compact with $r_{ess}(Q) = 0$ from (7). Now assume that $r \in (0, 1]$. Then, from Lemma 2, there exists a nontrivial nonnegative continuous linear form η on $\mathcal{B}_1$ such that $\eta \circ R = r\,\eta$ and $\eta(T1_{\mathbb{X}}) = (1 - r)\eta(1_{\mathbb{X}})$, from which we deduce that

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

Next, we have $RV = QV - TV = QV - \nu(V)1_S \leq \delta V + L1_S - \nu(V)1_S = \delta\, V + (L - \nu(V))\, 1_S$. Hence, setting $\tau := \max(0, L - \nu(V)) \geq 0$,

$$r\,\eta(V) = \eta(RV) \leq \delta\,\eta(V) + \tau\,\eta(1_S) \leq \delta\,\eta(V) + \tau\frac{(1-r)\eta(V)}{\nu(1_{\mathbb{X}})}. \tag{12}$$

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 Q is quasi-compact from (7) with $r_{ess}(Q) \leq r(Q - T) = r$. Then (12) gives (11). □

Remark 1 *If Conditions (9)-(10) are fulfilled for some iterate Q^N in place of Q (with parameters $\delta_N < 1$, $L_N > 0$ and positive measure $\nu_N(\cdot)$), then the conclusions of Theorem 2 hold true with (11) replaced by*

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

Remark 2 *Recall that $A \in \mathcal{X}$ is said to be an atom with respect to Q if*

$$\forall (a, a') \in A, \quad Q(a, \cdot) = Q(a', \cdot).$$

Any Markov model having a regenerative structure is concerned with such a property (e.g. see [Num84, Asm03]). If Q satisfies (9) with $S := A$, then

$$r_{ess}(Q) \leq \delta.$$

Indeed, note that the minorization condition (10) holds with A and $\nu(1.) := Q(a_0, \cdot)$ for any $a_0 \in A$. Choose $L := \sup_{x \in A}(QV)(x)$ in (9). Since A is an atom, we have $L = (QV)(a_0) = \nu(V)$ so that $\tau = 0$ in (11).

3 Truncation and V-geometric ergodicity

In this section, we consider a metric space $(\mathbb{X}, d)$ equipped with its Borel σ-algebra $\mathcal{X}$. Function $V : \mathbb{X} \to [1, +\infty)$ is assumed to be of the form

$$\forall x \in \mathbb{X}, \quad V(x) := v(d(x, x_0))$$

for some unbounded increasing function $v : [0, \infty) \to [1, +\infty)$ and some $x_0 \in \mathbb{X}$. Moreover P is a Markov kernel on $(\mathbb{X}, \mathcal{X})$ and we consider the associated truncation approximation family $\{P_k\}_{k \geq 1}$ (see Definition 1). Recall that P_k extends to a sub-Markov kernel $\widehat{P}_k$ on $(\mathbb{X}, \mathcal{X})$. For $k \geq 1$, the restriction V_{B_k} of V to $B_k := \{x \in \mathbb{X} \ : \ d(x, x_0) \leq k\}$ is denoted by V_k. We use the weak and strong drift conditions (**WD**) (which correspond to (WDN) with $N := 1$) and (**SD**). Note that (**SD**) implies (**WD**).

We denote by $\lfloor \cdot \rfloor$ the integer part function on $\mathbb{R}$. The next functional notations are used throughout the section.

Functional notations. *For $\beta \in [0, 1]$, $(\mathcal{B}_\beta, \| \cdot \|_\beta)$ denotes the Banach space of measurable functions $f : \mathbb{X} \to \mathbb{C}$ such that*

$$\|f\|_\beta := \sup_{x \in \mathbb{X}} V(x)^{-\beta} |f(x)| < \infty.$$

In particular $\mathcal{B}_0$ corresponds to the space of bounded measurable functions on $\mathbb{X}$, with $\|f\|_0 = \sup_{x \in \mathbb{X}} |f(x)|$, and $\mathcal{B}_1$ is the space defined in Introduction. We denote by $(\mathcal{L}(\mathcal{B}_\beta, \mathcal{B}_{\beta'}), \| \cdot \|_{\beta, \beta'})$ the space of all the bounded linear maps from $\mathcal{B}_\beta$ to $\mathcal{B}_{\beta'}$, equipped with its usual norm:

$$\|T\|_{\beta, \beta'} = \sup \big\{ \|Tf\|_{\beta'}, \, f \in \mathcal{B}_\beta, \, \|f\|_\beta \leq 1 \big\}.$$

We write $\mathcal{L}(\mathcal{B}_\beta)$ for $\mathcal{L}(\mathcal{B}_\beta, \mathcal{B}_\beta)$ and $\|T\|_\beta$ for $\|T\|_{\beta, \beta}$ which is a slight abuse of notation. Finally, under Conditions (V) and (V_k), we set

$$\|\widehat{\pi}_k - \pi\|_\beta := \sup_{\|f\|_\beta \leq 1} |\widehat{\pi}_k(f) - \pi(f)|. \tag{13}$$

3.1 Preliminary lemmas

The next lemma will be repeatedly used in this work. Under (**WD**) (thus also under (**SD**)), we set for every $\beta \in [0,1]$

$$M_\beta \equiv M_\beta(\delta, L) := \begin{cases} \dfrac{1-\delta^\beta + L^\beta}{1-\delta^\beta} & \text{if } \beta \in (0,1] \\ 1 & \text{if } \beta = 0. \end{cases}$$

Lemma 3 *Under* (**WD**) *with parameters* (δ, L)*, the two following assertions hold.*

1. *We have for every* $\beta \in [0,1)$*:*

$$\forall k \in \mathbb{N}^*, \quad \Delta_{\beta,1} := \|\widehat{P}_k - P\|_{\beta,1} \leq \frac{\max\left(2(\delta+L), (\delta+L)^\beta\right)}{v(k)^{1-\beta}}.$$

   In particular: $\lim_k \|\widehat{P}_k - P\|_{\beta,1} = 0$.

2. *For any* $k \in \mathbb{N}^*$*,* $\widehat{P}_k$ *also satisfies Condition* (**WD**) *with parameters* (δ, L)*:*

$$\forall k \in \mathbb{N}^*, \quad \widehat{P}_k V \leq \delta V + L\, 1_{\mathbb{X}}. \tag{14}$$

Proof. Let $f : \mathbb{X} \to \mathbb{C}$ be such that $\|f\|_\beta \leq 1$. First, using that $PV \leq (\delta + L)\, V$ from Condition (**WD**), we obtain for every $x \in B_k$:

$$\begin{aligned} \left|(Pf)(x) - (\widehat{P}_k f)(x)\right| &= \Big| \int_{B_k{}^c} f(y) P(x,dy) - \int_{B_k{}^c} f(x_k) P(x,dy) \Big| \\ &\leq 2 \int_{B_k{}^c} V(y)^\beta P(x,dy) = 2 \int_{B_k{}^c} V(y) \frac{1}{V(y)^{1-\beta}} P(x,dy) \\ &\leq \frac{2}{v(k)^{1-\beta}} (PV)(x) \\ &\leq \frac{2(\delta+L)}{v(k)^{1-\beta}} V(x). \end{aligned}$$

Second, for $x \in B_k{}^c$, we obtain from Jensen's inequality

$$\begin{aligned} \left|(Pf)(x) - (\widehat{P}_k f)(x)\right| = |(Pf)(x)| \leq (PV^\beta)(x) &\leq (PV(x))^\beta \\ &\leq V(x)^\beta (\delta+L)^\beta \\ &\leq \frac{(\delta+L)^\beta}{v(k)^{1-\beta}} V(x). \end{aligned}$$

This proves the first assertion of the lemma. Let us check that $\widehat{P}_k$ satisfies (14). If $x \in B_k$, then

$$(\widehat{P}_k V)(x) = \int_{B_k} V(y) P(x,dy) + \int_{B_k{}^c} V(x_k) P(x,dy) \leq \int_{B_k} V(y) P(x,dy) + \int_{B_k{}^c} V(y) P(x,dy)$$

so $(\widehat{P}_k V)(x) \leq (PV)(x) \leq \delta V(x) + L$. If $x \in B_k{}^c$, then: $(\widehat{P}_k V)(x) = 0$. □

Remark 3 *Let us explain why the continuity property* $(C_{1,1})$*, that is* $\lim_k \|P - \widehat{P}_k\|_1 = 0$*, does not hold under the drift condition* (**WD**) *or* (**SD**)*. As claimed in Introduction such a failure motivates the use of the weak perturbation method. Using* $\|Pf - \widehat{P}_k f\|_1 \le \|P - \widehat{P}_k\|_1$ *with* $f := V1_{B_k}$ *and* $f := V1_{B_k{}^c}$ *(recall that* $\widehat{P}_k f = 0$ *on* $B_k{}^c$*), we obtain*

$$\sup_{x\in B_k{}^c} \frac{(PV)(x)}{V(x)} \le \sup_{x\in B_k{}^c} \frac{P(V1_{B_k})(x)}{V(x)} + \sup_{x\in B_k{}^c} \frac{P(V1_{B_k{}^c})(x)}{V(x)} \le 2\|P - \widehat{P}_k\|_1.$$

Assume that $(C_{1,1})$ *holds. Then* $\lim_k \sup_{x\in B_k{}^c}(PV)(x)/V(x) = 0$ *which is impossible when* $V(x) \to +\infty$ *as* $d(x_0, x) \to +\infty$ *since this limit is equal to* δ *from* (**WD**) *or* (**SD**)*.*

The following lemma concerns the compactness property of $\widehat{P}_k : \mathcal{B}_0 \to \mathcal{B}_1$.

Lemma 4 *If* P *is compact from* $\mathcal{B}_0$ *to* $\mathcal{B}_1$*, then so is each extended sub-Markov operator* $\widehat{P}_k$*.*

Proof. For the sake of simplicity, the index k in $\widehat{P}_k$, B_k and x_k is omitted. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in the unit ball of $\mathcal{B}_0$. Since P is compact from $\mathcal{B}_0$ into $\mathcal{B}_1$ and $(f_n(x))_{n\in\mathbb{N}}$ is bounded in $\mathbb{C}$, there exists a subsequence $n_j \uparrow +\infty$ such that $\lim_k P(1_B f_{n_j}) = g$ in $\mathcal{B}_1$ and $\lim_k f_{n_j}(x) = z_0$ for some $g \in \mathcal{B}_1$ and $z_0 \in \mathbb{C}$. From $\widehat{P} f_{n_j} = 1_B[P(1_B f_{n_j}) + f_{n_j}(x)\, P(\cdot, B^c)]$, it follows that $\lim_k \widehat{P} f_{n_j} = 1_B(g + z_0 P(\cdot, B^c))$ in $\mathcal{B}_1$. Thus $\widehat{P}$ is compact from $\mathcal{B}_0$ into $\mathcal{B}_1$. □

Under Condition (PSD), a uniform control in k of $r_{ess}(\widehat{P}_k)$ is given in the next lemma.

Lemma 5 *Assume that* P *satisfies* (PSD)*. Let* $k_S \ge 1$ *be such that* $\forall k \ge k_S$*,* $S \subset B_k$ *and* $\nu(1_{B_k}) > 0$*. Then, with* $\tau_k := \max(0, L - \nu(V1_{B_k}))$ *for any* $k \ge 1$*, we have*

$$\forall k \ge k_S, \quad r_{ess}(\widehat{P}_k) \le \frac{\delta\, \nu(1_{B_k}) + \tau_k}{\nu(1_{B_k}) + \tau_k} \le \frac{\delta\, \nu(1_{B_{k_S}}) + \tau_{k_S}}{\nu(1_{B_{k_S}}) + \tau_{k_S}}. \tag{15}$$

Proof. P_k satisfies the strong drift condition w.r.t. V_k and set S. Indeed, let $x \in B_k$. First we have for any $F \in \mathcal{X}_k$: $P_k(x, F) \ge P(x, F) \ge \nu(1_F)1_S(x)$. Thus P_k satisfies the minorization condition with set S and restriction of ν to $(B_k, \mathcal{X}_k)$. Second, we obtain $P(V1_{B_k^c})(x) \ge V(x_k)P(x, B_k^c)$. Next, from

$$(P_k V_k)(x) = (PV)(x) - (P(V1_{B_k^c}))(x) + V(x_k)P(x, B_k^c),$$

it follows that $(P_k V_k)(x) \le (PV)(x) \le \delta V_k(x) + L1_S(x)$. Applying Theorem 2 to P_k gives $r_{ess}(P_k) \le (\delta\, \nu(1_{B_k}) + \tau_k)/(\nu(1_{B_k}) + \tau_k)$, where $r_{ess}(P_k)$ is the essential spectral radius of P_k acting on the Banach space composed of measurable functions $f : B_k \to \mathbb{C}$ such that $\sup_{x\in B_k} |f(x)|/V(x) < \infty$ (which defines the norm on this space). The last inequality in (15) follows from the monotonicity of maps $t \mapsto (\delta\, t + \tau_k)/(t + \tau_k)$ and $t \mapsto (\delta\, \nu(1_{B_{k_S}}) + t)/(\nu(1_{B_{k_S}}) + t)$. Then $r_{ess}(\widehat{P}_k)$ satisfies (15) from the definition of $\widehat{P}_k$. □

The next lemma is relevant to solve the last question of Issues (Q1)-(Q2).

Lemma 6 *Assume that Conditions* (**WD**)*,* (V) *and* (V_k) *are fulfilled. Then the following assertions hold.*

(a) For any $\beta \in [0,1)$ we have:

$$\forall R>1,\ \forall n\geq 1,\quad \|\widehat{\pi}_k-\pi\|_\beta \leq \frac{C_k\,{\rho_k}^n + C\rho^n}{\|1_{\mathbb{X}}\|_1} \ + \ \frac{d_\beta}{(R-1)^2}\,\frac{R^{n+1}}{v(k)^{1-\beta}} \tag{16a}$$

$$\text{with } d_\beta \equiv d_\beta(\delta,L) := \|1_{\mathbb{X}}\|_1^{-1} M_\beta M_1 \max\big(2(\delta+L),(\delta+L)^\beta\big). \tag{16b}$$

In particular, if (V_k) holds for every $k\geq 1$ with $\sup_k C_k<\infty$ and $\sup_k \rho_k<1$, then we have $\lim_k \|\widehat{\pi}_k-\pi\|_\beta=0$ for any $\beta\in[0,1)$.

(b) Setting $\hat{\rho}_k := \max(\rho,\rho_k)$, we have for any $\sigma\in(0,1]$:

$$\|\widehat{\pi}_k-\pi\|_{TV} \leq \left(\frac{C_k+C}{\hat{\rho}_k\|1_{\mathbb{X}}\|_1} + \frac{d_0\,\hat{\rho}_k^{\,\sigma}}{(1-\hat{\rho}_k^{\,\sigma})^2}\right)\frac{1}{v(k)^{1/(1+\sigma)}}. \tag{17}$$

Proof. Note that $\widehat{P}_k V^\beta \leq PV^\beta \leq \delta^\beta V^\beta + L^\beta$ for any $\beta\in(0,1]$ (use (**WD**) and Jensen's inequality when $\beta\in(0,1)$). Iterating this inequality gives

$$\forall \beta\in(0,1],\quad \max\big(\sup_{n\geq 0}\|P^n\|_\beta, \sup_{n\geq 0}\|\widehat{P}_k^{\,n}\|_\beta\big) \leq M_\beta. \tag{18}$$

Therefore the spectral radius of P and $\widehat{P}_k$ on both $\mathcal{B}_\beta$ and $\mathcal{B}_1$ is less than 1 (in fact the spectral radius of P is equal to 1 since P is Markov). If $\beta:=0$, then (18) is fulfilled with $M_0=1$ since $\widehat{P}_k 1_{\mathbb{X}} \leq P1_{\mathbb{X}} = 1_{\mathbb{X}}$.

Now let $R\in(1,+\infty)$ and let $\beta\in[0,1]$. From the equality $(zI-T)^{-1}=\sum_{n\geq 0} z^{-(n+1)}T^n$ applied to $T:=P$ and $T:=\widehat{P}_k$ with $|z|=R$, we obtain

$$\max\left(\sup_{z\in\mathbb{C}:|z|=R}\|(zI-\widehat{P}_k)^{-1}\|_\beta\ ,\ \sup_{z\in\mathbb{C}:|z|=R}\|(zI-P)^{-1}\|_\beta\right) \leq \frac{M_\beta}{R-1}.$$

Moreover the spectral theory gives

$$\widehat{P}_k^{\,n}-P^n = \frac{1}{2i\pi}\oint_{|z|=R} z^n\left[(zI-\widehat{P}_k)^{-1}-(zI-P)^{-1}\right]dz \tag{19}$$

and the following equality is obvious:

$$(zI-\widehat{P}_k)^{-1}-(zI-P)^{-1} = (zI-\widehat{P}_k)^{-1}(\widehat{P}_k-P)\,(zI-P)^{-1}.$$

Thus we obtain for every $f\in\mathcal{B}_\beta$

$$\|(\widehat{P}_k^{\,n}-P^n)f\|_1 \leq \frac{M_\beta M_1}{(R-1)^2}\,R^{n+1}\,\|\widehat{P}_k-P\|_{\beta,1}\,\|f\|_\beta.$$

Then using the triangle inequality, it follows that

$$\begin{aligned}
\big|\widehat{\pi}_k(f)-\pi(f)\big|\,\|1_{\mathbb{X}}\|_1 &= \|\widehat{\pi}_k(f)1_{\mathbb{X}}-\pi(f)1_{\mathbb{X}}\|_1 \\
&\leq \|\widehat{\pi}_k(f)1_{\mathbb{X}}-\widehat{P}_k^{\,n}f\|_1 + \|\widehat{P}_k^{\,n}f-P^nf\|_1 + \|P^nf-\pi(f)1_{\mathbb{X}}\|_1 \\
&\leq C_k{\rho_k}^n\,\|f\|_1 + \frac{M_\beta M_1}{(R-1)^2}\,R^{n+1}\,\|\widehat{P}_k-P\|_{\beta,1}\,\|f\|_\beta + C\rho^n\,\|f\|_1.
\end{aligned}$$

Inequality (16a) then follows from $\|f\|_1\leq\|f\|_\beta$ and Lemma 3. The last result of Assertion (a) is easily deduced from a usual "two-epsilon" argument (see also (17)). Inequality (17) can be obtained from (16a) with $\beta:=0$, $R:=\hat{\rho}_k^{\,-\sigma}$ and $n:=\lfloor -\ln v(k)/((1+\sigma)\ln\hat{\rho}_k)\rfloor$. □

3.2 Condition (P)

In the next subsections, Condition **(P)**, which is recalled for convenience, is assumed to hold.

Condition (P). *The Markov kernel P satisfies anyone of the two following conditions.*

(PWD) *P is compact from $\mathcal{B}_0$ into $\mathcal{B}_1$ and is such that*

$$\exists \delta \in (0,1),\ \exists L > 0, \quad PV \le \delta V + L\, 1_{\mathbb{X}}. \tag{\textbf{WD}}$$

In this case we set $\hat{\delta} := \delta$.

(PSD) *There exist a bounded set $S \in \mathcal{X}$ and a positive measure ν on $(\mathbb{X}, \mathcal{X})$ such that*

$$\exists \delta \in (0,1),\ \exists L > 0, \quad PV \le \delta\, V + L\, 1_S, \tag{\textbf{SD}}$$

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

In this case, we set

$$\hat{\delta} := \frac{\delta\, \nu(1_{B_{k_S}}) + \tau_{k_S}}{\nu(1_{B_{k_S}}) + \tau_{k_S}} \qquad \text{with } \tau_{k_S} := \max(0, L - \nu(V 1_{B_{k_S}})),$$

and we only consider the truncated Markov kernels P_k for $k \ge k_S$, where $k_S \ge 1$ is the smallest positive integer such that $S \subset B_{k_S}$ and $\nu(1_{B_{k_S}}) > 0$.

Note that the previous conditions only involve the kernel P. From (**WD**) (thus from (**SD**)), the kernels P and $\widehat{P}_k$ continuously act on $\mathcal{B}_1$ (see Lemma 3). We shall repeatedly use the following fact:

$$\text{Condition } \textbf{(P)} \implies r_{ess}(P) \le \hat{\delta} \ \text{ and } \ \sup_{k \ge \hat{k}} r_{ess}(\widehat{P}_k) \le \hat{\delta} \tag{20}$$

with $\hat{k} := 1$ under (PWD) and $\hat{k} := k_S$ under (PSD). Under (PWD), this follows from Theorem 1 and Lemmas 3 and 4, and under (PSD), from Theorem 2 and Lemma 5.

Finally observe that all the conditions (or conclusions) concerning the action of $\widehat{P}_k$ on $\mathcal{B}_1$ can be equivalently reformulated into the action of P_k on the Banach space of measurable functions $h : B_k \to \mathbb{C}$ such that $\sup_{x \in B_k} |h(x)|/V(x) < \infty$ (which defines the norm on this space). Actually, using our functional notations, Estimates (V)-(V_k) read as follows

$$\begin{aligned}
(V) \quad &\iff \quad \forall n \ge 1, \quad \sup_{f \in \mathcal{B}_1, \|f\|_1 \le 1} \|P^n f - \pi(f) 1_{\mathbb{X}}\|_1 \le C\, \rho^n \\
(V_k) \quad &\iff \quad \forall n \ge 1, \quad \sup_{f \in \mathcal{B}_1, \|f\|_1 \le 1} \|\widehat{P}_k^{\,n} f - \widehat{\pi}_k(f) 1_{B_k}\|_1 \le C_k\, {\rho_k}^n.
\end{aligned} \tag{21}$$

Although $\widehat{P}_k$ is not Markov, Property (V_k) shall be often expressed as the V-geometrical ergodicity of $\widehat{P}_k$.

3.3 From (V) to (V_k)

3.3.1 Solution to the issue (Q1)

Recall that $\|\cdot\|_1$ denotes the norm on $\mathcal{B}_1$ as well as the operator norm on $\mathcal{B}_1$. For any $(a,\theta)\in\mathbb{C}\times(0,+\infty)$, we denote by $D(a,\theta)$ (resp. $\overline{D}(a,\theta)$) the open (resp. closed) complex disk with center a and radius θ. 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

$$\mathcal{V}(r,\vartheta,T):=\big\{z\in\mathbb{C}:|z|\le r \text{ or } d(z,\sigma(T))\le\vartheta\big\} \quad\text{and}\quad \mathcal{V}(r,\vartheta,T)^c:=\mathbb{C}\setminus\mathcal{V}(r,\vartheta,T), \tag{22}$$

where $d(z,\sigma(T)):=\inf\{d(z,\lambda),\,\lambda\in\sigma(T)\}$.

The next constants are defined for any real numbers $r>\hat{\delta}$ and $\vartheta>0$:

$$H\equiv H(r,\vartheta,P):=\sup_{z\in\mathcal{V}(r,\vartheta,P)^c}\|(zI-P)^{-1}\|_1 \tag{23a}$$

$$n_1\equiv n_1(r):=\Big\lfloor\frac{\ln 2}{\ln(r/\hat{\delta})}\Big\rfloor \qquad n_2\equiv n_2(r,\vartheta,P):=\Big\lfloor\frac{3\,HL(L+3)\ln 2}{r^{n_1}\ln(r/\hat{\delta})}\Big\rfloor \tag{23b}$$

$$\varepsilon_1\equiv\varepsilon_1(r,\vartheta,P):=\frac{r^{n_1+n_2}}{8L\big(HL+(1-r)^{-1}\big)}. \tag{23c}$$

Theorem 3 below provides explicit parameters in (V_k) and bounds for $\|\widehat{\pi}_k-\pi\|_\beta$ under Conditions **(P)** and (V) for P and under the weak continuity condition $(C_{0,1})$ at rate ε_1

$$\|\widehat{P}_k-P\|_{0,1}:=\sup_{\|f\|_0\le 1}\|\widehat{P}_kf-Pf\|_1\le\varepsilon_1. \tag{$\mathcal{E}_{0,1}$}$$

Condition (V_k) is also assumed to hold in Theorem 3 with some mild information on the rate ρ_k. Of course, if (V_k) is satisfied with explicit rate and bound, Theorem 3 is not useful. Note that Inequality (25a) below gives: first a new rate r in (V_k) close to ρ; second an explicit constant c.

Theorem 3 *Assume that P satisfies Conditions* **(P)** *and (V) with rate $\rho\in(0,1)$. Let $(r,\vartheta)\in(0,1)^2$ be such that*

$$\max(\hat{\delta},\rho)+\vartheta<r<1-\vartheta. \tag{24}$$

Moreover suppose that, for some $k\in\mathbb{N}^*$, the k-th truncation Markov kernel P_k satisfies $(\mathcal{E}_{0,1})$ and (V_k) with some rate $\rho_k<1-\vartheta$. Then the following inequalities hold.

(a) For every measurable function $h:B_k\to\mathbb{C}$ such that $|h|\le V_{B_k}$, we have

$$\forall n\ge 1,\quad \sup_{x\in B_k}\frac{|({P_k}^nh)(x)-\pi_k(h)|}{V(x)}\le c\,r^{n+1} \tag{25a}$$

$$\text{with } c\equiv c(r,\vartheta,P):=\frac{4(L+1)}{r^{n_1}(1-r)}+\frac{1}{2\,\varepsilon_1}. \tag{25b}$$

(b) For any $\beta \in [0,1)$, we have the following estimate of $\|\widehat{\pi}_k - \pi\|_\beta$

$$\forall R > 1,\ \forall n \geq 1, \quad \|\widehat{\pi}_k - \pi\|_\beta \leq \frac{C\rho^n + c\, r^{n+1}}{\|1_{\mathbb{X}}\|_1} \ + \ \frac{d_\beta}{(R-1)^2}\, \frac{R^{n+1}}{v(k)^{1-\beta}}, \tag{26}$$

where the constants C and $d_\beta \equiv d_\beta(\delta, L)$ are given in (V) and (16b) respectively. In particular we have for any $\sigma \in (0,1]$ with $\xi := \max(\rho, r)$:

$$\|\widehat{\pi}_k - \pi\|_{TV} \leq \left(\frac{C + c\, r}{\xi \|1_{\mathbb{X}}\|_1} + \frac{d_0\, \xi^\sigma}{(1-\xi^\sigma)^2} \right) \frac{1}{v(k)^{1/(1+\sigma)}}. \tag{27}$$

Theorem 3 is only relevant when the rate ρ and the bound C in (V) are known. That ρ must be known is necessary to choose (r, ϑ) in (24). Constant C must also be known for an effective computation of an upper bound of $H \equiv H(r, \vartheta, P)$ in (23a) (see Appendix A), and so of bounds (26) and (27). Note that H is involved in the definition of $\varepsilon_1(r, \vartheta, P)$ and so in the computation of the bound $c(r, \vartheta, P)$ in (25b). Finally observe that, from the first assertion of Lemma 3, Condition $(\mathcal{E}_{0,1})$ holds for every $k \geq k_1$ with $k_1 \equiv k_1(r, \vartheta)$ defined as the smallest integer such that

$$v(k_1) \geq \frac{\max(2(\delta + L), 1)}{\varepsilon_1}. \tag{28}$$

Remark 4 (details in Appendix B) *Using the whole results of [Liv01] provides an explicit integer $k_0 \equiv k_0(r, \vartheta, P) \geq k_1$ such that, if Conditions* **(P)**, *(V) and (24) are fulfilled, then P_k satisfies (V_k) with $\rho_k < 1 - \vartheta$ for every $k \geq k_0$. Therefore the conclusions of Theorem 3 hold true for every $k \geq k_0$. Unfortunately, as mentioned in Introduction, k_0 is very large, in fact it is much larger that the integer k_1 (typically $k_0 \approx {k_1}^3$). Therefore it is more interesting to consider the above property on P_k as an assumption in Theorem 3, since it may be fulfilled for $k << k_0$. The bound in (25b) is then better than that involving the k_0, as in [Liv01] and in the works cited at the end of Introduction. Similarly the bounds (26) and (27) do not involve the integer k_0 since they will be derived from (25b) and Lemma 6.*

Moreover the use of the spectral rank-stability property [Liv01, cor. 3.1] yields bounds of the form $\|\widehat{\pi}_k - \pi\|_{TV} \leq D\, v(k)^{-\eta}$ with $D = O(H^2)$ and (typically) $\eta \approx 1/3$, while the constant in (27) is only $O(H)$ from (25b) and the definition of ε_1 in (23c).

It is proved in [Twe98] that $\sup_{|f| \leq V} |\widehat{\pi}_k(f) - \pi(f)| = \|\pi_k - \pi\|_1$ goes to 0 when $k \to +\infty$ for truncation approximation when $\mathbb{X}$ is discrete. Note that (26) does not give any rate of convergence for $\beta = 1$. Next, the expected rate of convergence $O(1/v(k))$ is approached in (27) when $\sigma \to 0$ but with constant increasing to $+\infty$. Finally, as mentioned in Introduction, Theorem 3 is interesting from a theoretical point of view, but solving the converse problem is more central. Namely, explicit rates and bounds in (V_k) are expected to be available in order to derive computable rate and bound in (V). This is the purpose of Subsection 3.4.

The proof of Theorem 3 is based on Assertion (a) of the next proposition.

Proposition 1 *Condition* **(P)** *and Condition $(\mathcal{E}_{0,1})$ for some $k \in \mathbb{N}^*$ are assumed to hold. Then:*

(a) we have

$$\sigma(\widehat{P}_k) \subset \mathcal{V}(r,\vartheta,P) \quad \text{and} \quad \sup_{z\in\mathcal{V}(r,\vartheta,P)^c} \|(zI-\widehat{P}_k)^{-1}\|_1 \leq \frac{4(L+1)}{r^{n_1}(1-r)} + \frac{1}{2\,\varepsilon_1}. \tag{29}$$

(b) If, in addition, P satisfies Condition (V) with rate ρ and if $\max(\hat{\delta},\rho) < 1-2\vartheta$, then, for k large enough, P_k satisfies (V_k) with any rate $\rho_k > \max(\hat{\delta},\rho)+\vartheta$.

Proposition 1 follows from the Keller-Liverani theorem [KL99] and the duality arguments introduced in [FHL11]. Using the explicit constants given in [Liv01, Prop. 3.1], the proof of Assertions (a) and (b) borrows the lines of [FHL11, Th. 1]. However some parts of the proof of [FHL11] must be revisited since $\widehat{P}_k$ is not Markov. The details are postponed to Subsection 3.3.2.

Proof of Theorem 3. The following inclusion holds from Condition (V)

$$\sigma(P) \subset D(0,\rho) \cup \{1\}.$$

Thus, it follows from (24) and the first inclusion in (29) 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

$$\sigma(\widehat{P}_k) \subset \overline{D}(0,r) \cup \{1\}. \tag{30}$$

We deduce from standard spectral calculus and Property (21) that for any $\tilde{r} \in (r, 1-\vartheta)$

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

Finally it follows from the second assertion of (29) that

$$\|\widehat{P}_k^{\,n} - \widehat{\pi}_k(\cdot)1_{B_k}\|_1 \leq c(r,\vartheta,P)\,(\tilde{r})^{n+1}.$$

Since $\tilde{r} \in (r, 1-\vartheta)$ is arbitrary, this gives (25a). Assertion (b) of Theorem 3 follows from Lemma 6. □

3.3.2 Proof of Proposition 1

We use the functional notations introduced at the beginning of the section as well as the duality notations of Subsection 2.1. The operator norm on the space $\mathcal{L}(\mathcal{B}_1')$ of bounded linear operators on $\mathcal{B}_1'$ is still denoted by $\|\cdot\|_1$. Recall that, for any $T \in \mathcal{L}(\mathcal{B}_1)$, we have $\|T'\|_1 = \|T\|_1$ and $\sigma(T) = \sigma(T')$. Given any real numbers $r > \hat{\delta}$ and $\vartheta > 0$, we use the notations introduced in (23a)-(23c). Note that H in (23a) 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 $\mathcal{V}(r,\vartheta,P) = \mathcal{V}(r,\vartheta,P')$. Assertion (a) in Proposition 1 then follows from the next result.

Proposition 2 *Condition* **(P)** *and Condition* $(\mathcal{E}_{0,1})$ *for some* $k \in \mathbb{N}^*$ *are assumed to hold. Then the following properties are valid with* $c(r,\vartheta,P)$ *given in (25b):*

$$\sigma(\widehat{P}_k{}') \subset \mathcal{V}(r,\vartheta,P) \quad \text{and} \quad \sup_{z\in\mathcal{V}(r,\vartheta,P)^c} \|(zI-\widehat{P}_k{}')^{-1}\|_1 \le c(r,\vartheta,P). \tag{31}$$

Proof. The Keller-Liverani perturbation theorem [KL99] is applied to the adjoint operators of P and $\widehat{P}_k$ acting on $\mathcal{B}_1'$ as in [FHL11, Th. 1]. Moreover we use the explicit constants given in [Liv01]. For convenience, we give a flavor of the duality-type arguments developed in [FHL11]. Define the following auxiliary norm on $\mathcal{B}_1'$:

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

Note that $\forall f' \in \mathcal{B}_1',\ \|f'\|_0 \le \|f'\|_1$. Moreover we have for any $f' \in \mathcal{B}_1'$

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

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

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

Next, we know from (**WD**) and Lemma 1 that $K' := P'$ and $K' := \widehat{P}_k{}'$ (for every $k \ge 1$) satisfy the following Doeblin-Fortet inequality on $\mathcal{B}_1'$:

$$\forall f' \in \mathcal{B}_1', \quad \|K'f'\|_1 \le \delta\|f'\|_1 + L\|f'\|_0. \tag{32}$$

Moreover we obtain by duality and from Inequality $(\mathcal{E}_{0,1})$

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

Finally, it follows from (20) that P' and $\widehat{P}_k{}'$ are quasi-compact on $\mathcal{B}_1'$ with essential spectral radius less than $\hat{\delta}$. The previous facts and the first part of [Liv01, Prop. 3.1] give (31). □

Let us introduce $\eta \equiv \eta(r) := 1 - \ln r/\ln\hat{\delta}$ (note that $\eta \in (0,1)$),

$$\varepsilon_2 \equiv \varepsilon_2(r,\vartheta,P) := \left\{\frac{r^{n_1}}{4L\big(H(2L+3)+2(1+L)+(1-r)^{-1}\big)}\right\}^{\frac{1}{\eta}}, \tag{33a}$$

$$\text{and } \varepsilon_0 \equiv \varepsilon_0(r,\vartheta,P) := \min(\varepsilon_1,\varepsilon_2). \tag{33b}$$

Using ε_0, the next statement specifies Assertion (b) of Proposition 1.

Proposition 3 *Assume that* P *satisfies Condition* **(P)** *and Condition* (V) *with rate* ρ. *Let* $\vartheta > 0$ *be such that* $\max(\hat{\delta},\rho) < 1-2\vartheta$ *and* r *be such that* $\max(\hat{\delta},\rho)+\vartheta < r < 1-\vartheta$. *Finally, suppose that* $\|\widehat{P}_k - P\|_{0,1} \le \varepsilon_0$ *for some* $k \in \mathbb{N}^*$. *Then* (V_k) *holds with* $\rho_k := r$.

The proof of Proposition 3 is based on the next lemma which follows by duality (as above) from the rank-stability property of spectral projections [Liv01, cor. 3.1].

Lemma 7 ([Liv01, cor. 3.1]) *Assume that Condition* **(P)** *holds and that, for $r > \hat{\delta}$ and $\vartheta > 0$, we have $\|\widehat{P}_k - P\|_{0,1} \le \varepsilon_0(r, \vartheta, P)$. Then the spectral projections of P' and $\widehat{P}'_k$ associated with any connected component of $\mathcal{V}(r, \vartheta, P)$ (not containing 0) have the same rank.*

Proof of Proposition 3. From (V) we have $\sigma(P') = \sigma(P) \subset D(0, \rho) \cup \{1\}$. Thus it follows from (31) that $\sigma(\widehat{P}'_k) \subset \overline{D}(0, r) \cup \overline{D}(1, \vartheta)$. Next, from $\|\widehat{P}_k - P\|_{0,1} \le \varepsilon_0$ and Lemma 7, we obtain that $\sigma(\widehat{P}_k{}') \cap \overline{D}(1, \vartheta) = \{\lambda\}$ for some eigenvalue $\lambda \in \mathbb{C}$ of $\widehat{P}_k{}'$ and that there exists an associated rank-one projection $\widehat{\Pi}'_{k,\lambda}$ on $\mathcal{B}'_1$ such that we have for any $\kappa \in (r, 1 - \vartheta)$:

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

It follows from (31) that

$$\forall n \ge 1, \quad \|\widehat{P}_k{}'^n - \lambda^n \widehat{\Pi}'_{k,\lambda}\|_1 \le c(r, \vartheta, P) \, \kappa^{n+1}. \tag{34}$$

Since κ is arbitrarily close to r, this gives the expected statement using duality and the next lemma. □

Lemma 8 *The eigenvalue λ in (34) is equal to 1. Moreover there exists a P_k-invariant probability measure π_k on $(B_k, \mathcal{X}_k)$ such that $\pi_k(V_k) < \infty$, and the rank-one projection $\widehat{\Pi}'_{k,\lambda}$ is the adjoint of the following rank-one projection $\widehat{\Pi}_k$ on $\mathcal{B}_1$:*

$$\forall f \in \mathcal{B}_1, \quad \widehat{\Pi}_k f := \pi_k(f_{B_k}) 1_{B_k}.$$

Proof. Since $|\lambda| > \kappa$, we deduce from $\widehat{P}_k 1_{B_k} = 1_{B_k}$ and (34) that, for any $f' \in \mathcal{B}'_1$, we have $\lim_n \lambda^{-n} f'(1_{B_k}) = \lim_n \lambda^{-n} (\widehat{P}_k{}'^n f')(1_{B_k}) = (\widehat{\Pi}'_{k,\lambda} f')(1_{B_k})$. Since there exists $f' \in \mathcal{B}'_1$ such that $f'(1_{B_k}) \neq 0$, $\{\lambda^{-n}\}_{n \in \mathbb{N}}$ converges in $\mathbb{C}$. Thus either $\lambda = 1$, or $|\lambda| > 1$. Moreover the sequence $\{\widehat{P}_k{}^n\}_{n \in \mathbb{N}}$ is bounded in $\mathcal{L}(\mathcal{B}_1)$ from (**WD**). Thus $\{\widehat{P}_k{}'^n\}_{n \in \mathbb{N}}$ is bounded in $\mathcal{L}(\mathcal{B}'_1)$, so that Inequality (34) implies that $\{\lambda^n\}_{n \in \mathbb{N}}$ is bounded in $\mathbb{C}$. Therefore $\lambda = 1$.

Now we omit λ in $\widehat{\Pi}'_{k,\lambda}$. We can prove as in [FHL11, 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:

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

Since $\widehat{\Pi}_k$ is rank-one, it has the form: $\forall f \in \mathcal{B}_1$, $\widehat{\Pi}_k f = \widehat{e}_k{}'(f) \, \phi_k$ for some nonnegative element $\widehat{e}_k{}' \in \mathcal{B}'_1$ and some nonnegative function $\phi_k \in \mathcal{B}_1$. From $\widehat{P}_k 1_{B_k} = 1_{B_k}$, it is easily checked that $\phi_k = 1_{B_k}$ (up to a multiplicative constant). Thus $\widehat{\Pi}_k$ may be assumed of the form $\widehat{\Pi}_k = \widehat{e}_k{}'(\cdot) \, 1_{B_k}$. More precisely, from $f = f1_{B_k} + f1_{B_k^c}$ and $\widehat{P}_k(f1_{B_k^c}) = 0$, we obtain $\widehat{\Pi}_k f = \widehat{\Pi}_k(f1_{B_k})$, hence: $\forall f \in \mathcal{B}_1$, $\widehat{e}_k{}'(f) = \widehat{e}_k{}'(f1_{B_k})$. Let us prove now that there exists a P_k-invariant probability measure π_k on B_k such that $\pi_k(V_k) < \infty$ and $\forall f \in \mathcal{B}_1$, $\widehat{e}_k{}'(f1_{B_k}) = \pi_k(f_{B_k})$. Fix $x_0 \in B_k$. Then, for every $F \in \mathcal{X}_k$, we obtain from the definition of $\widehat{e}_k{}'$ and $\widehat{P}_k$:

$$\widehat{e}_k{}'(1_F) = \lim_n (\widehat{P}_k{}^n 1_F)(x_0) = \lim_n P_k^n(x_0, F). \tag{35}$$

From the Vitali-Hahn-Saks theorem, since $P_k^n(x_0,\cdot)$ is a probability measure on B_k, so is π_k defined by: $\forall F \in \mathcal{X}_k,\ \pi_k(1_F) := \widehat{e}_k{}'(1_F)$. Next let $\widehat{\pi}_k$ be the probability measure on $(\mathbb{X},\mathcal{X})$ defined by $\widehat{\pi}_k(f) := \pi_k(f_{B_k})$ for any bounded measurable function $f : \mathbb{X} \to \mathbb{R}$. Then, for every $A \in \mathcal{X}$, we have $\widehat{\pi}_k(1_A) = \widehat{e}_k{}'(1_A)$, so that $\widehat{\pi}_k(\cdot)$ and $\widehat{e}_k{}'(\cdot)$ coincide on the set of simple functions on $\mathbb{X}$. Thus they coincide on $\mathcal{B}_0$: indeed any $f \in \mathcal{B}_0$ can be approached uniformly on $\mathbb{X}$ by a sequence $\{f_n\}_{n\in\mathbb{N}}$ of simple functions, thus $\widehat{\pi}_k(f) = \lim_n \widehat{\pi}_k(f_n) = \lim_n \widehat{e}_k{}'(f_n) = \widehat{e}_k{}'(f)$ since $\widehat{e}_k{}' \in \mathcal{B}_1'$ and $f_n \to f$ in $\mathcal{B}_1$. Therefore, it follows that, for every $n \geq 1$, we have $\widehat{\pi}_k(V1_{[V\leq n]}) = \widehat{e}_k{}'(V1_{[V\leq n]}) \leq \widehat{e}_k{}'(V) < \infty$ since $\widehat{e}_k{}'$ is nonnegative. We deduce from the monotone convergence theorem that $\pi_k(V_{B_k}) = \widehat{\pi}_k(V) < \infty$.

The previous fact implies that $\widehat{\pi}_k$ defines an element of $\mathcal{B}_1'$ since, for any $f \in \mathcal{B}_1$, we have $\widehat{\pi}_k(|f|) \leq \widehat{\pi}_k(V)\|f\|_1$. Moreover we deduce from the definition of $\widehat{e}_k{}'$ that: $\forall f \in \mathcal{B}_0,\ \widehat{\pi}_k(\widehat{P}_k f) = \widehat{e}_k{}'(\widehat{P}_k f) = \widehat{e}_k{}'(f) = \widehat{\pi}_k(f)$. Now let $g \in \mathcal{B}_1$. It easily follows from the last fact that $\widehat{\pi}_k(\widehat{P}_k g) = \widehat{\pi}_k(g)$ (use $f_n := g\,1_{[|g|\leq n]} \in \mathcal{B}_0$ and Lebesgue's theorem twice). Consequently we obtain $\widehat{\pi}_k(g) = \widehat{e}_k{}'(g)$ from the convergence $\lim_n \widehat{P}_k^{\,n} g = \widehat{e}_k{}'(g)\,1_{B_k}$ in $\mathcal{B}_1$ and from $\widehat{\pi}_k \in \mathcal{B}_1'$, $\widehat{\pi}_k(1_{B_k}) = 1$. □

3.4 From (V_k) to (V)

Let $\rho_V(P)$ denote the spectral gap of P, that is the infinum bound of the rates ρ in (V). Recall that $\rho_V(P)$ is unknown in general and that, even when it is known, finding an explicit constant C associated with some rate $\rho > \rho_V(P)$ in (V) is a difficult question. As mentioned in Introduction, there exist various methods [MT94, LT96, Bax05] providing rates and constants (ρ, C) in (V), but these methods yield a rate ρ which is much far from $\rho_V(P)$ (except in specific cases). The first purpose of Theorem 4 is to obtain a rate r_k in (V) which is all the more close to $\rho_V(P)$ that k is large and ρ_k in (V_k) is close to the spectral gap of P_k. Such a ρ_k is expected to be computable since the state space associated with P_k is a closed ball of $\mathbb{X}$ (a finite set when $\mathbb{X}$ is discrete). The second purpose of Theorem 4 is to provide an explicit constant $c_k \equiv c(r_k)$ associated with the rate r_k in (V). Then Lemma 6 gives a bound for $\|\widehat{\pi}_k - \pi\|_{TV}$. Theorem 4 can only be applied when a preliminary bound ρ in (V) (possibly far from $\rho_V(P)$) is known. Such a first rate in (V) can be found by using the above mentioned methods.

Let us briefly explain why the passage from the V_k-geometric ergodicity of P_k to the V-geometric ergodicity of P is theoretically more difficult than the converse one studied in the previous subsection. Assume that P satisfies Conditions **(P)**. Let $r > \hat{\delta}$ and $\vartheta > 0$. With $\varepsilon_1(r,\vartheta,P)$ defined in (23c), the first assertion of Proposition 1 reads as follows

$$\|\widehat{P}_k - P\|_{0,1} \leq \varepsilon_1(r,\vartheta,P) \quad \Longrightarrow \quad \left\{ \begin{array}{c} \sigma(\widehat{P}_k) \subset \mathcal{V}(r,\vartheta,P) \\ \sup\limits_{z\in\mathcal{V}(r,\vartheta,P)^c} \|(zI - \widehat{P}_k)^{-1}\|_1 < \infty. \end{array} \right. \tag{36}$$

Next, for any fixed $k \in \mathbb{N}^*$, exchanging the role of P and $\widehat{P}_k$ in Assertion (a) of Proposition 1

provides the following implication:

$$\|P-\widehat{P}_k\|_{0,1}\le\varepsilon_1(r,\vartheta,\widehat{P}_k)\quad\Longrightarrow\quad\left\{\begin{array}{c}\sigma(P)\subset\mathcal{V}(r,\vartheta,\widehat{P}_k)\\ \sup\limits_{z\in\mathcal{V}(r,\vartheta,\widehat{P}_k)^c}\|(zI-P)^{-1}\|_1<\infty.\end{array}\right.\tag{37}$$

There is a significant difference between the implications (36) and (37) since the inequality $\|\widehat{P}_k-P\|_{0,1}\le\varepsilon_1(r,\vartheta,P)$ in (36) is satisfied for k large enough from Lemma 3, while the inequality $\|P-\widehat{P}_k\|_{0,1}\le\varepsilon_1(r,\vartheta,\widehat{P}_k)$ in (37) could fail for every k. Fortunately, the previous failure cannot occur when P satisfies Condition **(P)** (see Lemma 9 below).

3.4.1 Solution to the issue (Q2)

Given any $\vartheta\in(0,1)$, we define the following subset of integers :

$$\mathcal{I}_\vartheta:=\big\{k\in\mathbb{N}^*:P_k\text{ satisfies }(V_k)\text{ with some rate }\rho_k<1-2\vartheta\big\}.$$

Under Condition **(P)**, for any $\vartheta\in(0,(1-\hat\delta)/2)$, for every $k\in\mathcal{I}_\vartheta$ and finally for every $r_k\in(0,1)$ such that

$$\max(\hat\delta,\rho_k)+\vartheta<r_k<1-\vartheta,\tag{38}$$

we introduce the following constants:

$$\varepsilon_1(k)\equiv\varepsilon_1(r_k,\vartheta,\widehat{P}_k):=\frac{{r_k}^{n_1(k)+n_2(k)}}{8L\big(H_kL+(1-r_k)^{-1}\big)}.\tag{39a}$$

with

$$H_k\equiv H(r_k,\vartheta,\widehat{P}_k):=\sup_{z\in\mathcal{V}(r_k,\vartheta,\widehat{P}_k)^c}\|(zI-\widehat{P}_k)^{-1}\|_1\tag{39b}$$

$$n_1(k)\equiv n_1(r_k):=\lfloor\frac{\ln 2}{\ln(r_k/\hat\delta)}\rfloor\qquad n_2(k)\equiv n_2(r_k,\vartheta,P_k):=\lfloor\frac{3\,H_kL(L+3)\ln 2}{{r_k}^{n_1(k)}\ln(r_k/\hat\delta)}\rfloor.\tag{39c}$$

The next lemma is inspired from [Liv01, Lem. 4.2].

Lemma 9 *Assume that P satisfies Condition* **(P)**. *Let $\vartheta\in(0,(1-\hat\delta)/2)$. Then there exists $\tilde k\equiv\tilde k(\vartheta)\in\mathbb{N}^*$ such that the following property holds for every $k\in[\tilde k,+\infty)\cap\mathcal{I}_\vartheta$:*

$$\|P-\widehat{P}_k\|_{0,1}\le\varepsilon_1(k).\tag{40}$$

A detailed proof of Lemma 9 will be provided in Subsection 3.4.2 since that of [Liv01, Lem. 4.2] is only sketched and the choice of r_k must be carefully examined. Now the proof of the following theorem borrows that of Theorem 3 by exchanging the role of P and $\widehat{P}_k$.

Theorem 4 *Assume that P satisfies Conditions* **(P)** *and (V) with some rate $\rho\in(0,1)$. Let ϑ be such that*

$$0<\vartheta<\frac{1-\max(\hat\delta,\rho)}{3}.$$

Moreover let $k\in\mathcal{I}_\vartheta$ be such that Condition (40) holds. Then, for any $r_k\in(0,1)$ satisfying $\max(\hat\delta,\rho_k)+\vartheta<r_k<1-\vartheta$, the next assertions hold true.

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

$$\forall n \geq 1, \quad \|P^n - \pi(\cdot)1_{\mathbb{X}}\|_1 \leq c_k \, {r_k}^{n+1} \tag{41a}$$

$$\textit{with} \qquad c_k \equiv c_k(r_k, \vartheta) := \frac{4(L+1)}{{r_k}^{n_1(k)}(1-r_k)} + \frac{1}{2\,\varepsilon_1(k)}. \tag{41b}$$

(b) The following inequality holds for any $\beta \in [0,1)$:

$$\forall R > 1, \ \forall n \geq 1, \quad \|\widehat{\pi}_k - \pi\|_\beta \leq \frac{C_k \, {\rho_k}^n + c_k \, {r_k}^{n+1}}{\|1_{\mathbb{X}}\|_1} \ + \ \frac{d_\beta}{(R-1)^2} \, \frac{R^{n+1}}{v(k)^{1-\beta}}, \tag{42}$$

where the constants C_k, and $d_\beta \equiv d_\beta(\delta, L)$ are given in (V_k) and (16b) respectively. In particular we have for any $\sigma \in (0,1]$ with $\hat{\rho}_k := \max(r_k, \rho_k)$

$$\|\widehat{\pi}_k - \pi\|_{TV} \leq \left(\frac{C_k + c_k r_k}{\hat{\rho}_k} + \frac{d_0 \, {\hat{\rho}_k}^\sigma}{(1 - {\hat{\rho}_k}^\sigma)^2} \right) \frac{1}{v(k)^{1/(1+\sigma)}}. \tag{43}$$

Assume that the assumptions on P and ϑ in Theorem 4 hold true. Then there exists $\tilde{k}_0 \equiv \tilde{k}_0(\vartheta) \in \mathbb{N}^*$ such that $\mathcal{I}_\vartheta = [\tilde{k}_0, +\infty)$. Indeed, it follows from Proposition 1(b) that there exists $\tilde{k}_0 \in \mathbb{N}^*$ such that, for every $k \geq \tilde{k}_0$, Property (V_k) holds for any rate $\rho_k > \max(\hat{\delta}, \rho) + \vartheta$, thus for any rate ρ_k such that $\max(\hat{\delta}, \rho) + \vartheta < \rho_k < 1 - 2\vartheta$. In particular we have: $\forall k \geq \tilde{k}_0, \ k \in \mathcal{I}_\vartheta$. The existence of both $\tilde{k}_0$ and $\tilde{k}$ (Lemma 9) is important since this ensures that the conclusions of Theorem 4 are valid for every $k \geq \max(\tilde{k}, \tilde{k}_0)$. Numerical upper bounds for $\tilde{k}$ and $\tilde{k}_0$ derived from [Liv01] are very large in general ($\tilde{k}$ is still greater than the integer k_0 already discussed in Remark 4). Since integer $k \in \mathcal{I}_\vartheta$ satisfying (40) exists for $k << \max(\tilde{k}, \tilde{k}_0)$, an effective application of Theorem 4 needs to find such a k with a value as small as possible. This can be done in testing the validity of the properties $k \in \mathcal{I}_\vartheta$ and (40) for increasing values of k.

Remark 5 (Further comments on the rate r_k in (41a)) *When the spectral gap $\rho_V(P)$ of P is known, Theorem 4 can be applied with any $\rho > \rho_V(P)$. If the value of ρ used in Theorem 4 is chosen to be close to $\rho_V(P)$, then the new rate r_k in (41a) may be a little bit larger than ρ. On the other hand, if ρ is close to 1 (this happens when $\rho_V(P)$ is either unknown, or known but close to 1), then the new rate r_k in (41a) is expected to be better than ρ, but this only happens when $\rho_V(P) << 1$. In any case, (41a) is of interest since an explicit bound c_k is provided. In fact, the choice of ρ in Theorem 4, which determines the value of ϑ, also affects the bound c_k.*

Remark 6 (Further comments on the constant c_k in (41b)) *If ρ is close to 1, then ϑ is small and the bound $c_k \equiv c_k(r_k, \vartheta)$ in (41b) may be large since $c_k \approx O(1/\varepsilon_1(k)) = O(H_k)$, with $H_k \geq \sup_{|z-1|=\vartheta} \|(zI - \widehat{P}_k)^{-1}\|_1 \to +\infty$ when $\vartheta \to 0$. When the rates ρ_k used in Theorem 4 are quite smaller than ρ, this means that $\rho_V(P) << \rho$. In this case, it is worth using again Theorem 4 with another larger value for ϑ, say ϑ'. The resulting new rate, say r'_k, in (41a) will be larger than the previous one since ϑ has been augmented (see (38)). But the new constant, say c'_k, associated to r'_k will be much smaller because so is the new bound $H'_k \equiv H(r'_k, \vartheta', \widehat{P}_k)$. More generally, an efficient use of Theorem 4 needs to find a good compromise between the rate and the associated constant in (41a).*

3.4.2 Proof of Lemma 9

All the spectral values of P and $\widehat{P}_k$ strictly larger than $\hat{\delta}$ are eigenvalues from (20). More precisely, for any $R > \hat{\delta}$, the operators P and $\widehat{P}_k$ have a finite number of eigenvalues of modulus larger than R.

Lemma 10 *Assume that P satisfies Conditions* **(P)**. *Let $R > \hat{\delta}$ and $\theta > 0$. Then there exists $k_0 \equiv k_0(R, \theta) \in \mathbb{N}^*$ such that, for any eigenvalue of P satisfying $|\lambda| > R$ and for every $k \geq k_0$, 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 $a := R - \hat{\delta}$, $b := \min\{|z - z'|,\, z, z' \in \mathcal{D}_R, z \neq z'\}$ and $c := \min\{|z| - R,\, z \in \mathcal{D}_R\}$. Without loss of generality we can suppose that $\theta < \min(a, b, c/2)$. Let $\vartheta \in (0, \theta)$ and $r := \hat{\delta} + \theta$. Let $k_0 \equiv k_0(r, \vartheta, P)$ be the smallest integer such that $v(k_0) \geq \max(2(\delta + L), 1)/\varepsilon_0$ with $\varepsilon_0 \equiv \varepsilon_0(r, \vartheta, P)$ given in (33b). From Lemma 3: $k \geq k_0 \Rightarrow \|\widehat{P}_k - P\|_{0,1} \leq \varepsilon_0$. Note that assumptions of Lemma 7 are satisfied.

Next, consider any $k \geq k_0$ and any $\lambda \in \mathcal{D}_R$. Let $\theta' \in (\vartheta, \theta)$ 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 θ' 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 if $z \in C(\lambda, \theta')$, then

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

thus $d(z, \overline{D}(0, R)) > \vartheta$. Moreover, if $z' \in \sigma(P)$ and $|z'| > R$, then $z' \in \mathcal{D}_R$, thus $|z - z'| \geq b > \vartheta$. We have proved that $d(z, \sigma(P)) > \vartheta$. Finally we have $|z| > r$ since $R > r$ (use $r = \hat{\delta} + \theta < \hat{\delta} + a = \hat{\delta} + R - \hat{\delta} = R$). 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 $k \geq k_0$ and Lemma 7. Since Π'_λ has a nonzero rank from $\lambda \in \sigma(P')$, so is $\Pi'_{k,\lambda}$. Thus we have $D(\lambda, \theta') \cap \sigma(\widehat{P}_k) \neq \emptyset$. □

To prove Lemma 9, we consider any $\vartheta \in (0, (1 - \hat{\delta})/2)$ and we set $\tilde{r} := \hat{\delta} + \vartheta/2$.

Lemma 11 *There exists $k_2 \equiv k_2(\tilde{r}, \vartheta) \in \mathbb{N}^*$ such that $\forall k \geq k_2,\ \mathcal{V}(\tilde{r}, \vartheta/4, P) \subset \mathcal{V}(\tilde{r}, \vartheta, \widehat{P}_k)$.*

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 $|\lambda| > \hat{\delta} + \vartheta/4$ from

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

Then it follows from Lemma 10 with $R := \hat{\delta} + \vartheta/4$ and $\theta := \vartheta/4$ that there exists $k_2 \equiv k_2(R,\theta) \in \mathbb{N}^*$ such that, for every $k \geq k_2$, the disk $D(\lambda, \vartheta/4)$ contains an eigenvalue of $\widehat{P}_k$, say λ_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 $u \in \mathcal{V}(\tilde{r}, \vartheta, \widehat{P}_k)$. □

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 $k \geq 1$:

$$\widetilde{H}_k := \sup_{z \in \mathcal{V}(\tilde{r}, \vartheta, \widehat{P}_k)^c} \|(zI - \widehat{P}_k)^{-1}\|_1. \tag{44}$$

Lemma 12 *The sequence $\{\widetilde{H}_k\}_{k \geq 1}$ is bounded.*

Proof. It follows from Lemma 11 that

$$\forall k \geq k_2, \quad \mathcal{V}(\tilde{r}, \vartheta, \widehat{P}_k)^c \subset \mathcal{V}(\tilde{r}, \vartheta/4, P)^c.$$

Apply Proposition 2 with $r := \tilde{r}$ and $\vartheta/4$ in place of ϑ. Let $\varepsilon_1(\tilde{r}, \vartheta/4, P)$ be given by (23c). From Lemma 3, there exists $k_3 \equiv k_3(\tilde{r}, \vartheta) \in \mathbb{N}^*$ such that

$$\forall k \geq k_3, \quad \|\widehat{P}_k - P\|_{0,1} \leq \varepsilon_1(\tilde{r}, \vartheta/4, P).$$

Set $k_4 := \max(k_2, k_3)$. Then we deduce from Lemma 11 and (31) that for every $k \geq k_4$:

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

This gives the expected assertion. □

Proof of Lemma 9. Let $H_k \equiv H(r_k, \vartheta, \widehat{P}_k)$ be defined by (39b). 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

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

It follows from Lemma 12 that $\{H_k\}_{k \geq 1}$ is bounded. Next, the sequences $\{n_1(k)\}_{k \geq 1}$ and $\{n_2(k)\}_{k \geq 1}$ given in (39c) are bounded since $\hat{\delta} + \vartheta < r_k < 1 - \vartheta$ and $\{H_k\}_{k \geq 1}$ is bounded. Therefore, the sequence $\{\varepsilon_1(k)\}_{k \geq 1}$ in (39a) is uniformly bounded away from zero, that is

$$\alpha_1 := \inf_{k \geq 1} \varepsilon_1(k) > 0.$$

Finally, from Lemma 3, 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 $k \geq \tilde{k}, \ \|\widehat{P}_k - P\|_{0,1} \leq \varepsilon_1(k)$. □

A Additional material on constant H in (23a) (or H_k in (39b))

Let Q be a Markov kernel on a measurable space $(\mathbb{Y}, \mathcal{Y})$. For some $W : \mathbb{Y} \to [1, +\infty)$, $\mathcal{B}_W$ denotes the space of measurable functions $g : \mathbb{Y} \to \mathbb{C}$ such that $\|g\|_W := \sup_{y \in \mathbb{Y}} |g(y)|/W(y) <$

∞. Q is said to be W-geometrically ergodicity if there exist an invariant probability measure π on $(\mathbb{Y}, \mathcal{Y})$ such that $\pi(W) < \infty$ and some constants $\kappa \in (0,1)$ and $K > 0$ such that

$$\forall n \geq 1, \quad \|Q^n - \pi(\cdot)1_{\mathbb{Y}}\|_W := \sup_{\|f\|_W \leq 1} \|Q^n f - \pi(f)1_{\mathbb{Y}}\|_W \leq K\,\kappa^n\,\|f\|_W. \tag{45}$$

From (45), the resolvent $(zI-Q)^{-1}$ is well-defined in $\mathcal{L}(\mathcal{B}_W)$ for any $z \in \mathbb{C}$ such that $|z| > \kappa$ and $z \neq 1$ (see the proof of Proposition A.1). In other words we have $\sigma(Q) \subset \overline{D}(0,\kappa) \cup \{1\}$. Moreover, for every $(r,\vartheta) \in (0,1)^2$ such that $\kappa + \vartheta < r < 1 - \vartheta$, we have

$$H(r,\vartheta,Q) := \sup_{z \in \overline{D}(0,r)^c \cap \overline{D}(1,\vartheta)^c} \|(zI-Q)^{-1}\|_W < \infty.$$

Computing or at least upper bounding $H(r,\vartheta,Q)$ is crucial in Theorems 3 and 4. However, this job is quite difficult in general. When the constants κ and K in (45) are known, the following estimate can be used.

Proposition A.1 *Condition* (45) *is assumed to hold. Then, for any* $(r,\vartheta) \in (0,1)^2$ *such that* $\kappa + \vartheta < r < 1 - \vartheta$*, the following inequality is valid:*

$$H(r,\vartheta,Q) \leq \frac{K}{r-\kappa} + \frac{\pi(W)\|1_{\mathbb{Y}}\|_W}{\vartheta} \leq \frac{\pi(W)\|1_{\mathbb{Y}}\|_W + K}{\vartheta}.$$

Since the invariant probability measure π may be unknown, it is worth noticing that $\pi(W) \leq L/(1-\delta)$ since $QW \leq \delta W + L$ with some $\delta \in (0,1)$ and $L > 0$.

Proof. For any $g \in \mathcal{B}_W$, write $g = (g - \pi(g)1_{\mathbb{Y}}) + \pi(g)1_{\mathbb{Y}}$. We have

$$(zI-Q)^{-1}\big(\pi(g)1_{\mathbb{Y}}\big) = \frac{\pi(g)}{z-1}1_{\mathbb{Y}},$$

while we obtain from Inequality (45) applied to the function $g - \pi(g)1_{\mathbb{Y}}$ and from $|z| > \kappa$:

$$(zI-Q)^{-1}(g - \pi(g)1_{\mathbb{Y}}) = \sum_{n \geq 0} z^{-(n+1)} Q^n (g - \pi(g)1_{\mathbb{Y}}).$$

It follows from $|\pi(g)| \leq \pi(W)\|g\|_W$ that

$$\begin{aligned}
\big\|(zI-Q)^{-1}g\big\|_W &= \Big\|\sum_{n\geq 0} z^{-(n+1)} Q^n (g - \pi(g)1_{\mathbb{Y}}) + \frac{\pi(g)}{z-1}1_{\mathbb{Y}}\Big\|_W \\
&\leq \left(K\sum_{n\geq 0}|z|^{-(n+1)}\kappa^n + \frac{\pi(W)}{|z-1|}\|1_{\mathbb{Y}}\|_W\right)\|g\|_W \\
&\leq \left(\frac{K}{|z|-\kappa} + \frac{\pi(W)\|1_{\mathbb{Y}}\|_W}{|z-1|}\right)\|g\|_W.
\end{aligned}$$

The final form is obtained using $|z| - \kappa > r - \kappa > \vartheta$ and $|z-1| > \vartheta$ in the definition of $H(r,\vartheta,Q)$. $\square$

B Further comments on the results of [Liv01]

As mentioned at the end of Introduction and in the remarks of Subsections 3.3.1 and 3.4.1, the explicit bounds in the second part of [Liv01, Prop. 3.1] and in the spectral rank-stability property [Liv01, Cor. 3.1] involve truncated Markov kernels for large k and yield large bounds for $\|\widehat{\pi}_k - \pi\|_{TV}$. The next comments complete this discussion. Specifically, for the sake of simplicity, we only consider here the simple context of Subsection 3.3.1, but all the next comments extend to the setting of Subsection 3.4.1 by considering the bounds $H_k \equiv H(r_k, \vartheta, \widehat{P}_k)$ of (39b) and by defining the corresponding $\varepsilon_0(k) \equiv \varepsilon_0(r_k, \vartheta, \widehat{P}_k)$.

Recall that $V(\cdot) \equiv v(d(\cdot, x_0))$ for some $x_0 \in \mathbb{X}$ and unbounded increasing function $v : [0, +\infty) \to [1, +\infty)$. We assume that P satisfies Conditions **(P)** and (V) with rate $\rho \in (0, 1)$. Given $(r, \vartheta) \in (0, 1)^2$ satisfying (24), that is $\max(\hat{\delta}, \rho) + \vartheta < r < 1 - \vartheta$, we define $k_0 \equiv k_0(r, \vartheta, P)$ as the smallest integer such that

$$v(k_0) \geq \frac{\max(2(\delta + L), 1)}{\varepsilon_0} \tag{46}$$

with $\varepsilon_0 \equiv \varepsilon_0(r, \vartheta, P)$ given in (33b). We obtain from Lemma 3

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

Thus, it follows from Proposition 3 that, for every $k \geq k_0$, the k-th truncation approximation kernel P_k satisfies Condition (V_k) with any $\rho_k > \max(\hat{\delta}, \rho) + \vartheta$ as rate of convergence, and in particular with some rate $\rho_k < 1 - \vartheta$. Consequently, under Conditions **(P)**, (V) and (24), the conclusions of Theorem 3 hold true for every $k \geq k_0$.

However k_0 is much larger than k_1 given in (28). Indeed, note that $k_1 \approx v^{-1}(H)$ where v^{-1} denotes the inverse function of v, while Inequality (46) and Formula (33a) give

$$k_0 \geq v^{-1}\left(\frac{2(\delta + L)}{\varepsilon_0}\right) \approx v^{-1}(H^{1/\eta})$$

with $H \equiv H(r, \vartheta, P)$ given in (23a) and $\eta := 1 - \ln r / \ln \hat{\delta}$. Since η is usually close to $1/3$, we (typically) obtain $k_0 \approx v^{-1}(H^3)$. This yields a very large integer k_0 because H is large.

Moreover the use of the second part of [Liv01, Prop. 3.1] provides larger bounds for $\|\widehat{\pi}_k - \pi\|_{TV}$ than in (27). In fact, the following inequality can be derived from [KL99, Liv01] (by using duality as in the proof of Proposition 2):

$$\sup_{z \in \mathcal{V}(r,\vartheta,T)^c} \|(zI - \widehat{P}'_k)^{-1} - (zI - P')^{-1}\|_{01} \leq d\, \|\widehat{P}_k - P\|_{0,1}^{\eta}, \tag{47}$$

with $d := aH + bH^2$, where $a \equiv a(r)$ and $b \equiv b(r)$ are defined by

$$a := \frac{1 + 8(1 + L)^2\big(2(1 + L) + (1 - r)^{-1}\big)}{(1 - r)r^{n_1}} \qquad b := 2\left\{L + \frac{4(1 + L)^2(2L + 3) + L}{(1 - r)r^{n_1}}\right\}.$$

Using standard spectral calculus, [Liv01, Cor. 3.1] (see Lemma 7) and finally Lemma 3, we obtain

$$\forall k \geq k_0, \quad \|\widehat{\pi}_k - \pi\|_{TV} \leq \frac{d\,\vartheta\, \max\big(2^{\eta}(\delta + L)^{\eta}, 1\big)}{v(k)^{\eta}}.$$

The previous bound is worst than in (27) since $d = O(H^2)$ while c in bound (27) is $c \equiv c(r, \vartheta, P) = O(H)$ (see (25b)).

References

[Asm03] S. Asmussen. *Applied probability and queues*. Springer-Verlag, New York, 2003.

[Bah06] W. Bahsoun. Rigorous numerical approximation of escape rates. *Nonlinearity*, 19(11):2529–2542, 2006.

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

[BB10] 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.

[BB11] W. Bahsoun and C. Bose. Invariant densities and escape rates: rigorous and computable approximations in the L^∞-norm. *Nonlinear Anal.*, 74(13):4481–4495, 2011.

[DL08] 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.

[FHL11] D. Ferré, L. Hervé, and J. Ledoux. Regular perturbation of V-geometrically ergodic Markov chains. Revised for publication in Applied Probability Journals, 2011.

[GHL11] D. Guibourg, L. Hervé, and J. Ledoux. Quasi-compactness of Markov kernels on weighted-supremum spaces and geometrical ergodicity. arXiv : 1110.3240, 2011.

[Hen93] H. Hennion. Sur un théorème spectral et son application aux noyaux lipchitziens. *Proc. Amer. Math. Soc.*, 118:627–634, 1993.

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

[KL99] G. Keller and C. Liverani. Stability of the spectrum for transfer operators. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze Sér. 4*, XXVIII:141–152, 1999.

[KM03] I. Kontoyiannis and S. P. Meyn. Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab.*, 13(1):304–362, 2003.

[Liu10] Y. Liu. Augmented truncation approximations of discrete-time Markov chains. *Oper. Res. Lett.*, 38(3):218–222, 2010.

[Liv01] C. Liverani. Rigorous numerical investigation of the statistical properties of piecewise expanding maps. A feasibility study. *Nonlinearity*, 14(3):463–490, 2001.

[Liv04] C. Liverani. Invariant measure and their properties. a functional analytic point of view. In Pisa Pubblicazioni della Classe di Scienze, Scuola Normale Superiore, editor, *Dynamical Systems. Part II: Topological Geometrical and Ergodic Properties of Dynamics*, Centro di Ricerca Matematica Ennio De Giorgi, 2004.

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

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

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

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

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

[Sch71] H. H. Schaefer. *Topological vector spaces*. Springer-Verlag, New York, 1971.

[Sen81] E. Seneta. *Nonnegative matrices and Markov chains*. Springer-Verlag, New York, 1981.

[Twe98] R. L. Tweedie. Truncation approximations of invariant measures for Markov chains. *J. Appl. Probab.*, 35(3):517–536, 1998.

[Wu04] L. Wu. Essential spectral radius for Markov semigroups. I. Discrete time case. *Probab. Theory Related Fields*, 128(2):255–321, 2004.

[Zhe10] Y. Zheng. Stochastic perturbations of fat solenoidal attractors. *Acta Math. Sin. (Engl. Ser.)*, 26(5):923–936, 2010.

[ZL96] Y. Q. Zhao and D. Liu. The censored Markov chain and the best augmentation. *J. Appl. Probab.*, 33(3):623–629, 1996.