Asymptotic of products of Markov kernels. Application to deterministic and random forward/backward products.

Loïc HERVÉ, and James LEDOUX [*]

version du Friday 19$^{\text{th}}$ November, 2021 – 16:37

Abstract

The asymptotic of products of general Markov/transition kernels is investigated using Doeblin's coefficient. We propose a general approximating scheme as well as a convergence rate in total variation of such products by a sequence of positive measures. These approximating measures and the control of convergence are explicit from the two parameters in the minorization condition associated with the Doeblin coefficient. This allows us to extend the well-known forward/backward convergence results for stochastic matrices to general Markov kernels. A new result for forward/backward products of random Markov kernels is also established.

AMS subject classification : 60J05, 60F99, 60B20

Keywords : Non-homogeneous Markov chains, Doeblin's coefficient

1 Introduction

There is a large literature on the asymptotic behaviour of non-homogeneous Markov chains. A main objective is to get convergence properties as well as rate of convergence of stochastic algorithms based on general Markov chains as, for instance, in Markov search for optimization or in stochastic simulation. Such an issue requires to analyse various products of transition kernels of the underlying Markov chain. In this paper the asymptotic of products of Markov/transition kernels is investigated using Doeblin's coefficient and the total variation norm. Let us introduce the basic material relevant to this work. Let $(\mathbb{X}, \mathcal{X})$ be a measurable space. We denote by $\mathcal{K}$ the set of all the Markov kernels on $(\mathbb{X}, \mathcal{X})$, and by $\mathcal{P}$ the set of all the probability measures on $(\mathbb{X}, \mathcal{X})$. If $K \in \mathcal{K}$ and if $(a, \nu) \in [0, 1] \times \mathcal{P}$, we write $K \geq a\,\nu$ when

$$\forall (x, A) \in \mathbb{X} \times \mathcal{X}, \quad K(x, A) \geq a\,\nu(A). \tag{1}$$

Obviously every $K \in \mathcal{K}$ satisfies (1) with $a = 0$. Doeblin's coefficient $\alpha(K)$ of any $K \in \mathcal{K}$ is defined as in [LC14] by

$$\alpha(K) := \sup \big\{ a \in [0, 1] : \exists \nu \in \mathcal{P}, \ K \geq a\,\nu \big\}. \tag{2}$$

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

When $\alpha(K) \in (0, 1]$, K satisfies the so-called minorization property [RR04]. We denote by $\mathcal{E}$ the set of all the positive measures μ on $(\mathbb{X}, \mathcal{X})$ such that $\mu(\mathbb{X}) \leq 1$. Let $(\mathcal{B}, \|\cdot\|_\infty)$ denote the space of bounded measurable real-valued functions on $(\mathbb{X}, \mathcal{X})$, equipped with the supremum norm: $\forall f \in \mathcal{B}$, $\|f\|_\infty := \sup_{x \in \mathbb{X}} |f(x)|$. Let $(\mathcal{L}(\mathcal{B}), \|\cdot\|)$ be the Banach space of all the bounded linear operators on $\mathcal{B}$ where $\|\cdot\|$ denotes the operator norm on $\mathcal{L}(\mathcal{B})$ defined by

$$\forall T \in \mathcal{L}(\mathcal{B}), \quad \|T\| := \sup\{\|Tf\|_\infty, \ \|f\|_\infty \leq 1\}.$$

Note that if T is non-negative (i.e. $f \geq 0 \Rightarrow Tf \geq 0$) then $\|T\| = \|T1_\mathbb{X}\|_\infty$. Throughout the paper, $K \in \mathcal{K}$ is identified with its functional action on $\mathcal{B}$ (still denoted by K) defined by

$$\forall f \in \mathcal{B}, \ \forall x \in \mathbb{X}, \quad (Kf)(x) := \int_\mathbb{X} f(y) \, K(x, dy).$$

Similarly any element $\mu \in \mathcal{E}$ acts on $\mathcal{B}$ according to:

$$\forall f \in \mathcal{B}, \quad \mu f = \mu(f) 1_\mathbb{X} \quad \text{where we shortly set} \ \ \mu(f) := \int_\mathbb{X} f(y) \, \mu(dy).$$

Obviously the maps $f \mapsto Kf$ and $f \mapsto \mu f$ are in $\mathcal{L}(\mathcal{B})$. Finally, if $(A, B) \in \mathcal{K}^2$ then $A \cdot B$ denotes the Markov kernel on $(\mathbb{X}, \mathcal{X})$ defined by the product of A by B, which is identified with its action $A \circ B$ on $\mathcal{B}$ (to simplify we only use the notation $A \cdot B$).

The following key statement (Theorem 2.2) is proved in Section 2. Let $(K_j)_{j \geq 1} \in \mathcal{K}^\mathbb{N}$ and, for every $j \geq 1$, let $(a_j, \nu_j) \in [0, 1] \times \mathcal{P}$ be chosen for K_j satisfying Inequality (1). For every $n \geq 1$ let σ_n be a permutation on the finite set $\{1, \ldots, n\}$, and introduce

$$\mathbf{K}_{\sigma_n} := \prod_{j=1}^n K_{\sigma_n(j)} \quad \text{and} \quad \mu_{\sigma_n} := \mathbf{K}_{\sigma_n} - \prod_{j=1}^n \big(K_{\sigma_n(j)} - a_{\sigma_n(j)} \nu_{\sigma_n(j)}\big). \tag{3}$$

Then, for all $n \geq 1$, we have $\mu_{\sigma_n} \in \mathcal{E}, \mu_{\sigma_n} \leq \mathbf{K}_{\sigma_n}$ and the following assertions are equivalent:

(a) $\displaystyle\sum_{i \geq 1} a_i = +\infty$.

(b) $\exists (\sigma_n)_{n \geq 1}, \ \lim_n \|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\| = 0$.

(c) $\forall (\sigma_n)_{n \geq 1}, \ \lim_n \|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\| = 0$.

As a result, Seneta's statements [Sen81] for the convergence of forward/backward products of finite stochastic matrices are extended in Section 3 to general Markov kernels via a condition of type (a) for some block-kernels. When $\mathbb{X}$ is finite, this condition is necessary and sufficient for the so-called weak ergodicity, see [Sen81, IS11]. Using Notation (3) with some fixed $(\sigma_n)_{n \geq 1}$, the weak ergodicity property writes as follows:

$$\lim_{n \to +\infty} \sup_{(x, x') \in \mathbb{X}^2} \sup_{\|f\|_\infty \leq 1} \big|(\mathbf{K}_{\sigma_n} f)(x) - (\mathbf{K}_{\sigma_n} f)(x')\big| = 0. \tag{4}$$

When $\mathbb{X}$ is infinite, condition of type (a) (for block-kernels) seems to be only sufficient for (4) to hold, as mentioned in [LC14] in the context of forward products (see Remark 3.6 for details and further comparisons with [LC14]). The novelty in our work is that the weak ergodicity condition (4) is replaced with the condition $\lim_n \|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\| = 0$ which implies that

$$d_{\mathrm{TV}}\big(\mathbf{K}_{\sigma_n}, \mathcal{E}\big) := \inf_{\mu \in \mathcal{E}} \sup_{x \in \mathbb{X}} \|\mathbf{K}_{\sigma_n}(x, \cdot) - \mu(\cdot)\|_{\mathrm{TV}} \longrightarrow 0 \quad \text{when } n \to +\infty, \tag{5}$$

where $\|\beta\|_{\mathrm{TV}} := \sup_{|f|\leq 1} \left| \int_{\mathbb{X}} f(x)\,\beta(dx) \right|$ denotes the total variation norm of any finite signed measure β on $(\mathbb{X}, \mathcal{X})$. The weak ergodicity property (4) directly implies that $\lim_n d_{\mathrm{TV}}\big(\mathbf{K}_{\sigma_n}, \mathcal{P}\big) = 0$. However the possibility in (5) of considering the distance with respect to the set $\mathcal{E}$ (in place of $\mathcal{P}$) provides much more flexibility in the proofs, knowing that the above positive measure μ_{σ_n} satisfies $\lim_n d_{\mathrm{TV}}(\mu_{\sigma_n}, \mathcal{P}) = 0$. The first interest of our approach is that the property $\lim_n \|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\| = 0$ is more explicit than weak ergodicity, since the sequence $(\mu_{\sigma_n})_n$ is simply defined from the kernels K_j and from the elements of the associated minorization conditions. The second interest is that Condition (b) is equivalent to Condition (a), contrarily to weak ergodicity (excepted in matrix case). The third interest is that the norm equality of Lemma 2.1 gives an accurate control of $\|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\|$, thus of $d_{\mathrm{TV}}\big(\mathbf{K}_{\sigma_n}, \mathcal{E}\big)$.

This new approach allows us to extend some classical results on products of stochastic matrices to general Markov kernels via very simple proofs. For instance, Corollary 4.1 and Theorem 4.3 extend the asymptotic results of [Wol63, CW08, Ste08] and of [HIV76] and provide explicit rates of convergence.

Our approach is also relevant to study the products of random Markov kernels. In particular simple criteria are presented in Theorem 5.1 for the convergence of forward/backward products when $(K_j)_{j\geq 1}$ is a sequence of independent and identically distributed random Markov kernels. To the best of our knowledge the results obtained in Section 5 in the random context are new too, even in matrix case.

Mention that the notion of weak ergodicity also leads to investigate stability issues for discrete and continuous time non-homogeneous Markov processes with respect to the parameters of the models (initial distributions, transition matrices in discrete time, intensity matrix-generator in continuous time). We refer to [LGM00] and to the recent review [ZKS20] (see also the references therein) for such questions in discrete or continuous time context. These issues are not addressed in our work.

The paper is organized as follows. General results concerning the convergence of products of Markov kernels in link with Doeblin's coefficient defined in (2) are presented in Section 2. The specific cases of backward and forward products are studied in Section 3. Complementary statements on the rate of convergence of forward/backward products are presented in Section 4. Applications to products of random Markov kernels are proposed in Section 5.

2 Convergence of products of Markov kernels

Let us consider any $p \in \mathbb{N}^*$ and any family $(T_j)_{1\leq j \leq p} \in \mathcal{K}^p$. For every $1 \leq j \leq p$, let $a_j \in [0, \alpha(T_j)]$ and $\nu_j \in \mathcal{P}$ such that $T_j \geq a_j \nu_j$. Set

$$\mathbf{T}_p := \prod_{j=1}^p T_j \quad \text{and} \quad \mu_p := \mathbf{T}_p - \prod_{j=1}^p \big(T_j - a_j\nu_j\big). \tag{6}$$

Lemma 2.1 *The element μ_p given in (6) belongs to $\mathcal{E}$ and we have $\mu_p \leq \mathbf{T}_p$. Moreover*

$$\big\|\mathbf{T}_p - \mu_p\big\| = \prod_{j=1}^p (1 - a_j). \tag{7}$$

Proof. Let us prove by induction on the integer p that $\mu_p \in \mathcal{E}$ and $\mu_p \leq \mathbf{T}_p$. If $p = 1$, then $\mu_1 = a_1\nu_1$, so that $\mu_1 \in \mathcal{E}$ and $\mu_1 \leq \mathbf{T}_1$. Now assume that the conclusions $\mu_p \in \mathcal{E}$ and $\mu_p \leq \mathbf{T}_p$ holds true for some $p \geq 1$. Let $(T_j)_{1\leq j\leq p+1} \in \mathcal{K}^{p+1}$ and, for every $1 \leq j \leq p+1$, let

$a_j \in [0, \alpha(T_j)]$ and $\nu_j \in \mathcal{P}$ be such that $T_j \geq a_j \nu_j$. Let $\mathbf{T}_p$ and μ_p be given in (6). Introduce $\mathbf{T}_{p+1} = \mathbf{T}_p \cdot T_{p+1}$ and

$$\mu_{p+1} = \mathbf{T}_{p+1} - \prod_{j=1}^{p+1} (T_j - a_j \nu_j) = \mathbf{T}_{p+1} - (\mathbf{T}_p - \mu_p) \cdot (T_{p+1} - a_{p+1} \nu_{p+1}). \tag{8}$$

Then we get $\mathbf{T}_{p+1} - \mu_{p+1} = (\mathbf{T}_p - \mu_p) \cdot (T_{p+1} - a_{p+1} \nu_{p+1})$ so that $\mathbf{T}_{p+1} - \mu_{p+1} \geq 0$ since $\mathbf{T}_p - \mu_p \geq 0$ by induction assumption and $T_{p+1} \geq a_{p+1} \nu_{p+1} \geq 0$. Moreover (8) gives

$$\mu_{p+1} = a_{p+1} \nu_{p+1} + \mu_p \cdot (T_{p+1} - a_{p+1} \nu_{p+1})$$

and we know from the induction assumption that $\mu_p \in \mathcal{E}$ with $\forall f \in \mathcal{B}$, $\mu_p f = \mu_p(f) 1_\mathbb{X}$. Thus

$$\mu_{p+1} f = \mu_{p+1}(f) 1_\mathbb{X} \text{ with } \mu_{p+1}(f) = a_{p+1} \nu_{p+1}(f) + \mu_p (T_{p+1} f - a_{p+1} \nu_{p+1}(f))$$

so that $\mu_{p+1}(\cdot)$ is defined as a finite signed measure on $(\mathbb{X}, \mathcal{X})$. But $\mu_{p+1}(\cdot)$ is a positive measure on $\mathbb{X}$ such that $\mu_{p+1}(\mathbb{X}) \leq 1$ since $a_{p+1} \nu_{p+1} \leq T_{p+1}, (T_{p+1} - a_{p+1} \nu_{p+1})(1_\mathbb{X}) = (1 - a_{p+1}) 1_\mathbb{X}$ and $\mu_p(\mathbb{X}) \leq 1$. We have proved that $\mu_{p+1} \in \mathcal{E}$. The first part of Lemma 2.1 is established.

Finally, to obtain (7), note that

$$\|\mathbf{T}_p - \mu_p\| = \|(\mathbf{T}_p - \mu_p) \cdot 1_\mathbb{X}\| = \|\prod_{j=1}^{p} (T_j - a_j \nu_j) \cdot 1_\mathbb{X}\| = \prod_{j=1}^{p} (1 - a_j)$$

since we have $\mathbf{T}_p - \mu_p \geq 0$, and $\|T_j - a_j \nu_j\| = \|(T_j - a_j \nu_j) \cdot 1_\mathbb{X}\|_\infty$ from $T_j - a_j \nu_j \geq 0$ and $(T_j - a_j \nu_j) \cdot 1_\mathbb{X} = (1 - a_j) 1_\mathbb{X}$. $\qquad\square$

Let Σ be the set of all the sequences $\sigma := (\sigma_n)_{n \geq 1}$, where σ_n is a permutation on the finite set $\{1, \ldots, n\}$. For any $(K_j)_{j \geq 1} \in \mathcal{K}^\mathbb{N}$ and $\sigma \in \Sigma$, let $(\mathbf{K}_{\sigma_n})_{n \geq 1} \in \mathcal{K}^\mathbb{N}$ be defined by

$$\forall n \geq 1, \quad \mathbf{K}_{\sigma_n} := \prod_{j=1}^{n} K_{\sigma_n(j)}. \tag{9}$$

For every $j \geq 1$, let $a_j \in [0, \alpha(K_j)]$ and let $\nu_j \in \mathcal{P}$ be such that $K_j \geq a_j \nu_j$. Finally define

$$\forall n \geq 1, \quad \mu_{\sigma_n} := \mathbf{K}_{\sigma_n} - \prod_{j=1}^{n} (K_{\sigma_n(j)} - a_{\sigma_n(j)} \nu_{\sigma_n(j)}). \tag{10}$$

The following theorem, which has its own interest, is crucial for the study of the forward and backward products in the next section.

Theorem 2.2 *For every $\sigma \in \Sigma$ and for every $n \geq 1$, we have $\mu_{\sigma_n} \in \mathcal{E}$, $\mu_{\sigma_n} \leq \mathbf{K}_{\sigma_n}$, and*

$$\forall n \geq 1, \quad \|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\| = \prod_{j=1}^{n} (1 - a_j). \tag{11}$$

Moreover the following assertions are equivalent:

(a) $\sum_{j \geq 1} a_j = +\infty$.

(b) $\exists \sigma \in \Sigma, \ \lim_n \|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\| = 0$.

(c) $\forall \sigma \in \Sigma, \ \lim_n \|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\| = 0$.

Proof. The first part follows from Lemma 2.1. Since Condition (a) does not depend on $\sigma \in \Sigma$, the equivalences hold true if we show that, for any $(K_j)_{j \geq 1} \in \mathcal{K}^{\mathbb{N}}$ and $\sigma \in \Sigma$,

$$\lim_n \|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\| = 0 \Leftrightarrow \sum_{j \geq 1} a_j = +\infty. \tag{12}$$

The following equivalences hold true from (11)

$$\lim_n \|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\| = 0 \Longleftrightarrow \lim_{n \to +\infty} \sum_{j=1}^n \ln(1 - a_j) = -\infty \Longleftrightarrow \sum_{j \geq 1} \ln(1 - a_j) = -\infty.$$

Moreover the last condition is equivalent to $\sum_{j \geq 1} a_j = +\infty$. Indeed, we have $\forall x \in [0, 1)$, $-x/(1 - x) \leq \ln(1 - x) \leq -x$ from Taylor's formula. Thus $\sum_{j \geq 1} a_j = +\infty$ implies that $\sum_{j \geq 1} \ln(1 - a_j) = -\infty$. Conversely assume that $\sum_{j \geq 1} a_j < +\infty$, and set $\tau_j = a_j/(1 - a_j)$. We have $\lim_j a_j = 0$, thus $\tau_j \sim a_j$ when $j \to +\infty$, so that $\sum_{j \geq 1} \tau_j < +\infty$. Therefore the series $\sum_{j \geq 1} \ln(1 - a_j)$ converges. The proof of (12) is complete. $\qquad\square$

Let us complete Theorem 2.2 with the following statement.

Proposition 2.3 *Let $(K_j)_{j \geq 1} \in \mathcal{K}^{\mathbb{N}}$ and let $\sigma \in \Sigma$. The following assertions are equivalent.*

(a) There exists $(\mathfrak{m}_n)_{n \geq 1} \in \mathcal{E}^{\mathbb{N}}$ such that $\forall n \geq 1$, $\mathfrak{m}_n \leq \mathbf{K}_{\sigma_n}$, and $\lim_n \|\mathbf{K}_{\sigma_n} - \mathfrak{m}_n\| = 0$.

(b) $\lim_n \alpha(\mathbf{K}_{\sigma_n}) = 1$.

Proof. Assume that Assertion (a) holds true. Write $\mathfrak{m}_n = b_n \beta_n$ with $(b_n, \beta_n) \in [0, 1] \times \mathcal{P}$. Then $b_n \beta_n \leq \mathbf{K}_{\sigma_n}$ implies that $b_n \leq \alpha(\mathbf{K}_{\sigma_n})$. Moreover $\lim_n b_n = 1$ from

$$1 - b_n = \|(\mathbf{K}_{\sigma_n} - \mathfrak{m}_n) 1_{\mathbb{X}}\|_\infty = \|\mathbf{K}_{\sigma_n} - \mathfrak{m}_n\|.$$

This gives $\lim_n \alpha(\mathbf{K}_{\sigma_n}) = 1$. Conversely assume that $\lim_n \alpha(\mathbf{K}_{\sigma_n}) = 1$. For every $n \geq 1$ there exists $(a_n, \nu_n) \in [0, 1] \times \mathcal{P}$ such that

$$\alpha(\mathbf{K}_{\sigma_n}) - \frac{1}{n} \leq a_n \leq \alpha(\mathbf{K}_{\sigma_n}) \quad \text{and} \quad a_n \nu_n \leq \mathbf{K}_{\sigma_n}$$

from the definition of $\alpha(\mathbf{K}_{\sigma_n})$. Moreover we have $a_n \nu_n \in \mathcal{E}$ and

$$\|\mathbf{K}_{\sigma_n} - a_n \nu_n\| = \|(\mathbf{K}_{\sigma_n} - a_n \nu_n) 1_{\mathbb{X}}\|_\infty = 1 - a_n$$

with $\lim_n a_n = 1$ since $\lim_n \alpha(\mathbf{K}_{\sigma_n}) = 1$. This gives (a) with $\mathfrak{m}_n = a_n \nu_n$. $\qquad\square$

3 Convergence of forward and backward products

Through this section we consider any sequence $(K_j)_{j \geq 1} \in \mathcal{K}^{\mathbb{N}}$. For every $1 \leq k \leq n$, set

- $\mathbf{K}_{k:n} := \prod_{j=k}^n K_j$; note that $\mathbf{K}_{k:n+1} = \mathbf{K}_{k:n} \cdot K_{n+1}$ (forward products).

- $\mathbf{K}_{n:k} := \prod_{j=n}^k K_j$; note that $\mathbf{K}_{n+1:k} = K_{n+1} \cdot \mathbf{K}_{n:k}$ (backward products).

In the next statements, the integer $k \geq 1$ is fixed, and the sequence of interest is then, either $(\mathbf{K}_{k:n})_{n \geq k}$ for forward products (Theorem 3.2), or $(\mathbf{K}_{n:k})_{n \geq k}$ for backward products (Theorem 3.3). The sequences $(\mathbf{K}_{k:n})_{n \geq k}$ and $(\mathbf{K}_{n:k})_{n \geq k}$ are both of the generic form (9) since we have $\mathbf{K}_{1:n} = \prod_{j=1}^{n} K_{\sigma_n(j)}$ with $\sigma_n(j) = j$, while $\mathbf{K}_{n:1} = \prod_{j=1}^{n} K_{\sigma_n(j)}$ with $\sigma_n(j) = n-j+1$ (setting $k = 1$ to simplify). The common basic property of both backward and forward products with respect to Doeblin's coefficient is the following.

Lemma 3.1 *For any $k \geq 1$, the sequences $(\alpha(\mathbf{K}_{k:i}))_{i \geq k}$ and $(\alpha(\mathbf{K}_{i:k}))_{i \geq k}$ are non-decreasing.*

Proof. These sequences are both non-decreasing provided that:

$$\forall (K, K') \in \mathcal{K}^2, \quad \alpha(K) \leq \min\big(\alpha(K \cdot K'), \alpha(K' \cdot K)\big).$$

Let $(a, \nu) \in [0,1] \times \mathcal{P}$ be such that $a\nu \leq K$. Then $(a\nu) \cdot K' = a(\nu \cdot K') \leq K \cdot K'$. Thus $a \leq \alpha(K \cdot K')$ since $\nu \cdot K' \in \mathcal{P}$. Hence $\alpha(K) \leq \alpha(K \cdot K')$. Similarly we deduce from $K \geq a\nu$ that $K' \cdot K \geq K' \cdot (a\nu) = a\nu$, thus $a \leq \alpha(K' \cdot K)$. Hence $\alpha(K) \leq \alpha(K' \cdot K)$. □

Theorem 3.2 (forward products) *The following assertions are equivalent.*

(a) *For every $k \geq 1$, there exists $(\mathfrak{m}_{k:n})_{n \geq k} \in \mathcal{E}^{\mathbb{N}}$ such that, for every $n \geq k$, $\mathfrak{m}_{k:n} \leq \mathbf{K}_{k:n}$, and such that $\lim_n \|\mathbf{K}_{k:n} - \mathfrak{m}_{k:n}\| = 0$.*

(b) *There exists a strictly increasing sequence $(\ell_j)_{j \geq 1}$ of positive integers such that we have $\sum_{j \geq 1} \alpha(Q_j) = +\infty$ with Q_j defined by $Q_j := \mathbf{K}_{\ell_j : \ell_{j+1}-1}$.*

(c) *$\forall k \geq 1, \; \lim_n \alpha(\mathbf{K}_{k:n}) = 1$.*

Theorem 3.2 extends the statement [Sen81, Th. 4.8] to general Markov kernels. Conditions (b) and (c) have been already proved to be equivalent in [LC14]. That Conditions (b) and (c) are both equivalent to Condition (a) is a new result to the best of our knowledge. Condition (a) then appears as a suitable alternative to the weak ergodicity definition to study the asymptotic behaviour of forward products of general Markov kernels, see Remark 3.6.

Proof. Assume that Assertion (a) holds. Set $\ell_1 := 1$. From Proposition 2.3 we know that $\lim_n \alpha(\mathbf{K}_{1:n}) = 1$. Thus there exists $\ell_2 > 1$ such that $\alpha(\mathbf{K}_{1:\ell_2-1}) \geq 1/2$. Similarly it follows from (a) that there exists $\ell_3 > \ell_2$ such that $\alpha(\mathbf{K}_{\ell_2:\ell_3-1}) \geq 1/2$. Iterating this fact shows that there exists an increasing sequence $(\ell_j)_{j \geq 1}$ of positive integers such that, for every $j \geq 1$, $\alpha(Q_j) \geq 1/2$ with $Q_j := \mathbf{K}_{\ell_j:\ell_{j+1}-1}$, so that $\sum_{j \geq 1} \alpha(Q_j) = +\infty$. Assume that Assertion (b) holds. Let $k \geq 1$, and let $i \geq 1$ such that $\ell_i \geq k$. For every $j \geq i$, it follows from the definition of $\alpha(Q_j)$ that there exists $(a_j, \nu_j) \in [0,1] \times \mathcal{P}$ such that

$$\alpha(Q_j) - \frac{1}{j^2} \leq a_j \leq \alpha(Q_j) \quad \text{and} \quad Q_j \geq a_j \nu_j.$$

Then $\sum_{j \geq i} a_j = +\infty$. Thus, by applying Theorem 2.2 to $(Q_j)_{j \geq i}$, we can define an explicit sequence $(\mu_{i:n})_{n \geq i} \in \mathcal{E}^{\mathbb{N}}$ such that $\mu_{i:n} \leq \mathbf{Q}_{i:n}$ and $\lim_n \|\mathbf{Q}_{i:n} - \mu_{i:n}\| = 0$. To that effect use the formulas (9)-(10) with $(Q_j)_{j \geq i} \in \mathcal{K}^{\mathbb{N}}$ and the above sequence $(a_j, \nu_j)_{j \geq i} \in ([0,1] \times \mathcal{P})^{\mathbb{N}}$. Thus $\lim_n \alpha(Q_{i:n}) = 1$ from Proposition 2.3. Therefore $\lim_n \alpha(\mathbf{K}_{\ell_i:n}) = 1$ since $(\alpha(\mathbf{K}_{\ell_i:n}))_{n \geq \ell_i}$ is non decreasing from Lemma 3.1 and contains the subsequence $(\alpha(Q_{i:n}))_{n \geq i}$. Then we obtain that $\lim_n \alpha(\mathbf{K}_{k:n}) = 1$ since $k \leq \ell_i \leq n$ implies that $\alpha(\mathbf{K}_{\ell_i:n}) \leq \alpha(\mathbf{K}_{k:n})$ from Lemma 3.1. We have proved that $(b) \Rightarrow (c)$. That $(c) \Rightarrow (a)$ follows from Proposition 2.3. □

The results of Theorem 3.2 easily extends to backward products. By contrast the strong ergodicity property (13) below is specific to backward products.

Theorem 3.3 (backward products) *The following assertions are equivalent.*

(a) *For every $k \geq 1$, there exists $(\mathfrak{m}_{n:k})_{n \geq k} \in \mathcal{E}^{\mathbb{N}}$ such that, for every $n \geq k$, $\mathfrak{m}_{n:k} \leq \mathbf{K}_{n:k}$, and such that $\lim_n \|\mathbf{K}_{n:k} - \mathfrak{m}_{n:k}\| = 0$.*

(b) *There exists a strictly increasing sequence $(\ell_j)_{j \geq 1}$ of positive integers such that we have $\sum_{j \geq 1} \alpha(Q_j) = +\infty$ with Q_j defined by $Q_j := \mathbf{K}_{\ell_{j+1}-1:\ell_j}$.*

(c) *$\forall k \geq 1$, $\lim_n \alpha(\mathbf{K}_{n:k}) = 1$.*

Moreover, for any sequence $(\mathfrak{m}_{n:k})_{n \geq k} \in \mathcal{E}^{\mathbb{N}}$ satisfying Condition (a), the following strong ergodicity property holds true: there exists $(\pi_k)_{k \geq 1} \in \mathcal{P}^{\mathbb{N}}$ such that

$$\forall k \geq 1, \quad \lim_n \|\mathfrak{m}_{n:k} - \pi_k\| = \lim_n \|\mathbf{K}_{n:k} - \pi_k\| = 0. \tag{13}$$

Theorem 3.3, which extends the statement [Sen81, Th. 4.18] to general Markov kernels, is new to the best of our knowledge. Note that the possibility of considering block-products in Condition (b) of both Theorems 3.2-3.3 may be relevant. The equivalence between (a), (b) and (c) in Theorem 3.3 can be established exactly as in Theorem 3.2. The strong ergodicity property (13) follows from the next lemma.

Lemma 3.4 *Let $(\mathfrak{m}_{n:k})_{n \geq k} \in \mathcal{E}^{\mathbb{N}}$ satisfying Assertion (a) of Theorem 3.3. Then there exists $(\pi_k)_{k \geq 1} \in \mathcal{P}^{\mathbb{N}}$ such that*

$$\forall k \geq 1, \ \forall n \geq k, \quad \max\left(\|\mathfrak{m}_{n:k} - \pi_k\|, \frac{1}{2}\|\mathbf{K}_{n:k} - \pi_k\|\right) \leq \|\mathbf{K}_{n:k} - \mathfrak{m}_{n:k}\|. \tag{14}$$

Proof. Let $k \geq 1$ be fixed and let $q > p \geq k$. Then

$$\|\mathfrak{m}_{q:k} - \mathfrak{m}_{p:k}\| \leq \|\mathfrak{m}_{q:k} - \mathbf{K}_{q:k}\| + \|\mathbf{K}_{q:k} - \mathfrak{m}_{p:k}\| \leq \|\mathfrak{m}_{q:k} - \mathbf{K}_{q:k}\| + \|\mathbf{K}_{p:k} - \mathfrak{m}_{p:k}\| \tag{15}$$

from $\mathbf{K}_{q:k} - \mathfrak{m}_{p:k} = K_q \cdots K_{p+1} \cdot (\mathbf{K}_{p:k} - \mathfrak{m}_{p:k})$ and $\|K_j\| = 1$. We deduce from the assumption that $(\mathfrak{m}_{n:k})_{n \geq k}$ is a Cauchy's sequence in the Banach space $\mathcal{M}$ of finite signed measures on $(\mathbb{X}, \mathcal{X})$ equipped with the total variation norm (e.g. see [DMPS18, App. D]). Consequently the sequence $(\mathfrak{m}_{n:k})_{n \geq k}$ converges in $\mathcal{M}$ to some $\pi_k \in \mathcal{M}$. When $q \to +\infty$, (15) gives: $\forall p \geq k$, $\|\pi_k - \mathfrak{m}_{p:k}\| \leq \|\mathbf{K}_{p:k} - \mathfrak{m}_{p:k}\|$. Then (14) follows from $\|\mathbf{K}_{p:k} - \pi_k\| \leq \|\mathbf{K}_{p:k} - \mathfrak{m}_{p:k}\| + \|\mathfrak{m}_{p:k} - \pi_k\|$. That $\pi_k \in \mathcal{P}$ is obvious since π_k is the limit of Markov kernels (thus $\pi_k \geq 0$ and $\pi_k(\mathbb{X}) = 1$). $\square$

Remark 3.5 (Finite case and link with Seneta's results) *Seneta introduced in [Sen81, Th. 4.8] the notion of proper coefficients of ergodicity to obtain a necessary and sufficient condition for the forward products of finite stochastic matrices to be weakly ergodic (see (16)). The same holds true in [Sen81, Th. 4.18] for the backward products, with the additional well-known fact that weak and strong ergodicity properties are equivalent in this case. Actually, in the matrix case, Theorems 3.2 and 3.3 corresponds to the statements [Sen81, Th. 4.8 and 4.18]*

in terms of Doeblin's coefficient of ergodicity. Indeed, if $K = (K(i,j))_{1 \le i,j \le d}$ is a stochastic $d \times d-$matrix, then the real number $\alpha(K)$ defined in (2) reduces to $\alpha(K) = \sum_{j=1}^{d} \alpha_j(K)$ with $\alpha_j(K) := \min_{i=1,\dots,d} K(i,j)$. Then Doeblin ergodicity coefficient of K in [Sen81] corresponds to $b(K) = 1 - \alpha(K)$. Let $(K_j)_{j \ge 1}$ be any sequence of stochastic $d \times d-$matrices. According to [Sen81, Def. 4.4], the weak ergodicity condition for forward products is

$$\forall k \ge 1, \ \forall i, i', j = 1, \dots, d, \quad \lim_{n \to +\infty} \left| \mathbf{K}_{k:n}(i,j) - \mathbf{K}_{k:n}(i',j) \right| = 0, \tag{16}$$

where $\mathbf{K}_{k:n} := (\mathbf{K}_{k:n}(i,j))_{1 \le i,j \le d}$. This condition is clearly equivalent to Condition (a) of Theorem 3.2 in the matrix case. The same conclusions hold true for backward products of stochastic $d \times d-$matrices, and the equivalence between weak and strong ergodicity properties is nothing else but the last assertion of Theorem 3.3. A proof of [Sen81, Th. 4.8] only based on the contraction property of the Doeblin ergodicity coefficient $b(\cdot)$ is addressed in [CL10]. This provides a proof of the weak ergodicity characterization given in Doeblin' work [Doe37] without proof. Our method provides another alternative way to establish [Sen81, Th. 4.8] via Doeblin's coefficient, and to prove [Sen81, 4.18] by the way.

Remark 3.6 *Although Theorems 3.2 and 3.3 are extensions of the matrix case, we point out that when $\mathbb{X}$ is infinite, none of the three equivalent conditions (a) (b) (c) in Theorems 3.2 is known to be equivalent to the following weak ergodicity condition*

$$\forall k \ge 1, \quad \lim_{n \to +\infty} \sup_{(x,x') \in \mathbb{X}^2} \sup_{\|f\|_\infty \le 1} \left| (\mathbf{K}_{k:n} f)(x) - (\mathbf{K}_{k:n} f)(x') \right| = 0 \tag{17}$$

which is a natural extension of (16). As mentioned in [LC14, p. 178], Condition (a) of Theorem 3.2 clearly implies that (17) holds. That (17) implies any of the three conditions (a) (b) (c) of Theorem 3.2 is an open question. Let us mention that in [Paz70, Ios72] the weak ergodicity property (17) is proved to be equivalent to the following condition: $\forall k \ge 1, \ \exists (\nu_{k,n})_{n \ge k} \in \mathcal{P}^{\mathbb{N}}, \ \lim_n \|\mathbf{K}_{k:n} - \nu_{k,n}\| = 0$. However this condition is quite different from Conditions (a) of Theorems 3.2 since $\nu_{k,n}$ is a probability measure which does not satisfy $\nu_{k,n} \le \mathbf{K}_{k:n}$ in general. By the way, finding a probability measure $\nu_{k,n}$ satisfying the above hypothesis of [Paz70, Ios72] is a difficult issue in practice. Mention that other norms can be considered to define weak ergodicity, for instance $\mathbb{L}^1$-type norm as in [Muk13].

4 Results on the rate of convergence

When the sequences $(\mathfrak{m}_{k:n})_{n \ge k} \in \mathcal{E}^{\mathbb{N}}$ or $(\mathfrak{m}_{n:k})_{n \ge k} \in \mathcal{E}^{\mathbb{N}}$ in Theorems 3.2 and 3.3 are computed thanks to Formulas (9)-(10), then the norm equality (11) provides an accurate control of $\|\mathbf{K}_{k:n} - \mathfrak{m}_{k:n}\|$. This is illustrated in Proposition 4.1 and Theorem 4.3 below.

Proposition 4.1 *Let $(K_j)_{j \ge 1} \in \mathcal{K}^{\mathbb{N}}$ be such that $\alpha_0 := \inf_{j \ge 1} \alpha(K_j) > 0$. Let $c \in (0, \alpha_0)$. Then for every $k \ge 1$ there exist sequences $(\mu_{k:n})_{n \ge k} \in \mathcal{E}^{\mathbb{N}}$ and $(\mu_{n:k})_{n \ge k} \in \mathcal{E}^{\mathbb{N}}$ such that*

$$\forall n \ge k, \quad \|\mathbf{K}_{k:n} - \mu_{k:n}\| \le (1-c)^{n-k+1} \quad and \quad \|\mathbf{K}_{n:k} - \mu_{n:k}\| \le (1-c)^{n-k+1}. \tag{18}$$

Moreover the following property holds true for backward products

$$\forall k \ge 1, \ \forall n \ge k, \quad \|\mathbf{K}_{n:k} - \pi_k\| \le 2(1-\alpha_0)^{n-k+1} \tag{19}$$

where $(\pi_k)_{k \ge 1} \in \mathcal{P}^{\mathbb{N}}$ is the sequence of Theorem 3.3 associated with the $\mu_{n:k}$'s.

Similar results to (18) are obtained in [Wol63] when $\{K_j, \ j \geq 1\}$ is a finite set of finite stochastic matrices. The results of [Wol63] are extended in [CW08] to the case when $\{K_j, \ j \geq 1\}$ is a compact set of finite stochastic matrices. In the homogeneous case Property (19) is a well-known result (e.g. see [RR04] and Remark 4.2). In the non-homogeneous case Property (19) is proved in [Ste08] when $\mathbb{X}$ is finite. The statement for a complete separable metric space $\mathbb{X}$ is stated in [Ste08] with the indication that a proof could be provided by using an iterated function system. Note that the properties (18) and (19) are obtained here for general state spaces and without any topological assumption on the set $\{K_j, \ j \geq 1\}$.

Proof. To simplify assume that $k = 1$. For every $j \geq 1$ there exist $a_j \in [c, \alpha_0]$ and $\nu_j \in \mathcal{P}$ such that $K_j \geq a_j \nu_j$. Using $\sigma_n(j) = j$ and $\sigma_n(j) = n - j + 1$ respectively, Formula (10) can be used to define $(\mu_{1:n})_{n \geq 1} \in \mathcal{E}^{\mathbb{N}}$ and $(\mu_{n:1})_{n \geq 1} \in \mathcal{E}^{\mathbb{N}}$. Inequalities in (18) follow from the norm equality (11) since $c \leq a_j$. To prove (19), note that the second inequality of (18) and Property (14) applied to the sequence $(\mu_{n:1})_{n \geq 1}$ give: $\forall n \geq 1, \ \|\mathbf{K}_{n:1} - \pi_1\| \leq 2(1 - c)^n$. Inequality (19) then holds since c is arbitrarily closed to α_0. $\qquad \square$

Remark 4.2 *If $K \in \mathcal{K}$ satisfies $K \geq \alpha_0 \nu$ for some $(\alpha_0, \nu) \in (0, 1] \times \mathcal{P}$ and if K is assumed to admit an invariant probability measure π, then it is well-known that $\|K^n - \pi\| \leq (1 - \alpha_0)^n$, see [RR04, Th. 8]. Property (19) in case $K_j = K$ and $k = 1$ provides the same result, up to the factor 2 which is due to the functional definition of the total variation norm. The existence of π is not assumed here (it is a by-product of Theorem 3.3). Obviously, if some iterate K^{n_0} satisfies the above minorization condition (in place of K), the same conclusion holds by replacing the rate $(1 - \alpha_0)^n$ by $(1 - \alpha_0)^{\lfloor n/n_0 \rfloor}$ where $\lfloor \cdot \rfloor$ is the integer part function.*

As an application of Assertion (a) of Theorem 3.2, the following statement extends to general Markov kernels the result of [HIV76] concerning forward products of stochastic matrices, see Remark 4.5.

Theorem 4.3 *Let $K \in \mathcal{K}$ be strongly ergodic, namely*

$$\exists \pi \in \mathcal{P}, \ \exists c \in (0, +\infty), \ \exists \beta \in (0, 1), \ \forall m \geq 1, \quad \|K^m - \pi\| \leq c\,\beta^m. \tag{20}$$

Let $(K_n)_{n \geq 1} \in \mathcal{K}^{\mathbb{N}}$ be such that $\lim_n \|K_n - K\| = 0$ and

$$\exists \alpha_0 \in (0, 1), \ \exists i_0 \geq 1, \ \forall i \geq i_0, \quad \alpha(K_i) > \alpha_0. \tag{21}$$

Then the following uniform convergence holds:

$$\lim_{n \to +\infty} \sup_{k \geq 1} \|\mathbf{K}_{k:k+n} - \pi\| = 0. \tag{22}$$

More precisely, setting $d_{i_0} := (1 - \alpha_0)^{1 - i_0}$, we have for all $k \geq 1$, $n \geq i_0$ and $m \geq 1$

$$\|\mathbf{K}_{k:k+n+m} - \pi\| \leq 2\,d_{i_0}\,(1 - \alpha_0)^n \left(1 + (1 - \alpha_0)^m\right) + m\,\gamma_n + c\,\beta^m \tag{23}$$

with $\gamma_n := \sup_{j \geq n+1} \|K_n - K\|$.

Proof. Note that $\sum_{i \geq 1} \alpha(K_j) = +\infty$ from Assumption (21). It follows from Assertion (a) of Theorem 3.2 that, for every $k \geq 1$, there exists $(\mu_{k,k+n})_{n \geq 1} \in \mathcal{E}^{\mathbb{N}}$ such that, for every $n \geq 1$, $\mu_{k,k+n} \leq \mathbf{K}_{k:k+n}$, and such that

$$\lim_{n \to +\infty} \Delta_{k,k+n} = 0 \quad \text{where} \quad \Delta_{k,k+n} := \|\mathbf{K}_{k:k+n} - \mu_{k,k+n}\|.$$

Actually the sequence $(\mu_{k,k+n})_{n\geq k} \in \mathcal{E}^{\mathbb{N}}$ is provided by Theorem 2.2, from which we deduce the following inequality by using (21)

$$\forall k \geq 1, \ \forall n \geq i_0, \quad \Delta_{k,k+n} \leq d_{i_0} \left(1-\alpha_0\right)^n \quad \text{with} \quad d_{i_0} := \left(1-\alpha_0\right)^{1-i_0}. \qquad (24)$$

Note that $\mu_{k,k+n} \cdot \pi = \mu_{k,k+n}(1_{\mathbb{X}})\,\pi$. We have for $n, m \geq 1$

$$
\begin{aligned}
\mu_{k,k+n+m} - \mu_{k,k+n}(1_{\mathbb{X}})\,\pi \ = \ & \left(\mu_{k,k+n+m} - \mathbf{K}_{k:k+n+m}\right) + \mathbf{K}_{k:k+n} \cdot \left(\mathbf{K}_{k+n+1:k+n+m} - K^m\right) \\
& + \left(\mathbf{K}_{k:k+n} - \mu_{k,k+n}\right) \cdot K^m + \mu_{k,k+n} \cdot \left(K^m - \pi\right).
\end{aligned}
$$

Moreover an easy induction based on the triangular inequality gives

$$\forall i \geq 1, \ \forall m \geq 1, \quad \|\mathbf{K}_{i+1:i+m} - K^m\| \leq m\,\gamma_i \quad \text{with} \quad \gamma_i := \sup_{j\geq i+1} \|K_j - K\|.$$

Thus: $\forall k, n, m \geq 1$, $\|\mathbf{K}_{k+n+1:k+n+m} - K^m\| \leq m\,\gamma_n$ since $\gamma_{k+n} \leq \gamma_n$. From these remarks and from (20) and (24) we deduce that for all $k \geq 1$, $n \geq i_0$ and $m \geq 1$

$$\|\mu_{k,k+n+m} - \mu_{k,k+n}(1_{\mathbb{X}})\,\pi\| \leq d_{i_0}\left(1-\alpha_0\right)^{n+m} + m\,\gamma_n + d_{i_0}\left(1-\alpha_0\right)^n + c\,\beta^m.$$

Next observe that

$$\forall k, n \geq 1, \quad 1 - \mu_{k,k+n}(1_{\mathbb{X}}) = \|\mathbf{K}_{k:k+n}1_{\mathbb{X}} - \mu_{k,k+n}1_{\mathbb{X}}\|_\infty = \|\mathbf{K}_{k:k+n} - \mu_{k,k+n}\| = \Delta_{k,k+n}$$

since $\mathbf{K}_{k:k+n} - \mu_{k,k+n} \geq 0$. Consequently we have for all $k \geq 1$, $n \geq i_0$ and $m \geq 1$

$$
\begin{aligned}
\|\mu_{k,k+n+m} - \pi\| \ &\leq \ \|\mu_{k,k+n+m} - \mu_{k,k+n}(1_{\mathbb{X}})\,\pi\| + \|\left(\mu_{k,k+n}(1_{\mathbb{X}}) - 1\right)\pi\| \\
&\leq \ \|\mu_{k,k+n+m} - \mu_{k,k+n}(1_{\mathbb{X}})\,\pi\| + 1 - \mu_{k,k+n}(1_{\mathbb{X}}) \\
&\leq \ d_{i_0}\left(1-\alpha_0\right)^{n+m} + m\,\gamma_n + 2\,d_{i_0}\left(1-\alpha_0\right)^n + c\,\beta^m
\end{aligned}
$$

from which we deduce that

$$
\begin{aligned}
\|\mathbf{K}_{k:k+n+m} - \pi\| \ &\leq \ \|\mathbf{K}_{k:k+n+m} - \mu_{k,k+n+m}\| + \|\mu_{k,k+n+m} - \pi\| \\
&\leq \ 2\,d_{i_0}\left(1-\alpha_0\right)^{n+m} + m\,\gamma_n + 2\,d_{i_0}\left(1-\alpha_0\right)^n + c\,\beta^m. \qquad (25)
\end{aligned}
$$

This proves (23). Now let $\varepsilon > 0$. First fix $m_0 \geq 1$ such that $c\,\beta^{m_0} \leq \varepsilon/2$. Then

$$\exists n_0 \geq i_0, \ \forall n \geq n_0, \quad 2\,d_{i_0}\left(1-\alpha_0\right)^{n+m_0} + m_0\,\gamma_n + 2\,d_{i_0}\left(1-\alpha_0\right)^n \leq \varepsilon/2$$

since $\alpha_0 \in (0,1]$ and $\lim_n \gamma_n = 0$ by the assumption $\lim_n \|K_n - K\| = 0$. It follows that: $\forall q \geq m_0 + n_0$, $\forall k \geq 1$, $\|\mathbf{K}_{k:k+q} - \pi\| \leq \varepsilon$. This proves (22). $\square$

Remark 4.4 *If Assumption (21) of Theorem 4.3 is replaced with the following one*

$$\exists s \geq 1, \ \exists i_0 \geq 1, \ \inf_{i\geq i_0} \alpha(K_{is+1} \cdots K_{is+s}) > 0,$$

then the results of Theorem 4.3 can be extended by considering suitable block-products. More precisely the proof of Theorem 4.3 applies to the sequence $(K'_j)_{i\geq 0}$ defined by $K'_j := K_{is+1:is+s}$ since K^s is strongly ergodic and $\lim_n \|K'_n - K^s\| = 0$.

Remark 4.5 *When $(K_n)_{n \geq 1}$ is a sequence of infinite stochastic matrices (i.e. $\mathbb{X} = \mathbb{N}$), a qualitative control of $\sup_{k \geq 1} \|\mathbf{K}_{k:k+n} - \pi\|$ was obtained in [HIV76]. Theorem 4.3 not only extends this result to general Markov kernels, but Inequality (23) also provides an accurate (uniform in $k \geq 1$) control of $\|\mathbf{K}_{k:k+n+m} - \pi\|$. In particular, if the Markov kernel K in Theorem 4.3 satisfies $K \geq a\nu$ for some $(a, \nu) \in (0, 1] \times \mathcal{P}$, then Condition (20) holds with $c = 2$ and $\beta = 1 - a$, see Remark 4.2. In this case, Inequality (23) holds true with explicit constants only depending on the data a, i_0, α_0 and γ_n. Note that Assumption (21) is not assumed in [HIV76] because it is automatically fulfilled in case $\mathbb{X} = \mathbb{N}$. Indeed, in this case, the map $\alpha(\cdot)$ is easily proved to be lower semi-continuous on the metric space $(\mathcal{K}, d)$, with d defined by: $\forall (K, K') \in \mathcal{K}^2$, $d(K, K') = \|K - K'\|$. Therefore if $\alpha := \alpha(K) > 0$, then Assumption (21) holds true for any $\alpha_0 \in (0, \alpha)$ from the assumption $\lim_n \|K_n - K\| = 0$ and from the lower semi-continuity of the map $\alpha(\cdot)$ since $\liminf_n \alpha(K_n) \geq \alpha(K)$. If $\alpha(K) = 0$, the proof of Theorem 4.3 can be adapted by considering suitable block-products of some (fixed) length. Indeed it follows from the strong ergodicity of K that $\lim_n \alpha(K^n) = 1$. Thus: $\exists s \geq 1$, $\alpha(K^s) \geq 1/2$. From the assumption $\lim_n \|K_n - K\| = 0$, there exists $i_0 \geq 1$ such that, for every $i \geq i_0$, $\alpha(K_{is+1} \cdots K_{is+s}) \geq 1/4$ since $\lim_i \|K_{is+1} \cdots K_{is+s} - K^s\| = 0$ and $\alpha(\cdot)$ is lower semi-continuous. Then, as already mentioned in the previous remark, the proof of Theorem 4.3 can be applied to the sequence $(K'_j)_{i \geq 0}$ defined by $K'_j := K_{is+1:is+s}$.*

5 Applications to products of random Markov kernels

Let $(K_j)_{j \geq 1}$ be a sequence of random variables (r.v.) defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking their values in $\mathcal{K}$. For the sake of simplicity, if $n \geq k \geq 1$, we still denote by $\mathbf{K}_{k:n}$ and $\mathbf{K}_{n:k}$ the following $\mathcal{K}$-valued random variables:

$$\forall \omega \in \Omega, \quad \mathbf{K}_{k:n}(\omega) := \prod_{j=k}^{n} K_j(\omega) \quad \text{and} \quad \mathbf{K}_{n:k}(\omega) = \prod_{j=n}^{k} K_j(\omega).$$

If $\omega \in \Omega$ is such that $\alpha(\mathbf{K}_{k:n}(\omega)) > 0$, then it follows from the definition of $\alpha(\mathbf{K}_{k:n}(\omega))$ that, for every $a \in (0, \alpha(\mathbf{K}_{k:n}(\omega)))$, there exists $a_{k:n}(\omega) \in [a, \alpha(\mathbf{K}_{k:n}(\omega))]$ and $\nu_{k:n}(\omega) \in \mathcal{P}$ such that $\mathbf{K}_{k:n}(\omega) \geq a_{k:n}(\omega)\nu_{k:n}(\omega)$. The previous proofs are compatible with the present random context assuming that a choice may be done so that the maps $\omega \mapsto a_{k:n}(\omega)$ and $\omega \mapsto \nu_{k:n}(\omega)$ define random variables from $(\Omega, \mathcal{F})$ to $[0, 1]$ and to $\mathcal{P}$ respectively. Note that these assumptions hold in the finite state space case.

The use of Doeblin's coefficient seems to be quite relevant in this random context, as shown by the next theorem which provides necessary and sufficient conditions for the convergence of the forward/backward random products when the sequence $(K_j)_{j \geq 1}$ is assumed to be independent and identically distributed (i.i.d.).

Theorem 5.1 *If $(K_j)_{j \geq 1}$ is i.i.d. then*

1. **(forward products)** *the two following statements are equivalent*

    (a) *there exists an integer number $q \geq 1$ such that $\mathbb{P}\big(\alpha(\mathbf{K}_{1:q}) > 0\big) > 0$*

    (b) *for every $k \geq 1$, there exists a sequence $(\mathfrak{m}_{k:n})_{n \geq k}$ of $\mathcal{E}$-valued r.v. such that we have $\mathbb{P}-$almost surely: $\forall n \geq k$, $\mathfrak{m}_{k:n} \leq \mathbf{K}_{k:n}$, and $\lim_n \|\mathbf{K}_{k:n} - \mathfrak{m}_{k:n}\| = 0$.*

2. **(backward products)** *the two following statements are equivalent*

    (a) there exists an integer number $q \geq 1$ such that $\mathbb{P}\big(\alpha(\mathbf{K}_{q:1}) > 0\big) > 0$,

    (b) There exists a sequence $(\pi_k)_{k \geq 1}$ of $\mathcal{P}$-valued r.v. such that we have $\mathbb{P}-$almost surely: $\forall k \geq 1$, $\lim_n \|\mathbf{K}_{n:k} - \pi_k\| = 0$.

To the best of our knowledge, the results in Theorem 5.1 and in Proposition 5.2 below are new, even in the finite case. When the K_j's are r.v. taking their values in the set $\mathcal{K}_d$ of stochastic $d \times d-$matrices for some $d \geq 1$, Assertion (1b) is a well-known result under the following stronger assumptions, see for instance [CL94, McK14]: the K_j's are i.i.d and $\mathbb{P}(K_1 \in \mathcal{K}_d^*) > 0$, where $\mathcal{K}_d^*$ denotes the subset of $\mathcal{K}_d$ composed of matrices with strictly positive entries. Note that the assumption $\mathbb{P}(K_1 \in \mathcal{K}_d^*) > 0$ in [CL94, McK14] is more restrictive than $\mathbb{P}(\alpha(K_1) > 0) > 0$ since $K_1 \in \mathcal{K}_d^* \Rightarrow \alpha(K_1) > 0$.

    Proof. Assume that (1a) holds true: for some $q \geq 1$, $\mathbb{P}\big(\alpha(\mathbf{K}_{1:q}) > 0\big) > 0$. For every $j \in \mathbb{N}^*$ define: $Q_j = \mathbf{K}_{q(j-1)+1\,:\,qj}$. The sequence $(Q_j)_{j \geq 1}$ is i.i.d. and we have $\mathbb{P}\big(\alpha(Q_1) > 0\big) > 0$ by hypothesis. From $\mathbb{P}(\alpha(Q_1) > 0) > 0$, there exists $p \geq 1$ such that $\mathbb{P}(\alpha(Q_1) \geq 1/p) > 0$. Thus $\sum_{j \geq 1} \mathbb{P}(\alpha(Q_j) \geq 1/p) = +\infty$ since the Q_j's are i.d.. Let $\Omega_0 = \cap_{n \geq 1} \cup_{j \geq n} [\alpha(Q_j) \geq 1/p]$. From the independence of the events $[\alpha(Q_j) \geq 1/p]$, $j \geq 1$, the Borel-Cantelli lemma ensures that $\mathbb{P}(\Omega_0) = 1$ so that $\sum_{j \geq 1} \alpha(Q_j) = +\infty$ $\mathbb{P}-$a.s. Then Property (1b) follows from Theorem 3.2.

    Now assume that $\forall q \geq 1$, $\alpha(\mathbf{K}_{1:q}) = 0$ $\mathbb{P}-$a.s.. Define $\Omega_1 := \cap_{q \geq 1} \big[\alpha(\mathbf{K}_{1:q}) = 0\big]$. Then $\mathbb{P}(\Omega_1) = 1$, and $\forall \omega \in \Omega_1$, $\forall q \geq 1$, $\alpha(\mathbf{K}_{1:q})(\omega) = 0$. It follows from Assertion (c) of Theorem 3.2 that Assertion (1b) of Theorem 5.1 does not hold (in fact, for every $\omega \in \Omega_1$, the expected conclusion for the sequence $(\mathbf{K}_{k:n}(\omega))_n$ does not hold).

    Equivalence (2a) $\Leftrightarrow$ (2b) can be proved similarly from Theorem 3.3.   □

    Let us propose alternative assumptions to obtain that $\sum_{j \geq 1} \alpha(K_j) = +\infty$ $\mathbb{P}-$a.s., so that the statements (1b) and (2b) of Theorem 5.1 hold true.

Proposition 5.2 *Statements (1b) and (2b) of Theorem 5.1 hold true when any of the two following conditions is fulfilled:*

    (i) the r.v. $(K_j)_{j \geq 1}$ are pairwise independent and i.d., and $\mathbb{P}(\alpha(K_1) > 0) > 0$.

    (ii) $(K_j)_{j \geq 1}$ is stationary, $(\alpha(K_j))_{j \geq 1}$ is ergodic, and $\mathbb{P}(\alpha(K_1) > 0) > 0$.

    Proof. From Theorems 3.2-3.3 it is sufficient to prove that any of the two sets of assumption *(i)* or *(ii)* ensures that $\sum_{j \geq 1} \alpha(K_j) = +\infty$ $\mathbb{P}-$a.s. Under Assumption *(i)* such a convergence can be obtained as in the proof of (1a) $\Rightarrow$ (1b) in Theorem 5.1: replace Q_j with K_j and apply the Borel-Cantelli lemma for pairwise independent r.v.. Under Assumption *(ii)* note that the sequence $(\alpha(K_j))_{j \geq 1}$ is stationary and ergodic. Then we obtain that $\sum_{j \geq 1} \alpha(K_j) = +\infty$ $\mathbb{P}-$a.s. under the assumption $\mathbb{P}(\alpha(K_1) > 0) > 0$ from the strong law of large numbers for ergodic stationary sequences.   □

References

[CL94]   J.-F. Chamayou and G. Letac. A transient random walk on stochastic matrices with Dirichlet distributions. *Ann. Probab.*, 22(1):424–430, 1994.

[CL10]   S. Chestnut and M. E. Lladser. Occupancy distributions in Markov chains via Doeblin's ergodicity coefficient. In *21st International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms (AofA'10)*, Discrete Math. Theor. Comput. Sci. Proc., AM, pages 79–92. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2010.

[CW08]   D. Coppersmith and C. W. Wu. Conditions for weak ergodicity of inhomogeneous Markov chains. *Statist. Probab. Lett.*, 78(17):3082–3085, 2008.

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

[Doe37]   W. Doeblin. Le cas dicontinu des probabilité en chaîne. *Publ. Fac. Sci. Univ. Masaryk (Brno)*, 1937.

[HIV76]   C. C. Huang, D. Isaacson, and B. Vinograde. The rate of convergence of certain nonhomogeneous Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 35(2):141–146, 1976.

[Ios72]   M. Iosifescu. On two recent papers on ergodicity in nonhomogeneous Markov chains. *Ann. Math. Statist.*, 43:1732–1736, 1972.

[IS11]   I. C. F. Ipsen and T. M. Selee. Ergodicity coefficients defined by vector norms. *SIAM J. Matrix Anal. Appl.*, 32(1):153–200, 2011.

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

[LGM00]   F. Le Gland and L. Mével. Basic properties of the projective product with application to products of column-allowable nonnegative matrices. *Math. Control Signals Systems*, 13:41–62, 2000.

[McK14]   S. McKinlay. A characterisation of transient random walks on stochastic matrices with Dirichlet distributed limits. *J. Appl. Probab.*, 51(2):542–555, 2014.

[Muk13]   F. Mukhamedov. On L_1-weak ergodicity of nonhomogeneous discrete Markov processes and its applications. *Rev. Mat. Complut.*, 26(2):799–813, 2013.

[Paz70]   A. Paz. Ergodic theorems for infinite probabilistic tables. *Ann. Math. Statist.*, 41:539–550, 1970.

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

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

[Ste08]   Ö. Stenflo. Perfect sampling from the limit of deterministic products of stochastic matrices. *Electron. Commun. Probab.*, 13:474–481, 2008.

[Wol63]   J. Wolfowitz. Products of indecomposable, aperiodic, stochastic matrices. *Proc. Amer. Math. Soc.*, 14:733–737, 1963.

[ZKS20]   A. Zeifman, V. Korolev, and Y. Satin. Two approaches to the construction of perturbation bounds for continuous-time markov chains. *Mathematics*, 8(2), 2020.