



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

Loïc Hervé, James Ledoux

► To cite this version:

Loïc Hervé, James Ledoux. Asymptotic of products of Markov kernels. Application to deterministic and random forward/backward products. 2019. hal-02354594v1

HAL Id: hal-02354594

<https://hal.science/hal-02354594v1>

Preprint submitted on 7 Nov 2019 (v1), last revised 15 Jul 2021 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

Loïc HERVÉ, and James LEDOUX *

version du Thursday 7th November, 2019 – 18:53

Abstract

The asymptotic of products of general Markov/transition kernels is investigated using Doeblin's coefficient. We propose a very 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 explicitly defined from the two parameters in the minorization condition associated with the Doeblin coefficient. This allows us to extend various standard convergence results to general Markov chains and to propose a new result on forward/backward products of random Markov kernels.

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 require 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. Let us introduce the basic material relevant to the description of the main result of this paper (Theorem 3.1). 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 the following condition holds:

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

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]

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

*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 is said to satisfy 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 $\mu(\mathbb{X}) \leq 1$. In other words $\mu \in \mathcal{E}$ if there exists $(b, \beta) \in [0, 1] \times \mathcal{P}$ such that $\mu = b\beta$. Let $(\mathcal{B}, \|\cdot\|_\infty)$ denote the space of bounded measurable real-valued functions on $(\mathbb{X}, \mathcal{X})$, equipped with its usual 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:

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

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 main contribution of this paper is the following result (Theorem 3.1). 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 $\{1, \dots, n\}$, and introduce

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

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

- (a) $\sum_{i \geq 1} a_i = +\infty$.
- (b) $\exists (\sigma_n)_{n \geq 1}, \lim_n \|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\| = 0$.
- (c) $\forall (\sigma_n)_{n \geq 1}, \lim_n \|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\| = 0$.

The basic norm equality $\|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\| = \prod_{i=1}^n (1 - a_i)$ from Lemma 2.1 is central to derive such statements (specifically, that statement (b) implies statement (c)). As a result, the standard cases of forward/backward products are investigated in Theorem 4.1 and Theorem 4.2 respectively. Actually, when \mathbb{X} is finite (i.e. each K_i is a $d \times d$ -stochastic matrix for some $d \geq 1$), condition of type (a) applied to some block-matrices is known to be a necessary and sufficient condition for the forward/backward products to be weakly ergodic, see [Sen81]. Mention that Seneta's proof involves several proper ergodicity coefficients, and that a proof only based on Doeblin's coefficient is addressed in [CL10]. When \mathbb{X} is a general measurable space, Condition (a) seems to be just sufficient for the weak ergodicity of forward products to hold true, see [LC14]. Actually, with the above notations, the weak ergodicity condition writes as follows

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

In our work this condition is replaced with Condition (b), which implies that

$$d_{\text{TV}}(\mathbf{K}_{\sigma_n}, \mathcal{E}) := \inf_{\mu \in \mathcal{E}} \sup_{x \in \mathbb{X}} \|\mathbf{K}_{\sigma_n}(x, \cdot) - \mu(\cdot)\|_{\text{TV}} \longrightarrow 0 \quad \text{when } n \rightarrow +\infty, \quad (4)$$

where for any signed measure β on $(\mathbb{X}, \mathcal{X})$ we denote by $\|\beta\|_{\text{TV}} := \sup_{|f| \leq 1} \left| \int_{\mathbb{X}} f d\beta \right|$ its total variation norm. The weak ergodicity property directly implies that $\lim_n d_{\text{TV}}(\mathbf{K}_{\sigma_n}, \mathcal{P}) = 0$. However the possibility in (4) of considering as a first step the distance with respect to the set \mathcal{E} (in place of \mathcal{P}) seems to provide much more flexibility in the proofs, knowing that as a second step the positive measure μ_{σ_n} in the above properties (b)-(c) clearly satisfies $\lim_n d_{\text{TV}}(\mu_{\sigma_n}, \mathcal{P}) = 0$. More specifically, the first interest of our approach is that Condition (b) is more explicit than the weak ergodicity condition, since the above sequence $(\mu_{\sigma_n})_n \in \mathcal{E}^{\mathbb{N}}$ is simply defined from the kernels K_j and from the elements of the associated minorization conditions. The second interest is that Condition (b) can be proved to be equivalent to Condition (a), contrarily to weak ergodicity (excepted in finite case). The third interest is that the above norm equality gives an accurate control of $\|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\|$, and consequently of $d_{\text{TV}}(\mathbf{K}_{\sigma_n}, \mathcal{E})$.

This new approach allows us to extend some classical results on products of stochastic matrices to general Markov kernels via very simple proofs (even simpler than most of the former ones known in the finite case). For instance, Corollaries 2.1 and 2.2 simply extend the results of [Ste08]. Similarly Corollary 4.1 generalizes the result of [HIV76] to the case of general state spaces. Our approach is also relevant to study the products of random Markov kernels. In particular simple criteria are presented in Corollary 5.2 for the convergence of the forward/backward products in the case when $(K_j)_{j \geq 1}$ is an independent and identically distributed sequence of random Markov kernels. To the best of our knowledge the results obtained in this random context (Section 5) are new too, even in the finite state space case.

The paper is organized as follows. The basic Lemma 2.1 for analysing the products of Markov kernels as well as some direct consequences are presented in Section 2. The main contribution of the paper on the convergence of products of Markov kernels (Theorem 3.1) is proved in Section 3. The specific cases of backward and forward products are studied in Section 4. Applications to products of random Markov kernels are addressed in Section 5.

2 A key lemma

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

$$\mathbf{T}_p := \prod_{j=1}^n T_j \quad \text{and} \quad \mu_p := \mathbf{T}_p - \prod_{j=1}^n (T_j - a_j \nu_j) \quad (5)$$

(here μ_p is defined as a real-valued kernel on (X, \mathbb{X}) , or as an element of $\mathcal{L}(\mathcal{B})$). Note that

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

since $\|T_j - a_j \nu_j\| = \|(T_j - a_j \nu_j) \cdot 1_{\mathbb{X}}\|_{\infty}$ from $T_j - a_j \nu_j \geq 0$ and $(T_j - a_j \nu_j) \cdot 1_{\mathbb{X}} = (1 - a_j)1_{\mathbb{X}}$. Since any finite product of nonnegative kernels is a nonnegative kernel, this inequality is

actually an equality

$$\|\mathbf{T}_p - \mu_p\| = \|(\mathbf{T}_p - \mu_p) \cdot 1_{\mathbb{X}}\| = \left\| \prod_{j=1}^p (T_j - a_j \nu_j) \cdot 1_{\mathbb{X}} \right\| = \prod_{j=1}^p (1 - a_j) \quad (6)$$

and this will be crucial to get Theorem 3.1, the main result of this paper.

Lemma 2.1 *The element μ_p given in (5) belongs to \mathcal{E} (i.e. $\exists(b, \beta) \in [0, 1] \times \mathcal{P}$, $\mu_p = b\beta$) and we have $\mu_p \leq \mathbf{T}_p$. Moreover Equality (6) holds true, that is $\|\mathbf{T}_p - \mu_p\| = \prod_{j=1}^p (1 - a_j)$.*

Proof. First consider μ_p in (5) as an element of $\mathcal{L}(\mathcal{B})$, and note that $0 \leq \mu_p \leq \mathbf{T}_p$. Indeed

$$\mathbf{T}_p - \mu_p = \prod_{j=1}^p (T_j - a_j \nu_j) \geq 0 \quad \text{and} \quad \prod_{j=1}^p (T_j - a_j \nu_j) \leq \mathbf{T}_p$$

since, for every $j = 1, \dots, p$, we have $0 \leq T_j - a_j \nu_j \leq T_j$. Next, in order to prove that $\mu_p \in \mathcal{E}$ holds true, let us proceed by induction on the integer p . If $p = 1$, then $\mu_1 = a_1 \nu_1$, so that $\mu_1 \in \mathcal{E}$. Now assume that the conclusion $\mu_p \in \mathcal{E}$ in Lemma 2.1 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 $\nu_j \in \mathcal{P}$ such that $T_j \geq a_j \nu_j$. Let \mathbf{T}_p and μ_p be given in (5) and introduce

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

We know that $0 \leq \mu_{p+1} \leq \mathbf{T}_{p+1}$ from the above remark. Now note that

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

so that

$$\mu_{p+1} = a_{p+1} \nu_{p+1} + \mu_p \cdot T_{p+1} - a_{p+1} \mu_p \cdot \nu_{p+1}. \quad (7)$$

By induction assumption we know that there exists $\mu_p \in \mathcal{E}$ such that $\forall f \in \mathcal{B}$, $\mu_p f = \mu_p(f) 1_{\mathbb{X}}$. Then Equality (7) rewrites as

$$\mu_{p+1} f = \mu_{p+1}(f) 1_{\mathbb{X}} \quad \text{with} \quad \mu_{p+1}(f) = a_{p+1} \nu_{p+1}(f) + \mu_p(T_{p+1} f) - a_{p+1} \mu_p(1_{\mathbb{X}}) \nu_{p+1}(f). \quad (8)$$

Here $\mu_{p+1}(\cdot)$ is defined as a signed measure on $(\mathbb{X}, \mathcal{X})$. However it follows from $0 \leq \mu_{p+1} \leq \mathbf{T}_{p+1}$ that $\mu_{p+1}(\cdot)$ is in fact a positive measure on \mathbb{X} such that $\mu_{p+1}(\mathbb{X}) \leq 1$, so that $\mu_{p+1} \in \mathcal{E}$. \square

Remark 2.1 *The statements of Lemma 2.1 are new in the framework of non-homogeneous products of Markov kernels to the best of our knowledge (even if \mathbb{X} is finite). They can be thought of as a generalization of a specific decomposition of the finite product of a kernel satisfying Doeblin's condition. A detailed discussion is postponed to Remark 2.4.*

Let us provide a first simple application of Lemma 2.1. Let \mathcal{N} be any subset of \mathcal{K} such that $\alpha_0 := \inf_{K \in \mathcal{N}} \alpha(K) > 0$. For every $n \geq 1$, let $(K_{n,j})_{1 \leq j \leq n} \in \mathcal{N}^n$. Let $c \in (0, \alpha_0)$ be fixed. Then, for every $n \geq 1$ and for every $1 \leq j \leq n$, there exists $a_{n,j} \in [c, \alpha_0]$ and $\nu_{n,j} \in \mathcal{P}$ such that $K_{n,j} \geq a_{n,j} \nu_{n,j}$. Define

$$\forall n \geq 1, \quad \mathbf{T}_n := \prod_{j=1}^n K_{n,j} \quad \text{and} \quad \mu_n := \mathbf{T}_n - \prod_{j=1}^n (K_{n,j} - a_{n,j} \nu_{n,j}).$$

Corollary 2.1 *Under the previous assumptions, we have $\mu_n \in \mathcal{E}$, $\mu_n \leq \mathbf{T}_n$. Moreover $\lim_n \|\mathbf{T}_n - \mu_n\| = 0$ with $\forall n \geq 1$, $\|\mathbf{T}_n - \mu_n\| = \prod_{j=1}^n (1 - a_{n,j}) \leq (1 - c)^n$.*

Let us give an application to the convergence of backward products of Markov kernels.

Corollary 2.2 *Let $(K_j)_{j \geq 1} \in \mathcal{K}^{\mathbb{N}}$ be such that $\alpha_0 := \inf_{j \geq 1} \alpha(K_j) > 0$, and for every $n \geq 1$ let us introduce the backward product $\mathbf{T}_n := \prod_{j=p}^1 K_j$. Then there exists $\pi \in \mathcal{P}$ such that $\lim_n \|\mathbf{T}_n - \pi\| = 0$ with moreover $\forall n \geq 1$, $\|\mathbf{T}_n - \pi\| \leq 2(1 - \alpha_0)^n$.*

Proof. Let $c \in (0, \alpha_0)$ be fixed, and for every $j \geq 1$, let $a_j \in [c, \alpha_0]$ and $\nu_j \in \mathcal{P}$ such that $K_j \geq a_j \nu_j$. Define

$$\forall n \geq 1, \quad \mu_n = \mathbf{T}_n - \prod_{j=p}^1 (K_j - a_j \nu_j).$$

From Corollary 2.1 applied to $\mathcal{N} = \{K_j, j \geq 1\}$ it follows that $\|\mathbf{T}_n - \mu_n\| \leq (1 - c)^n$. Let $q > p \geq 1$. Then

$$\|\mu_q - \mu_p\| \leq \|\mu_q - \mathbf{T}_q\| + \|\mathbf{T}_q - \mu_p\| \leq \|\mu_q - \mathbf{T}_q\| + \|\mathbf{T}_p - \mu_p\| \quad (9)$$

from $\mathbf{T}_q - \mu_p = K_q \cdots K_{p+1}(\mathbf{T}_p - \mu_p)$ and $\|K_j\| = 1$. Then $(\mu_n)_{n \geq 1}$ 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. Consequently the sequence $(\mu_n)_{n \geq 1}$ converges in \mathcal{M} to some $\pi \in \mathcal{M}$. From $\|\mathbf{T}_n - \pi\| \leq \|\mathbf{T}_n - \mu_n\| + \|\mu_n - \pi\|$, we deduce that $\lim_n \|\mathbf{T}_n - \pi\| = 0$. That $\pi \in \mathcal{P}$ is then obvious since π is the limit of Markov kernels (thus $\pi \geq 0$ and $\pi(\mathbb{X}) = 1$). Finally it follows from (9) and from $\lim_q \mu_q = \pi$ that $\forall p \geq 1$, $\|\pi - \mu_p\| \leq (1 - c)^p$. Hence $\|\mathbf{T}_n - \pi\| \leq 2(1 - c)^n$. The last inequality of Corollary 2.2 then holds since c is arbitrarily closed to α_0 . \square

Remark 2.2 *Let $\mathcal{N} = \{K_1, \dots, K_p\}$ be a finite set of stochastic $d \times d$ -matrices for some $d \geq 1$. The following statement is proved in [Wol63]: if any stochastic matrix A obtained as a finite product of matrices in \mathcal{N} (repetitions permitted) is aperiodic and irreducible, then for every $\varepsilon > 0$, there exists an integer number $n_\varepsilon \geq 1$ such that for all $n \geq n_\varepsilon$ every stochastic matrix $B = (b_{i,j})_{1 \leq i,j \leq d}$ obtained as a product of n matrices in \mathcal{N} (repetitions permitted) satisfies : $\max_j \max_{i_1, i_2} |b_{i_1,j} - b_{i_2,j}| < \varepsilon$. Corollary 2.1 provides this conclusion when the above aperiodicity/irreducibility hypothesis of [Wol63] is replaced with the following one : $\forall i = 1, \dots, p$, $\alpha(K_i) > 0$. Note that this assumption cannot be compared with that of [Wol63] since an aperiodic and irreducible stochastic matrix A may satisfy $\alpha(A) = 0$, while a stochastic matrix A satisfying $\alpha(A) > 0$ is not necessarily aperiodic and irreducible. The result of [Wol63] is extended in [CW08] to the case when \mathcal{N} is a compact set of stochastic $d \times d$ -matrices. Note that no topological assumption on \mathcal{N} is required in Corollary 2.1 and that the state space may be general.*

Remark 2.3 *Corollary 2.2 is proved in [Ste08] when \mathbb{X} is finite. The statement for a complete separable metric space \mathbb{X} is stated in [Ste08]. It is mentioned in [Ste08] that a proof could be provided by using an iterated function system. Note that our proof via Lemma 2.1 is quite simple and is valid for general state spaces. The convergence of backward products is investigated in Theorem 4.1 in a more general setting by relaxing assumption $\inf_{j \geq 1} \alpha(K_j) > 0$.*

Remark 2.4 (The case of finite products of a Markov kernel) *In the homogeneous case Corollary 2.2 is a well-known result (e.g. see [RR04]). Actually it can be directly proved using an explicit computation of μ_n from the iterates of $K - a\nu$. Indeed let $K \in \mathcal{K}$ such that $\alpha(K) \in (0, 1]$. Let $a \in (0, \alpha(K)]$ and $\nu \in \mathcal{P}$ such that $K \geq a\nu$. Define*

$$\forall n \geq 1, \quad \mu_n := K^n - (K - a\nu)^n.$$

It follows from Lemma 2.1 that $\mu_n \in \mathcal{E}$, $\mu_n \leq K^n$ and that

$$\forall n \geq 1, \quad \|K^n - \mu_n\| = (1 - a)^n. \quad (10)$$

Since $p \cdot \mu_n = \mu_n$ for any $p \in \mathcal{P}$, it follows that π is a K -invariant probability iff π is a stochastic solution to equation $\pi \cdot (I - (K - \mu_1)) = \mu_1$. Since $\|K - \mu_1\| = 1 - a \in (0, 1)$ from (10), we know that the last equation has a unique solution. Let π be the invariant distribution of K . We deduce from $0 \leq \mu_n \leq K^n$ and from $\pi \cdot K = \pi$ and $\pi \cdot \mu_n = \mu_n$ that

$$\forall n \geq 1, \quad 0 \leq \mu_n \leq \pi, \quad \text{that is: } \forall f \in \mathcal{B}, f \geq 0, \quad 0 \leq \int f d\mu_n \leq \int f d\pi.$$

Moreover it follows from (10) that

$$\forall n \geq 1, \quad \|\pi - \mu_n\| \leq (1 - a)^n. \quad (11)$$

We know from Formula (7) that $\mu_1 = a\nu$ and $\mu_{i+1} = a\nu + \mu_i \cdot (K - a\nu)$ for any $i \geq 1$. Then we obtain by induction that for any $i \geq 1$, $\mu_i = a\nu \sum_{j=0}^{i-1} (K - a\nu)^j$. Thus the sequence $(\mu_i)_{i \geq 1}$ is non-decreasing, that is $\forall i \geq 1$, $\mu_i \leq \mu_{i+1}$. Note that we have obtained the following decomposition of the Markov kernel K^n

$$K^n = a\nu \sum_{j=0}^{n-1} (K - a\nu)^j + (K - a\nu)^n = a\nu \sum_{j=0}^{n-1} (1 - a)^j S^j + (1 - a)^n S^n$$

where $S := (1/(1 - a))(K - a\nu)$ is a Markov kernel. Such a decomposition is known for a Markov kernel K satisfying the Doeblin condition with $\alpha(K) \in (0, 1]$ and any $a \in (0, \alpha(K)]$ (e.g. see [LC14]).

3 Convergence of products of Markov kernels

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 $\{1, \dots, n\}$. 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

$$\forall n \geq 1, \quad \mathbf{K}_{\sigma_n} := \prod_{j=1}^n K_{\sigma_n(j)}. \quad (12)$$

Proposition 3.1 *Let $(K_j)_{j \geq 1} \in \mathcal{K}^{\mathbb{N}}$ and let $\sigma \in \Sigma$. The following assertions are equivalent.*

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

Proof. Assume that Assertion (a) holds true. Write $\mathbf{m}_n = b_n \beta_n$ 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

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

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 $\mathbf{m}_n = a_n \nu_n$. \square

In the next statement we consider any $(K_j)_{j \geq 1} \in \mathcal{K}^{\mathbb{N}}$. For every $j \geq 1$, let $a_j \in [0, \alpha(K_j)]$ and let $\nu_j \in \mathcal{P}$ be such that $K_j \geq a_j \nu_j$. If $\sigma \in \Sigma$, let $(\mathbf{K}_{\sigma_n})_{n \geq 1} \in \mathcal{K}^{\mathbb{N}}$ be given by (12), and $(\mu_{\sigma_n})_{n \geq 1} \in \mathcal{E}^{\mathbb{N}}$ be defined by

$$\mu_{\sigma_n} := \mathbf{K}_{\sigma_n} - \prod_{j=1}^n (K_{\sigma_n(j)} - a_{\sigma_n(j)} \nu_{\sigma_n(j)}). \quad (13)$$

Theorem 3.1 *For every $\sigma \in \Sigma$ and for every $n \geq 1$, we have $\mu_{\sigma_n} \in \mathcal{E}$ and $\mu_{\sigma_n} \leq \mathbf{K}_{\sigma_n}$. 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) $\forall \sigma \in \Sigma, \lim_n \|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\| = 0$.

Finally, under any of these three conditions, the following norm equality holds:

$$\forall n \geq 1, \quad \|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\| = \prod_{j=1}^n (1 - a_j). \quad (14)$$

Proof. That $\mu_{\sigma_n} \in \mathcal{E}$ and $\mu_{\sigma_n} \leq \mathbf{K}_{\sigma_n}$ follow from Lemma 2.1. The equivalences in Theorem 3.1 follow from the next lemma. Note that the condition $\sum_{j \geq 1} a_j = +\infty$ does not depend on $\sigma \in \Sigma$. \square

Lemma 3.2 *Let $(K_j)_{j \geq 1} \in \mathcal{K}^{\mathbb{N}}$ and $\sigma \in \Sigma$. Then: $\lim_n \|\mathbf{K}_{\sigma_n} - \mu_{\sigma_n}\| = 0 \Leftrightarrow \sum_{j \geq 1} a_j = +\infty$.*

Proof. From the norm equality (6) it follows that

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

from which we deduce (14). Then the following equivalences hold true

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

Now let us prove that the last condition is equivalent to $\sum_{j \geq 1} a_j = +\infty$. Note that

$$\forall x \in [0, 1), \quad \frac{-x}{1-x} \leq \ln(1-x) \leq -x$$

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 $j \rightarrow +\infty$, so that $\sum_{j \geq 1} \tau_j < +\infty$. Therefore the series $\sum_{j \geq 1} \ln(1 - a_j)$ converges. This proves Lemma 3.2. \square

4 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 $k \geq 1$ has to be seen as fixed, and the sequence of interest is then, either $(\mathbf{K}_{k:n})_{n \geq k}$ for forward products (Theorem 4.1), or $(\mathbf{K}_{n:k})_{n \geq k}$ for backward products (Theorem 4.2). The sequences $(\mathbf{K}_{k:n})_{n \geq k}$ and $(\mathbf{K}_{n:k})_{n \geq k}$ are both of the form given in (12). More precisely, considering $k = 1$ to simplify, 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 $\sigma_n(j) = n - j + 1$. The common basic property of both backward and forward products with respect to Doeblin's coefficient is the following.

Proposition 4.1 *For every $j \geq 1$, the sequences $(\alpha(\mathbf{K}_{j:i}))_{i \geq j}$ and $(\alpha(\mathbf{K}_{i:j}))_{i \geq j}$ are both non decreasing.*

This proposition easily follows from the next lemma.

Lemma 4.2 *Let $(K, K') \in \mathcal{K}^2$. We have $\alpha(K) \leq \min(\alpha(K \cdot K'), \alpha(K' \cdot K))$.*

Proof. 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 $\alpha(K) \leq \alpha(K' \cdot K)$. \square

Theorem 4.1 (forward products) *The following assertions are equivalent.*

- For every $k \geq 1$, there exists $(\mathbf{m}_{k:n})_{n \geq k} \in \mathcal{E}^{\mathbb{N}}$ such that, for every $n \geq k$, $\mathbf{m}_{k:n} \leq \mathbf{K}_{k:n}$, and such that $\lim_n \|\mathbf{K}_{k:n} - \mathbf{m}_{k:n}\| = 0$.*
- There exists an increasing sequence $(\ell_j)_{j \geq 1}$ of positive integers such that $\sum_{j \geq 1} \alpha(Q_j) = +\infty$ with Q_j defined by $Q_j := \mathbf{K}_{\ell_j: \ell_{j+1}-1}$.*
- $\forall k \geq 1$, $\lim_n \alpha(\mathbf{K}_{k:n}) = 1$.*

In [LC14], Conditions (b) and (c) have been already proved to be equivalent, and to be a sufficient condition for the weak ergodicity of the sequence $(K_j)_{j \geq 1} \in \mathcal{K}^{\mathbb{N}}$, see Remarks 4.3 and 4.5. 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. That any of the three conditions of Theorem 4.1 implies the weak ergodicity property is straightforward by using Condition (a), see Remark 4.3. Moreover note that, under the assumption $\sum_{j \geq 1} \alpha(K_j) = +\infty$, the sequence $(\mathbf{m}_{k:n})_{n \geq k} \in \mathcal{E}^{\mathbb{N}}$ in Theorem 4.1 can be computed by using the formulas (12)-(13) and that the norm equality (14) provides an accurate control of $\|\mathbf{K}_{k:n} - \mathbf{m}_{k:n}\|$. This remark is applied in the forthcoming Corollary 4.1. Under the more general condition (b) in Theorem 4.1, the same facts hold by considering suitable block-products (see the proof below). The previous comments obviously extend to the backward products in Theorem 4.2.

Proof of Theorem 4.1. Assume that Assertion (a) holds. Set $\ell_1 = 1$. From Proposition 3.1 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 a strictly 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 $\sum_{j \geq 1} \alpha(Q_j) = +\infty$. Conversely 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 $\sum_{j \geq i} a_j = +\infty$. Consequently, by applying Theorem 3.1 to the sequence $(Q_j)_{j \geq i}$, 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 $\lim_n \|\mathbf{Q}_{i:n} - \mu_{i:n}\| = 0$, where $\mathbf{Q}_{i:n} = Q_i \cdots Q_n$. To that effect use the formulas (12)-(13) in case $\sigma_n(j) = j$ with the sequence $(Q_j)_{j \geq i} \in \mathcal{K}^{\mathbb{N}}$ 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 3.1. Therefore $\lim_n \alpha(\mathbf{K}_{\ell_i:n}) = 1$ since the sequence $(\alpha(\mathbf{K}_{\ell_i:n}))_{n \geq \ell_i}$ is non decreasing from Proposition 4.1 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 Proposition 4.1. Then Assertion (a) follows from Proposition 3.1 applied to $(K_j)_{j \geq k}$. We have proved that Conditions (a) and (b) are equivalent. The equivalence between (a) and (c) follows from Proposition 3.1. \square

The results of Theorem 4.1 easily extends to backward products. By contrast the strong ergodicity property (15) below is specific to backward products.

Theorem 4.2 (backward products) *The following assertions are equivalent.*

- (a) *For every $k \geq 1$, there exists $(\mathbf{m}_{n:k})_{n \geq k} \in \mathcal{E}^{\mathbb{N}}$ such that, for every $n \geq k$, $\mathbf{m}_{n:k} \leq \mathbf{K}_{n:k}$, and such that $\lim_n \|\mathbf{K}_{n:k} - \mathbf{m}_{n:k}\| = 0$.*
- (b) *There exists a strictly increasing sequence $(\ell_j)_{j \geq 1}$ of positive integers such that $\sum_{j \geq 1} \alpha(Q_j) = +\infty$ 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 any of these three conditions implies that the following strong ergodicity property holds: there exists $(\pi_k)_{k \geq 1} \in \mathcal{P}^{\mathbb{N}}$ such that

$$\forall k \geq 1, \quad \lim_n \|\mathbf{K}_{n:k} - \pi_k\| = 0. \quad (15)$$

Proof. The equivalence between (a), (b) and (c) can be established exactly as in Theorem 4.1. The last assertion of Theorem 4.2 can be proved as in Corollary 2.2: more precisely, for any fixed $k \geq 1$ fixed, prove that $(\mathbf{m}_{n:k})_{n \geq k}$ is a Cauchy sequence (proceed as in Corollary 2.2 with $\mathbf{K}_{n:k}$, $\mathbf{m}_{n:k}$ and π_k in place of \mathbf{T}_n , μ_n and π respectively). \square

Remark 4.1 (Finite case and link with Seneta's results) *Seneta in [Sen81, Th. 4.8.] introduced the notion of proper coefficients of ergodicity to obtain a necessary and sufficient condition for the forward products of stochastic matrices to be weakly ergodic. The same holds true in [Sen81, Th. 4.18.] for the backward products of stochastic matrices, with the additional well-known fact that weak and strong ergodicity properties are equivalent in this case. Actually, in the matrix case, Theorems 4.1 and 4.2 corresponds to the statements [Sen81, Th. 4.8. and 4.18.] when they are stated using Doeblin's coefficient of ergodicity. More precisely, note that, if $K = (K(i, j))_{1 \leq i, j \leq d}$ is a stochastic $d \times d$ -matrix, then the real number $\alpha(K)$ defined in (2) is given by*

$$\alpha(K) := \sum_{j=1}^d \alpha_j(K) \quad \text{with } \forall j = 1, \dots, d, \quad \alpha_j(K) := \min_{i=1, \dots, d} K(i, j).$$

Indeed we clearly have

$$K \geq \mathfrak{A}(K) := \begin{pmatrix} \alpha_1(K) & \cdots & \alpha_d(K) \\ \vdots & \cdots & \vdots \\ \alpha_1(K) & \cdots & \alpha_d(K) \end{pmatrix} \quad (16)$$

and any $d \times d$ -matrix L having d identical non-negative rows and satisfying $L \leq K$ is such that $L \leq \mathfrak{A}(K)$. This gives the claimed statement. Recall that Doeblin's ergodicity coefficient of K is defined in [Sen81] by $b(K) = 1 - \alpha(K)$, and that $b(\cdot)$ is a proper coefficient of ergodicity according to [Sen81, Chap. 4]. Now let $(K_j)_{j \geq 1}$ be any sequence of stochastic $d \times d$ -matrices. The weak ergodicity condition for forward products writes as (see [Sen81, Def. 4.4.])

$$\forall k \geq 1, \forall i, i', j = 1, \dots, d, \quad \lim_{n \rightarrow +\infty} |\mathbf{K}_{k:n}(i, j) - \mathbf{K}_{k:n}(i', j)| = 0, \quad (17)$$

where $\mathbf{K}_{k:n} := (\mathbf{K}_{k:n}(i, j))_{1 \leq i, j \leq d}$. Obviously (17) implies that $\lim_n \|\mathbf{K}_{k:n} - \mathfrak{A}(\mathbf{K}_{k:n})\| = 0$, so that the weak ergodicity condition (17) is equivalent to Condition (a) of Theorem 4.1 in the matrix case. The same holds 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 4.2.

Remark 4.2 *The possibility of considering block-products in Condition (b) of Theorems 4.1 and 4.2 may be relevant. For instance we can easily find stochastic 3×3 -matrices A_1 and A_2 such that $\alpha(A_i) = 0$ for $i = 1, 2$, $\alpha(A_1 A_2) > 0$, and such that the product $A_1 A_2$ is aperiodic and irreducible (thus there exists $\pi \in \mathcal{P}$ such that $\lim_n \|(A_1 A_2)^n - \pi\| = 0$). Let $(K_j)_{j \geq 1}$ be defined by $K_j = A_1$ if j is odd and $K_j = A_2$ if j is even. Then Condition (a) of Theorem 4.1 is fulfilled since $\lim_n \|\mathbf{K}_{k:n} - \mathbf{m}_{k:n}\| = 0$ with $\mathbf{m}_{k:n} = \pi$ if n is even, and $\mathbf{m}_{k:n} = \pi \cdot A_1$ if n is odd. Note that $\sum_{j \geq 1} \alpha(K_j) = 0$, but $\sum_{j \geq 1} \alpha(K_{2\ell-1} K_{2\ell}) = +\infty$.*

Remark 4.3 *Theorems 4.1 and 4.2 are extensions of the matrix case, up to the following fact: none of the three (equivalent) conditions (a) (b) (c) in Theorems 4.1 or 4.2 is known*

to be equivalent to the weak ergodicity condition (excepted in finite case). More precisely, as a natural extension of (17), the forward products associated with any sequence $(K_j)_{j \geq 1} \in \mathcal{K}^{\mathbb{N}}$ are said to be weakly ergodic if (see [Jos72])

$$\forall k \geq 1, \quad \lim_{n \rightarrow +\infty} \sup_{(x, x') \in \mathbb{X}^2} \sup_{\|f\|_{\infty} \leq 1} |(\mathbf{K}_{k:n}f)(x) - (\mathbf{K}_{k:n}f)(x')| = 0. \quad (18)$$

Any of the three conditions (a) (b) (c) of Theorem 4.1 implies that (18) holds true. Indeed, if Condition (a) of Theorem 4.1 is fulfilled, then it follows from the triangular inequality that, for every $(x, x') \in \mathbb{X}^2$ and for every $f \in \mathcal{B}$ such that $\|f\|_{\infty} \leq 1$

$$|(\mathbf{K}_{k:n}f)(x) - (\mathbf{K}_{k:n}f)(x')| \leq 2\|\mathbf{K}_{k:n} - \mathbf{m}_{k:n}\|$$

from which we deduce (18). That the weak ergodicity property implies any of the three conditions (a) (b) (c) of Theorem 4.1 is an open question. In this regard note that Property (18) stated with any fixed $x' = a \in \mathbb{X}$ implies that $\lim_n \|\mathbf{K}_{k:n} - \mathbf{m}_{k:n}\| = 0$ with $\mathbf{m}_{k:n}f = (\mathbf{K}_{k:n}f)(a)$. However, since such a $\mathbf{m}_{k:n}$ is a probability measure, the condition $\mathbf{m}_{k:n} \leq \mathbf{K}_{k:n}$ is rather restrictive (i.e. in general there is no $a \in \mathbb{X}$ such that $(\mathbf{K}_{k:n}f)(a) \leq \mathbf{K}_{k:n}$). The same remarks hold for backward products.

Remark 4.4 In [Paz70] (countable state space) and [Jos72] (general state space) it is proved that the weak ergodicity property (18) is equivalent to the analog of Condition (b) of Theorem 4.1, where $\alpha(\cdot)$ is replaced with $\beta(\cdot)$ defined by

$$\forall K \in \mathcal{K}, \quad \beta(K) := 1 - \delta(K) \quad \text{with} \quad \delta(K) = \sup_{(x, x') \in \mathbb{X}^2} \sup_{\|f\|_{\infty} \leq 1} |(Kf)(x) - (Kf)(x')|.$$

In [Paz70, Jos72] the weak ergodicity property (18) is also proved to be equivalent to the following condition :

$$\forall k \geq 1, \quad \exists (\nu_{k,n})_{n \geq k} \in \mathcal{P}^{\mathbb{N}}, \quad \lim_n \|\mathbf{K}_{k:n} - \nu_{k,n}\| = 0. \quad (19)$$

The equivalence between (18) and (19) is straightforward. Indeed, if (18) holds (i.e. $\forall k \geq 1, \lim_n \delta(\mathbf{K}_{k:n}) = 0$), then (19) is satisfied with $\nu_{k,n}(f) = (\mathbf{K}_{k:n}f)(a)$ for any fixed $a \in X$. That (19) implies (18) is obvious from the triangular inequality. However it is worth noticing that Condition (19) is much less relevant than Condition (a) of Theorem 4.1 since the $\nu_{k,n}$'s in (19) are probability measures (finding $\nu_{k,n}$ satisfying (19) is a difficult issue), while the $\mathbf{m}_{k:n}$'s in Condition (a) of Theorem 4.1 are provided by Theorem 3.1, with moreover an explicit estimate of $\|\mathbf{K}_{k:n} - \mathbf{m}_{k:n}\|$ due to (14).

Remark 4.5 If $(K, K') \in \mathcal{K}^2$ then $\alpha(K \cdot K') \geq \alpha(K) \alpha(K')$. Indeed let $(a, \nu) \in [0, 1] \times \mathcal{P}$ and $(a', \nu') \in [0, 1] \times \mathcal{P}$ be such that $a\nu \leq K$ and $a'\nu' \leq K'$. Then $(a\nu) \cdot (a'\nu') \leq K \cdot K'$. This gives $aa'\nu' \leq K \cdot K'$ since $(a\nu) \cdot (a'\nu') = aa'\nu'$. Hence $aa' \leq \alpha(K \cdot K')$, from which we deduce the desired inequality. Setting $b(K) = 1 - \alpha(K)$ as a natural extension of Doeblin's ergodicity coefficient (see Remark 4.1), we obtain that:

$$\forall (K, K') \in \mathcal{K}^2, \quad b(K \cdot K') \leq b(K) b(K'). \quad (20)$$

Surprisingly the contraction property (20) is not used in our work, contrarily to the papers [CL10, LC14] where (20) plays an important role. More precisely, in the finite case, Property

(20) is used in [CL10] to propose a new proof of [Sen81, Th. 4.8.] in the specific case of Doeblin's coefficient. This corresponds to Doeblin's characterization of weak ergodicity which has been stated without proof in [Doe37], see [Sen73, pages 509-510] for details about Doeblin's paper. Note that, in the finite case, the proof of Theorems 4.1, completed by the straightforward arguments of Remark 4.1 giving the equivalence with the weak ergodicity property, also provides a simple alternative way to establish [Sen81, Th. 4.8.] via Doeblin's coefficient.

As an application of Assertion (a) of Theorem 4.1, the following statement extends to general Markov kernels the result of [HIV76] concerning forward products (the infinite matrix case considered in [HIV76] is discussed in Remark 4.7).

Corollary 4.1 *Let $K \in \mathcal{K}$ be strongly ergodic, namely*

$$\exists \pi \in \mathcal{P}, \exists c \in (0, +\infty), \exists \beta \in (0, 1), \forall m \geq 1, \quad \|K^m - \pi\| \leq c\beta^m. \quad (21)$$

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

$$\exists \alpha_0 \in (0, 1), \exists i_0 \geq 1, \forall i \geq i_0, \quad \alpha(K_i) > \alpha_0. \quad (22)$$

Then the following uniform (in $k \geq 1$) convergence holds:

$$\lim_{n \rightarrow +\infty} \sup_{k \geq 1} \|\mathbf{K}_{k:k+n} - \pi\| = 0. \quad (23)$$

More precisely there exists $d \in (0, +\infty)$ such that for all $k \geq 1$, $n \geq i_0$ and $m \geq 1$

$$\|\mathbf{K}_{k:k+n+m} - \pi\| \leq d(1 + (1 - \alpha_0)^m)(1 - \alpha_0)^n + m\gamma_n + c\beta^m, \quad (24)$$

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

Proof. Note that $\sum_{i \geq 1} \alpha(K_i) = +\infty$ from Assumption (22). It follows from Assertion (a) of Theorem 4.1 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

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

Actually the sequence $(\mu_{k,k+n})_{n \geq k} \in \mathcal{E}^{\mathbb{N}}$ is provided by Theorem 3.1, from which we deduce the following inequality by using (22)

$$\forall k \geq 1, \forall n \geq i_0, \quad \Delta_{k,n} \leq d_{i_0} (1 - \alpha_0)^n \quad \text{with} \quad d_{i_0} := (1 - \alpha_0)^{-i_0}. \quad (25)$$

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

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

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 (21) and (25) we deduce that for all $k \geq 1$, $n \geq i_0$ and $m \geq 1$

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

Finally observe that

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

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 $m \geq 1$

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

from which we deduce that

$$\begin{aligned} \|\mathbf{K}_{k:k+n+m} - \pi\| &\leq \|\mathbf{K}_{k:k+n+m} - \mu_{k,k+n+m}\| + \|\mu_{k,k+n+m} - \pi\| \\ &\leq 2 d_{i_0} (1 - \alpha_0)^{n+m} + m \gamma_n + 2 d_{i_0} (1 - \alpha_0)^n + c \beta^m. \end{aligned} \quad (26)$$

This proves (24). Now let $\varepsilon > 0$. First fix $m_0 \geq 1$ such that $c \beta^{m_0} \leq \varepsilon/2$. Then

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

since $\alpha_0 \in (0, 1]$ and $\lim_n \gamma_n = 0$ by the assumption $\lim_n \|K_n - K\| = 0$. It follows that: $\forall k \geq 1, \forall q \geq m_0 + n_0, \|\mathbf{K}_{k:k+q} - \pi\| \leq \varepsilon$. This proves (23). \square

Remark 4.6 *If Assumption (22) of Corollary 4.1 is replaced with the following one*

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

then the results of Corollary 4.1 can be extended by considering suitable block-products. More precisely the proof of Corollary 4.1 applies to the sequence $(K'_j)_{j \geq 0}$ defined by $K'_j := K_{is+1:is+s}$ since K^s is strongly ergodic and $\lim_n \|K'_n - K^s\| = 0$.

Remark 4.7 *When $(K_n)_{n \geq 1}$ is a sequence of infinite stochastic matrices (i.e. $\mathbb{X} = \mathbb{N}$), Corollary 4.1 is proved in [HIV76] without assuming (22). The reason why Assumption (22) is not necessary in the case $\mathbb{X} = \mathbb{N}$ comes from the fact that the map $\alpha(\cdot)$ is lower semi-continuous on the metric space (\mathcal{K}, d) , with: $\forall (K, K') \in \mathcal{K}^2, d(K, K') = \|K - K'\|$. First assume that this property holds and let us explain why Assumption (22) is useless. If $\alpha := \alpha(K) > 0$, then Assumption (22) 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 Corollary 4.1 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 Corollary 4.1 can be applied to the sequence $(K'_j)_{j \geq 0}$ defined by $K'_j := K_{is+1:is+s}$.*

Finally prove that $K \mapsto \alpha(K)$ is lower semi-continuous on (\mathcal{K}, d) when $\mathbb{X} = \mathbb{N}$. Let $K = (k(i, j))_{(i, j) \in \mathbb{N}^2}$ and $K' = (k'(i, j))_{(i, j) \in \mathbb{N}^2}$. For every $(i, j) \in \mathbb{N}^2$ we have $|k(i, j) - k'(i, j)| \leq \|K - K'\|$. Thus, for every fixed $j \in \mathbb{N}$, we obtained that

$$|\inf_{i \in \mathbb{N}} k(i, j) - \inf_{i \in \mathbb{N}} k'(i, j)| \leq \|K - K'\|.$$

It follows that, for every $j \in \mathbb{N}$, the function $\phi_j : K \mapsto \inf_{i \in \mathbb{N}} k(i, j)$ is continuous on (\mathcal{K}, d) . We obtain that $\alpha(\cdot)$ is lower semi-continuous on (\mathcal{K}, d) since

$$\alpha(K) := \sum_{j \geq 0} \phi_j(K) = \sup_{n \geq 0} \sum_{j=0}^n \phi_j(K)$$

and since, for every $n \geq 0$, the function $\sum_{j=0}^n \phi_j$ is continuous on (\mathcal{K}, d) .

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 $\mathbf{K}_{k:n}(\omega) \geq a_{k:n}(\omega) \nu_{k:n}(\omega)$. In order for the previous proofs to be compatible with the present random context, we assume that such 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 (Ω, \mathcal{F}) to $[0, 1]$ and to \mathcal{P} respectively. The same is assumed for backward products.

Note that these assumptions hold in the finite state space case since every stochastic matrix K such that $\alpha(K) > 0$ satisfies $K \geq \alpha(K) \tilde{\mathfrak{A}}(K)$, where $\tilde{\mathfrak{A}}(K)$ is the stochastic matrix with identical rows defined by $\tilde{\mathfrak{A}}(K) = \alpha(K)^{-1} \mathfrak{A}(K)$ (see (16)). Consequently the above mentioned choices may be $a_{k:n} := \alpha(\mathbf{K}_{k:n})$ and $\nu_{k:n} = \alpha(\mathbf{K}_{k:n})^{-1} \mathfrak{A}(\mathbf{K}_{k:n})$, which clearly both define random variables. To the best of our knowledge, the next results in Corollaries 5.1 and 5.2 are new, even in the finite case.

Corollary 5.1 *If the r.v. K_j , $j \geq 1$, are pairwise independent and identically distributed, and if $\mathbb{P}(\alpha(K_1) > 0) > 0$, then the following assertions hold true.*

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

In the finite state space case, that is 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 (b) is a well-known result under

the following stronger assumptions, see for instance [CL94]: the r.v. K_j 's are independent and identically distributed (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 r.v. K_j 's are just assumed to be pairwise independent in Corollary 5.1, and that the assumption $\mathbb{P}(K_1 \in \mathcal{K}_d^*) > 0$ is more restrictive than our assumption $\mathbb{P}(\alpha(K_1) > 0) > 0$ (since $K_1 \in \mathcal{K}_d^* \Rightarrow \alpha(K_1) > 0$).

Corollary 5.1 follows from Theorems 4.1-4.2 and from the following lemma.

Lemma 5.1 *Under the assumptions of Corollary 5.1, we have $\sum_{j \geq 1} \alpha(K_j) = +\infty$ \mathbb{P} -a.s..*

Proof. From $\mathbb{P}(\alpha(K_1) > 0) > 0$ there exists $p \geq 1$ such that $\mathbb{P}(\alpha(K_1) \geq 1/p) > 0$. Thus $\sum_{j \geq 1} \mathbb{P}(\alpha(K_j) \geq 1/p) = +\infty$ since the K_j 's are i.i.d.. Let $\Omega_0 = \cap_{n \geq 1} \cup_{j \geq n} [\alpha(K_j) \geq 1/p]$. Since the events $[\alpha(K_j) \geq 1/p]$, $j \geq 1$, are pairwise independent, the second Borel-Cantelli lemma ensures that $\mathbb{P}(\Omega_0) = 1$. Finally we have: $\forall \omega \in \Omega_0$, $\sum_{j \geq 1} \alpha(K_j(\omega)) = +\infty$. \square

Alternative assumptions may be proposed to obtain that $\sum_{j \geq 1} \alpha(K_j) = +\infty$ \mathbb{P} -a.s.. For instance, if the sequence $(K_j)_{j \geq 1}$ is stationary, then so is $(\alpha(K_j))_{j \geq 1}$. If moreover the sequence $(\alpha(K_j))_{j \geq 1}$ is ergodic and if $\mathbb{P}(\alpha(K_1) > 0) > 0$, then the conclusions (a) and (b) of Corollary 5.1 hold true. This follows from the strong law of large numbers for ergodic stationary sequences, which implies that $\sum_{j \geq 1} \alpha(K_j) = +\infty$ \mathbb{P} -a.s. under the assumption $\mathbb{P}(\alpha(K_1) > 0) > 0$. Actually the use of Doeblin's coefficient seems to be quite relevant in this random context since, as shown in the next corollary, it provides necessary and sufficient conditions for the convergence (as previously stated) 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.).

Corollary 5.2 *If $(K_j)_{j \geq 1}$ is i.i.d., then the two following assertions are fulfilled:*

- (i) *Property (a) of Corollary 5.1 holds if, and only if, there exists an integer number $q \geq 1$ such that $\mathbb{P}(\alpha(\mathbf{K}_{1:q}) > 0) > 0$,*
- (ii) *Property (b) of Corollary 5.1 holds true if, and only if, there exists an integer number $q \geq 1$ such that $\mathbb{P}(\alpha(\mathbf{K}_{q:1}) > 0) > 0$,*

Proof. Assume that, for some $q \geq 1$, we have $\mathbb{P}(\alpha(\mathbf{K}_{1:q}) > 0) > 0$. For every $j \in \mathbb{N}^*$ define: $Q_j = \mathbf{K}_{q(j-1)+1:qj}$. Since the sequence $(Q_j)_{j \geq 1}$ is i.i.d. and since $\mathbb{P}(\alpha(Q_1) > 0) > 0$ by hypothesis, we deduce from Lemma 5.1 that $\sum_{j \geq 1} \alpha(Q_j) = +\infty$ \mathbb{P} -a.s.. Then Property (a) of Corollary 5.1 follows from Theorem 4.1. Now assume that $\forall q \geq 1$, $\alpha(\mathbf{K}_{1:q}) = 0$ \mathbb{P} -a.s.. Define $\Omega_0 := \cap_{q \geq 1} [\alpha(\mathbf{K}_{1:q}) = 0]$. Then $\mathbb{P}(\Omega_0) = 1$, and $\forall \omega \in \Omega_0$, $\forall q \geq 1$, $\alpha(\mathbf{K}_{1:q})(\omega) = 0$. From Assertion (c) of Theorem 4.1 it follows that Property (a) of Corollary 5.1 does not hold (in fact, for every $\omega \in \Omega_0$, the expected conclusion for the sequence $(\mathbf{K}_{k:n}(\omega))_n$ does not hold). Assertion (ii) can be proved similarly. \square

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. R. Chestnut and S. 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.
- [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. Vinograd. 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.
- [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.
- [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.
- [Sen73] E. Seneta. On the historical development of the theory of finite inhomogeneous Markov chains. *Proc. Cambridge Philos. Soc.*, 74:507–513, 1973.
- [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.