LE MATEMATICHE
Vol. LXXVI (2021) – Issue I, pp. 243–276
doi: 10.4418/2021.76.1.13

PERTURBATION AND STABILITY BOUNDS FOR ERGODIC GENERAL STATE MARKOV CHAINS WITH RESPECT TO VARIOUS NORMS

Z. MOUHOUBI

This paper provides new perturbation bounds for general state Markov chains with respect to various norms. The transition and stationary characteristics estimates are given in terms of the generalized norm ergodicity coefficient, the norm of a residual kernel, or the parameters given in some drift condition. In fact, those results improve and generalize, for general state-space with respect to various norms and for a more large amplitude of perturbation of the transition kernel, the bounds obtained in [1, 32, 40]. Furthermore, we improve some other inequalities established in [22, 39, 41]. Theoretical comparison and on the basis of some examples show the quality of the results obtained in this paper.

1. Introduction

Let $X = (X_t, t = 0, 1, \dots)$ an homogeneous Markov chain on the measurable space $(E, \mathcal{E})$ with a transition kernel P and having a unique and finite invariant probability measure π. Suppose that X is perturbed to be another Markov chain $Y = (Y_t, t = 0, 1, \dots\}$ with transition kernel Q. In regular perturbation theory, we usually assume that the perturbed chain Y has also a unique invariant probability

Received on September 12, 2020

AMS 2010 Subject Classification: 60J05, 60J10

Keywords: Perturbation bound, stability bound, norm ergodicity coefficient, residual kernel, transition distribution, stationary distribution

measure ν and we are interested in deriving the perturbation bounds (norm-wise bounds) for the deviation $\|\nu-\pi\|$ in term of the difference $\|Q-P\|$ with respect to a suitable norm. The perturbation analysis investigates the quantitative estimates on the effect of switching from P to Q on the stationary distribution of the chain. More precisely, we study the bounds of the type

$$\|\nu-\pi\| \leq \kappa\|Q-P\| \tag{1}$$

where κ is the so called condition number. We have to point out that the most results, in discrete state, are expressed in terms of the potential (fundamental matrix) $R=(I-P+\Pi)^{-1}$ or the group inverse $\mathtt{A}^{\#}$ of the matrix $I-P$; where Π is the ergodic projector of P, i.e., the matrix with rows identical to π^T and π is the unique stationary distribution vector. We recall that the group inverse $\mathtt{K}^{\#}=\left(\mathtt{k}_{ij}^{\#}\right)$ for a finite matrix $\mathtt{K}=(k_{ij})$, is the unique square matrix satisfying the following equalities : $KK^{\#}K=K$, $K^{\#}KK^{\#}=K^{\#}$ and $KK^{\#}=K^{\#}K$. We should notice that the group inverse is a particular case of the Drazin inverse (see [7]). Let D denote the deviation matrix of P defined by :

$$D=(I-P+\Pi)^{-1}-\Pi=\sum_{k=0}^{\infty}\left(P^k-\Pi\right)=R-\Pi$$

provided that it exists. So from theorem 3.1 in [34], we have $\mathtt{A}^{\#}=D=R-\Pi$. According to [14], the generalized inverse plays a major role in perturbation analysis for finite Markov chains and has computational advantages than the deviation matrix $D=R-\Pi$. For more details on the potential matrix and the group inverse, we can consult respectively [26, 48, 54] and [34]. The ergodicity coefficient is also used in inequality (1). For a matrix $B=(b_{ij})$ its ergodicity coefficient is defined as follows :

$$\tau(B)=\frac{1}{2}\sup_{i,j\in E}\sum_{k\in E}|b_{ik}-b_{jk}|.$$

Some excellent summaries about the coefficient ergodicity for finite matrices can be found in [18, 50] and for infinite case in [19, 45–47]. Let us review some known results on perturbation and stability bounds.

The first type of results concerns the sensitivity of stationary distribution for finite Markov chains established with respect to the L^p norm for $p=1$ (total variation norm : the sum of absolute values of the vector components) or $p=\infty$ (the maximum absolute value). Here the matrix analysis is the main tool to obtain different inequalities. The conditional number κ given in (1) is investigated in [9, 10, 15, 16, 28–31, 35, 36, 48, 50–53]. In the most framework, the condition number is expressed in terms of the potential $R=(I-P+\Pi)^{-1}$ (fundamental matrix : see [26, 48]) of the unperturbed transition matrix P, the group inverse $A^{\#}$ of the square matrix $I-P$ (see [8, 11, 17, 35, 43]) and the ergodicity coefficients $\tau(P)$, $\tau(R)$ or $\tau\left(\mathtt{A}^{\#}\right)$ (see [51, 53]). Especially, the following norm-wise bounds are derived respectively in [51, 52] :

$$\|\nu-\pi\| \leq \frac{\|\Delta\|}{1-\tau(P)} \tag{2}$$

if $\tau(P)<1$ and

$$\|\nu-\pi\| \leq \tau(\mathtt{A}^{\#})\|\Delta\|. \tag{3}$$

Note that $\tau(\mathtt{A}^{\#})=\tau(R)$. We may consult the main review of those results with a comparison between different bounds in [6]. It is shown in [30] that (3) is the best norm-wise bound. More recently, the graph approach theory is used in [55] in order to obtain a specific norm-wise bound.

For denumerable Markov chains, the use of the weighted norms (v-norms) is the rule. Thus the more general v-norm-wise perturbation bounds of the type (1) are investigated deeply during the last decade. Indeed, under the drift condition $\mathbf{D_2}(v,C,\lambda,b)$ (will be exposed in subsection 3.2), a series expansion of the unperturbed stationary distribution is established for regular and singular perturbation in [2]. Under the same condition, explicit bounds are obtained in [32] and extended for continuous time Markov chains. Different type of bounds are investigated in [40] and a strong stability estimates are derived under the condition $\mathbf{D_1}(n,T,h,\alpha)$ (will be developed in subsection 3.2). Recently, a comparison of some bounds found in the literature for finite and denumerable Markov chains has been done in [1] and also introduced a new bound based on series expansions developed in earlier [12, 13].

The third type of results are devoted to general state Markov chains. Unfortunately there are few published research concerning the perturbed bounds of type (1) since the problem is more hard to derive those corresponding bounds. Using the operator perturbation theory, some transition and stationary estimates are established in [22, 23] with respect to various norms and in [24] with respect to the weighted norm. In [39, 41], a renewal theory is used in order to obtain the transition and stationary bounds with respect to various norms. For nonhomogeneous Markov chains, transition characteristics and stationary estimates are considered in [3] and established with respect to the total variation norm. The norm-wise bound (2) is extended for the general state space in [38, Inequality (3.11)] with respect to the total variation norm. However, the perturbation bounds are established in [22, 23, 39, 41] for a wide class of various norms where the total and weighted variation norm are particular cases. We point out that the need of perturbation bounds with respect to various norms comes from the fact that most of the results are established for a class of geometrically Markov chains with respect to the total variation norm. However, a wide class of Markov chains are not geometrically or uniformly ergodic with respect to the total variation norm like several processes in the queuing systems or other (see e.g. [37]). For example, the simple Bernoulli random walk on $E = \mathbb{Z}_+ = \{0, 1, \dots\}$ with absorption at zero is not uniformly ergodic with respect to the total variation norm. The choice of the appropriate norm depends on the structure of the process. For example, in queuing theory with an infinite capacity of the queues and risk theory, we often use the weighted norms. For quantum Markov processes we use the Schatten p-norm or any unitarily invariant norm. For the most models described by finite-state Markov chains, we use total variation norm (for example for some inventory models and finite queues).

The first goal of this paper is to establish new perturbation and stability bounds for the transition and stationary characteristics as well as to extend for general state and to various norms some existing results obtained either for discrete state or for total variation norm. The second goal is to improve other stability estimates in a more wide stability domain of the unperturbed kernel P.

This paper is organized as follows. Section 2 contains the necessary definitions and notations. In section 3 we expose the main results, with respect to a large class of norms, concerning the transition characteristics. In section 4, we provide the norm-wise bounds for the stationary distribution in terms of the norm of the residual kernel, introduced in the mixing condition $\mathbf{D_1}(n, T, h, \alpha)$, or of the norm ergodicity coefficient. Moreover, we derive explicit bounds under the drift condition $\mathbf{D_2}(v, C, \lambda, b)$. In section 5 we expose examples showing the quality of our results. Finally, to make reading easier for the reader, we present in the appendix the most important results obtained in [32] for which comparisons were made with respect to some results reported in this article.

2. Preliminaries and notations

Let us consider $X = (X_t, t = 0, 1, \dots)$ an homogeneous Markov chain taking values in a measurable space $(E, \mathcal{E})$, where E is the state space of the Markov chain X and $\mathcal{E}$ is a countably generated σ-algebra. Moreover, we assume that X is given by a regular transition kernel $P(x,A),\ x \in E,\ A \in \mathcal{E}$ and having a unique and finite invariant probability measure π.

Denote by $\mathfrak{m}\mathcal{E}$, $f\mathcal{E}$, and $b\mathcal{E}$ the spaces of finite signed measures on $\mathcal{E}$, measurable functions on E, and measurable bounded functions on E, respectively. Let $\mathfrak{m}\mathcal{E}^+$, $f\mathcal{E}^+$, and $\mathfrak{b}\mathcal{E}^+$ the cone of nonnegative elements in these spaces.

For all kernel $K(x,A)$, $x \in E$, $A \in \mathcal{E}$, all measurable function $g \in f\mathcal{E}$, and all signed measure $\mu \in \mathfrak{m}\mathcal{E}$, we introduce the following well known operations: $K g(x) = \int_E g(y) K(x, dy)$, $\mu K(A) = \int_E K(x,A) \mu(dx)$ and $\mu g = \int_E g(x) \mu(dx)$ provided that these integrals are well defined. Assume that a Banach space $\mathfrak{M}$

is given in $\mathfrak{m}\mathcal{E}$ with a norm $\|\cdot\|$ satisfying the following condition

$$|\mu|(E) \leq k\|\mu\| \text{ for } \mu \in \mathfrak{M}. \tag{4}$$

Where $|\mu|$ is the variation of the measure μ and k is a finite positive constant. We introduce the following consistency condition on the norm with the order structure and the uniform topology in $m\mathcal{E}$:

$$\|\mu_1\| \leq \|\mu_1 + \mu_2\| \, \forall \mu_i \in \mathfrak{M}^+ \;\; \forall i = 1,2 \tag{5}$$

$$\|\mu_1\| \leq \|\mu_1 - \mu_2\| \;\; \forall \mu_i \in \mathfrak{M}^+ \text{ and } \mu_1 \perp \mu_2. \tag{6}$$

The conjugate norm is defined on the dual space $\mathfrak{N}$ of functions with finite norm as follows: $\|f\| = \sup\{|\mu f|, \|\mu\| \leq 1\}$. Furthermore, we introduce the space $\mathfrak{B}$ of operators K such that $\mathfrak{M}K \subset \mathfrak{M}$ and the following corresponding operator norm is finite $\|K\| = \sup\{\|\mu K\|, \|\mu\| \leq 1\}$. In the sequel of this paper, we use the same notation $\|\cdot\|$ for the measure, function and operator norms. Moreover, transition kernels and corresponding linear operators are denoted by the same symbols. By definition, $\mu \leq \lambda$ if and only if $\lambda - \mu \in \mathfrak{M}^+$, and $K_1 \leq K_2$ if and only if $\mu(K_2 - K_1) \in \mathfrak{M}^+$ for $\mu \in \mathfrak{M}^+$. We define the product K_1K_2 of two kernels K_1 and K_2 as follows

$$\forall (x,A) \in E \times \mathcal{E} : K_1K_2(x,A) = \int_E K_1(x,dy)K_2(y,A).$$

We denote by K^t the t-times product of K by itself and $K^0 = I$, where here $I(x,A) = \mathbb{1}_{\{x \in A\}}$ is the unit operator in $\mathfrak{M}$. We assume $P \in \mathfrak{B}$, that is,

$$\mathfrak{M}P \subset \mathfrak{M} \text{ and } \|P\| < \infty.$$

For a function $f \in \mathfrak{N}$ and measure $\mu \in \mathfrak{M}$, we denote $f \otimes \mu$ the kernel defined by $(f \otimes \mu)(x,A) = f(x)\mu(A)$ for all $x \in E$ and $A \in \mathcal{E}$. The stationary projector Π of the kernel P on $(E,\mathcal{E})$, is defined as a kernel which verifies $\Pi P = P\Pi = \Pi^2 = \Pi$ and $\mu\Pi = \mu$ provides that $\mu P = \mu$ for $\mu \in \mathfrak{M}$. If a kernel P admits a unique invariant measure π, then P has a unique stationary projector Π of the form $\Pi = \mathbb{1} \otimes \pi$ where $\mathbb{1}$ is a function identically equals to the unit (see [24] for more details). Observe that (4) ensures that $\|\mathbb{1}\| \leq k$ is finite and so if $\pi \in \mathfrak{M}$ then $\|\Pi\| = \|\mathbb{1} \otimes \pi\| \leq \|\mathbb{1}\|\,\|\pi\|$ is also finite.

Let a measurable function v on E such that $\hbar = \inf_E v(x) > 0$. Define, for all $\mu \in \mathfrak{m}\mathcal{E}$, a weighted variation norm

$$\|\mu\|_v = \int_E v(x)|\mu|(dx). \tag{7}$$

This class of norms will be considered in examples of section 7. Le denote $\mathfrak{M}_v = \{\mu \in \mathfrak{m}\mathcal{E} : \|\mu\|_v < \infty\}$. The corresponding norms in $\mathfrak{N}_v$ and $\mathfrak{B}_v$ are of the form

$$\|f\|_v = \sup_{x\in E}\left(\frac{|f(x)|}{v(x)}\right) \text{ and } \|K\|_v = \sup_{x\in E}\left(\frac{\int_E v(y)|P(x,dy)|}{v(x)}\right). \tag{8}$$

For a signed measure $\mu \in \mathfrak{M}_v$, the v-norm, for $v \equiv \mathbb{1}$, coincides with the total variation norm of μ.

The following definition is a generalization, for any norms verifying (4)-(6), of the classical uniform ergodicity of a Markov chain usually defined for either weighted or total variation norm.

Definition 2.1. The chain X is called uniformly ergodic with respect to a given norm $\|\cdot\|$ if there exist positive constants $\rho < 1$ and $C < \infty$ such that for all $t \in \mathbb{Z}_+$, we have

$$\|P^t - \Pi\| \le C\rho^t. \tag{9}$$

An alternative definition of the uniform ergodicity states: $\|P^t - \Pi\| \le \rho^t$ for $t \ge N$. We point out that we do not know the length of the transition phase. Some explicit estimates of the rate of convergence in (9) for general state space are established earlier in [23], and recently in [41] for norms other than the weighted or total variation norms. Bounds for some class of Markov chains (Doeblin's Chains, reversible chains, monotone chains, etc.) have been established with respect to the weighted and total variation norms (see e.g. [37, Chapter 16] and the references therein).

For a probability measure $\pi^{(0)}$, we define the distribution of X over t steps by $\pi^{(t)} = \pi^{(0)}P^t$ (transition distribution) where $t \ge 1$. We introduce an other Markov chain $Y = \{Y_t, t \in \mathbb{Z}_+\}$ given on the same phase state $(E,\mathcal{E})$. Let denote Q the transition operator of Y and $\nu^{(t)} = \nu^{(0)}Q^t$ the distribution of Y over t steps. Thus X is regarded as the unperturbed chain, whereas Y is considered as the perturbed one.

It is well known in general perturbation theory that the first goal is to establish upper bounds for the deviation of transition or stationary characteristics in term of the deviation of the transition operators $\|P - Q\|$. For this, we denote $\Delta = Q - P$, $\Delta_t = Q^t - P^t$ and $\Delta^{(t)} = \nu^{(t)} - \pi^{(t)}$ for $t \ge 0$.

To establish the perturbation bounds, we will use the generalization of the concept of ergodicity coefficient of an operator. Thus, for all operator $K \in \mathfrak{B}$ and an integer $m \ge 1$, we define the norm ergodicity coefficient of order m of K as follows

$$\Lambda_m(K) = \sup\{\|\mu K^m\| : \|\mu\| \le 1, \mu \in \mathfrak{M}_0\}$$

with $\mathfrak{M}_0 = \{\mu \in \mathfrak{M} : \mu(E) = 0\}$.

For the total variation norm and discrete space, the norm ergodicity coefficient $\Lambda_m(K)$ coincides with the well known ergodicity coefficient denoted by $\tau_m(K)$ (eg. [49]). It is easy to show that for transition operators K_1, K_2 we have $\Lambda_1(K_1^m + K_2^m) \leq \Lambda_m(K_1) + \Lambda_m(K_2)$ and $\Lambda_1(K_1^m.K_2^m) \leq \Lambda_m(K_1)\Lambda_m(K_2)$. Moreover, and in contrast to the total variation norm, we may have (for example for the weighted variation norm) $\Lambda_1(K) > 1$ for a transition operator K. It is of interest to precise that for the norm which verifies (4)-(6), the uniform ergodicity of the chain as it is stated in definition 2.1, is equivalent to the existence of an integer $m \geq 1$ such that $\Lambda_m(P) < 1$. Finally, in the sequel of this paper, we denote $\Lambda_m = \Lambda_m(P)$, $\mathbf{q}_s = \max_{0\leq i\leq s} \|Q^i\|$, $\mathbf{p}_s = \max_{0\leq i\leq s} \|P^i\|$, $\mathbf{q} = \sup_{s\geq 0} \|Q^s\|$, $\mathbf{p} = \sup_{s\geq 0} \|P^s\|$, $\overline{\Delta}_s = \max_{1\leq i\leq s} \|\Delta_i\|$ and $\overline{\Lambda}_s = \max_{1\leq i\leq s} \Lambda_i(P)$. Furthermore, we denote $\lfloor x \rfloor$ the largest integer less than or equal to x. Henceforth, we use the generic $\|\cdot\|$ sign whenever the result holds for any norm verifying (4). In the case of the v-norm (weighted variation norm) or the total variation norm we use $\|\cdot\|_v$ and $\|\cdot\|_1$ respectively.

3. Perturbation bounds of transition characteristics

3.1. General perturbation and stability bounds

We start our investigation by stating the following lemma which allows us to establish the main results of this section and theorem 4.7 in section 4. For this, we denote by $\mathfrak{B}_0 = \{K \in \mathfrak{B} : K\mathbb{1} = 0\}$ the set of all weak derivatives of transition kernels and $\Omega = \{K \in \mathfrak{B} : K\mathbb{1} = a\mathbb{1}, \text{ for some } a \in \mathbb{R}\} \supset \mathfrak{B}_0$ the set of all Markov chains up to some normalizing constant. Note that any transition operator belongs to Ω.

Lemma 3.1. *Let $L \in \mathfrak{B}_0$, $K \in \Omega$ and $\mu \in \mathfrak{M}_0$. Then, for all $t \geq 0$ and $m \geq 1$, we get*

$$\|LK^{tm}\| \leq \|L\| \, (\Lambda_m(K))^t \tag{10}$$

and

$$\|\mu K^{tm}\| \leq \|\mu\| \, (\Lambda_m(K))^t. \tag{11}$$

Proof. Consider a fixed integer $m \geq 1$ and the operators sequence $\{\mathbb{H}_t\}_{t\geq 0}$ defined for all $t \geq 0$ as follows

$$\mathbb{H}_t = \frac{1}{\|LK^{tm}\|} LK^{tm}.$$

Notice that $\mathbb{H}_t \in \mathfrak{B}_0$ for $t \geq 0$ and $\|LK^m\| = \|L\| \, \|\mathbb{H}_0 K^m\|$. Thus, assuming the following identity is true

$$\|LK^{(t-1)m}\| = \|L\| \prod_{k=0}^{t-2} \|\mathbb{H}_k K^m\|.$$

Then we have

$$\|LK^{tm}\| = \|LK^{(t-1)m}\|\,\|\mathbb{H}_{t-1}K^m\| = \left(\|L\|\prod_{k=0}^{t-2}\|\mathbb{H}_k K^m\|\right)\|\mathbb{H}_{t-1}K^m\|$$
$$= \|L\|\prod_{k=0}^{t-1}\|\mathbb{H}_k K^m\|.$$

Hence, we obtain by induction that for all $t \geq 0$,

$$\|LK^{tm}\| = \|L\|\prod_{k=0}^{t-1}\|\mathbb{H}_k K^m\|. \tag{12}$$

Moreover, for all measure $\mu \in \mathfrak{M}$ and $t \geq 1$, the measure $\nu = \mu\mathbb{H}_t \in \mathfrak{M}_0$ and verifies $\|\nu\| = \|\mu\mathbb{H}_t\| \leq \|\mu\|$. This yields the following assertion

$$\begin{aligned}\|\mathbb{H}_t K^m\| &= \sup\{\|\mu(\mathbb{H}_t K^m)\| : \|\mu\| \leq 1\}\\ &\leq \sup\{\|\overline{\mu}K^m\| : \|\overline{\mu}\| \leq 1, \overline{\mu}(E) = 0, \overline{\mu} \in \mathfrak{M}\} = \Lambda_m(K).\end{aligned}$$

Substituting the latter estimate in the identity (12), we obtain inequality (10).

Let denote

$$\mathbb{B}_t = \frac{1}{\|\mu K^{tm}\|}\,\mu K^{tm}.$$

So in the same way, we obtain the following identity

$$\|\mu K^{tm}\| = \|\mu\|\prod_{k=0}^{t-1}\|\mathbb{B}_k K^m\|. \tag{13}$$

Furthermore, for all $k \geq 0$, we have $\mathbb{B}_k \in \mathfrak{M}_0$, and $\|\mathbb{B}_k\| = 1$ and consequently $\|\mathbb{B}_k K^m\| \leq \Lambda_m(K)$. Thus,by utilizing the previous estimate in the equality (13), we derive (11). □

The first estimate concerns the upper perturbation bound for the deviation of the transition operators over t-steps where $t \geq 0$.

Theorem 3.2. *For all $t \geq 0$ and fixed $m \geq 1$, we have*

$$\|Q^{tm} - P^{tm}\| \leq \mathbf{q}_{(t-1)m}\,\|\Delta_m\|\,\frac{1-\Lambda_m^t}{1-\Lambda_m} \tag{14}$$

and

$$\|Q^t - P^t\| \leq \mathbf{q}_t\,\|\Delta_m\|\,\frac{1-\Lambda_m^{\lfloor \frac{t}{m}\rfloor}}{1-\Lambda_m} + \Lambda_m^{\lfloor \frac{t}{m}\rfloor}\,\|\Delta_{t-\lfloor \frac{t}{m}\rfloor m}\|. \tag{15}$$

Moreover, if we assume $\Lambda_m < 1$ and $\mathbf{q} < \infty$, then the following inequality holds

$$\sup_{t\geq 0}\|Q^t - P^t\| \leq \mathbf{q}\,\frac{\|\Delta_m\|}{1-\Lambda_m} + \overline{\Delta}_{m-1}. \tag{16}$$

Proof. For all $t \geq 1$, we have $\Delta_{tm} = Q^m \Delta_{(t-1)m} + \Delta_m P^{(t-1)m}$. Utilizing the induction procedure, we obtain the following identity

$$\Delta_{tm} = Q^{(t-1)m}\Delta_m + Q^{(t-2)m}\Delta_m P^m + \ldots + \Delta_m P^{(t-1)m}. \tag{17}$$

This yields to the following inequality

$$\|\Delta_{tm}\| \leq \|Q^{(t-1)m}\| \, \|\Delta_m\| + \|Q^{(t-2)m}\| \, \|\Delta_m P^m\| + \ldots + \|\Delta_m P^{(t-1)m}\|.$$

So using lemma 3.1, we obtain

$$\begin{aligned}\|\Delta_{tm}\| &\leq \|Q^{(t-1)m}\| \, \|\Delta_m\| + \|Q^{(t-2)m}\| \, \|\Delta_m\| \, \Lambda_m + \ldots + \|\Delta_m\| \, \Lambda_m^{t-1} \\ &\leq \|\Delta_m\| \sup_{0 \leq s \leq (t-1)m} \|Q^s\| \left(1 + \Lambda_m + \ldots + \Lambda_m^{t-1}\right) = \mathbf{q}_{(t-1)m} \|\Delta_m\| \frac{1-\Lambda_m^t}{1-\Lambda_m}.\end{aligned}$$

So, the estimate (14) is established.

Set $t = km + s$ where $s = \overline{0, m-1}$ and $k \geq 1$, then we have the following recursive equation $\Delta_t = Q^s \Delta_{km} + \Delta_s P^{km}$. Therefore, using (17), we get

$$\begin{aligned}Q^t - P^t &= Q^s \left(Q^{km} - P^{km}\right) + \left(Q^s - P^s\right) P^{km} = Q^s \sum_{i=0}^{k-1} Q^{im} \Delta_m P^{(k-1-i)m} + \Delta_s P^{km} \\ &= \sum_{i=0}^{k-1} Q^{im+s} \Delta_m P^{(k-1-i)m} + \left(Q^s - P^s\right) P^{km}.\end{aligned}$$

According to lemma 3.1, we obtain

$$\begin{aligned}\|\Delta_t\| &\leq \mathbf{q}_t \, \|\Delta_m\| \sum_{i=0}^{k-1} \Lambda_m^{k-1-i} + \Lambda_m^{\lfloor \frac{t}{m} \rfloor} \, \|Q^{t-\lfloor \frac{t}{m} \rfloor m} - P^{t-\lfloor \frac{t}{m} \rfloor m}\| \\ &\leq \mathbf{q}_t \, \|\Delta_m\| \frac{1-\Lambda_m^{\lfloor \frac{t}{m} \rfloor}}{1-\Lambda_m} + \Lambda_m^{\lfloor \frac{t}{m} \rfloor} \, \|\Delta_{t-\lfloor \frac{t}{m} \rfloor m}\|.\end{aligned}$$

Then inequality (15) is obtained. Moreover, the inequality (16) follows from the condition $\Lambda_m < 1$. □

Remark 3.3. It is easy to derive corresponding estimates directly in term of the deviation $\|\Delta\|$. Indeed, we have the relation $\Delta_m = \sum_{k=0}^{m-1} Q^k \Delta P^{m-k-1}$. Then, we obtain

$$\|\Delta_m\| \leq m \, \mathbf{q}_{m-1} \, \overline{\Lambda}_{m-1} \, \|\Delta\| \leq m \, \mathbf{q}_{m-1} \, \mathbf{p}_{m-1} \|\Delta\|.$$

The need to consider a wide class of norms is explained in the introduction of this paper and is related to the fact that several processes are not necessarily uniformly ergodic for the usual total variation norm and consequently have

$\Lambda_m = \tau_m = 1$ for all m. Hence the bounds [38, Bounds (3.11) and (3.15)] are not applicable. A simple example is given by the Bernoulli random walk on $E = \mathbb{Z}_+ = \{0, 1, \ldots\}$ with absorption at zero which will be investigated in example 5.3 of section 7. For this aim, theorems 3.4 and 4.7 extend, for various norms, [38, Bounds (3.11) and (3.15)].

Theorem 3.4. *Let us assume $\Lambda_m < 1$ for some $m \geq 1$ and $\mathbf{q} < \infty$. Then, for all $t \geq 0$, the following estimates are fulfilled*

$$\|\Delta^{(t)}\| \leq \Theta_m \Lambda_m^{\lfloor \frac{t}{m} \rfloor} + \mathbf{q}_t \, \|\nu^{(0)}\| \, \frac{1 - \Lambda_m^{\lfloor \frac{t}{m} \rfloor}}{1 - \Lambda_m} \, \|\Delta_m\| \tag{18}$$

and

$$\sup_{t \geq 0} \|\Delta^{(t)}\| \leq \|\Delta^{(0)}\| \, \mathbf{p}_{m-1} + \|\nu^{(0)}\| \, \overline{\Delta}_{m-1} + \frac{\|\nu^{(0)}\| \, \mathbf{q}}{1 - \Lambda_m} \, \|\Delta_m\|, \tag{19}$$

where $\Theta_m = \|\Delta^{(0)}\| \, \overline{\Lambda}_{m-1} + \|\nu^{(0)}\| \, \overline{\Delta}_{m-1} \leq \|\Delta^{(0)}\| \, \mathbf{p}_{m-1} + \|\nu^{(0)}\| \, \overline{\Delta}_{m-1}$.

Proof. From (17), we have for every integer $t = km + s$ where $k \geq 1$,

$$\begin{aligned} \nu^{(t)} - \pi^{(t)} &= \nu^{(s)} \left(Q^{km} - P^{km} \right) + \left(\nu^{(s)} - \pi^{(s)} \right) P^{km} \\ &= \nu^{(s)} \sum_{i=0}^{k-1} Q^{im} \Delta_m P^{(k-1-i)m} + \left(\nu^{(s)} - \pi^{(s)} \right) P^{km} \\ &= \nu^{(0)} \sum_{i=0}^{k-1} Q^{im+s} \Delta_m P^{(k-1-i)m} + \left(\nu^{(s)} - \pi^{(s)} \right) P^{km}. \end{aligned}$$

Therefore, utilizing lemma 3.1 we derive

$$\begin{aligned} \|\Delta^{(t)}\| \quad &\leq \quad \|\nu^{(0)}\| \, \mathbf{q}_t \|\Delta_m\| \sum_{i=0}^{\lfloor \frac{t}{m} \rfloor - 1} \Lambda_m^{\lfloor \frac{t}{m} \rfloor - 1 - i} + \|\Delta^{(t - \lfloor \frac{t}{m} \rfloor m)}\| \, \Lambda_m^{\lfloor \frac{t}{m} \rfloor} \\ &= \quad \|\nu^{(0)}\| \, \mathbf{q}_t \, \frac{1 - \Lambda_m^{\lfloor \frac{t}{m} \rfloor}}{1 - \Lambda_m} \, \|\Delta_m\| + \|\Delta^{(t - \lfloor \frac{t}{m} \rfloor m)}\| \, \Lambda_m^{\lfloor \frac{t}{m} \rfloor}. \end{aligned} \tag{20}$$

Furthermore, we have for all $s = \overline{0, m-1}$, the identity $\Delta^{(s)} = \Delta^{(0)} P^s + \nu^{(0)} \Delta_s$, which yields the following inequality $\|\Delta^{(s)}\| \leq \|\Delta^{(0)}\| \, \Lambda_s(P) + \|\nu^{(0)}\| \, \|\Delta_s\|$. It follows the following inequality

$$\sup_{0 \leq s \leq m-1} \|\Delta^{(s)}\| \leq \|\Delta^{(0)}\| \, \overline{\Lambda}_{m-1} + \|\nu^{(0)}\| \, \overline{\Delta}_{m-1} \leq \Delta^{(0)}\| \, \mathbf{p}_{m-1} + \|\nu^{(0)}\| \, \overline{\Delta}_{m-1}. \tag{21}$$

The bound (19) derives straightforwardly from (20) and (21). $\square$

Remark 3.5. Observe that for the total variation norm, the inequalities (18) and (19) become

$$\|\Delta^{(t)}\|_1 \leq \left(\|\Delta^{(0)}\|_1 + \overline{\Delta}_{m-1}\right) \Lambda_m^{\lfloor \frac{t}{m} \rfloor} + \frac{1-\Lambda_m^{\lfloor \frac{t}{m} \rfloor}}{1-\Lambda_m} \|\Delta_m\|_1 \tag{22}$$

and

$$\sup_{t\geq 0} \|\Delta^{(t)}\|_1 \leq \|\Delta^{(0)}\|_1 + \overline{\Delta}_{m-1} + \frac{\|\Delta_m\|_1}{1-\Lambda_m}. \tag{23}$$

It is worth noting that the bound (22) coincides with [38, Bound (3.15)].

Theorem 3.6. *Let us assume $\Lambda_m < 1$ for $m \geq 1$ and $\mathbf{p} = \sup_{t\geq 0} \|P^t\| < \infty$. Then, for all $t \geq 0$ and $\Lambda_m + \|\Delta_m\| < 1$, the following estimates are fulfilled*

$$\|\Delta^{(t)}\| \leq \Theta_m \left(\Lambda_m + \|\Delta_m\|\right)^{\lfloor \frac{t}{m} \rfloor} + \|\pi^{(0)}\| \, \mathbf{p}_t \, \frac{1-(\Lambda_m + \|\Delta_m\|)^{\lfloor \frac{t}{m} \rfloor}}{1-\Lambda_m - \|\Delta_m\|} \|\Delta_m\| \tag{24}$$

and

$$\sup_{t\geq 0} \|\Delta^{(t)}\| \leq \|\Delta^{(0)}\| \, \mathbf{p}_{m-1} + \|\nu^{(0)}\| \, \overline{\Delta}_{m-1} + \frac{\|\pi^{(0)}\| \, \mathbf{p}}{1-\Lambda_m - \|\Delta_m\|} \|\Delta_m\|. \tag{25}$$

Where Θ_m is defined in theorem 3.4.

Proof. We have for every integer $t = km + s$ where $k \geq 1$,

$$\begin{aligned}
\nu^{(t)} - \pi^{(t)} &= \pi^{(s)} \left(Q^{km} - P^{km}\right) + \left(\nu^{(s)} - \pi^{(s)}\right) Q^{km} \\
&= \pi^{(s)} \sum_{i=0}^{k-1} P^{im} \Delta_m Q^{(k-1-i)m} + \left(\nu^{(s)} - \pi^{(s)}\right) Q^{km} \\
&= \pi^{(0)} \sum_{i=0}^{k-1} P^{im+s} \Delta_m Q^{(k-1-i)m} + \left(\nu^{(s)} - \pi^{(s)}\right) Q^{km}.
\end{aligned}$$

According to lemma 3.1 we get:

$$\begin{aligned}
\|\Delta^{(t)}\| &\leq \|\pi^{(0)}\| \, \mathbf{p}_t \|\Delta_m\| \sum_{i=0}^{k-1} \left(\Lambda_m + \|\Delta_m\|\right)^{k-1-i} + \|\Delta^{(s)}\| \left(\Lambda_m + \|\Delta_m\|\right)^k \\
&= \|\pi^{(0)}\| \, \mathbf{p}_t \, \frac{1-(\Lambda_m + \|\Delta_m\|)^k}{1-\Lambda_m - \|\Delta_m\|} \|\Delta_m\| + \|\Delta^{(s)}\| \left(\Lambda_m + \|\Delta_m\|\right)^k.
\end{aligned}$$

The assertion follows from the latter and (21) which complete the proof. □

Remark 3.7. If we consider the usual total variation norm, we must point out that those bounds are useless if are greater than 2 since $\sup_{t\geq 0} \|\Delta_t\| \leq 2$ and $\sup_{t\geq 0} \|\Delta^{(t)}\| \leq 2$.

3.2. Estimates of the parameters q and p

The previous bounds are expressed in terms of the parameters $\mathbf{q}$ and $\mathbf{p}$. In general situations, the structure of the transition kernel P of the unperturbed Markov chain is generally simpler than that of the perturbed chain Y. Consequently, the spectral decomposition of P allows us to estimate $\mathbf{p}$ more easier than $\mathbf{q}$. Especially, the index $\mathbf{q}$ is finite if, in particular, the Markov chain X is uniformly ergodic and aperiodic for all sufficiently small $\|\Delta\|$. In the case of the total variation norm, we have $\mathbf{p} = \mathbf{q} = 1$ and the previous bounds are sufficiently explicit.

For this purpose, we establish in this subsection some estimates of these parameters under some conditions. First we consider the following mixing condition for the unperturbed Markov chain X.

$\mathbf{D_1}(n, T, h, \alpha)$:

I) $\|P\| < \infty$.

II) There exist a natural integer n, nonnegative measurable function $h \in \mathfrak{N}^+$, measure $\alpha \in \mathfrak{M}^+$ such that: $\alpha h > 0$, $\pi h > 0$, $\alpha \mathbb{1} = 1$ and the operator $T = P^n - h \otimes \alpha$ is nonnegative.

III) $\|T\| < 1$ (T is called proper or quasi-compact for the norm $\|\cdot\|$).

Remark 3.8. Condition (II) is fulfilled for all Harris Markov chains. Furthermore, conditions (II) and (III) for the usual total variation norm $\|\mu\| = |\mu|(E)$ are equivalent to the well known Doeblin condition for the kernel P to be quasi compact. For more details we may consult [44, Chapitre 6.3]. Observe that calculating $\|T\|$ is not easy to do. For this, we use the following equivalent condition for (III):

IV) There exists $\rho \in [0, 1[$ such that $\|T\| \leq \rho$.

For the weighted variation norm, condition (III) or (IV) is equivalent to the following condition: there exists $\rho \in [0, 1[$ such that

$$\forall x \in E : Tv(x) \leq \rho\, v(x). \tag{26}$$

We point out that the kernel T is called the residual kernel (see [39]). In some situations it represents a degenerate kernel that avoids entering some specific set state A (Markov kernel conditioned on the even that the chain does not reach A). In this case, T is called a taboo kernel and denoted by ${}_AP$ (see [37] for details and the references therein).

It is worth noting that finding the appropriate substochastic kernel T which verifying condition (II) may be done by using the first input and last exit formula (eg. see [37]). This technique have been used in an elegant and flexible way in

[24]. The form of the kernel T is not unique. Unfortunately, condition (III) is more difficult to be satisfied and consequently choosing appropriate kernel T verifying conditions (II)-(III) is more hard to do. Note that the choice of a kernel T is equivalent to determining a function h and a measure α. Unfortunately, the optimal choice of the kernel T, and consequently the choice of the function h and the measure α, remains an open problem.

The two following results establish the estimate of $\mathbf{p}$ and $\mathbf{q}$ under the mixing condition $\mathbf{D_1}(n,T,h,\alpha)$.

Proposition 3.9. *Assume that condition $\mathbf{D_1}(n,T,h,\alpha)$ holds. Then, the following estimates are fulfilled for all perturbed transition kernel Q such that $\|\Delta_n\| < 1-\|T\|$:*

$$\sup_{i\geq 0}\|Q^{in}\| = \mathbf{q}_{(n)} \leq \frac{\|h\|\,\|\alpha\|}{(1-\|T\|-\|\Delta_n\|)^2} \tag{27}$$

and

$$\sup_{i\geq 0}\|Q^{i}\| = \mathbf{q} \leq \frac{\|h\|\,\|\alpha\|\left(\overline{\Delta}_{n-1}+\mathbf{p}_{n-1}\right)}{(1-\|T\|-\|\Delta_n\|)^2}. \tag{28}$$

Proof. Since T, h and α are nonnegative, so from the sub-condition (II) of the condition $\boldsymbol{D_1}(n,h,\alpha,T)$, we get for all $x\in E$: $1 = P^n(x,E) \geq h(x)\alpha(E)$. This yields, for all $x\in E$ and $A\in\mathcal{E}$, $h(x)\alpha(A)\leq 1$. Hence, for all $i\geq 0$, we obtain

$$|\alpha Q^{in}h| = |\int_E\int_E \alpha(dx)Q^{in}(x,dy)h(y)| \leq 1. \tag{29}$$

Moreover, we have for all $i\geq 1$

$$\begin{aligned} Q^{in} &= Q^{in} - Q^{(i-1)n}P^n + Q^{(i-1)n}T + Q^{(i-1)n}h\otimes\alpha \\ &= Q^{(i-1)n}(\Delta_n+T) + Q^{(i-1)n}h\otimes\alpha. \end{aligned}$$

Therefore, we get

$$\begin{aligned} \|\alpha Q^{in}\| &= \|\alpha\left(Q^{(i-1)n}(\Delta_n+T)\right) + (\alpha Q^{(i-1)n}h)\alpha\| \\ &\leq \|\alpha Q^{(i-1)n}\|\,\|\Delta_n+T\| + |\alpha Q^{(i-1)n}h|\,\|\alpha\|. \end{aligned}$$

This implies the following inequality

$$\sup_{i\geq 1}\|\alpha Q^{in}\| \leq \frac{\|\alpha\|}{1-\|\Delta_n+T\|}|\alpha Q^{(i-1)n}h|.$$

Since $\|\Delta_n+T\| \leq \|\Delta_n\|+\|T\| < 1$ and from (29), we obtain

$$\sup_{i\geq 1}\|\alpha Q^{in}\| \leq \frac{\|\alpha\|}{1-\|T\|-\|\Delta_n\|}. \tag{30}$$

Note that for $i=0$, the last inequality (30) holds also true. Furthermore, from the trivial recursive relation $Q^{in} = (\Delta_n + T)Q^{(i-1)n} + h \otimes \alpha Q^{(i-1)n}$, we obtain

$$\|Q^{in}\| \leq \|Q^{(i-1)n}\| \, \|\Delta_n + T\| + \|\alpha Q^{(i-1)}\| \, \|h\|.$$

This yields the following

$$\sup_{i\geq 0} \|Q^{in}\| \leq \frac{1}{1 - \|\Delta_n + T\|} \|h\| \sup_{t\geq 0} \|\alpha Q^{tn}\|. \tag{31}$$

By using the latter inequality and (30), we derive (27). Moreover, for $t = kn + s$ where $0 \leq s \leq n-1$, we have

$$\|Q^t\| \leq \|Q^{kn}\| \|Q^s\| \leq \mathbf{q}_{(n)} \, \mathbf{q}_{n-1} \leq \mathbf{q}_{(n)} \, \left(\overline{\Delta}_{n-1} + \mathbf{p}_{n-1}\right). \tag{32}$$

Therefore, by using (27) and (32), the inequality (28) is derived. □

Proposition 3.10. *Assume that condition* $\mathbf{D_1}(n,T,h,\alpha)$ *holds. Then, we get*

$$\sup_{k\geq 1} \|P^{kn}\| = \boldsymbol{p}_{(n)} \leq \|T\| + \frac{\|h\| \, \|\alpha\|}{(1 - \|T\|)^2} \tag{33}$$

and

$$\boldsymbol{p} = \sup_{i\geq 1} \|P^i\| \leq \left(\|T\| + \frac{\|h\| \, \|\alpha\|}{(1-\|T\|)^2} \right) \boldsymbol{p}_{n-1}. \tag{34}$$

Proof. We consider the probability distribution $(p_n; n = 1,2,\ldots)$ defined for all $n \geq 1$ by $p_n = \alpha T^{n-1} h$. We denote by d the positive integer called the period of the Markov chain defined by $d = gcd\{n \geq 1 : p_n > 0\} \geq 1$. Under the condition $\mathbf{D_1}(n,T,h,\alpha)$, we have $d = 1$ since $p_1 = \alpha h > 0$. In this case P is called aperiodic transition kernel. It is well known that $(p_n; n = 1,2,\ldots)$ verifies the following renewal equation

$$\begin{cases} \lambda(t) = \sum_{k=1}^{t-1} \lambda(k) p_{t-k}, & t \geq 2; \\ \lambda(t) = 0, & t \leq 0; \\ \lambda(1) = 1. \end{cases} \tag{35}$$

According to [41, Relation (1)], we have $P^t = T^t + \sum_{i\geq 0} \sum_{j\geq 0} \lambda_{ij}^t T^i h \otimes \alpha T^j$, where $\lambda_{ij}^t = \lambda(t-i-j)$. From the latter equation, it follows that for $t \geq 1$:

$$\|P^t\| \leq \|T^t\| + \sup_{s\geq 0} \lambda(s) \, \|h\| \, \|\alpha\| \sum_{i\geq 0} \sum_{j\geq 0} \|T\|^{i+j} \leq \|T\| + \frac{\|h\| \, \|\alpha\|}{(1-\|T\|)^2}.$$

The inequality (34) derives straightforwardly from (33) and the obvious inequality $\boldsymbol{p} \leq \boldsymbol{p}_{(n)} \boldsymbol{p}_{n-1}$. □

Remark 3.11. It is worth noting that under the mixing condition $\mathbf{D_1}(n,T,h,\alpha)$, we can easily establish $(\pi h)\,\alpha \le \pi = (\pi h)\,\alpha\,(I-T)^{-1}$ and $\|\pi\| \le \dfrac{(\pi h)\,\|\alpha\|}{1-\|T\|}\cdot$

In the sequel of this section, we derive explicit bounds, with respect to the weighted variation norm, under the following drift condition.

$\mathbf{D_2}(v,C,\lambda,b)$: There exists a finite function v bounded away from zero, a measurable set C and positive constants $\lambda < 1$, $b < \infty$ such that:

$$Pv \le \lambda v + b\mathbb{1}_C. \tag{36}$$

It is well known that if X is an aperiodic Markov chain with transition kernel P, then X is geometrically ergodic if and only if P satisfies the drift condition $\mathbf{D_2}(v,C,\lambda,b)$ for an unbounded test function $v \ge 1$ (see [37]).

Proposition 3.12. *Under the condition* $\mathbf{D_2}(v,C,\lambda,b)$*, we have the following estimate*

$$\sup_{t\ge 1}\|P^t\|_v = \boldsymbol{p} \le \lambda + \frac{b}{\hbar\,(1-\lambda)}\cdot \tag{37}$$

Where $\hbar = \inf_{x\in E} v(x) > 0$.

Proof. Using the induction procedure from the inequality (36), we obtain:

$$P^t v = \lambda v + b\sum_{k=0}^{t}\lambda^k = \lambda v + b\frac{1-\lambda^{t+1}}{1-\lambda} \le \lambda v + \frac{b}{1-\lambda}\cdot$$

According to the latter and the definition of the v-norm (see relation (8)), we get:

$$\begin{aligned}\|P^t\|_v = \sup_{x\in E}\left(\frac{P^t v(x)}{v(x)}\right) &\le \sup_{x\in E}\left(\lambda + \frac{1}{v(x)}\,\frac{b}{1-\lambda}\right)\\ &\le \lambda + \frac{b}{\left(\inf_{x\in E} v(x)\right)\,(1-\lambda)}\cdot\end{aligned}$$

This yields the desired result. □

Remark 3.13. By integrating both sides of the inequality (36) with respect to the invariant probability measure π, we get $\pi P v \le \lambda\,\pi v + b\,\pi(C)$ and using the invariance of π with respect to P, we obtain the following estimate

$$\pi v = \|\pi\|_v \le \frac{b\,\pi\,(C)}{1-\lambda}\cdot \tag{38}$$

4. Perturbation bounds for stationary distributions

In order to preserve the irreducibility of the unperturbed chain, we have to consider the problem of the stability effects. For this, we consider the perturbed kernel Q with an invariant probability measure ν. Since ν is an invariant measure of Q ($\nu = \nu Q$) and $\nu \mathbb{1} = 1$, we get easily

$$\nu(I-P+\Pi) = \nu(Q-P)+\nu\Pi = \nu(Q-P)+\pi.$$

Moreover, $R = (I-P+\Pi)^{-1}$ is well defined according to the basic theory of the linear operators (see [21]) and $\pi R = \pi$. Hence, we have

$$\nu = \nu(Q-P)R+\pi R = \nu(Q-P)R+\pi.$$

However, $\Delta\Pi = \Delta(\mathbb{1}\otimes\pi) = 0$. Hence, $\nu = \nu\Delta(R-\Pi)+\pi$, and consequently $\nu-\pi = \nu\Delta R = \nu-\pi = \nu\Delta(R-\Pi)$. This yields $\nu(I-\Delta(R-\Pi)) = \pi$. If $\|\Delta(R-\Pi)\| < 1$, then the operator $(I-\Delta(R-\Pi))^{-1}$ exists and norm bounded and its Neumann series converges. So we have straightforwardly the following relation $\nu = \pi\,(I-\Delta(R-\Pi))^{-1} = \pi\sum_{k\geq 0}[\Delta(R-\Pi)]^k$. Or, explicitly,

$$\nu-\pi = \pi\sum_{k\geq 1}[\Delta(R-\Pi)]^k = \pi\sum_{k\geq 1}[\Delta(R-\Pi)]^k. \tag{39}$$

In the sequel, we consider norms satisfying (4) and we assume that $\|R\| < \infty$ and $\|\Pi\| < \infty$ which implies that the deviation operator $D = R-\Pi$ has a finite norm ($\|D\| = \|R-\Pi\| < \infty$). Indeed, if the considered norm satisfies also the assertions (5)-(6) as the weighted and total variation norms, it follows from [21], then the chain is uniformly ergodic (see definition 2.1) with respect to this given norm if and only if the operator $I-P+\Pi$ is invertible and $\|R\| < \infty$ where $R = (I-P+\Pi)^{-1}$. Further, $\|\Pi\| < \infty$. Consequently, this implies that $\|D\| = \|R-\Pi\| < \infty$ (see more details in [21]).

Now we state a result that establishes a sufficient condition in order that a perturbed chain admits an invariant probability measure. The sketch of proof is similar to the one used in [1, Theorem 1], for the discrete state, but with differences from the subtle issues arising from dealing with general state space.

Theorem 4.1. *Let X a Markov chain with regular transition kernel P with a unique stationary distribution π and Q is the transition kernel of the perturbed chain. Assume that the Neumann series $\pi\sum_{k\geq 0}[\Delta(R-\Pi)]^k$ converges to some finite nonnegative measure $\nu = \pi\,(I-\Delta(R-\Pi))^{-1} = \pi\,(I-\Delta R)^{-1}$.*

1. *The limit ν is an invariant probability measure of the kernel Q.*

2. *If we assume, additionally, that the perturbed chain is ergodic, then ν is the unique stationary probability measure of the transition kernel Q.*

Proof. It is well known for the potential R, we have $(I-P)R = I-\Pi$. Multiplying this equation from the left by ν, that is taking integral with respect to the measure ν, we get

$$\nu - \pi = \nu\,(I-P)\,R. \tag{40}$$

Further, we have

$$\nu = \pi \sum_{k=0}^{+\infty} (\Delta R)^k = \pi + \pi \sum_{k=1}^{+\infty} (\Delta R)^k = \pi + \pi \left(\sum_{k=0}^{+\infty} (\Delta R)^k \right) \Delta R = \pi + \nu \Delta R. \tag{41}$$

By subtracting (41) from (40), we obtain : $\nu\,(I-P-\Delta)\,R = \nu\,(I-Q)\,R = 0$. Since R exists, so its inverse also exists (see this result in [21] and the remarks above). Multiplying this latter equation from the right by the regular kernel $I-P+\Pi = R^{-1}$, we obtain $\nu\,(I-Q) = 0$, which yields $\nu Q = \nu$. This proves that ν is an invariant measure of Q. Moreover, by multiplying the relation (40) from the left by the function $\mathbb{I}$ and using the fact that $R\,\mathbb{I} = P\,\mathbb{I} = I\,\mathbb{I} = \mathbb{I}$ and $\pi\,\mathbb{I} = 1$, we obtain: $\nu\,\mathbb{I} - \pi\,\mathbb{I} = \nu\,(I-P)R\,\mathbb{I} = 0 \Leftrightarrow \nu\mathbb{I} = 1$. Hence ν is an invariant probability measure of Q. Consequently, the first assertion is proved.

Since the transition kernel Q is ergodic, then $\lim_{t\to+\infty} Q^t = \widehat{\Pi}$ where $\widehat{\Pi}$ is the stationary projector of Q, that is by definition the stochastic kernel verifying $Q\widehat{\Pi} = \widehat{\Pi}\,Q = \widehat{\Pi}^2 = \widehat{\Pi}$ and $\mu\,\widehat{\Pi} = \mu$ provided that $\mu\,Q = \mu$ for $\mu \in \mathfrak{M}$. In this case, $\widehat{\Pi} = \mathbb{I} \otimes \hat{\pi}$ where $\hat{\pi}$ is the stationary distribution of Q. Since ν is an invariant measure of Q as it is proved above, we have $\nu = \nu Q^t$ for all t, this yields $\nu = \lim_{t\to+\infty} \nu Q^t = \nu\,\widehat{\Pi} = \nu\,(\mathbb{I} \otimes \hat{\pi}) = (\nu\,\mathbb{I})\,\hat{\pi} = \hat{\pi}$. This proves that ν is the unique stationary distribution of the kernel Q. □

The following theorem exposes general upper stability bound under the mixing condition $\mathbf{D_1}(n,T,h,\alpha)$ for $n=1$. This bound is better than those obtained in [24] (see proposition 7.2 in the appendix) and valid in a more large neighborhood of the unperturbed kernel P.

Theorem 4.2. *Assume that the condition $\mathbf{D_1}(1,T,h,\alpha)$ holds.*

1. *If $\|\Delta\| < \dfrac{1-\|T\|}{1+(\sigma-1)\,\|T\|}$, then the perturbed kernel Q has an invariant probability measure ν such that*

$$\|\nu - \pi\| \leq \|\pi\| \frac{1+(\sigma-1)\,\|T\|}{1-\|T\|-(1+(\sigma-1)\,\|T\|)\,\|\Delta\|} \|\Delta\| \tag{42}$$

where $\sigma = \|I-\Pi\|$.

Proof. According to the mixing condition $\mathbf{D_1}(1,T,h,\alpha)$, we get $\|T\| < 1$. From [25], we get the relation $R-\Pi = (I-\Pi)R_T(I-\Pi)$, where $R_T = (I-T)^{-1}$. Hence, by using $\Delta\Pi = \Delta(\mathbb{1}\otimes\pi) = (\Delta\mathbb{1})\otimes\pi = 0$, we derive

$$\Delta R = \Delta(R-\Pi) = \Delta(I-\Pi)R_T(I-\Pi) = \Delta R_T(I-\Pi).$$

Moreover, we have $\Delta R_T(I-\Pi) = \Delta(I+TR_T)(I-\Pi) = \Delta(I+TR_T(I-\Pi))$. Hence, we get:

$$\begin{aligned}\|\Delta R_T(I-\Pi)\| &= \|\Delta(I+TR_T(I-\Pi))\| \leq \|\Delta\|(1+\|T\|\|R_T\|\|I-\Pi\|)\\ &\leq \|\Delta\|\left(1+\frac{\|T\|\|I-\Pi\|}{1-\|T\|}\right) = \|\Delta\|\left(\frac{1-\|T\|+\|T\|\|I-\Pi\|}{1-\|T\|}\right)\\ &\leq \|\Delta\|\left(\frac{1+(\sigma-1)\|T\|}{1-\|T\|}\right). \end{aligned} \tag{43}$$

The condition $\sigma \leq 1+\|\mathbb{1}\|\|\pi\|$ involves $\|\Delta R\| = \|\Delta R_T(I-\Pi)\| < 1$. Consequently the Neumann series $\pi\sum_{k\geq 0}[\Delta(R-\Pi)]^k$ converges and according to theorem 4.1, the kernel Q admits an invariant probability measure ν such that $\nu = \pi\sum_{k\geq 0}[\Delta(R-\Pi)]^k$. Substituting the expression of $\Delta(R-\Pi)$ in the equality (39), we obtain: $\nu-\pi = \pi\sum_{k\geq 1}[\Delta(R-\Pi)]^k = \pi\sum_{k\geq 1}[\Delta R_T(I-\Pi)]^k$. By taking the norm both sides of this last equality, we get:

$$\|\nu-\pi\| \leq \|\pi\|\sum_{k\geq 1}\|\Delta R_T(I-\Pi)\|^k = \frac{\|\pi\|\,\|\Delta R_T(I-\Pi)\|}{1-\|\Delta R_T(I-\Pi)\|}. \tag{44}$$

The bound (42) follows from combining (43) and (44). □

Remark 4.3. Observe that the bound (42) is better than those obtained in [24, Theorem 2] and holds for more large magnitude of perturbation for the unperturbed kernel P (see inequality (67) of proposition 7.2 in the appendix), since $\sigma \leq 1+\|\mathbb{1}\|\|\pi\|$ implies $1+(\sigma-1)\|T\| \leq 1+\|\mathbb{1}\|\|\pi\|\|T\|$. Moreover, if we consider the weighted variation norm $\|\cdot\|_v$, where v is a measurable function on E such that $\hbar = \inf_{x\in E} v(x) = 1$, as assumed in [1], the bound (42) provides the following estimate :

$$\begin{aligned}\|\nu-\pi\| &\leq \|\pi\|\frac{1+(\sigma-1)\|T\|}{1-\|T\|-(1+(\sigma-1)\|T\|)\|\Delta\|}\|\Delta\|\\ &\leq \|\pi\|\frac{1+\|\pi\|\|T\|}{1-\|T\|-(1+\|\pi\|\|T\|)\|\Delta\|}\|\Delta\|. \end{aligned} \tag{45}$$

While the bound [1, Inequality 38] is established for discrete state and for the specific function $v(n) = \beta^n$ where $\beta > 1$ and $n \in S \subseteq \mathbb{N}$, is given as follows :

$$\|\nu - \pi\| \leq \|\pi\| \frac{1 + \|\pi\|}{1 - \|T\| - (1 + \|\pi\|)\|\Delta\|} \|\Delta\|. \tag{46}$$

It is clear that the bound (45) is better than (46) and valid in a more large neighborhood.

The two following theorems derive explicit bounds, with respect to the weighted variation norm, under the drift condition $\mathbf{D_2}(v, C, \lambda, b)$.

Theorem 4.4. *Let us assume that P verifies the drift condition $\mathbf{D_2}(v, C, \lambda, b)$ and $\delta = \lambda + \sup_{x \in C} \left(\frac{b - \alpha v}{v(x)} \right) < 1$ where $\alpha(A) = \inf_{x \in C} P(x, A)$ is a non trivial measure and $v \geq 1$.*

1. *If $\|\Delta\|_v < \frac{1-\rho}{1+\rho\|\pi\|_v}$, then the perturbed transition kernel Q has an invariant probability measure ν such that:*

$$\|\nu - \pi\|_v \leq \frac{\|\pi\|_v \, (1+\rho\|\pi\|_v) \, \|\Delta\|_v}{1-\rho-(1+\rho\|\pi\|_v) \, \|\Delta\|_v} \tag{47}$$

   where $\rho = \max(\lambda, \delta)$.

2. *Any transition kernel Q such that $\|\Delta\|_v < \frac{(1-\rho)(1-\lambda)}{1-\lambda+\rho\, b\, \pi(C)}$ has an invariant probability measure and we have the following inequality*

$$\|\nu - \pi\|_v \leq \frac{b\,\pi(C)\,(1-\lambda+\rho\, b\, \pi(C))\,\|\Delta\|_v}{(1-\rho)(1-\lambda)^2 - (1-\lambda)(1-\lambda+\rho\, b\, \pi(C))\,\|\Delta\|_v}. \tag{48}$$

   In particular, if $\|\Delta\|_v < \frac{(1-\rho)(1-\lambda)}{1-\lambda+\rho\, b}$, then Q has an invariant probability measure ν such that

$$\|\nu - \pi\|_v \leq \frac{b\,(1-\lambda+\rho\, b)\,\|\Delta\|_v}{(1-\rho)(1-\lambda)^2 - (1-\lambda)(1-\lambda+\rho\, b)\,\|\Delta\|_v}. \tag{49}$$

Proof. We consider the residual kernel defined for $x \in E$ and $A \in \mathcal{E}$ by:

$$T(x, A) = P(x, A) - h(x)\alpha(A) = \begin{cases} P(x, A), & x \notin C \\ P(x, A) - \inf_{x \in C} P(x, A), & x \in C, \end{cases}$$

where $h(x) = \mathbb{I}_C(x)$ and $\alpha(A) = \inf_{x\in C} P(x,A)$. Hence, $T(x,A) \geq 0$ for all $x \in E$ and $A \in \mathcal{E}$. Therefore,

$$Tv(x) = \begin{cases} Pv(x), & x \notin C \\ Pv(x) - \alpha v, & x \in C. \end{cases}$$

Utilizing the drift condition, we obtain:

$$Tv(x) \leq \begin{cases} \lambda v(x), & x \notin C \\ \lambda v(x) + b - \alpha v, & x \in C \end{cases} \Longrightarrow \frac{Tv(x)}{v(x)} \leq \begin{cases} \lambda, & x \notin C \\ \lambda + \frac{b-\alpha v}{v(x)}, & x \in C, \end{cases}$$

which implies that $\|T\|_v \leq \rho < 1$. Further, we have $\|P\|_v \leq \lambda + \dfrac{b}{\inf_{x\in C} v(x)} < \infty$. Therefore, the unperturbed chain verifies the condition $\mathbf{D_1}(m,T,h,\alpha)$. Therefore, according to the inequality (38), we derive

$$\pi v = \|\pi\|_v \leq \frac{b\pi(C)}{1-\lambda} \leq \frac{b}{1-\lambda}.$$

Since $\sigma \leq 1 + \|\pi\|_v$, then we have

$$1 + (\sigma - 1)\,\|T\|_v \leq 1 + \|\pi\|_v\,\|T\|_v \leq 1 + \rho\,\|\pi\|_v$$

Therefore, each of the conditions $\|\Delta\| < \dfrac{1-\rho}{1+\rho\,\|\pi\|_v}$, $\|\Delta\|_v < \dfrac{(1-\rho)(1-\lambda)}{1-\lambda+\rho\, b\,\pi(C)}$ or $\|\Delta\|_v < \dfrac{(1-\rho)(1-\lambda)}{1-\lambda+\rho\, b}$ involves that the hypothesis of theorem 4.2 is satisfied, which therefore implies the existence of an invariant probability measure for the perturbed kernel Q. The first claim (47) follows directly from (42). The bounds (48) and (49) derive immediately by replacing $\|\pi\|_v$, in the condition $\|\Delta\| < \dfrac{1-\rho}{1+\rho\|\pi\|_v}$ and in (47), with their upper bounds $\dfrac{b\pi(C)}{1-\lambda}$ and $\dfrac{b}{1-\lambda}$ respectively. □

In the next theorem we consider the test set $C = \{x_0\}$, where $x_0 \in E$ and we denote $\pi_{x_0} = \pi(\{x_0\})$. So we can obtain similar results like (47)-(49) under weaker conditions. The proof is quite similar to those of theorem 4.4, and consequently we omit some details.

Theorem 4.5. *Let x_0 any fixed state in E and assume that the drift condition $\mathbf{D_2}(v,C,\lambda,b)$ is verified where $C = \{x_0\}$ and $v \geq 1$.*

1. *If* $\|\Delta\|_v < \dfrac{1-\lambda}{1+\lambda\,\|\pi\|_v}$, *then any Markov chain with a transition kernel* Q *has a unique stationary distribution* ν *such that*
$$\|\nu-\pi\|_v \le \frac{\|\pi\|_v\,(\hbar+\lambda\,\|\pi\|_v)\,\|\Delta\|_v}{1-\lambda-(1+\lambda\,\|\pi\|_v)\,\|\Delta\|_v}. \tag{50}$$

2. *Any transition kernel* Q *such that* $\|\Delta\|_v < \dfrac{(1-\lambda)^2}{1-\lambda+\lambda\,b\,\pi_{x_0}}$ *has a unique stationary distribution* ν *such that*
$$\|\nu-\pi\|_v \le \frac{b\,\pi_{x_0}\,(1-\lambda+\lambda\,b\,\pi_{x_0})\,\|\Delta\|_v}{(1-\lambda)^3-(1-\lambda)\,(1-\lambda+\lambda\,b\,\pi_{x_0})\,\|\Delta\|_v}. \tag{51}$$
In particular, if $\|\Delta\|_v < \dfrac{(1-\lambda)^2}{1+\lambda\,(b-1)}$, *then* Q *has a unique stationary distribution* ν *with*
$$\|\nu-\pi\|_v \le \frac{b\,(1+\lambda\,(b-1))\,\|\Delta\|_v}{(1-\lambda)^3-(1-\lambda)\,(1+\lambda\,(b-1))\,\|\Delta\|_v}. \tag{52}$$

Proof. The existence of an invariant probability measure for the perturbed transition kernel Q under the condition $\|\Delta\|_v < \dfrac{\hbar\,(1-\lambda)}{\hbar+\lambda\,\|\pi\|_v}$, $\|\Delta\|_v < \dfrac{\hbar\,(1-\lambda)^2}{\hbar+\lambda\,(b-\hbar)}$ or $\|\Delta\|_v < \dfrac{\hbar\,(1-\lambda)^2}{\hbar(1-\lambda)+\lambda\,b\,\pi_{x_0}}$ follows from theorem 4.2. Let consider the residual kernel defined for $x\in E$ and $A\in\mathcal{E}$ by:
$$T(x,A)=P(x,A)-h(x)\alpha(A)=\begin{cases} P(x,A), & x\neq x_0\\ 0, & x=x_0,\end{cases}$$
where $h(x)=\mathbb{I}_C(x)$ and $\alpha(A)=P(x_0,A)$. Hence, $T(x,A)\ge 0$ for all $x\in E$ and $A\in\mathcal{E}$. Therefore, $Tv(x)=\begin{cases} Pv(x), & x\neq x_0\\ 0, & x=x_0.\end{cases}$

Using the drift condition, we derive
$$Tv(x)\le\begin{cases}\lambda v(x), & x\neq x_0\\ 0, & x=x_0\end{cases} \Longrightarrow \frac{Tv(x)}{v(x)}\le\begin{cases}\lambda, & x\neq x_0\\ 0, & x=x_0,\end{cases}$$
which yields $\|T\|_v\le\lambda$. Moreover, it is easy to get $\|P\|_v\le\lambda+\dfrac{b}{v(x_0)}<\infty$. Hence, the unperturbed chain verifies the condition $\mathbf{D_1}(1,T,h,\alpha)$. Taking into account that $\pi(C)=\pi_{x_0}$ and the estimate
$$\pi v=\|\pi\|_v\le\frac{b\pi_{x_0}}{1-\lambda}\le\frac{b}{1-\lambda}, \tag{53}$$

we derive simply (50)-(52) by following similar sketch of proof than those of theorem 4.4. □

Remark 4.6. The bounds (50) and (52) extend to general state space and improve [32, Bounds (3.3)-(3.4)] (see the bounds (68)-(69) of proposition 7.3 in the appendix). Moreover, they are valid for more large magnitude of perturbation. Note that the bounds (50) and (52) improve (68)-(69) in both numerator and denominator. Let us point out that the estimate of $\|\pi\|_v = \pi v$ in theorem 4.5 may be improved from (53) and by taking into account that $\pi h = \pi_{x_0}$. We obtain the following estimate:

$$\pi v = \|\pi\|_v \leq \frac{\min(b, \alpha v)\, \pi_{x_0}}{1-\lambda}\cdot$$

The stability bounds obtained in theorem 4.5 and 4.4 are given in terms of the parameters λ and b of the drift condition $\mathbf{D_2}(v, C, \lambda, b)$. Since the geometric ergodicity is often established for many complex processes, which are not necessarily random walks, by using the drift condition $\mathbf{D_2}(v, C, \lambda, b)$ (see e.g. [37] and the references therein), it follows that the obtained perturbation bounds may be applied for more complex processes other than random walks. We point out that recently in [42], it is proved that the waiting process in queuing systems with impatient units (a process which is not random walk in the classical meaning) is strongly stable and some stability estimates were obtained.

The next theorem gives the upper bound perturbation for the stationary distribution in term of the norm ergodicity coefficient.

Theorem 4.7. *Let assume $\Lambda_m < 1$ for $m \geq 1$. Then, for all transition kernel Q such that $\|\Delta_m\| < 1 - \Lambda_m$ has an invariant probability measure ν. Further, the following estimate*

$$\|\nu - \pi\| \leq \frac{\mathbf{q}_{(m)}\, \|\pi\|}{1-\Lambda_m}\, \|\Delta_m\| \tag{54}$$

is fulfilled provided that $\mathbf{q}_{(m)} = \sup_{t\geq 0} \|Q^{tm}\| < \infty$.

Proof. For the skeleton X^m of the chain X, its potential R_m is given by the relation

$$R_m = (I - P^m + \Pi)^{-1} = \sum_{k\geq 0} (P^m - \Pi)^k.$$

It follows that $R_m - \Pi = \sum_{k\geq 0}\left(P^{km} - \Pi\right)$. By using lemma 3.1, we get

$$\begin{aligned}\|\Delta_m (R_m - \Pi)\| &\leq \|\Delta_m\|\, \Lambda(R_m - \Pi) = \|\Delta_m\|\, \Lambda(R_m) \\ &\leq \|\Delta_m\| \sum_{k\geq 0} \Lambda\left(P^{km}\right) \leq \|\Delta_m\| \sum_{k\geq 0} \Lambda_m^k = \frac{\|\Delta_m\|}{1-\Lambda_m}\cdot\end{aligned}$$

Therefore, the condition $\|\Delta_m\| < 1-\Lambda_m$ involves $\|\Delta_m(R_m-\Pi)\| < 1$, and consequently the series $\pi\sum_{k\geq 0}[\Delta_m(R_m-\Pi)]^k$ converges. Hence, from theorem 4.1 the transition kernel Q^m has an invariant probability measure ν such that $\nu = \pi\sum_{k\geq 0}[\Delta_m(R_m-\Pi)]^k$. Furthermore, for $t \geq 0$, we have

$$\pi^{(tm)} - \nu^{(tm)} = \pi^{(0)}(P^{tm} - Q^{tm}) + \left(\pi^{(0)} - \nu^{(0)}\right) Q^{tm}.$$

According to lemma 3.1, since $\Delta_m \mathbb{I} = 0$, and from

$$\Lambda(Q^m) = \Lambda(\Delta_m + P^m) \leq \Lambda(\Delta_m) + \Lambda(P^m) \leq \|\Delta_m\| + \Lambda(P^m) < 1,$$

we obtain

$$\begin{aligned}
\|\Delta^{(tm)}\| &\leq \|\pi^{(0)}\|\,\|\Delta_{tm}\| + \|\pi^{(0)} - \nu^{(0)}\|\,(\Lambda(Q^m))^t \\
&\leq \|\pi^{(0)}\|\,\|\Delta_{tm}\| + \|\pi^{(0)} - \nu^{(0)}\|\,(\|\Delta_m\| + \Lambda_m)^t.
\end{aligned}$$

Utilizing (14), we get

$$\begin{aligned}
\|\Delta^{(tm)}\| &\leq \mathbf{q}_{(t-1)m}\,\|\pi^{(0)}\|\,\|\Delta_m\|\,\frac{1-\Lambda_m^t}{1-\Lambda_m} + \|\pi^{(0)} - \nu^{(0)}\|\,(\|\Delta_m\| + \Lambda_m)^t \\
&\leq \|\Delta_m\|\,\frac{\mathbf{q}_{(m)}\,\|\pi^{(0)}\|}{1-\Lambda_m}\| + \|\pi^{(0)} - \nu^{(0)}\|\,(\|\Delta_m\| + \Lambda_m)^t.
\end{aligned}$$

Finally, substituting $\pi^{(0)} = \pi$ and $\nu^{(0)} = \nu$ in the latter inequality, and taking the limit as $t \longrightarrow +\infty$, we obtain the bound (54). □

Remark 4.8. Observe that if we substitute $\pi^{(0)} = \pi$ and $\nu^{(0)} = \nu$ in the inequality (18), and taking the limit as $t \longrightarrow 0$, we derive straightly the estimate of the the deviation $\|\nu - \pi\|$ as follows

$$\|\nu - \pi\| \leq \frac{\mathbf{q}\,\|\nu\|}{1-\Lambda_m}\,\|\Delta_m\|.$$

Unfortunately, this upper perturbation bound depends on the unknown distribution ν which we want to estimate. This is why we have considered an other way to obtain an upper bound of $\|\nu - \pi\|$ in theorem 4.7. For the total variation norm, the bound (54) becomes

$$\|\nu - \pi\| \leq \frac{\|\Delta_m\|}{1-\tau_m} \tag{55}$$

which coincides with [38, Theorem 3.2, Inequality (3.11)]. Furthermore, we don't need the condition $\|\Delta_m\| < 1-\Lambda_m$ but in return we must assume that the perturbed kernel has a unique stationary distribution and further the proof must be modified. Indeed for the total variation norm, we have $\mathbf{q}_{t-m} = \|\nu^{(0)}\| = 1$. Therefore, by substituting $\pi^{(0)} = \pi$ and $\nu^{(0)} = \nu$ in the inequality (18), and taking the limit as $t \longrightarrow +\infty$, we obtain (55), without assuming $\|\Delta_m\| < 1-\Lambda_m$.

5. Examples

In this section, we will study two examples with continuous space and one for the denumerable case, with respect to the total and weighted variation norms. In order to show the quality of the estimates established in this paper, comparison with other existing results in the literature is carried out.

Example 5.1. We borrow this random walk example from [25, Example 3.1] and [22, Example 3]. Let $(\xi_t, t \geq 1)$ a sequence of independent identically distributed random variables with values in $E = [0,1)$ and with a common distribution F, defined for all $B \in \mathcal{B} = \mathcal{B}[0,1)$ by $F(B) = \mathbb{P}(\xi_1 \in B)$. Moreover, we assume that F has an absolutely continuous component and its density is not less than θ on the measurable set $C \in \mathcal{B}[0,1)$. On $(E,\mathcal{B})$, we consider the Markov chain $X = (X_t, t \geq 0)$ defined by $X_{t+1} = X_t + \xi_{t+1} (\mathrm{mod}\, 1)$. It has transition kernel $P(x,B) = F(B-x)$ and with the stationary projector $\Pi(x,B) = L(B)$ for all $(x,B) \in [0,1] \times \mathcal{B}[0,1)$, where $L = \mu^{\mathrm{leb}}$ is the Lebesgue measure on $E = [0,1)$ and $B - x = \{(b-x)(\mathrm{mod}\, 1), b \in B\}$, i.e the Markov chain X has a unique invariant measure $\pi = L$. Suppose that $c = L(C) > \dfrac{1}{2}$ and consider the space $\mathcal{M}$ of finite measures with the total variation norm $\|\mu\| = |\mu|(E)$ (uniform norm). Therefore, from [25, Example 3.1] it is shown that $\tau_1 \leq \rho = 1 - \theta(2c-1) < 1$. Let us set $\varepsilon = \|\Delta\|_1$. Hence, from (18) we derive for all transition kernel Q on $([0,1),\mathcal{B}[0,1))$ and $t \geq 1$, the following uniform estimate with respect to $x \in E$, $B \in \mathcal{B}[0,1)$,

$$|\mathbb{P}_x(X_t \in A) - \mathbb{P}_x(Y_t \in A)| \leq \frac{1-\rho^t}{\theta(2c-1)}\, \varepsilon. \tag{56}$$

Further, we get

$$\sup_{t \geq 0} \sup_{x \in E, A \in \mathcal{B}} |\mathbb{P}_x(X_t \in A) - \mathbb{P}_x(Y_t \in A)| \leq \frac{\varepsilon}{\theta(2c-1)}. \tag{57}$$

It follows from (55) that for all transition kernel Q admitting a unique invariant measure ν, we have for all A the following inequality

$$\sup_{A \in \mathcal{B}} |\pi(A) - \nu(A)| \leq \frac{\varepsilon}{\theta(2c-1)}. \tag{58}$$

Observe that the estimates [25, Example 3.1; Inequality (3.22)] and [22, Example 3; Inequality (20)] are given respectively as follows

$$|\mathbb{P}_x(X_t \in A) - \mathbb{P}_x(Y_t \in A)| \leq \frac{\varepsilon}{\theta(2c-1) - 2\varepsilon} \tag{59}$$

and

$$\sup_{A \in \mathcal{B}} |\pi(A) - \nu(A)| \leq \frac{\varepsilon}{\theta(2c-1) - 2\varepsilon}, \tag{60}$$

for all transition kernel Q belonging to a neighborhood

$$\varepsilon = \sup_{x\in E, A\in\mathcal{B}} |Q(x,x+A) - F(A)| < \theta(c - \frac{1}{2}). \tag{61}$$

Consequently, it is obviously seen that the bounds (56) and (58) are strictly sharper than (59)-(60) respectively. Moreover, the inequalities (56)-(57) are valid for every transition kernel Q and (58) is also valid for all transition kernel provided it has an invariant measure ν, while (59)-(58) hold only for all transition kernel Q verifying the inequality (61). Observe that in this case and from [25, Theorem 3.5], the perturbed chain has a unique invariant measure and it is aperiodic and uniformly ergodic.

Example 5.2 (Random walk on half line)**.** Let us consider the Markov chain $X = \{X_t; t \in \mathbb{Z}_+\}$ defined by the recursive equation

$$X_{t+1} = (X_t + \xi_{t+1})^+ = \max(0, X_t + \xi_{t+1})$$

for $t \geq 0$, and taking values in $\mathbb{R}_+$. Here $(\xi_t)_{t\geq 0}$ is a sequence of independent random variables taking values in $\mathbb{R}$ and identically distributed with a common distribution function $\mathbb{F}$. The perturbation and stability inequalities have been considered recently in [39, 41] and the strong stability of this process have been investigated more earlier in [25]. Assume that $\mathbb{E}[\xi_1] < \infty$. Consequently, from [33], the Markov chain is ergodic if and only if $\mathbb{E}[\xi_1] < 0$, i.e., X has a unique probability distribution if and only if $\mathbb{E}[\xi_1] < 0$ with $\mathbb{E}[\xi_1] < \infty$. Let us consider the weighted function $v(x) = e^{\gamma x}$ where γ is reel parameter such that $\gamma > 1$. The v-norm on $\mathfrak{M}$ have the form $\|\mu\|_\gamma = \|\mu\|_v = \int_0^{+\infty} e^{\gamma x}|\mu|(dx)$. The norms in the space $\mathfrak{N}$ and $\mathfrak{B}$ are defined, respectively, by $\|f\|_\gamma = \sup_{x\geq 0} e^{-\gamma x}|f(x)|$ and $\|K\|_\gamma = \sup_{x\geq 0} e^{-\gamma x} \int_0^{+\infty} |Q|(x,dy)e^{\gamma y}$. Recall that, for all $(x,dy) \in (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ and $t \geq 0$, we have $P^t(x,dy) = \mathbb{P}(X_t \in dy | X_0 = x)$. Observe that the transition kernel can be decomposed as follows

$$P(x,A) = \mathbb{P}(0 < x + \xi_1 \in A) + \mathbb{P}(x + \xi_1 \leq 0) \cdot \delta_0(A) \tag{62}$$

where δ_0 is the degenerate measure concentrated on $\{0\}$. Let denote the substochastic kernel $T(x,A) = \mathbb{P}(0 < x + \xi_1 \in A)$ for $(x,A) \in (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$, the measurable function $h(x) = \mathbb{P}(\xi_1 + x \leq 0)$ for $x \in \mathbb{R}_+$ and the measure α on $\mathcal{B}(\mathbb{R}_+)$ defined for $A \in \mathcal{B}_{\mathbb{R}_+}$ by $\alpha(dy) = \delta_0(dy)$. It follows from (62), the following equation $P(x,A) = T(x,A) + h(x) \cdot \alpha(A)$. Notice that $\alpha \mathbb{1} = 1$ and $\alpha h = h(0) = \mathbb{P}(\xi_1 \leq 0) > 0$, $\pi h = \pi(\{0\}) > 0$. Further, we get

$$\begin{aligned} Tv(x) &= \int_{\mathbb{R}^+} P(0 < x + \xi_1 \in dy)\, e^{\gamma y} = \mathbb{E}\left[e^{\gamma(x+\xi_1)}, x + \xi_1 > 0\right] \\ &\leq \mathbb{E}\left[e^{\gamma(x+\xi_1)}\right] = \rho(\gamma)\, v(x) \end{aligned}$$

where $\rho = \rho(\gamma) = \mathbb{E}\left[e^{\gamma\xi_1}\right]$. Since $\mathbb{E}[\xi_1] < 0$ and $\rho(0) = 1$, the convexity of $\rho(\gamma)$ involves the existence of $\gamma_0 > 0$ such that for all $\gamma \in]0, \gamma_0[$, we have $\rho(\gamma) < 1$. This proves that $\|T\|_\gamma \le \rho(\gamma) = \mathbb{E}\left[e^{\gamma\xi_1}\right] < 1$ and

$$\|P\|_\gamma = \|T + h\otimes\alpha\|_\gamma \le \|T\|_\gamma + \|h\|_\gamma\|\alpha\|_\gamma < \rho(\gamma) + 1 < 2 < \infty$$

for all $\gamma \in]0, \gamma_0[$. Hence the chain verifies the condition $\mathbf{D_1}(1, T, h, \alpha)$. Since $\|\alpha\|_\gamma = 1$, we derive $\|\pi\|_\gamma \le \dfrac{(\pi h)\ (\alpha v)}{1-\|T\|_\gamma} = \dfrac{\pi(\{0\})}{1-\|T\|_\gamma}$.

Assume that a control sequence $(\xi_i)_{i\ge 0}$ is perturbed to yield another sequence $(\zeta_n)_{n\ge 0}$ of independent random variables taking values in $\mathbb{R}$ and identically distributed with a common distribution function $\mathbb{G} \ne \mathbb{F}$. This involves that the Markov process Y which taking values in $\mathbb{R}_+$ and generated by the recursive equation $Y_{t+1} = (Y_t + \zeta_{t+1})^+$, for $t \ge 0$, is the corresponding perturbed chain. Let denote Q, ν the transition kernel and the stationary distribution of the Markov chain Y, respectively. From the definition of the v-norm, we have

$$\varepsilon = \|Q - P\|_\gamma = \sup_{x\ge 0} e^{-\gamma x}\int_0^{+\infty} |Q(x,dy) - P(x,dy)|\, e^{\gamma y}.$$

Let us extend the measures $\mu \in \mathcal{B}(\mathbb{R}_+)$ to $\mathcal{B}(\mathbb{R})$ such that $\mu(A) = 0$ for all $A \in \mathcal{B}(\mathbb{R}_-)$. Since $\|T\|_\gamma \le \rho$, it follows that $\|\pi\|_\gamma \le \dfrac{\pi(\{0\})}{1-\rho}$.

Therefore, from (42), for $\varepsilon < \dfrac{(1-\rho)^2}{(1-\rho\,(1-\pi(\{0\})))}$, we have

$$\|\nu - \pi\|_\gamma \le \frac{\pi(\{0\})}{1-\rho}\,\frac{(1-\rho\,(1-\pi(\{0\})))\,\varepsilon}{(1-\rho)^2 - (1-\rho\,(1-\pi(\{0\})))\,\varepsilon} = C_1.$$

According to [24], it is proved that for all γ such that $\rho = \rho(\gamma) = \mathbb{E}[e^{\gamma\xi_1}] < 1$ and $0 \le \varepsilon(\mathbb{F},\mathbb{G}) \le \frac{(1-\rho)^2}{2}$, we have the following strong stability bound (see definition 7.1 in the appendix)

$$\|\pi - \nu\|_\gamma \le \frac{2\varepsilon(\mathbb{F},\mathbb{G})}{1-\rho}\,\frac{1}{(1-\rho)^2 - 2\varepsilon(\mathbb{F},\mathbb{G})} = C_2$$

with $\varepsilon(\mathbb{F},\mathbb{G}) = \sup_{x\ge 0}\left(e^{-\gamma x}|\mathbb{F}(-x) + \mathbb{G}(-x) + \int_{-x}^{+\infty} e^{\gamma y}|\mathbb{F}-\mathbb{G}|(dy)\right)$.

It is obvious that $\|Q - P\|_\gamma = \varepsilon \le \varepsilon(\mathbb{F},\mathbb{G})$. Consequently, it is clear that $C_1 < C_2$. In fact, the improvement is made both in the numerator and in the denominator. Further, the upper bound C_1 for the deviation of stationary distributions is not only better but is also valid on an optimal stability domain

$$\varepsilon < \frac{(1-\rho)^2}{(1-\rho\,(1-\pi(\{0\})))} > (1-\rho)^2 > \frac{(1-\rho)^2}{2}.$$

Example 5.3. (denumerable state and bounds with respect to weighted norm) Let the Bernoulli random walk chain X on $\mathbb{Z}_+$ considered in [40, Example 1] and defined by its transition matrix $P = (p_{i,j})_{i,j\in\mathbb{Z}_+}$ where $p_{0,0} = 1$, $p_{i,i-1} = q > p$, $p_{i,i+1} = p = 1-q$ for $i \geq 1$ and $p_{i,j} = 0$ else. It is well known that this process is note uniformly ergodic with respect to the total variation norm (see eg. [40]). Let us introduce the test function $v(n) = \beta^n$ for $n \in \mathbb{Z}_+$ and we consider the measure weighted v-norm $\|\mu\|_v = \|\mu\|_\beta$:

$$\|\mu\|_v = |\mu|v = \sum_{k\geq 0} v(k)|\mu|(\{k\}) = \sum_{k\geq 0} \beta^k |\mu_k|$$

where $\mu = (\mu_k)_{k\in\mathbb{Z}_+}$. This induces a norm on the vector function and matrix spaces defined respectively as follows: $\|f\|_\beta = \sup_{k\geq 0} \beta^{-k}|f(k)|$ and

$$\|K\|_\beta = \sup_{k\geq 0} \beta^{-k}|K|v(k) = \sup_{k\geq 0} \beta^{-k} \sum_{j\geq 0} v(j)|K(k,\{j\})| = \sup_{k\geq 0} \beta^{-k} \sum_{j\geq 0} \beta^j |K_{kj}|.$$

For $1 < \beta < \dfrac{q}{p}$, we get easily that $q\beta^{-1} + p\beta < 1$ and it follows

$$\begin{aligned}\|P\|_\beta &= \sup_{k\geq 0} \beta^{-k} \sum_{j\geq 0} \beta^j p_{kj} = \max\left(1, \sup_{k\geq 1} \beta^{-k} P v(k)\right) \\ &= \max\left(1, q\beta^{-1} + p\beta\right) = 1.\end{aligned}$$

Observe that the spectrum $\sigma_\beta(P)$ of P is $\sigma_\beta(P) = \{pz + qz^{-1}, |z| \geq \beta\} \cup \{1\}$ for each $1 < \beta < \dfrac{q}{p}$ (the necessary and sufficient condition for the uniform ergodicity of the chain) which is easy to derive (see for example [23]). At the same time, if we consider the boundary of the spectrum, i.e., the points $\theta_\beta = p\beta + q\beta$, then we can easily prove that the smallest value of θ_β is $\theta = 2\sqrt{pq}$ and reached for $\beta = \sqrt{\dfrac{q}{p}}$. That means that $\theta = \theta_\beta$. In the sequel of this example we consider $\beta = \sqrt{\dfrac{q}{p}}$, the measure $\alpha = (p_{0,j})$ for all $j \in \mathbb{Z}_+$ and the measurable indicator function $h = \mathbb{1}_{\{i=0\}}$. Therefore, it is obvious that the residual kernel $T = (T_{ij})$ defined for all $(i,j) \in \mathbb{Z}_+^2$ by: $T(i,\{j\}) = T_{ij} = p_{ij} - h_i\alpha_j = \begin{cases} 0, & i = 0; \\ p_{ij}, & i \geq 1 \end{cases}$

is non negative. Let us compute the norm $\|T\|_v$. For this, it is equivalent from (26) to estimate $Tv(k)$ for all $k \in \mathbb{Z}_+$.

1. For $k = 0$, we have $Tv(0) = 0$.
2. For $k \geq 1$, we have $Tv(k) = \beta^k(q\beta^{-1} + p\beta) = 2\sqrt{pq}v(k)$.

This yields $\|T\|_\beta \leq \theta < 1$, where $\theta = 2\sqrt{pq}$. Hence, the condition $\mathbf{D_1}(1,T,h,\alpha)$ is satisfied. Moreover, $\pi = \delta_0$ is the unique stationary distribution of the kernel P and $\|\pi\|_\beta = \|\Pi\|_\beta = \|\mathbb{1}\|_\beta = 1$. It is worth noting that the condition $\mathbf{D_2}(v,C,\lambda,b)$ is also verified for the set $C = \{0\}$, $\lambda = \theta$ and $b = 1$. Indeed, according to the decomposition $P = T + h \otimes \alpha$, we obtain

$$Pv = Tv + (\alpha v)\, \mathbb{1}_{\{i=0\}} = Tv + \mathbb{1}_{\{i=0\}} \leq \theta v + \mathbb{1}_{\{i=0\}}.$$

From (42), (50) or (52), for $\|\Delta\|_\beta < \dfrac{1-\theta}{1+\theta}$, we have the same following stability bound

$$\|\nu - \pi\|_\beta \leq \frac{1+\theta}{1-\theta-(1+\theta)\,\|\Delta\|_\beta}\|\Delta\|_\beta. \tag{63}$$

While the bound expressed in [32, Inequality (3.3)] and [1, Inequality (38)], for $\|\Delta\|_\beta < \dfrac{1-\theta}{2}$, yields the following estimate

$$\|\nu - \pi\|_\beta \leq \frac{2}{1-\theta-2\|\Delta\|_\beta}\|\Delta\|_\beta. \tag{64}$$

According to (51), we have for $\|\Delta\|_\beta < (1-\theta)^2$ the following inequality

$$\|\nu - \pi\|_\beta \leq \frac{\|\Delta\|_\beta}{(1-\theta)^3 - (1-\theta)\,\|\Delta\|_\beta}. \tag{65}$$

On the other hand, and from [32, Inequalitiy (3.4)], we get for $\|\Delta\|_\beta < \dfrac{(1-\theta)^2}{2-\theta}$ the following estimate

$$\|\nu - \pi\|_\beta \leq \frac{2-\theta}{(1-\theta)^3 - (1-\theta)\,(2-\theta)\,\|\Delta\|_\beta}\|\Delta\|_\beta. \tag{66}$$

So it is clearly that the bounds (63) and (65) are better and hold under more intense perturbation (more large neighborhood of the unperturbed kernel P) than (64) and (66). An alternative bound can be expressed in term of the norm coefficient. Indeed, we have established in [40, Example 1] that $\Lambda_1 \leq \dfrac{2\theta}{1+\theta}$.

According to (28), for $n = 1$, and since $\|\alpha\|_\beta = \|h\|_\beta = 1$, we get for all transition kernel in the neighborhood $\{Q \in \mathfrak{B} : \|Q-P\|_\beta = \|\Delta\|_\beta < 1-\theta\}$, the following inequality $\mathbf{q} \leq \dfrac{1}{\left(1-\theta-\|\Delta\|_\beta\right)^2}$.

Therefore, according to (54), we derive

$$\|\nu - \pi\|_\beta \leq \frac{(1+\theta)\,\|\Delta\|_\beta}{(1-\theta)\left(1-\theta-\|\Delta\|_\beta\right)^2}.$$

This latter bound is valid under a more large magnitude of perturbation, but less accurate than (63)-(66).

6. Concluding remarks

Few results are considered in the last two decades for the perturbation bounds for general state Markov chains. Moreover, most results are established with respect to the total variation norm. In this paper, we have established new perturbation bounds for the transition and stationary characteristics of general state Markov chains, with respect to a wide class of norms, in terms of the generalized norm ergodicity coefficient or the residual kernel given in the condition $\mathbf{D_1}(n,T,h,\alpha)$. In this case, the unperturbed chain is strongly stable with respect to the given norm. Namely, for a small parameter disturbance, then the perturbed chain inherits some suitable characteristics of the unperturbed chain (see [22, 25, 39, 41]). Explicit bounds are established under the drift condition $\mathbf{D_2}(v,C,\lambda,b)$. We have shown by a theoretical comparison and on the basis of examples the quality of the inequalities obtained in this paper. More precisely, some estimates improve some specifically bounds in [1, 3, 32], hold true for general state and various norms, and applicable for a more large magnitude of the perturbation.

For scale models, that is scale perturbation (linear perturbation), the perturbed kernel $Q(\theta)$ is given by the convex combination of the two transition kernels P and R defined explicitly by

$$Q(\theta) = (1-\theta)P + \theta R, \theta \in [0,1]$$

where $Q(0) = P$ and $Q(1) = R$. Here θ is the scale parameter. Let denote $\Delta(\theta) = Q(\theta) - P = \theta \|R-P\|$. It follows that all results obtained in this paper remain valid where we substitute $\|Q-P\|$ by $\theta \|R-P\|$, which allows us to scale the size of the perturbation via the scalar control parameter θ.

It is worth noting that we can improve the stability bounds obtained in [32, Inequalities (4.10)-(4.11)] for continuous-time Markov chains (CTMCs) with the discrete state by following a similar process of proof as that of [32] and based on the improved bounds (50)-(52). Similarly, we can establish bounds similar to (48)-(49) for (CTMCs) under the drift condition. But that is outside the scope of this article and will be the subject of another paper.

7. Appendix

We set out some results for which comparisons have been made. First we clarify the definition of the strong stability of Markov chains. We set out some results for which comparisons have been made in this paper. First, we clarify the definition of the strong stability of Markov chains.

Definition 7.1. Let X a Markov chain with transition kernel P and unique invariant probability measure π. The chain X is called strongly stable with respect to the norm $\|\cdot\|$ if the following two conditions are fullfilled.

1. $\|P\| < \infty$.
2. Any transition kernel $Q \in \mathfrak{B}$ in some neighborhood $\{\|Q-P\| < \varepsilon\}$ admits a unique invariant probability measure ν and $\|\nu - \pi\| \longrightarrow 0$ when $\|Q-P\| \longrightarrow 0$ uniformly in this neighborhood.

The next proposition is taken straightforwardly from [24, Theorem 1 and 2] and [24, Corollary 2].

Proposition 7.2. *Let v a finite measurable function on E bounded away from zero. The Markov chain X taking values in a measurable space $(E,\mathcal{E})$ with transition kernel P and the unique invariant probability measure π is v-strongly stable (strongly stable with respect to the v-norm $\|\cdot\|_v$) if and only if the following condition holds:* $\mathbf{D_1}(n,T,h,\alpha)$

I) $\|P\|_v < \infty$.

II) There exist a natural integer n, measurable function $h \in \mathfrak{N}^+$ and measure $\alpha \in \mathfrak{M}^+$ such that: $\alpha h > 0$, $\pi h > 0$, $\alpha \mathbf{1} = 1$ and the residual kernel $T = P^n - h \otimes \alpha$ is nonnegative.

III) There exists a real number $\rho \in [0,1[$ such that $Tv(x) \le \rho\, v(x)$ for all $x \in E$.

Furthermore, if Q is a transition kernel of a perturbed Markov chain Y such that $\|\Delta\|_v = \|Q-P\|_v < \dfrac{1-\rho}{1+\|\mathbf{1}\|_v \|\pi\|_v}$, then the perturbed Markov chain Y have a unique probability measure ν and we have:

$$\|\nu - \pi\|_v \le \frac{\|\pi\|_v\,(1+\|\mathbf{1}\|_v\,\|\pi\|_v)\,\|\Delta\|_v}{1-\rho-(1+\|\mathbf{1}\|_v\,\|\pi\|_v)\,\|\Delta\|_v}. \tag{67}$$

The following proposition resumes [32, Corollary 3.1].

Proposition 7.3. *Let i_0 be any fixed state in E. Suppose that the unperturbed transition kernel P satisfies $\mathbf{D_2}(v,C,\lambda,b)$ for $C=\{i_0\}$.*

(i) Let $c = 1+\|\mathbf{1}\|_v\,\|\pi\|_v$. If $\|\Delta\|_v < \dfrac{1-\lambda}{c}$, then X is positive recurrent and

$$\|\nu - \pi\|_v \le \frac{c\,\|\pi\|_v\,\|\Delta\|_v}{1-\lambda-c\,\|\Delta\|_v}. \tag{68}$$

(ii) If $v \geq 1$ *and* $\|\Delta\|_v < \dfrac{(1-\lambda)^2}{b+1-\lambda}$*, then X is positive recurrent and*

$$\|\nu - \pi\|_v \leq \frac{b\,(b+1-\lambda)\,\|\Delta\|_v}{(1-\lambda)^3 - (1-\lambda)\,(b+1-\lambda)\,\|\Delta\|_v}. \tag{69}$$

REFERENCES

[1] K. Abbas - B. Heidergott - B. Joost, *A Critical Account of Perturbation Analysis of Markov Chains*, Markov Process. relat. Fields 22 (2006), 227–266.

[2] E. Altman - C. E Avrachenkov - R. Nunez-Queija, it Perturbation analysis for denumerable Markov chains with applications to queueing models, Adv. Appl. Prob. 36 (2004), 839–853.

[3] V. V. Anisimov, it Estimates for the deviation of the transition characteristics of nonhomogeneous Markov processes, Ukranian Math. J. 40 (1988), 588–592.

[4] H. Caswell, *Sensitivity analysis of discrete Markov chains via matrix calculus*, Linear Algebra Appl. 438 (2013), 1727–1745.

[5] G. E. Cho - C. D. Meyer, *Markov chain sensitivity measured by mean first passage time*, Linear Algebra Appl. 316 (2000), 21–28.

[6] G. E. Cho - C. D. Meyer, *Comparison of perturbation bounds for the stationary probabilities of a finite Markov chain*, Linear Algebra Appl. 335 (2001), 137–150.

[7] M. P. Drazin, *Pseudo-inverses in associative rings and semigroups*, Amer. Math. Monthly 65 (7) (1958), 506–514.

[8] R. E. Meyer- C. D. Meyer, *Sensitivity of the stationary vector for an ergodic Markov chains*, Linear Algebra Appl. 76 (1986), 1–17.

[9] G. H. Golub - C. D. Meyer, *Using the QR factorization and group inversion to compute, differentiate, and estimate the sensitivity of stationary probabilities for Markov chains*, SIAM J. Algebraic Discrete Methods 7 (2) (1986), 273–281.

[10] D. J. Hartfiel, *How a row change in a stochastic matrix affects the corresponding stochastic eigenvalues*, Linear and Multilinear Algebra 10 (1981), 205–216.

[11] M. Haviv - L. Van Der Heiden, *Perturbation bounds for the stationary probabilities of a finite Markov chains*, Adv. Appl. Prob. 16 (1984), 804–818.

[12] B. Heidergott - A. Hordijk, *Taylor expansions for stationary Markov chains*, Adv. Appl. Prob. 35 (2003), 1046–1070.

[13] B. Heidergott - A. Hordijk - M. van Uitert, *Series expansions for finite-state Markov chains*, Probability in Engineering and Informational Sciences 21 (2007), 381–400.

[14] J. J. Hunter, *Generalized inverses and their application to applied probability problems*, Linear Algebra Appl. 45 (1982), 157–198.

[15] J. J. Hunter, *Stationary distributions and mean first passage times of perturbed Markov chains*, Linear Algebra Appl. 410 (2005), 217–243.

[16] J. J. Hunter, *Mixing times with applications to perturbed Markov chains*, Linear Algebra Appl. 417 (2006), 108–123.

[17] C. Ipsen - C. D. Meyer, *Uniform stability of Markov Chains*, SIAM J. Matrix Anal. Appl. 15 (4) (1994), 1061–1074.

[18] I. Ipsen - T. Selee, *Ergodicity coefficients defined by vector norms*, SIAM J. Matrix Anal. Appl. 32 (2011), 153–200.

[19] D. Isaacson - R. Madsen, *Positive columns for stochastic matrices*, J. Appl. Prob. 11 (1974), 829–835.

[20] N. V. Kartashov, *On ergodicity and stability properties of Markov chains with respect to various norms in the measure space*, Theor. Prob. Appl. 1 (1981), 18–39.

[21] N. V. Kartashov, *Criteria for uniform ergodicity and strong stability of Markov chains with a common phase space*, Theor. Prob. Math. Stat. 30 (1985), 71–89.

[22] N. V. Kartashov, *Inequalities in Theorems of Ergodicity and Stability For Markov Chains With Common Phase space I*, Theor. Prob. Appl. 30 (2) (1985), 247–259.

[23] N. V. Kartashov, *Inequalities in Theorems of Ergodicity and Stability For Markov Chains With Common Phase space II*, Theor. Prob. Appl. 30 (3) (1985), 507–515.

[24] N. V. Kartashov, *Strongly Stable Markov Chains*, J. Sovi. Math. 34 (1986), 1493–1498.

[25] N. V. Kartashov, *Strong Stable Markov Chains*, Edition VSP, Utrecht TBIMC Scientific Publishers, Kiev, 1996.

[26] J. G. Kemeny - J. L. Snell, *Finite Markov chains*, The University Series in Undergraduate Mathematics. D. Van Nostrand Co., Inc., Princeton, N.J.-Toranto-London-New York, 1960.

[27] S. J. Kirkland - M. Neumann - B. L. Shader, *Applications of Paz's inequality to perturbation bounds for Markov chains*, Linear Algebra Appl. 268 (1998), 183–196.

[28] S. J. Kirkland, *On question concerning condition numbers for Markov chains*, SIAM J. Matrix Anal. Appl. 23 (2002), 1109–1119.

[29] S. J. Kirkland, *Conditioning properties of the stationary distribution for a Markov chain*, Electron. J. Linear Algebra 10 (2003), 1–15.

[30] S. J. Kirkland - M. Neumann - N. Sze, *On optimal condition numbers for Markov chains*, Numer. Math. 110 (2008), 521–537.

[31] W. Li - L. Jiang - W. K. - Ching - L. B. Cui, *On perturbation bounds for the joint stationary distribution of multivariate Markov chain models*, East Asian Journal on Applied Mathematics 3 (2013), 1–17.

[32] Y. Liu, *Perturbation bounds for the stationary distributions of Markov chains*, SIAM J. Matrix Anal. Appl. 33 (2012), 1057–1074.

[33] R. M. Loynes, *The stability of queue with non-independent inter-arrival and ser-*

vice times, Proc. Cambridge Philos. Soc. 58 (3) (1962), 497–520.

[34] C. D. Meyer, *The Role of the Group Generalize Inverse in the Theory of Finite Markov Chains*, SIAM Rev. 17 (1975), 443–464.

[35] C. D. Meyer, *The Condition of Finite Markov Chain and Perturbation bounds for the Limiting distribution*, SIAM J. Algebraic Discrete Methods 1 (1980), 273–287.

[36] C. D. Meyer, *Sensitivity of the stationary distribution of the Markov chains*, SIAM J. Matrix Anal. Appl. 15 (1994), 715–728.

[37] S. P Meyn - R. L. Tweedie, *Markov Chains and Stochastic Stability*, Cambridge University Press, 2nd Edition, 2009.

[38] A. Yu. Mitrophanov, *Sensitivity and convergence of uniformly ergodic Markov chains*, J. Appl. Prob. 42 (2005), 1003–1014.

[39] Z. Mouhoubi - D. Aïssani, *Some Inequalities of the Uniform Ergodicity and Strong Stability of Homogeneous Markov Chains*, Pliska Stud. Math. Bulgar. 17 (2005), 171–189.

[40] Z. Mouhoubi - D. Aïssani, *New perturbation bounds for denumerable Markov chains*, Linear Algebra Appl. 432 (2010), 1627–1649.

[41] Z. Mouhoubi - D. Aïssani, *Uniform Ergodicity and Strong Stability Estimates of Homogeneous Markov Chains*, J. Math. Sci. 20 (2014), 452–461.

[42] Z. Mouhoubi - D. Aïssani, *Strong stability and uniform ergodicity estimates for the waiting process in queuing models with impatient customers*, Markov Process. relat. Fields 25 (2019), 329-358.

[43] M. Neumann - J. Xu, *Improved bounds for a condition number of Markov chains*, Linear Algebra Appl. 386 (2004), 225–241.

[44] J. Neveu, *Base Mathématiques du Calcul des Probabilités*, Masson et Cie, Paris, 1964.

[45] A. Paz - M. Reichaw, *Ergodic theorems for sequences of infinite stochastic matrices*, Proc. Camb. Phil. Soc. 63 (1967), 777–784.

[46] A. Paz, *Ergodic theorems for infinite probabilistic tables*, Ann. Math. Statist. 41 (1970), 539–550.

[47] A. Rhodius, *On ergodicity coefficients of infinite stochastic matrices*, Z. Anal. Anwendungen 19 (2000), 873–887.

[48] P. J.Schweitzer, *Perturbation Theory and Finite Markov Chains*, J. Appl. Prob. 5 (1968), 401–413.

[49] E. Seneta, *Coefficients of ergodicity: Structure and applications*, Adv. in Appl. Prob. 11 (1979), 576–590.

[50] E. Seneta, *Sensitivity to perturbation of the stationary distribution: Some refinements*, Linear Algebra Appl. 108 (1988), 121–126.

[51] E. Seneta, *Perturbation of the stationary distribution measured by the ergodicity coefficients*, Adv. Appl. Prob. 20 (1988), 228–230.

[52] E. Seneta, *Sensitivity analysis, ergodicity coefficient, and rank-one updates for*

finite Markov chain. In: *Numerical Solution of Markov Chains, Probability: Pure and Applied*, ed W.J. Stewart, 8, Marcel-Dekker, New York, 1991, 121–129.

[53] E. Seneta, *Sensitivity of a finite Markov chains under perturbation*, Statist. Probab. Lett. 17 (1993), 163–168.

[54] R. Syski, *Ergodic potential*, Stoch. Proc. Appl. 7 (1978), 311–336.

[55] E. Solan - N. Vieille, *Perturbed Markov chains*, J. Appl. Prob. 40 (2003), 107–122.

Z. MOUHOUBI
Department of Mathematics
Édouard-Montpetit College
945 Chemin de Chambly
Longueuil, QC J4H 3M6, Canada
e-mail: `zahir.mouhoubi@cegepmontpetit.ca`