Basic Properties of the Projective Product with Application to Products of Column–Allowable Nonnegative Matrices *

François LE GLAND and Laurent MEVEL
IRISA / INRIA Rennes, France

corresponding author :

François LE GLAND
IRISA / INRIA, Campus de Beaulieu
35042 RENNES Cédex, France

e–mail :	`legland@irisa.fr`
telephone :	+33 2 99 84 73 62
Fax :	+33 2 99 84 71 71

Abstract

We use basic properties of the projective product, to obtain exponential bounds for the Lipschitz constant associated with the projective product of column–allowable nonnegative matrices. We obtain similar bounds for the associated linear tangent maps.

Keywords : projective product, Birkhoff contraction coefficient, cocycle, exponential forgetting, product of matrices.

*This work was partially supported by the Commission of the European Communities, under the SCIENCE project *System Identification*, project number SC1*–CT92–0779, and under the HCM project *Statistical Inference for Stochastic Processes*, project number CHRX–CT92–0078, by the French–Hungarian cooperation program BALATON, project number 96–036, and by the Army Research Office, under grant DAAH04–95–1–0164.

Contents

1 Introduction

The purpose of this paper is to obtain exponential bounds for the difference

$$\| M_{n,m} \cdot p - M_{n,m} \cdot p' \| ,$$

where $M_{n,m} = M_n \cdots M_m$ is a product of column–allowable nonnegative $N \times N$ matrices, p and p' are N–dimensional probability vectors, and $\cdot$ denotes the projective product in $\mathbf{R}^N$.

Similar bounds have been obtained in Kaijser [4, Lemma 6.2], and in Arapostathis and Marcus [1, Lemma 2.2], using earlier results in Furstenberg and Kesten [3] for product of positive matrices. Our main contribution is to improve these bounds, in a more general context where matrices are only primitive, and to consider related problems for the linear tangent equation.

Our motivation for this problem is the study of large time asymptotic properties in hidden Markov models with finite state space, and arbitrary observation conditional densities. Indeed, the dependence of the prediction filter p_n w.r.t. its initial condition $p_m = p$ can be expressed by means of the relation

$$p_n = M_{n,m} \cdot p \qquad \text{with} \qquad M_{n,m} = Q^* B(Y_n) \cdots Q^* B(Y_m) ,$$

where Q is the transition probability matrix of the unobserved Markov chain, and for any integer l such that $m \leq l \leq n$ the matrix $B(Y_l)$ is diagonal with positive diagonal entries. Using the bounds obtained in this paper it is possible to extend the results of Arapostathis and Marcus [1], i.e. to prove exponential forgetting properties for the prediction filter and its gradient w.r.t. some parameter, and to prove geometric ergodicity properties for the extended Markov chain whose components are the unobserved Markov chain, the observation sequence, the prediction filter and its gradient. Details can be found in the companion paper [5].

In Section 2, we recall the definition of the Birkhoff contraction coefficient, and we give some basic useful properties of the projective product. These results are used in Section 3 to obtain an exponential forgetting property for products of matrices. In Section 4, we give some first–order properties of the projective product, i.e. properties of an associated cocycle. These results are used in Section 5 to obtain linear tangent stability and exponential forgetting properties for products of matrices.

2 Contraction coefficients and basic properties of the projective product

Let $S = \{1, \cdots, N\}$, let $\mathcal{P}$ denote the set of N–dimensional probability vectors, and let $\| \cdot \|$ denote the L_1–norm, i.e. for any $u = (u^i)$ in $\mathbf{R}^N$

$$\|u\| = \sum_{i \in S} |u^i| .$$

We recall some results on coefficients of ergodicity, taken from Seneta [6] and [7, Chapter 3].

For any column–allowable nonnegative $N \times N$ matrix $M = (M^{j,i})$ (i.e. with at least one positive element in each column) the Birkhoff contraction coefficient $\tau_{\rm B}(M)$ is defined by

$$\tau_{\rm B}(M) = \frac{1 - \sqrt{\Phi(M)}}{1 + \sqrt{\Phi(M)}} \leq 1 - \sqrt{\Phi(M)} \; ,$$

where

- if each row of M is either zero–free or all–zero

$$\Phi(M) = \min_{i_1,i_2,j_1,j_2 \in S} \frac{M^{j_1,i_1} \; M^{j_2,i_2}}{M^{j_1,i_2} \; M^{j_2,i_1}} \; ,$$

  where the indices $j_1, j_2 \in S$ are restricted to the zero–free rows,

- otherwise (i.e. if there is a row with both zero and positive elements), $\Phi(M) = 0$,

see Seneta [7, Theorems 3.10 and 3.12], and Birkhoff [2, Chapter XVI]. The following *multiplicative* property holds : for any column–allowable nonnegative matrices M and L

$$\tau_{\rm B}(M \, L) \leq \tau_{\rm B}(M) \; \tau_{\rm B}(L) \; ,$$

and if in addition the matrix L is diagonal (with positive diagonal entries), then $\tau_{\rm B}(L) = 1$, and

$$\tau_{\rm B}(M \, L) = \tau_{\rm B}(L \, M) = \tau_{\rm B}(M) \; .$$

For any stochastic matrix $P = (P^{i,j})$, and any probability vectors $p, q \in \mathcal{P}$

$$\|P^* \, (p - q)\| \leq \tau_1(P) \; \|p - q\| \; ,$$

where the coefficient of ergodicity $\tau_1(P)$ is defined by

$$\tau_1(P) = \tfrac{1}{2} \max_{i,i' \in S} \sum_{j \in S} |P^{i,j} - P^{i',j}| \; ,$$

see Seneta [6]. Moreover $\tau_1(P) \leq \tau_{\rm B}(P^*)$, and if in addition the stochastic matrix P is allowable (i.e. both row– and column–allowable), then $\tau_1(P) \leq \tau_{\rm B}(P^*) = \tau_{\rm B}(P)$, see Seneta [7, Theorem 3.13].

For any column–allowable nonnegative $N \times N$ matrix $M = (M^{j,i})$ and for any $i \in S$, define the sum of the i–th column entries of M as

$$M^{\bullet,i} = \sum_{j \in S} M^{j,i} \; ,$$

and the ratio

$$\omega(M) = \frac{\max_{i \in S} M^{\bullet,i}}{\min_{i \in S} M^{\bullet,i}} \; .$$

Lemma 2.1 *Any column–allowable nonnegative matrix M can be factorized into the product of a column stochastic matrix P^*, and a diagonal matrix D with positive diagonal entries, i.e.*

$$M = P^* \, D \ ,$$

with the additional properties

$$\tau_{\rm B}(M) = \tau_{\rm B}(P^*) \qquad \textit{and} \qquad \omega(M) = \omega(D) \ .$$

Proof. For any $i,j \in S$

$$M^{j,i} = \frac{M^{j,i}}{M^{\bullet,i}} \, M^{\bullet,i} \ ,$$

hence the factorization $M = P^* \, D$ holds where $P^* = (P^{j,i})$ is the column stochastic matrix obtained by normalizing the columns of M, i.e.

$$P^{j,i} = \frac{M^{j,i}}{M^{\bullet,i}} \ ,$$

for any $i,j \in S$, and where $D = \text{diag}\,[d^1, \cdots, d^N]$ is the diagonal matrix whose diagonal entries are the normalization coefficients of the columns of M, i.e.

$$d^i = M^{\bullet,i} \ ,$$

for any $i \in S$. The two additional properties are obvious. □

For any column–allowable nonnegative $N \times N$ matrix M, and any $p \in \mathcal{P}$, define the projective product

$$M \cdot p = \frac{M \, p}{e^* \, M \, p} \ ,$$

with values in $\mathcal{P}$, where $e = (1, \cdots, 1)^*$ denotes the N–dimensional vector with all entries equal to 1.

Lemma 2.2 *For any column–allowable nonnegative $N \times N$ matrix M, and any $p, p' \in \mathcal{P}$*

$$\| M \cdot p - M \cdot p' \, \| \leq \omega(M) \; \tau_{\rm B}(M) \; \| p - p' \| \ ,$$

and

$$\| M \cdot p - M \cdot p' \, \| \leq 2 \; \tau_{\rm B}(M) \ .$$

Proof. Using the factorization of Lemma 2.1 yields

$$M \cdot p = (P^* \, D) \cdot p = P^* \; (D \cdot p) \ ,$$

for any $p \in \mathcal{P}$. Therefore

$$M \cdot p - M \cdot p' = P^* \; (D \cdot p - D \cdot p') \ ,$$

and

$$\| M \cdot p - M \cdot p' \, \| \leq \tau_{\rm B}(P^*) \; \| D \cdot p - D \cdot p' \, \| \ ,$$

for any $p, p' \in \mathcal{P}$. Recall that

$$\tau_{\rm B}(P^*) = \tau_{\rm B}(M) \qquad \text{and} \qquad \omega(D) = \omega(M) \ ,$$

hence using the estimates of Lemma A.2 finishes the proof. □

3 Application to products of matrices

For any integer $n \geq 0$, let $M_n = (M_n^{j,i})$ be a column–allowable nonnegative $N \times N$ matrix, and for any integers n, m such that $n \geq m$, define the forward product

$$M_{n,m} = (M_{n,m}^{j,i}) = M_n \cdots M_m \ .$$

Througout the paper, the convention $M_{n,n+1} = M_{m-1,m} = I$ will be used.

In this special case, the estimates of Lemma 2.2 can be reformulated as

Proposition 3.1 *For any $p, p' \in \mathcal{P}$, and any integers n, m such that $n \geq m$*

$$\| M_{n,m} \cdot p - M_{n,m} \cdot p' \| \leq \omega(M_{n,m}) \ \tau_{\rm B}(M_{n,m}) \ \| p - p' \| \ ,$$

and

$$\| M_{n,m} \cdot p - M_{n,m} \cdot p' \| \leq 2 \ \tau_{\rm B}(M_{n,m}) \ .$$

Recall from Lemma 2.1 that for any integer $n \geq 0$, the matrix M_n can be factorized into the product of a column–stochastic matrix Q_n^*, and a diagonal matrix B_n with positive entries, i.e.

$$M_n = Q_n^* \, B_n \ .$$

In order to apply the results of Proposition 3.1, we need estimates on the constants $\omega(M_{n,m})$ and $\tau_{\rm B}(M_{n,m})$, which we will obtain in the following important special case, where the stochastic matrix $Q_n = Q$ is time–independent, and *primitive.*

Special case : $M_n = Q^* \, B_n$ with Q primitive

For any integer $n \geq 0$, assume that $M_n = (M_n^{j,i}) = Q^* \, B_n$, where $Q = (q^{i,j})$ is a primitive (or equivalently, irreducible and aperiodic) $N \times N$ stochastic matrix, and $B_n = \mathrm{diag}\,[b_n^1, \cdots, b_n^N]$ is a diagonal $N \times N$ matrix with positive diagonal entries. Define

$$1 \leq \delta_n = \frac{\displaystyle\max_{i \in S} b_n^i}{\displaystyle\min_{i \in S} b_n^i} < \infty \qquad \text{and} \qquad \varepsilon = \min_{i,j \in S}{}^{+} \, q^{i,j} > 0 \ ,$$

where the notation $\min^+$ denotes the minimum over positive elements. Let r denote the index of primitivity of Q, i.e. the smallest integer r such that the stochastic matrix Q^r is positive.

Lemma 3.2 *For any integers n, m such that $n \geq m$*

$$\omega(M_{n,m}) \leq \delta_m \cdots \delta_n \ .$$

If $n \geq m + r - 1$, then

$$\omega(M_{n,m}) \leq \varepsilon^{-r} \ \delta_m \cdots \delta_{m+r-1} \ .$$

Remark 3.3 These estimates can be merged into the following single estimate

$$\omega(M_{n,m}) \leq \varepsilon^{-r}\ \delta_m \cdots \delta_{\min(m+r-1,n)}\ ,$$

which holds for any integers n, m such that $n \geq m$.

PROOF. In the general case, for any $i, i' \in S$

$$\frac{M^{\bullet,i}_{n,m}}{M^{\bullet,i'}_{n,m}} = \frac{\displaystyle\sum_{j\in S} M^{j,i}_{n,m}}{\displaystyle\sum_{j\in S} M^{j,i'}_{n,m}} = \frac{\displaystyle\sum_{i_{m+1},\cdots,i_n,j\in S} q^{i,i_{m+1}}\, b^i_m \cdots q^{i_n,j}\, b^{i_n}_n}{\displaystyle\sum_{i_{m+1},\cdots,i_n,j\in S} q^{i',i_{m+1}}\, b^{i'}_m \cdots q^{i_n,j}\, b^{i_n}_n}$$

$$\leq\ \delta_m \cdots \delta_n\ \frac{\displaystyle\sum_{i_{m+1},\cdots,i_n,j\in S} q^{i,i_{m+1}} \cdots q^{i_n,j}}{\displaystyle\sum_{i_{m+1},\cdots,i_n,j\in S} q^{i',i_{m+1}} \cdots q^{i_n,j}} = \delta_m \cdots \delta_n\ .$$

In the particular case where $n \geq m + r - 1$, for any $i, i' \in S$

$$\frac{M^{\bullet,i}_{n,m}}{M^{\bullet,i'}_{n,m}} = \frac{\displaystyle\sum_{j\in S} M^{\bullet,j}_{n,m+r}\ M^{j,i}_{m+r-1,m}}{\displaystyle\sum_{j\in S} M^{\bullet,j}_{n,m+r}\ M^{j,i'}_{m+r-1,m}} \leq \max_{j\in S}\ \frac{M^{j,i}_{m+r-1,m}}{M^{j,i'}_{m+r-1,m}}$$

$$\leq \max_{j\in S}\ \frac{\displaystyle\sum_{i_{m+1},\cdots,i_{m+r-1}\in S} q^{i,i_{m+1}}\, b^i_m \cdots q^{i_{m+r-1},j}\, b^{i_{m+r-1}}_{m+r-1}}{\displaystyle\sum_{i_{m+1},\cdots,i_{m+r-1}\in S} q^{i',i_{m+1}}\, b^{i'}_m \cdots q^{i_{m+r-1},j}\, b^{i_{m+r-1}}_{m+r-1}}$$

$$\leq\ \delta_m \cdots \delta_{m+r-1}\ \max_{j\in S}\ \frac{\displaystyle\sum_{i_{m+1},\cdots,i_{m+r-1}\in S} q^{i,i_{m+1}} \cdots q^{i_{m+r-1},j}}{\displaystyle\sum_{i_{m+1},\cdots,i_{m+r-1}\in S} q^{i',i_{m+1}} \cdots q^{i_{m+r-1},j}}$$

$$\leq\ \varepsilon^{-r}\ \delta_m \cdots \delta_{m+r-1}$$

(if $m = n - r + 1$, then $M_{n,m+r} = M_{n,n+1} = I$ by convention). □

Lemma 3.4 *For any integers n, m such that $n \geq m + r - 1$*

$$\tau_{\mathrm{B}}(M_{n,m}) \leq \prod_{\kappa=1}^{[n,m]} (1 - \varepsilon^r\ [\,\delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}\,]^{-1})\ ,$$

where $[n,m] = \lfloor \dfrac{n-m+1}{r} \rfloor$ is the maximum number of disjoint blocks of length r in the set $\{m, \cdots, n\}$.

PROOF. Any product of at least r matrices is positive, and its Birkhoff contraction coefficient is stictly less than 1. In particular

$$\tau_{\rm B}(M_{m+r-1,m}) = \tau_{\rm B}(M_{m+r-1,m+1}\, Q^*\, B_m) = \tau_{\rm B}(M_{m+r-1,m+1}\, Q^*) \ ,$$

and computations similar to those in the proof of Lemma 3.2 yield

$$\sqrt{\Phi(M_{m+r-1,m+1}\, Q^*)} \geq \varepsilon^r\, [\, \delta_{m+1} \cdots \delta_{m+r-1}\,]^{-1} \ ,$$

hence

$$\tau_{\rm B}(M_{m+r-1,m}) \leq \frac{1-\varepsilon^r\, [\, \delta_{m+1} \cdots \delta_{m+r-1}\,]^{-1}}{1+\varepsilon^r\, [\, \delta_{m+1} \cdots \delta_{m+r-1}\,]^{-1}} \leq 1-\varepsilon^r\, [\, \delta_{m+1} \cdots \delta_{m+r-1}\,]^{-1} \ ,$$

for any integer $m \geq 0$, and the following lower bound holds

$$1-\varepsilon^r\, [\, \delta_{m+1} \cdots \delta_{m+r-1}\,]^{-1} \geq 1-\varepsilon^r \ . \tag{1}$$

Splitting the set $\{m, \cdots, n\}$ into $[n,m]$ disjoint consecutive blocks of length r yields

$$M_{n,m} = M_{n,m+[n,m]r}\, [\, \prod_{\kappa=1}^{[n,m]} M_{m+\kappa r-1,m+(\kappa-1)r}\,] \ ,$$

and the multiplicative property of the Birkhoff contraction coefficient yields

$$\begin{aligned}\tau_{\rm B}(M_{n,m}) \ &\leq \ \prod_{\kappa=1}^{[n,m]} \tau_{\rm B}(M_{m+\kappa r-1,m+(\kappa-1)r}) \\ &\leq \ \prod_{\kappa=1}^{[n,m]} (1-\varepsilon^r\, [\, \delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}\,]^{-1}\,) \ . \qquad \square\end{aligned}$$

In this special case, the estimates of Proposition 3.1 become

Theorem 3.5 *If the stochastic matrix Q is primitive, with index of primitivity r, then for any $p, p' \in \mathcal{P}$, and any integers n, m such that $n \geq m+r-1$*

$$\begin{aligned}&\| M_{n,m} \cdot p - M_{n,m} \cdot p'\, \| \\ &\qquad \leq \varepsilon^{-r}\, \delta_m \cdots \delta_{m+r-1} \prod_{\kappa=1}^{[n,m]} (1-\varepsilon^r\, [\, \delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}\,]^{-1})\ \|p-p'\| \ ,\end{aligned} \tag{2}$$

and

$$\| M_{n,m} \cdot p - M_{n,m} \cdot p'\, \| \leq 2 \prod_{\kappa=1}^{[n,m]} (1-\varepsilon^r\, [\, \delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}\,]^{-1}) \ ,$$

where $[n,m] = \lfloor \dfrac{n-m+1}{r} \rfloor$ is the maximum number of disjoint blocks of length r in the set $\{m, \cdots, n\}$.

Remark 3.6 If $r=1$, i.e. if the stochastic matrix Q is positive, then the estimate (2) reduces to

$$\| M_{n,m} \cdot p - M_{n,m} \cdot p'\, \| \leq \varepsilon^{-1}\, \delta_m\, (1-\varepsilon)^{n-m+1}\, \|p-p'\| \ ,$$

which improves estimates obtained in Kaijser [4, Lemma 6.2], and in Arapostathis and Marcus [1, Lemma 2.2], using earlier results in Furstenberg and Kesten [3].

4 First–order properties of the projective product

Let $\Sigma = \{w \in \mathbf{R}^N \, : \, e^* \, w = 0\}$ denote the linear tangent space to $\mathcal{P}$.

For any $p \in \mathcal{P}$, and any column–allowable nonnegative $N \times N$ matrix M

$$V[M,p] = [I - \frac{M \, p \, e^*}{e^* \, M \, p}] \, \frac{M}{e^* \, M \, p} = [I - (M \cdot p) \, e^* \,] \, \frac{M}{e^* \, M \, p} \ ,$$

is the Jacobian matrix at the point $p \in \mathcal{P}$ of the mapping

$$p \mapsto \frac{M \, p}{e^* \, M \, p} = M \cdot p \ .$$

Notice that $V[I,p] = [I - p \, e^*]$ for any $p \in \mathcal{P}$, hence $V[I,p] \, u = u$ for any $u \in \Sigma$. The following *cocycle* property holds.

Lemma 4.1 *For any $p \in \mathcal{P}$, and any column–allowable nonnegative $N \times N$ matrices M and L*

$$V[M, L \cdot p] \, V[L,p] = V[M \, L, p] \ . \tag{3}$$

Proof. Notice that

$$V[M,p] \, p = [I - \frac{M \, p \, e^*}{e^* \, M \, p}] \, \frac{M \, p}{e^* \, M \, p} = 0 \ ,$$

for any $p \in \mathcal{P}$, hence

$$\begin{aligned} V[M, L \cdot p] \, V[L,p] &= V[M, L \cdot p] \, [I - (L \cdot p) \, e^* \,] \, \frac{L}{e^* \, L \, p} \\ &= V[M, L \cdot p] \, \frac{L}{e^* \, L \, p} \\ &= [I - (M \cdot (L \cdot p)) \, e^* \,] \, \frac{M}{e^* \, M \, (L \cdot p)} \, \frac{L}{e^* \, L \, p} \\ &= [I - ((M \, L) \cdot p) \, e^* \,] \, \frac{M \, L}{e^* \, M \, L \, p} \\ &= V[M \, L, p] \ . \qquad \square \end{aligned}$$

Remark 4.2 An alternate, more direct and less algebraic proof of Lemma 4.1 can be given, by observing that $V[M \, L, p]$ is the Jacobian matrix at the point $p \in \mathcal{P}$ of the mapping

$$p \mapsto (M \, L) \cdot p = M \cdot (L \cdot p) \ ,$$

and by using the chain rule for the derivative of the composition of two mappings.

Lemma 4.3 *For any $p \in \mathcal{P}$, any $u \in \Sigma$, and any column–allowable nonnegative $N \times N$ matrix M*

$$\| \, V[M,p] \, u \| \leq \omega(M) \ \tau_{\mathrm{B}}(M) \ \|u\| \ .$$

PROOF. For any $p \in \mathcal{P}$, using the factorization of Lemma 2.1 yields

$$V[M,p] = V[P^* D, p] = P^* V[D,p] \ ,$$

and notice also that $e^* V[D,p] = 0$, hence $V[D,p]\, u \in \Sigma$ for any $u \in \mathbf{R}^N$. Therefore

$$\| V[M,p]\, u\| \leq \tau_{\mathrm{B}}(P^*)\; \| V[D,p]\, u\| \ ,$$

for any $u \in \Sigma$, and using the second estimate of Lemma A.1 finishes the proof. □

Remark 4.4 Since $V[M,p]$ is the Jacobian matrix at the point $p \in \mathcal{P}$ of the mapping

$$p \mapsto M \cdot p \ ,$$

the Lipschitz estimate of Lemma 2.2 can be used alternatively to prove Lemma 4.3.

Lemma 4.5 *For any $p, p' \in \mathcal{P}$, any $u \in \Sigma$, and any column–allowable nonnegative $N \times N$ matrices M and L*

$$\| \,(\, V[M, L\cdot p] - V[M, L\cdot p']\,)\; u\|$$

$$\leq \tfrac{1}{2}\, [\omega(M) - 1]\; [\omega(M\, L) + \omega(M)\, \omega(L)\,]\; \tau_{\mathrm{B}}(M)\; \tau_{\mathrm{B}}(L)\; \|p - p'\|\; \|u\| \ ,$$

and

$$\| \,(\, V[M, L\cdot p] - V[M, L\cdot p']\,)\; u\| \leq [\omega^2(M) - 1]\; \tau_{\mathrm{B}}(M)\; \tau_{\mathrm{B}}(L)\; \|u\| \ .$$

PROOF. Simple algebra yields

$$\begin{aligned}
V[M,p] - V[M,p'] &= [I - \frac{M\,p\,e^*}{e^*\,M\,p}]\; \frac{M}{e^*\,M\,p} - [I - \frac{M\,p'\,e^*}{e^*\,M\,p'}]\; \frac{M}{e^*\,M\,p'} \\
&= [\, \frac{M\,p'\,e^*}{e^*\,M\,p'} - \frac{M\,p\,e^*}{e^*\,M\,p}\,]\; \frac{M}{e^*\,M\,p} \\
&\qquad + [I - \frac{M\,p'\,e^*}{e^*\,M\,p'}]\; [\, \frac{M}{e^*\,M\,p} - \frac{M}{e^*\,M\,p'}\,] \\
&= [\, \frac{M\,p'}{e^*\,M\,p'} - \frac{M\,p}{e^*\,M\,p}\,]\; \frac{e^*\,M}{e^*\,M\,p} + \frac{e^*\,M\,(p' - p)}{e^*\,M\,p}\; V[M,p'] \\
&= [\, M\cdot p' - M\cdot p\,]\; \frac{e^*\,M}{e^*\,M\,p} + \frac{e^*\,M\,(p' - p)}{e^*\,M\,p}\; V[M,p'] \ ,
\end{aligned}$$

for any $p, p' \in \mathcal{P}$, hence

$$\begin{aligned}
&V[M, L\cdot p] - V[M, L\cdot p'] \\
&= [\, M\cdot(L\cdot p') - M\cdot(L\cdot p)\,]\; \frac{e^*\,M}{e^*\,M\,(L\cdot p)} + \frac{e^*\,M\,(L\cdot p' - L\cdot p)}{e^*\,M\,(L\cdot p)}\; V[M, L\cdot p'] \\
&= [\,(M\;L)\cdot p' - (M\;L)\cdot p\,]\; \frac{e^*\,M}{e^*\,M\,(L\cdot p)} + \frac{e^*\,M\,(L\cdot p' - L\cdot p)}{e^*\,M\,(L\cdot p)}\; V[M, L\cdot p'] \ .
\end{aligned}$$

If the first (Lipschitz) estimate in Lemma 2.2 is used, then the estimates in Lemmas 4.3 and A.3 yield

$$\|\,(\,V[M, L\cdot p] - V[M, L\cdot p']\,)\,u\|$$

$$\leq \omega(M\;L)\,\tau_{\rm B}(M\;L)\;\|p - p'\|\;\tfrac{1}{2}\,[\omega(M) - 1]\;\|u\|$$

$$+\;\tfrac{1}{2}\,[\omega(M) - 1]\;\omega(L)\,\tau_{\rm B}(L)\;\|p - p'\|\;\omega(M)\,\tau_{\rm B}(M)\;\|u\|$$

$$\leq \tfrac{1}{2}\,[\omega(M) - 1]\;[\omega(M\;L) + \omega(M)\,\omega(L)\,]\;\tau_{\rm B}(M)\;\tau_{\rm B}(L)\;\|p - p'\|\;\|u\|\ .$$

If the second (rough) estimate in Lemma 2.2 is used instead, then

$$\|\,(\,V[M, L\cdot p] - V[M, L\cdot p']\,)\,u\|$$

$$\leq 2\,\tau_{\rm B}(M\;L)\;\tfrac{1}{2}\,[\omega(M) - 1]\;\|u\| + \tfrac{1}{2}\,[\omega(M) - 1]\;2\,\tau_{\rm B}(L)\;\omega(M)\,\tau_{\rm B}(M)\;\|u\|$$

$$\leq [\omega(M) - 1]\;[1 + \omega(M)]\;\tau_{\rm B}(M)\;\tau_{\rm B}(L)\;\|u\|$$

$$= [\omega^2(M) - 1]\;\tau_{\rm B}(M)\;\tau_{\rm B}(L)\;\|u\|\ . \qquad \Box$$

5 Application to products of matrices : linear tangent equations

We consider simultaneously the following two equations for $\{p_n\ ,\ n \geq 0\}$ and $\{w_n\ ,\ n \geq 0\}$

$$p_{n+1} = \frac{M_n\,p_n}{e^*\,M_n\,p_n} = M_n \cdot p_n\ ,$$

and

$$w_{n+1} = V[M_n, p_n]\;w_n + u_n[p_n]\ ,$$

for any integer $n \geq 0$, where

$$e^*\,u_n[p] = 0\ ,$$

for any $p \in \mathcal{P}$. We assume that for any integer $n \geq 0$, the mapping $p \mapsto u_n[p]$ with values in Σ, is Lipschitz continuous (hence bounded since $\mathcal{P}$ is compact), i.e.

$$\|u_n[p] - u_n[p']\,\| \leq {\rm Lip}(u_n)\;\|p - p'\|\ ,$$

for any $p, p' \in \mathcal{P}$, and by definition

$${\rm K}(u_n) = \sup_{p\in\mathcal{P}} \|u_n[p]\,\|\ .$$

Iterating the cocycle identity (3) yields

$$\prod_{k=l}^{n} V[M_k, M_{k-1,l}\cdot p] = V[M_n, M_{n-1,l}\cdot p]\cdots V[M_l, p] = V[M_{n,l}, p]\ ,$$

for any $p \in \mathcal{P}$, and any integers n, l such that $n \geq l$, hence

$$\prod_{k=l}^{n} V[M_k, p_k] = \prod_{k=l}^{n} V[M_k, M_{k-1,l}\cdot p_l] = V[M_{n,l}, p_l]\ .$$

Therefore, the *variation of parameters* formula yields

$$\begin{aligned} w_{n+1} &= \prod_{k=m}^{n} V[M_k, p_k]\ w_m + u_n[p_n] + \sum_{l=m}^{n-1}\ \prod_{k=l+1}^{n} V[M_k, p_k]\ u_l[p_l] \\ &= V[M_{n,m}, p_m]\ w_m + \sum_{l=m}^{n} V[M_{n,l+1}, p_{l+1}]\ u_l[p_l] \\ &= V[M_{n,m}, p_m]\ w_m + \sum_{l=m}^{n} V[M_{n,l+1}, M_{l,m}\cdot p_m]\ u_l[M_{l-1,m}\cdot p_m] \\ &= W_{n,m}[u_n, \cdots, u_m, p_m, w_m]\ , \end{aligned}$$

with the convention $M_{n,n+1} = M_{m-1,m} = I$.

Using the estimate of Lemma 4.3, it is immediate to obtain the following *stability* property.

Proposition 5.1 *For any $p \in \mathcal{P}$, any $w \in \Sigma$, and any integers n, m such that $n \geq m$*

$$\|W_{n,m}[u_n, \cdots, u_m, p, w]\,\| \leq \omega(M_{n,m})\,\tau_{\mathrm{B}}(M_{n,m})\ \|w\| + \sum_{l=m}^{n} \mathrm{K}(u_l)\ \omega(M_{n,l+1})\,\tau_{\mathrm{B}}(M_{n,l+1})\ ,$$

with the convention $M_{n,n+1} = I$, hence the rough estimate

$$\|W_{n,m}[u_n, \cdots, u_m, p, w]\,\| \leq \omega(M_{n,m})\ \|w\| + \sum_{l=m}^{n} \mathrm{K}(u_l)\ \omega(M_{n,l+1})\ .$$

For any $p, p' \in \mathcal{P}$, any $w, w' \in \Sigma$, the following decomposition holds

$$
\begin{aligned}
& W_{n,m}[u_n, \cdots, u_m, p, w] - W_{n,m}[u_n, \cdots, u_m, p', w'] \\
& \quad = V[M_{n,m}, p]\; w + \sum_{l=m}^{n} V[M_{n,l+1}, M_{l,m} \cdot p]\; u_l[M_{l-1,m} \cdot p] \\
& \qquad - V[M_{n,m}, p']\; w' - \sum_{l=m}^{n} V[M_{n,l+1}, M_{l,m} \cdot p']\; u_l[M_{l-1,m} \cdot p'] \\
& \quad = \tfrac{1}{2} \left(V[M_{n,m}, p] + V[M_{n,m}, p'] \right) (w - w') \\
& \qquad + \tfrac{1}{2} \left(V[M_{n,m}, p] - V[M_{n,m}, p'] \right) (w + w') \\
& \qquad + \sum_{l=m}^{n} V[M_{n,l+1}, M_{l,m} \cdot p] \left(u_l[M_{l-1,m} \cdot p] - u_l[M_{l-1,m} \cdot p'] \right) \\
& \qquad + \sum_{l=m}^{n-1} \left(V[M_{n,l+1}, M_{l,m} \cdot p] - V[M_{n,l+1}, M_{l,m} \cdot p'] \right) u_l[M_{l-1,m} \cdot p'] \;,
\end{aligned}
\tag{4}
$$

with the convention $M_{n,n+1} = M_{m-1,m} = I$.

Proposition 5.2 *For any $p, p' \in \mathcal{P}$, any $w, w' \in \Sigma$, and any integers n, m such that $n \geq m$,*

$$
\begin{aligned}
& \| W_{n,m}[u_n, \cdots, u_m, p, w] - W_{n,m}[u_n, \cdots, u_m, p', w'] \| \\
& \quad \leq \tfrac{1}{2} [\omega^2(M_{n,m}) + 1]\; \tau_{\mathrm{B}}(M_{n,m}) \left[\|w - w'\| + \|p - p'\| \left(\|w\| + \|w'\| \right) \right] \\
& \qquad + \Big[\sum_{l=m}^{n} \mathrm{Lip}(u_l)\; \omega(M_{n,l+1})\, \omega(M_{l-1,m})\, \tau_{\mathrm{B}}(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{l-1,m}) \\
& \qquad\qquad + \sum_{l=m}^{n-1} \mathrm{K}(u_l)\; \tfrac{1}{2} [\omega(M_{n,l+1}) - 1]\; [\omega(M_{n,m}) + \omega(M_{n,l+1})\, \omega(M_{l,m})] \\
& \qquad\qquad\qquad \tau_{\mathrm{B}}(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{l,m}) \Big] \;\; \|p - p'\| \;,
\end{aligned}
$$

and

$$
\begin{aligned}
& \| W_{n,m}[u_n, \cdots, u_m, p, w] - W_{n,m}[u_n, \cdots, u_m, p', w'] \| \\
& \quad \leq \tfrac{1}{2} [\omega^2(M_{n,m}) + 1]\; \tau_{\mathrm{B}}(M_{n,m}) \left(\|w\| + \|w'\| \right)
\end{aligned}
$$

$$+\, 2 \sum_{l=m}^{n} \mathrm{Lip}(u_l)\, \omega(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{l-1,m})$$

$$+ \sum_{l=m}^{n-1} \mathrm{K}(u_l)\, [\omega^2(M_{n,l+1}) - 1]\, \tau_{\mathrm{B}}(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{l,m}) \ ,$$

with the convention $M_{n,n+1} = M_{m-1,m} = I$.

The proof of Proposition 5.2 is given in Appendix B.

Special case (continued) : $M_n = Q^* B_n$ with Q primitive

In order to apply the results of Proposition 5.2 in this special case, we need estimates on the product $\tau_{\mathrm{B}}(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{l-1,m})$.

Lemma 5.3 *For any integers n, l, m such that $n \geq m + r - 1$, and $m \leq l \leq n$*

$$\tau_{\mathrm{B}}(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{l-1,m}) \leq (1 - \varepsilon^r)^{-1} \prod_{\kappa=1}^{[n,m]} (1 - \varepsilon^r\, [\, \delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}\,]^{-1}\,) \ .$$

Proof. For any integer l such that $m + 1 \leq l \leq n - 1$, let κ_0 be uniquely determined by the condition

$$m + (\kappa_0 - 1)r \leq l \leq m + \kappa_0 r - 1 \ ,$$

i.e. that l belongs to the κ_0–th block of length r in the set $\{m, \cdots, n\}$. Therefore

$$\tau_{\mathrm{B}}(M_{l-1,m}) \leq \tau_{\mathrm{B}}(M_{m+(\kappa_0-1)r-1,m}) \leq \prod_{\kappa=1}^{\kappa_0-1} \tau_{\mathrm{B}}(M_{m+\kappa r-1, m+(\kappa-1)r}) \ ,$$

and

$$\tau_{\mathrm{B}}(M_{n,l+1}) \leq \tau_{\mathrm{B}}(M_{n,m+\kappa_0 r}) \leq \prod_{\kappa=\kappa_0+1}^{[n,m]} \tau_{\mathrm{B}}(M_{m+\kappa r-1, m+(\kappa-1)r}) \ .$$

Combining these two estimates and using the lower bound (1) yields

$$\begin{aligned}
&\tau_{\mathrm{B}}(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{l-1,m}) \\
&\leq \prod_{\kappa=1}^{\kappa_0-1} \tau_{\mathrm{B}}(M_{m+\kappa r-1, m+(\kappa-1)r}) \prod_{\kappa=\kappa_0+1}^{[n,m]} \tau_{\mathrm{B}}(M_{m+\kappa r-1, m+(\kappa-1)r}) \\
&\leq \prod_{\kappa=1}^{\kappa_0-1} (1 - \varepsilon^r\, [\, \delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}\,]^{-1}\,) \prod_{\kappa=\kappa_0+1}^{[n,m]} (1 - \varepsilon^r\, [\, \delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}\,]^{-1}\,) \\
&\leq (1 - \varepsilon^r)^{-1} \prod_{\kappa=1}^{[n,m]} (1 - \varepsilon^r\, [\, \delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}\,]^{-1}\,) \ .
\end{aligned}$$

If $l = m$, then

$$\tau_{\mathrm{B}}(M_{n,m+1}) \leq \tau_{\mathrm{B}}(M_{n,m+r}) \leq \prod_{\kappa=2}^{[n,m]} (1 - \varepsilon^r [\, \delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}]^{-1})$$

$$\leq (1-\varepsilon^r)^{-1} \prod_{\kappa=1}^{[n,m]} (1 - \varepsilon^r [\, \delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}]^{-1}) \ ,$$

and if $l = n$, then

$$\tau_{\mathrm{B}}(M_{n-1,m}) \leq \tau_{\mathrm{B}}(M_{m+([n,m]-1)r-1,m}) \leq \prod_{\kappa=1}^{[n,m]-1} (1 - \varepsilon^r [\, \delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}]^{-1})$$

$$\leq (1-\varepsilon^r)^{-1} \prod_{\kappa=1}^{[n,m]} (1 - \varepsilon^r [\, \delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}]^{-1}) \ .$$

□

Remark 5.4 For any integers n, m such that $n \geq m + 2r - 1$, the following estimates hold depending on the position of l relatively to m and n :

- If $m \leq l \leq n - 2\,r$, there are at least two disjoint blocks of length r in the set $\{l+1, \cdots, n\}$, and
$$\tau_{\mathrm{B}}(M_{n,l+1}) \leq \prod_{\kappa=2}^{[n,l+1]} (1 - \varepsilon^r [\, \delta_{l+(\kappa-1)r+2} \cdots \delta_{l+\kappa r+1}]^{-1}) \ ,$$
otherwise if $n - 2\,r + 1 \leq l \leq n$
$$\tau_{\mathrm{B}}(M_{n,l+1}) \leq 1 \ .$$

- If $m + r \leq l \leq n$, there is at least one block of length r in the set $\{m, \cdots, l-1\}$, and
$$\tau_{\mathrm{B}}(M_{l-1,m}) \leq \prod_{\kappa=1}^{[l-1,m]} (1 - \varepsilon^r [\, \delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r}]^{-1}) \ ,$$
otherwise, if $m \leq l \leq m + r - 1$
$$\tau_{\mathrm{B}}(M_{l-1,m}) \leq 1 \ .$$

In this special case, the stability estimates of Proposition 5.1 become

Theorem 5.5 *If the stochastic matrix Q is primitive, with index of primitivity r, then for any $p \in \mathcal{P}$, any $w \in \Sigma$, and any integers n, m such that $n \geq m + r - 1$*

$$\| W_{n,m}[u_n, \cdots, u_m, p, w] \, \|$$

$$\leq \varepsilon^{-r}\,\delta_m\cdots\delta_{m+r-1}\prod_{\kappa=1}^{[n,m]}(1-\varepsilon^r\,[\,\delta_{m+(\kappa-1)r+1}\cdots\delta_{m+\kappa r-1}\,]^{-1}\,)\;\|w\|$$

$$+\,\mathrm{K}(u_n)+\sum_{l=m}^{n-1}\varepsilon^{-r}\,\delta_{l+1}\cdots\delta_{\min(l+r,n)}\,\mathrm{K}(u_l)$$

$$\prod_{\kappa=1}^{[n,l+1]}(1-\varepsilon^r\,[\,\delta_{l+(\kappa-1)r+2}\cdots\delta_{l+\kappa r}\,]^{-1}\,)\ ,$$

with the convention that products over empty sets are equal to 1*, hence the rough estimate*

$$\|W_{n,m}[u_n,\cdots,u_m,p,w]\|\ \leq\ \varepsilon^{-r}\,\delta_m\cdots\delta_{\min(m+r-1,n)}\;\|w\|$$

$$+\,\mathrm{K}(u_n)+\sum_{l=m}^{n-1}\varepsilon^{-r}\,\delta_{l+1}\cdots\delta_{\min(l+r,n)}\,\mathrm{K}(u_l)\ .$$

It follows from the rough estimates in Propositions 3.1 and 5.1 that

$$\|M_{n,m}\cdot p'-M_{n,m}\cdot p''\|\;\|W_{n,m}[u_n,\cdots,u_m,p,w]\|$$

$$\leq 2\,\omega(M_{n,m})\,\tau_{\mathrm{B}}(M_{n,m})\;\|w\|+2\sum_{l=m}^{n}\mathrm{K}(u,y_l)\,\omega(M_{n,l+1})\,\tau_{\mathrm{B}}(M_{n,m}),$$

with the convention $M_{n,n+1}=I$, and using the bounds of Remark 5.4 yields the following estimate, which is used in the companion paper [5] on hidden Markov models.

Proposition 5.6 *If the stochastic matrix Q is primitive, with index of primitivity r, then for any $p,p',p''\in\mathcal{P}$, any $w\in\Sigma$, and any integers n,m such that $n\geq m+3r-1$*

$$\|M_{n,m}\cdot p'-M_{n,m}\cdot p''\|\;\|W_{n,m}[u_n,\cdots,u_m,p,w]\|$$

$$\leq 2\,\varepsilon^{-r}\,\delta_m\cdots\delta_{m+r-1}\prod_{\kappa=2}^{[n,m]}(1-\varepsilon^r\,[\,\delta_{m+(\kappa-1)r+1}\cdots\delta_{m+\kappa r-1}\,]^{-1}\,)\;\|w\|$$

$$+\,2\,\mathrm{K}(u_n)\prod_{\kappa=1}^{[n-1,m]}(1-\varepsilon^r\,[\,\delta_{m+(\kappa-1)r+1}\cdots\delta_{m+\kappa r-1}\,]^{-1}\,)$$

$$+\,2\sum_{l=m}^{n-1}\mathrm{K}(u_l)\,\varepsilon^{-r}\,\delta_{l+1}\cdots\delta_{\min(l+r,n)}$$

$$\prod_{\kappa=2}^{[n,l+1]}(1-\varepsilon^r\,[\,\delta_{l+(\kappa-1)r+2}\cdots\delta_{l+\kappa r}\,]^{-1}\,)$$

$$\prod_{\kappa=1}^{[l-1,m]}(1-\varepsilon^r\,[\,\delta_{m+(\kappa-1)r+1}\cdots\delta_{m+\kappa r-1}\,]^{-1}\,)\ ,$$

with the convention that products over empty sets are equal to 1.

Similarly, the estimates of Proposition 5.2 become

Theorem 5.7 *If the stochastic matrix Q is primitive, with index of primitivity r, then for any $p, p' \in \mathcal{P}$, any $w, w' \in \Sigma$, and any integers n, m such that $n \geq m + r - 1$*

$$\| W_{n,m}[u_n, \cdots, u_m, p, w] - W_{n,m}[u_n, \cdots, u_m, p', w'] \|$$

$$\leq \tfrac{1}{2} \, (\, [\, \varepsilon^{-r} \, \delta_m \cdots \delta_{m+r-1} \,]^2 + 1 \,) \prod_{\kappa=1}^{[n,m]} (1 - \varepsilon^r \, [\, \delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1} \,]^{-1} \,)$$

$$[\, \|w - w'\| + \|p - p'\| \, (\, \|w\| + \|w'\| \,) \,]$$

$$+ \, \varepsilon^{-r} \, \delta_m \cdots \delta_{m+r-1} \, \Big[\, \mathrm{Lip}(u_n) + \sum_{l=m}^{n-1} \varepsilon^{-r} \, \delta_{l+1} \cdots \delta_{\min(l+r,n)} \, \mathrm{Lip}(u_l)$$

$$+ \, \tfrac{1}{2} \sum_{l=m}^{n-1} (\, [\, \varepsilon^{-r} \, \delta_{l+1} \cdots \delta_{\min(l+r,n)} \,]^2 - 1 \,) \, \mathrm{K}(u_l) \, \Big]$$

$$(1 - \varepsilon^r)^{-1} \prod_{\kappa=1}^{[n,m]} (1 - \varepsilon^r \, [\, \delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1} \,]^{-1} \,) \, \|p - p'\| \ ,$$

Remark 5.8 If $r = 1$, i.e. if the stochastic matrix Q is positive, then the estimate in Theorem 5.7 reduces to

$$\| W_{n,m}[u_n, \cdots, u_m, p, w] - W_{n,m}[u_n, \cdots, u_m, p', w'] \|$$

$$\leq \tfrac{1}{2} \, (\, \varepsilon^{-2} \, \delta_m^2 + 1 \,) \, (1 - \varepsilon)^{n-m+1} \, [\, \|w - w'\| + \|p - p'\| \, (\, \|w\| + \|w'\| \,) \,]$$

$$+ \, \varepsilon^{-1} \, \delta_m \, \Big[\, \mathrm{Lip}(u_n) + \sum_{l=m}^{n-1} \varepsilon^{-1} \, \delta_{l+1} \, \mathrm{Lip}(u_l)$$

$$+ \, \tfrac{1}{2} \sum_{l=m}^{n-1} (\, \varepsilon^{-2} \, \delta_{l+1}^2 - 1 \,) \, \mathrm{K}(u_l) \, \Big] \, (1 - \varepsilon)^{n-m} \, \|p - p'\| \ .$$

Using the rough estimate of Proposition 5.2, and the bounds of Remark 5.4 yields the following estimate, which is used in the companion paper [5] on hidden Markov models.

Proposition 5.9 *If the stochastic matrix Q is primitive, with index of primitivity r, then for any $p, p' \in \mathcal{P}$, any $w, w' \in \Sigma$, and any integers n, m such that $n \geq m + 3r - 1$*

$$\| W_{n,m}[u_n, \cdots, u_m, p, w] - W_{n,m}[u_n, \cdots, u_m, p', w'] \|$$

$$
\begin{aligned}
&\leq \tfrac{1}{2}\,(\,[\varepsilon^{-r}\,\delta_m \cdots \delta_{m+r-1}\,]^2 + 1)\ \prod_{\kappa=2}^{[n,m]} (1 - \varepsilon^r\,[\,\delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}\,]^{-1}\,)\ (\,\|w\| + \|w'\|\,) \\
&\quad + 2\,\mathrm{Lip}(u_n)\ \prod_{\kappa=1}^{[n-1,m]}\ (1 - \varepsilon^r\,[\,\delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}\,]^{-1}\,) \\
&\quad + 2\sum_{l=m}^{n-1}\ \mathrm{Lip}(u_l)\ \varepsilon^{-r}\,\delta_{l+1} \cdots \delta_{\min(l+r,n)} \\
&\qquad\qquad \prod_{\kappa=2}^{[n,l+1]} (1 - \varepsilon^r\,[\,\delta_{l+(\kappa-1)r+2} \cdots \delta_{l+\kappa r}\,]^{-1}\,) \\
&\qquad\qquad\qquad \prod_{\kappa=1}^{[l-1,m]} (1 - \varepsilon^r\,[\,\delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}\,]^{-1}\,) \\
&\quad + \sum_{l=m}^{n-1}\ \mathrm{K}(u_l)\ (\,[\,\varepsilon^{-r}\,\delta_{l+1} \cdots \delta_{\min(l+r,n)}\,]^2 - 1) \\
&\qquad\qquad \prod_{\kappa=2}^{[n,l+1]} (1 - \varepsilon^r\,[\,\delta_{l+(\kappa-1)r+2} \cdots \delta_{l+\kappa r}\,]^{-1}\,) \\
&\qquad\qquad\qquad \prod_{\kappa=1}^{[l-1,m]} (1 - \varepsilon^r\,[\,\delta_{m+(\kappa-1)r+1} \cdots \delta_{m+\kappa r-1}\,]^{-1}\,)\ ,
\end{aligned}
$$

with the convention that products over empty sets are equal to 1.

Remark 5.10 The advantage of the estimate in Proposition 5.9 is that each term in the sum is a product of factors which span disjoint index sets. The same remark holds for the estimate in Proposition 5.6. In the context of products of matrices with Markovian dependence, it would be then straightforward to obtain estimates in mean absolute sense, see the companion paper [5]. On the other hand, the advantage of the estimate in Theorem 5.7 is that it yields immediately the exponential forgetting property in pathwise sense.

Acknowledgement

The authors gratefully acknowledge Jan van Schuppen for bringing the paper [1] to their attention, László Gerencsér and György Michaletzky for their contribution to the refined estimates of Appendix A, and Dan Ocone and Amarjit Budhiraja for very interesting discussion, in particular for their suggestion of using the Birkhoff contraction coefficient.

References

[1] A. ARAPOSTATHIS and S.I. MARCUS. Analysis of an identification algorithm arising in the adaptive estimation of Markov chains. *Mathematics of Control, Signals, and Systems*, 3(1):1–29, 1990.

[2] G. BIRKHOFF. *Lattice Theory*, volume XXV of *Colloquium Publications.* AMS, Providence, RI, 3rd edition, 1967.

[3] H. FURSTENBERG and H. KESTEN. Product of random matrices. *The Annals of Mathematical Statistics*, 31(2):457–469, June 1960.

[4] T. KAIJSER. A limit theorem for partially observed Markov chains. *The Annals of Probability*, 3(4):677–696, August 1975.

[5] F. LE GLAND and L. MEVEL. Exponential forgetting and geometric ergodicity in hidden Markov models. (submitted).

[6] E. SENETA. Coefficients of ergodicity : structure and applications. *Advances in Applied Probability*, 11:576–590, 1979.

[7] E. SENETA. *Non–negative Matrices and Markov Chains.* Springer Series in Statistics. Springer Verlag, New York, 2nd edition, 1981.

A Rough and refined estimates

Lemma A.1 *For any $p \in \mathcal{P}$, and any $N \times N$ diagonal matrix D with positive diagonal entries, the following two estimates hold : (i) for any $u \in \mathbf{R}^N$*

$$\| \, [I - \frac{D\,p\,e^*}{e^*\,D\,p}] \; u \, \| \leq 2 \, \|u\| \ , \tag{5}$$

and (ii) for any $u \in \Sigma$

$$\| \, [I - \frac{D\,p\,e^*}{e^*\,D\,p}] \; \frac{D\,u}{e^*\,D\,p} \, \| \leq \omega(D) \, \|u\| \ . \tag{6}$$

Proof. For any $u \in \mathbf{R}^N$

$$\|D\,p\,e^*\,u\| = \|D\,p\| \; |e^*\,u| \leq \|D\,p\| \; \|u\| \ ,$$

and

$$e^*\,D\,p = \|D\,p\| \ .$$

Therefore

$$\|\frac{D\,p\,e^*}{e^*\,D\,p} \; u\| \leq \|u\| \ ,$$

which proves the rough estimate (5).

Using Lemma 2.4 in Seneta [7] — which simply exploits that any point in a compact convex subset of $\mathbf{R}^N$ is a linear convex combination of extremal points — any $u \in \Sigma$ such that $\|u\| = 1$ can be represented as

$$u = \sum_{i,j \in S} K^{i,j}(u)\, f_{i,j} \;, \tag{7}$$

where $(K^{i,j}(u))$ is a probability distribution on $S \times S$, and $f_{i,j} = \frac{1}{2}\,(e_i - e_j)$ for any $i,j \in S$. Therefore, introducing the following $N \times N$ matrix

$$V = V[D,p] = [I - \frac{D\,p\,e^*}{e^*\,D\,p}]\,\frac{D}{e^*\,D\,p} \;,$$

it holds

$$\max_{u \in \Sigma\,:\,\|u\|=1} \|V\,u\| \le \max_{i,j \in S} \|V\,f_{i,j}\| \;.$$

Notice that

$$V\,e_i = [I - \frac{D\,p\,e^*}{e^*\,D\,p}]\,\frac{d^i\,e_i}{e^*\,D\,p} = [e_i - \frac{D\,p}{e^*\,D\,p}]\,\frac{d^i}{e^*\,D\,p} \;,$$

hence

$$V\,f_{i,j} = \tfrac{1}{2}\,\frac{1}{e^*\,D\,p}\,[d^i\,e_i - d^j\,e_j - (d^i - d^j)\,\frac{D\,p}{e^*\,D\,p}] \;,$$

and

$$\|V\,f_{i,j}\| = \tfrac{1}{2}\,\frac{1}{e^*\,D\,p}\,[\,|d^i - \frac{(d^i - d^j)\,d^i\,p^i}{e^*\,D\,p}| + |d^j + \frac{(d^i - d^j)\,d^j\,p^j}{e^*\,D\,p}| + |d^i - d^j| \sum_{k \in S\,:\,k \neq i\,,\,k \neq j} \frac{d^k\,p^k}{e^*\,D\,p}] \;.$$

Introducing

$$s = \sum_{k \in S\,:\,k \neq i\,,\,k \neq j} d^k\,p^k \;,$$

it holds

$$d^i\,e^*\,D\,p - (d^i - d^j)\,d^i\,p^i = d^i\,(d^i\,p^i + d^j\,p^j + s) - (d^i - d^j)\,d^i\,p^i = d^i\,d^j\,(p^i + p^j) + d^i\,s \ge 0 \;,$$

and

$$d^j\,e^*\,D\,p + (d^i - d^j)\,d^j\,p^j = d^j\,(d^i\,p^i + d^j\,p^j + s) + (d^i - d^j)\,d^j\,p^j = d^i\,d^j\,(p^i + p^j) + d^j\,s \ge 0 \;,$$

hence

$$\begin{aligned}
\|V\,f_{i,j}\| &= \tfrac{1}{2}\,\frac{1}{(e^*\,D\,p)^2}\,[2\,d^i\,d^j\,(p^i + p^j) + d^i\,s + d^j\,s + |d^i - d^j|\,s] \\
&= \frac{1}{(e^*\,D\,p)^2}\,[d^i\,d^j\,(p^i + p^j) + \max(d^i, d^j)\,s] \\
&= \frac{1}{(e^*\,D\,p)^2}\,[d^i\,d^j\,(p^i + p^j) + \max(d^i, d^j)\,(e^*\,D\,p - d^i\,p^i - d^j\,p^j)] \\
&= \frac{\max(d^i, d^j)}{e^*\,D\,p}\,[1 - \frac{d^i\,p^i + d^j\,p^j - \min(d^i, d^j)\,(p^i + p^j)}{e^*\,D\,p}] \\
&\le \frac{\max(d^i, d^j)}{e^*\,D\,p} \le \omega(D) \;,
\end{aligned}$$

which proves the refined estimate (6). □

Lemma A.2 *For any $p, p' \in \mathcal{P}$, and any $N \times N$ diagonal matrix D with positive diagonal entries*

$$\| D \cdot p - D \cdot p' \| \leq \omega(D) \, \| p - p' \| \ ,$$

and

$$\| D \cdot p - D \cdot p' \| \leq 2 \ .$$

PROOF. The second crude estimate is obvious. To prove the first estimate, notice that the mapping

$$p \mapsto \frac{D\, p}{e^* \, D\, p} = D \cdot p$$

is differentiable in $\mathcal{P}$, and its Jacobian matrix at the point $p \in \mathcal{P}$ is

$$[I - \frac{D\, p\, e^*}{e^* \, D\, p}] \, \frac{D}{e^* \, D\, p} \ .$$

For any $p, p' \in \mathcal{P}$, the Taylor formula yields

$$D \cdot p - D \cdot p' = \int_0^1 [I - \frac{D\, p(\lambda)\, e^*}{e^* \, D\, p(\lambda)}] \, \frac{D\, (p - p')}{e^* \, D\, p(\lambda)} \, d\lambda \ ,$$

where $p(\lambda) = \lambda\, p + (1 - \lambda)\, p'$ belongs to $\mathcal{P}$ for any $0 \leq \lambda \leq 1$. Since $(p - p')$ belongs to Σ, the estimate (6) yields

$$\| D \cdot p - D \cdot p' \| \leq \int_0^1 \| \, [I - \frac{D\, p(\lambda)\, e^*}{e^* \, D\, p(\lambda)}] \, \frac{D\, (p - p')}{e^* \, D\, p(\lambda)} \, \| \, d\lambda \leq \omega(D) \, \| p - p' \| \ . \qquad \square$$

Lemma A.3 *For any $p \in \mathcal{P}$, and any column–allowable nonnegative $N \times N$ matrix M, the following two estimates hold : (i) for any $u \in \mathbf{R}^N$*

$$| \frac{e^* \, M\, u}{e^* \, M\, p} | \leq \omega(M) \, \| u \| \ , \tag{8}$$

and (ii) for any $u \in \Sigma$

$$| \frac{e^* \, M\, u}{e^* \, M\, p} | \leq \tfrac{1}{2} \, [\omega(M) - 1] \, \| u \| \ . \tag{9}$$

PROOF. For any $u \in \mathbf{R}^N$

$$| e^* \, M\, u | \leq \sum_{i \in S} M^{\bullet,i} \, | u^i | \leq \max_{i \in S} M^{\bullet,i} \, \| u \| \ ,$$

whereas for any $p \in \mathcal{P}$

$$e^* \, M\, p = \sum_{i \in S} M^{\bullet,i} \, p^i \geq \min_{i \in S} M^{\bullet,i} \ ,$$

which proves the rough estimate (8).

Using again the representation (7) it holds

$$\max_{u \in \Sigma \, : \, \| u \| = 1} | \frac{e^* \, M\, u}{e^* \, M\, p} | \leq \max_{i,j \in S} | \frac{e^* \, M\, f_{i,j}}{e^* \, M\, p} | \ .$$

Notice that

$$e^* M f_{i,j} = \tfrac{1}{2} (M^{\bullet,i} - M^{\bullet,j}) ,$$

hence

$$|e^* M f_{i,j}| = \tfrac{1}{2} |M^{\bullet,i} - M^{\bullet,j}| \leq \tfrac{1}{2} [\max_{i \in S} M^{\bullet,i} - \min_{i \in S} M^{\bullet,i}] ,$$

which proves the refined estimate (9). □

B Proof of Proposition 5.2

► Using the first estimate in Lemma 2.2 and the estimate in Lemma 4.3 yields

$$\begin{aligned}
&\sum_{l=m}^{n} \| V[M_{n,l+1}, M_{l,m} \cdot p] \, (u_l[M_{l-1,m} \cdot p] - u_l[M_{l-1,m} \cdot p']) \| \\
&\leq \sum_{l=m}^{n} \omega(M_{n,l+1}) \, \tau_{\mathrm{B}}(M_{n,l+1}) \, \| u_l[M_{l-1,m} \cdot p] - u_l[M_{l-1,m} \cdot p'] \| \\
&\leq \sum_{l=m}^{n} \mathrm{Lip}(u_l) \, \omega(M_{n,l+1}) \, \tau_{\mathrm{B}}(M_{n,l+1}) \, \| M_{l-1,m} \cdot p - M_{l-1,m} \cdot p' \| \\
&\leq \sum_{l=m}^{n} \mathrm{Lip}(u_l) \, \omega(M_{n,l+1}) \, \omega(M_{l-1,m}) \, \tau_{\mathrm{B}}(M_{n,l+1}) \, \tau_{\mathrm{B}}(M_{l-1,m}) \, \| p - p' \| ,
\end{aligned}$$

with the convention $M_{n,n+1} = M_{m-1,m} = I$, and similarly using the first estimate in Lemma 4.5 yields

$$\begin{aligned}
&\sum_{l=m}^{n-1} \| (V[M_{n,l+1}, M_{l,m} \cdot p] - V[M_{n,l+1}, M_{l,m} \cdot p']) \, u_l[M_{l-1,m} \cdot p'] \| \\
&\leq \sum_{l=m}^{n-1} \tfrac{1}{2} [\omega(M_{n,l+1}) - 1] \, [\omega(M_{n,m}) + \omega(M_{n,l+1}) \, \omega(M_{l,m})] \\
&\qquad\qquad \tau_{\mathrm{B}}(M_{n,l+1}) \, \tau_{\mathrm{B}}(M_{l,m}) \, \|p - p'\| \, \|u_l[M_{l-1,m} \cdot p'] \| \\
&\leq \sum_{l=m}^{n-1} \mathrm{K}(u_l) \, \tfrac{1}{2} [\omega(M_{n,l+1}) - 1] \, [\omega(M_{n,m}) + \omega(M_{n,l+1}) \, \omega(M_{l,m})] \\
&\qquad\qquad \tau_{\mathrm{B}}(M_{n,l+1}) \, \tau_{\mathrm{B}}(M_{l,m}) \, \|p - p'\| .
\end{aligned}$$

With the decomposition introduced in (4), using the above estimates, the estimate in Lemma 4.3 and the first estimate in Lemma 4.5 yields

$$\|W_{n,m}[u_n, \cdots, u_m, p, w] - W_{n,m}[u_n, \cdots, u_m, p', w'] \|$$

$$
\begin{aligned}
&\leq \tfrac{1}{2} \,\|\, (\, V[M_{n,m}, p] + V[M_{n,m}, p'] \,)\; (w - w') \,\| + \tfrac{1}{2} \,\|\, (\, V[M_{n,m}, p] - V[M_{n,m}, p'] \,)\; (w + w') \,\| \\
&\qquad + \sum_{l=m}^{n} \;\|\, V[M_{n,l+1}, M_{l,m} \cdot p] \;(\, u_l[M_{l-1,m} \cdot p] - u_l[M_{l-1,m} \cdot p'] \,) \,\| \\
&\qquad + \sum_{l=m}^{n-1} \;\|\, (\, V[M_{n,l+1}, M_{l,m} \cdot p] - V[M_{n,l+1}, M_{l-1,m} \cdot p'] \,)\; u_l[M_{l-1,m} \cdot p'] \,\| \\
&\leq \omega(M_{n,m})\, \tau_{\mathrm{B}}(M_{n,m})\; \|w - w'\| + \tfrac{1}{2}\, \omega(M_{n,m})\; [\omega(M_{n,m}) - 1]\; \tau_{\mathrm{B}}(M_{n,m})\; \|p - p'\| \;(\, \|w\| + \|w'\| \,) \\
&\qquad + \sum_{l=m}^{n} \; \mathrm{Lip}(u_l)\; \omega(M_{n,l+1})\, \omega(M_{l-1,m})\, \tau_{\mathrm{B}}(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{l-1,m})\; \|p - p'\| \\
&\qquad + \sum_{l=m}^{n-1} \; \mathrm{K}(u_l)\; \tfrac{1}{2}\, [\omega(M_{n,l+1}) - 1]\; [\omega(M_{n,m}) + \omega(M_{n,l+1})\, \omega(M_{l,m}) \,] \\
&\qquad\qquad\qquad \tau_{\mathrm{B}}(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{l,m})\; \|p - p'\| \\
&\leq \tfrac{1}{2}\, [\omega^2(M_{n,m}) + 1]\; \tau_{\mathrm{B}}(M_{n,m})\; [\, \|w - w'\| + \|p - p'\| \;(\, \|w\| + \|w'\| \,) \,] \\
&\qquad + \Big[\sum_{l=m}^{n} \; \mathrm{Lip}(u_l)\; \omega(M_{n,l+1})\, \omega(M_{l-1,m})\, \tau_{\mathrm{B}}(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{l-1,m}) \\
&\qquad\qquad + \sum_{l=m}^{n-1} \; \mathrm{K}(u_l)\; \tfrac{1}{2}\, [\omega(M_{n,l+1}) - 1]\; [\omega(M_{n,m}) + \omega(M_{n,l+1})\, \omega(M_{l,m}) \,] \\
&\qquad\qquad\qquad\qquad \tau_{\mathrm{B}}(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{l,m}) \,\Big] \;\; \|p - p'\| \;,
\end{aligned}
$$

with the convention $M_{n,n+1} = M_{m-1,m} = I$, since $\max(\omega, \frac{1}{2}\,\omega\,[\omega - 1]) \leq \frac{1}{2}\,[\omega^2 + 1]$ for any $\omega \geq 1$.

▶ Using the second estimate in Lemma 2.2 and the estimate in Lemma 4.3 yields

$$
\begin{aligned}
&\sum_{l=m}^{n} \;\|\, V[M_{n,l+1}, M_{l,m} \cdot p] \;(\, u_l[M_{l-1,m} \cdot p] - u_l[M_{l-1,m} \cdot p'] \,) \,\| \\
&\qquad \leq \sum_{l=m}^{n} \; \omega(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{n,l+1}) \;\|\, u_l[M_{l-1,m} \cdot p] - u_l[M_{l-1,m} \cdot p'] \,\| \\
&\qquad \leq \sum_{l=m}^{n} \; \mathrm{Lip}(u_l)\; \omega(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{n,l+1}) \;\|\, M_{l-1,m} \cdot p - M_{l-1,m} \cdot p' \,\| \\
&\qquad \leq 2 \sum_{l=m}^{n} \; \mathrm{Lip}(u_l)\; \omega(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{n,l+1})\, \tau_{\mathrm{B}}(M_{l-1,m}) \;,
\end{aligned}
$$

with the convention $M_{n,n+1} = M_{m-1,m} = I$, and similarly using the second estimate in Lemma 4.5

yields

$$\sum_{l=m}^{n-1} \| \left(V[M_{n,l+1}, M_{l,m} \cdot p] - V[M_{n,l+1}, M_{l,m} \cdot p'] \right) u_l[M_{l-1,m} \cdot p'] \|$$

$$\leq \sum_{l=m}^{n-1} [\omega^2(M_{n,l+1}) - 1] \, \tau_{\mathrm{B}}(M_{n,l+1}) \, \tau_{\mathrm{B}}(M_{l,m}) \, \|u_l[M_{l-1,m} \cdot p'] \|$$

$$\leq \sum_{l=m}^{n-1} \mathrm{K}(u_l) \, [\omega^2(M_{n,l+1}) - 1] \, \tau_{\mathrm{B}}(M_{n,l+1}) \, \tau_{\mathrm{B}}(M_{l,m}) \; .$$

With the decomposition introduced in (4), using the above estimates, the estimate in Lemma 4.3 and the second estimate in Lemma 4.5 yields

$$\|W_{n,m}[u_n, \cdots, u_m, p, w] - W_{n,m}[u_n, \cdots, u_m, p', w'] \|$$

$$\leq \tfrac{1}{2} \| \left(V[M_{n,m}, p] + V[M_{n,m}, p'] \right) (w - w') \| + \tfrac{1}{2} \| \left(V[M_{n,m}, p] - V[M_{n,m}, p'] \right) (w + w') \|$$

$$+ \sum_{l=m}^{n} \| V[M_{n,l+1}, M_{l,m} \cdot p] \left(u_l[M_{l-1,m} \cdot p] - u_l[M_{l-1,m} \cdot p'] \right) \|$$

$$+ \sum_{l=m}^{n-1} \| \left(V[M_{n,l+1}, M_{l,m} \cdot p] - V[M_{n,l+1}, M_{l,m} \cdot p'] \right) u_l[M_{l-1,m} \cdot p'] \|$$

$$\leq \omega(M_{n,m}) \, \tau_{\mathrm{B}}(M_{n,m}) \, \|w - w'\| + \tfrac{1}{2} [\omega^2(M_{n,m}) - 1] \, \tau_{\mathrm{B}}(M_{n,m}) \, \left(\|w\| + \|w'\| \right)$$

$$+ 2 \sum_{l=m}^{n} \mathrm{Lip}(u_l) \, \omega(M_{n,l+1}) \, \tau_{\mathrm{B}}(M_{n,l+1}) \, \tau_{\mathrm{B}}(M_{l-1,m})$$

$$+ \sum_{l=m}^{n-1} \mathrm{K}(u_l) \, [\omega^2(M_{n,l+1}) - 1] \, \tau_{\mathrm{B}}(M_{n,l+1}) \, \tau_{\mathrm{B}}(M_{l,m})$$

$$\leq \tfrac{1}{2} [\omega^2(M_{n,m}) + 1] \, \tau_{\mathrm{B}}(M_{n,m}) \, \left(\|w\| + \|w'\| \right)$$

$$+ 2 \sum_{l=m}^{n} \mathrm{Lip}(u_l) \, \omega(M_{n,l+1}) \, \tau_{\mathrm{B}}(M_{n,l+1}) \, \tau_{\mathrm{B}}(M_{l-1,m})$$

$$+ \sum_{l=m}^{n-1} \mathrm{K}(u_l) \, [\omega^2(M_{n,l+1}) - 1] \, \tau_{\mathrm{B}}(M_{n,l+1}) \, \tau_{\mathrm{B}}(M_{l,m}) \; ,$$

with the convention $M_{n,n+1} = M_{m-1,m} = I$, since $\max(\omega, \frac{1}{2} [\omega^2 - 1]) \leq \frac{1}{2} [\omega^2 + 1]$ for any $\omega \geq 1$.
□