Convergence of products of matrices in projective spaces *

Shmuel Friedland †
Department of Mathematics, Statistics and Computer Science
University of Illinois at Chicago
Chicago, Illinois 60607-7045‡
and
Institute for Mathematics and its Applications
400 Lind Hall, 207 Church St., S.E.
Minneapolis, MN 55455-0436

July 28, 2004

Abstract

Let $A_k, k \in \mathbb{N}$ be a sequence of $n \times n$ complex valued matrices which converge to a matrix A. If A and each A_k is positive then the product $\frac{A_k A_{k-1} \ldots A_2 A_1}{||A_k A_{k-1} \ldots A_2 A_1||}$ converges to a rank one matrix positive matrix $\mathbf{u}\mathbf{w}^{\mathrm{T}}$, where $\mathbf{u}$ is a positive column eigenvector of A. If each A_k is nonsingular and A has exactly one simple eigenvalue λ of the maximal modulus with the corresponding eigenvector $\mathbf{u}$, then $e^{\sqrt{-1}\theta_k} \frac{A_k A_{k-1} \ldots A_2 A_1}{||A_k A_{k-1} \ldots A_2 A_1||}, \theta_k \in \mathbb{R}$ converges to a rank one matrix $\mathbf{u}\mathbf{w}^{\mathrm{T}}$.

2000 Mathematical Subject Classification: 15A48, 47H09, 65L20
Keywords: Positive matrices, contractions, convergence of products of matrices, projective spaces.

1 Introduction

For $\mathbb{F} = \mathbb{R}, \mathbb{C}$ denote by $\mathbb{F}^n, \mathrm{M}_n(\mathbb{F}), \mathrm{GL}_n(\mathbb{F})$ the n-dimensional column vector space, the algebra of $n \times n$ matrices and the subgroup of $n \times n$ invertible matrices over the field $\mathbb{F}$. Denote by $|| \cdot ||$ any vector norm on $\mathbb{F}^n$ or on $\mathrm{M}_n(\mathbb{F})$. Let $|| \cdot ||_2$ be the ℓ_2

*Dedicated to Richard Brualdi on his birthday

†*Partially supported by NSF*

‡Tel.: +1-312-996-3041; fax: +1-312-996-1491; email: friedlan@uic.edu

norm on $\mathbb{F}^n$ induced by the standard inner product $\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{y}^*\mathbf{x}$ on $\mathbb{F}^n$ and denote by $|| \cdot ||_2$ the induced operator norm on $\mathrm{M}_n(\mathbb{F})$. Consider an iteration scheme

$$\mathbf{x}_k := A_k \mathbf{x}_{k-1}, \quad \mathbf{x}_0 \in \mathbb{F}^n,\ A_k \in \mathrm{M}_n(\mathbb{F}), k \in \mathbb{N}. \tag{1.1}$$

This system is called *convergent* if $\mathbf{x}_k, k \in \mathbb{N}$ is a convergent sequence for each $\mathbf{x}_0 \in \mathbb{F}^n$. This is equivalent to the convergence of the infinite product $...A_k A_{k-1}...A_2 A_1$, which is defined as the limit of $A_k A_{k-1}...A_2 A_1$ as $k \to \infty$. For the stationary case $A_k = A, k \in \mathbb{N}$ the necessary and sufficient conditions for convergency are well known. First, the spectral radius $\rho(A)$ can not exceed 1. Second, if $\rho(A) = 1$, then 1 is an eigenvalue of A and all its Jordan blocks have size 1. Third all other eigenvalues λ of A different from 1 satisfy $|\lambda| < 1$.

In some instances, as Lyapunov exponents in dynamical systems [11], one interested if the line spanned by the vector $\mathbf{x}_i$ converges for all $\mathbf{x}_0 \neq 0$ in some homogeneous open Zariski set in $\mathbb{F}^n$ [2]. If this condition holds we call (1.1) *projectively convergent.*

For the stationary case $0 \neq A_k = A \in \mathrm{M}_n(\mathbb{C})$ it is straightforward to show that (1.1) is projectively convergent if and only if among all the eigenvalues λ of A satisfying $|\lambda| = \rho(A)$, there is exactly one eigenvalue λ_0 which has Jordan blocks of the maximal size. See for example the arguments in [4, Thm 2.2].

A variation of projectively convergent iterations was considered in the literature for the nonnegative matrices under the name nonhomogeneous matrix products [7], [12] and [8]. Let $\mathbb{R}_+ := (0, \infty)$ and denote by $\mathbb{R}^n_+ \subset \mathbb{R}^n, \mathrm{M}_n(\mathbb{R}_+) \subset \mathrm{M}_n(\mathbb{R})$ the cone of positive vectors and the semialgebra of positive matrices. Denote by $\mathbb{P}\mathbb{R}^n_+$ and $\mathbb{P}\mathrm{M}_n(\mathbb{R}_+)$ the projective space formed by the rays spanned by $\mathbf{x} \in \mathbb{R}^n_+$ and $A \in \mathrm{M}_n(\mathbb{R}_+)$. Then $\mathbb{P}\mathbb{R}^n_+$ has the Hilbert (hyperbolic) metric. Furthermore each $A \in \mathrm{M}_n(\mathbb{R}_+)$ acts on $\mathbb{P}\mathbb{R}^n_+$, where this action is denoted $\hat{A} : \mathbb{P}\mathbb{R}^n_+ \to \mathbb{P}\mathbb{R}^n_+$, and $\hat{A}$ is a contraction [1]. That is the Lipschitz constant $L(\hat{A})$ of $\hat{A}$ is less than 1. Let $A_k \in \mathrm{M}_n(\mathbb{R}_+), k \in \mathbb{N}$ be a sequence of positive matrices. Then the condition $\lim_{k\to\infty} L(\widehat{A_1...A_k}) = 0$, which is equivalent to the notion of weak ergodicity of the products $A_1...A_k, k \in \mathbb{N}$ [12], implies that for each $\mathbf{x}_0 \in \mathbb{R}^n_+$ the ray spanned by $A_1...A_k\mathbf{x}_0$ converges to a fixed ray in $\mathbb{P}\mathbb{R}^n_+$.

Clearly $A_k...A_1, \in \mathrm{M}_n(\mathbb{R}_+), k \in \mathbb{N}$ is projectively convergent if

$$\lim_{k\to\infty} \frac{A_k A_{k-1}...A_2 A_1}{||A_k A_{k-1}...A_2 A_1||} = E, \quad (\text{where } ||E|| = 1) \tag{1.2}$$

and $E \in \mathrm{M}_n(\mathbb{R}_+)$. We show that the assumption

$$\lim_{k\to\infty} L(\widehat{A_k...A_1}) = 0 \quad (\iff \lim_{k\to\infty} L(\widehat{A_1^{\mathrm{T}}...A_k^{\mathrm{T}}}) = 0) \tag{1.3}$$

does not imply (1.2).

The aim of this paper is to show

Theorem 1.1 *Let $A_k \in \mathrm{M}_n(\mathbb{R}_+), k \in \mathbb{N}$ be a sequence of positive matrices which converges to a positive matrix $A \in \mathrm{M}_n(\mathbb{R}_+)$. Then (1.2) holds. Furthermore*

$$E = \mathbf{u}\mathbf{w}^{\mathrm{T}}, \quad \mathbf{u}, \mathbf{w} \in \mathbb{R}_+^n, \ \ A\mathbf{u} = \rho(A)\mathbf{u}. \tag{1.4}$$

One can view the above Theorem as an improvement of [12, Thm 3.6].

Theorem 1.2 [1] *Let $A_k \in \mathrm{GL}_n(\mathbb{C}), k \in \mathbb{N}$ be a sequence of matrices which converges to a matrix $0 \neq A \in \mathrm{M}_n(\mathbb{C})$. Assume furthermore that $\rho(A) > 0$ and the circle $\{z :\in \mathbb{C}, \ |z| = \rho(A)\}$ contains exactly one eigenvalue λ of A, which is a simple root of its characteristic polynomial. Let $A\mathbf{u} = \lambda\mathbf{u}, 0 \neq \mathbf{u} \in \mathbb{C}^n$. Then the complex line spanned by $A_k ... A_1 \in \mathrm{M}_n(\mathbb{C})$ converges to the complex line spanned by $\mathbf{u}\mathbf{w}^{\mathrm{T}} \in \mathrm{M}_n(\mathbb{C})$, for some $0 \neq \mathbf{w} \in \mathbb{C}^n$. Hence for each $\mathbf{x}_0 \in \mathbb{C}^n$ such that $\mathbf{w}^{\mathrm{T}}\mathbf{x}_0 \neq 0$, the complex line spanned by $\mathbf{x}_k$ given by (1.1) converges to the complex line spanned by $\mathbf{u}$.*

We now list briefly the contents of the paper. In §2 we recall basic results on the real and complex projective spaces used in this paper. In §3 we discuss Lipschitz continuous maps and contractions, and simple conditions for pointwise convergence of the products of Lipschitzian maps to a constant map. In §4 we prove Theorem 1.1 and use it to prove Theorem 1.2 in the real case. In §5 we prove Theorem 1.2 in the complex case by using directly the results of §3 and Theorem 1.2 in the real case. In §6 we extend Theorem 1.1 to strictly totally positive matrices (of order p). We also extend Theorem 1.2 to the case where the limit matrix A, has p simple eigenvalues $\lambda_1, ..., \lambda_p$, such that $|\lambda_1| > ... > |\lambda_p| > 0$ and all other eigenvalues of A lie in $|z| < |\lambda_p|$.

2 Projective spaces

In this section we recall the well known notions and results about projective spaces used here. Recall that for $\mathbb{F} = \mathbb{R}, \mathbb{C}$ the spaces $\mathbb{P}\mathbb{F}^n, \mathbb{P}\mathrm{M}_n(\mathbb{F}), \mathbb{P}\mathrm{GL}_n(\mathbb{F})$ are obtained by identifying the orbits of the action of $\mathbb{F}^* := \mathbb{F}\backslash\{0\}$ on the nonzero elements of the corresponding sets. ($\mathbb{F}^*$ acts by multiplication.) Then $\mathbb{P}\mathbb{R}^n, \mathbb{P}\mathrm{M}_n(\mathbb{R})$ and $\mathbb{P}\mathbb{C}^n, \mathbb{P}\mathrm{M}_n(\mathbb{C})$ are compact real and complex manifolds respectively. (For the reason that will be seen later our notation for $\mathbb{P}\mathbb{F}^n$ is slightly different from the standard notation.) Note that we can view $\mathbb{P}\mathrm{M}_n(\mathbb{F})$ as isomorphic to $\mathbb{P}\mathbb{F}^{n^2}$. For any $U \subset \mathbb{F}^n$

[1] *Acknowledgement* A variant of this theorem was suggested by Boris Shapiro.

we denote by $\hat{U} \subset \mathbb{PF}^n$ the set generated by the orbits of $\mathbb{F}^*(U\backslash\{0\})$. $(\widehat{\{0\}} = \emptyset.)$ A set $V \subset \mathbb{PF}^n$ is called a (projective) variety if $V = \hat{U}$, where U is the zero set of a finite number of homogeneous polynomials over $\mathbb{F}$ in $\mathbb{F}^n$. $H \subset \mathbb{PF}^n$ is called a hyperplane if $H = \hat{U}$, where U is a subspace of $\mathbb{F}^n$ of codimension 1. $V \subset \mathbb{PF}^n$ is called a linear space if it is an intersection of a finite number of hyperplanes. $W \subset \mathbb{PR}^n$ is called Zariski open if $W = \mathbb{PF}^n\backslash V$ for some variety V.

For $\mathbf{x} \in \mathbb{F}^n\backslash\{0\}, A \in \mathrm{M}_n(\mathbb{F})\backslash\{0\}$ denote by $\hat{\mathbf{x}}, \hat{A}$ the corresponding elements of $\mathbb{PF}^n, \mathbb{PM}_n(\mathbb{F})$ respectively. Let $A \in \mathrm{GL}_n(\mathbb{F})$. Then A acts on $\mathbb{F}^n\backslash\{0\}$, so $\widehat{A\mathbf{x}} = \hat{A}\hat{x}$ for any $\mathbf{x} \in \mathbb{F}^n\backslash\{0\}$. That is $\hat{A}$ acts on $\mathbb{PF}^n$. Let $A \in \mathrm{M}_n(\mathbb{F})\backslash\{0\}$ Then $\hat{A}$ acts on Zariski open set $\mathbb{PF}^n\backslash\widehat{\ker A}$.

Since $\mathbb{PF}^n, \mathbb{PM}_n(\mathbb{F})$ are compact for any sequences

$$\mathbf{x}_k \in \mathbb{F}^n\backslash\{0\}, A_k, B_k \in \mathrm{M}_n(\mathbb{F})\backslash\{0\}, k \in \mathbb{N}$$

we can find a subsequence $k_l, l \in \mathbb{N}$ and corresponding $\mathbf{x} \in \mathbb{F}^n\backslash\{0\}, A, B \in \mathrm{M}_n(\mathbb{F})\backslash\{0\}$, depending on $k_l, l \in \mathbb{N}$ such that

$$\lim_{l\to\infty} \hat{\mathbf{x}}_{k_l} = \hat{\mathbf{x}}, \quad \lim_{l\to\infty} \hat{A}_{k_l} = \hat{A}, \quad \lim_{l\to\infty} \hat{B}_{k_l} = \hat{B}.$$

Note also

$$\lim_{l\to\infty} \widehat{A_{k_l}\mathbf{x}_{k_l}} = \widehat{A\mathbf{x}} = \hat{A}\hat{\mathbf{x}} \quad \text{if } A\mathbf{x} \neq 0,$$
$$\lim_{l\to\infty} \widehat{A_{k_l}B_{k_l}} = \widehat{AB} = \hat{A}\hat{B} \quad \text{if } AB \neq 0.$$

The convergence of sequences in $\mathbb{PF}^n$ and $\mathbb{PM}_n(\mathbb{F})$ are equivalent to the following statement:

Proposition 2.1 *Let $\mathbf{x}_k \in \mathbb{F}^n\backslash\{0\}, A_k \in \mathrm{M}_n(\mathbb{F})\backslash\{0\}, k \in \mathbb{N}$. Then sequences $\hat{\mathbf{x}}_k, \hat{A}_k, k \in \mathbb{N}$ converge in $\mathbb{PF}^n, \mathbb{PM}_n(\mathbb{F})$ respectively if and only if there exist two sequences $\mu_k, \nu_k \in \{z \in \mathbb{C} : |z| = 1\} \cap \mathbb{F}, k \in \mathbb{N}$ such that the sequences $\mu_k \frac{\mathbf{x}_k}{||\mathbf{x}_k||}$, $\nu_k \frac{A_k}{||A_k||}$, $k \in \mathbb{N}$ converge in $\mathbb{F}^n, \mathrm{M}_n(\mathbb{F})$ respectively.*

Note that for $\mathbb{F} = \mathbb{R}$ $\mu_k, \nu_k \in \{1, -1\}$. Thus if $\mathbf{x}_k \in \mathbb{R}^n_+, A_k \in \mathrm{M}_n(\mathbb{R}_+)$ it is clear that in Proposition 2.1 we may assume that $\mu_k = \nu_k = 1$. Hence for $A_k \in \mathrm{M}_n(\mathbb{R}_+), k \in \mathbb{N}$, $\widehat{A_k...A_1}$ converges in $\mathbb{PM}_n(\mathbb{R})$ if and only if (1.2) holds.

Let $\mathbb{PR}^n_+, \mathbb{PM}_n(\mathbb{R}_+)(\approx \mathbb{PR}^{n^2}_+)$ be the set of orbits in $\mathbb{R}^n_+, \mathrm{M}_n(\mathbb{R}_+)$ under the action of $\mathbb{R}_+$ (by multiplication). We view $\mathbb{PR}^n_+, \mathbb{PM}_n(\mathbb{R}_+)$ as corresponding subsets of $\mathbb{PR}^n, \mathbb{PM}_n(\mathbb{R})$ respectively. Note that $\mathbb{PM}_n(\mathbb{R}_+)$ acts on $\mathbb{PR}^n_+$. Sometime it is convenient to identify $\mathbb{PR}^n_+$ and $\mathbb{PM}_n(\mathbb{R}_+)$ with the open set of positive probability vectors and the open set of positive matrices whose sum of coordinates is equal to 1 respectively.

Recall the notion of Hilbert (hyperbolic) metric on $\mathbb{PR}_+^n$ [9], which is not equivalent to the metric induced by the standard Riemannian metric on the compact manifold $\mathbb{PR}^n$. Let

$$d(\hat{\mathbf{x}}, \hat{\mathbf{y}}) = \log \frac{\max_i \frac{x_i}{y_i}}{\min_i \frac{x_i}{y_i}}, \quad \mathbf{x} = (x_1, ..., x_n)^{\mathrm{T}}, \mathbf{y} = (y_1, ..., y_n)^{\mathrm{T}} \in \mathbb{R}_+^n. \tag{2.1}$$

It is straightforward to show that $d(\cdot, \cdot)$ is a metric on $\mathbb{PR}_+^n$, $\mathbb{PR}_+^n$ is a complete separable metric space with respect to $d(\cdot, \cdot)$, which has an infinite diameter. Moreover, $\mathcal{Y} \subset \mathbb{PR}_+^n$ is compact with respect to the above metric if and only if $\mathcal{Y}$ is compact with respect to the standard metric on $\mathbb{PR}^n$.

3 Convergence of contractions

Let $\mathcal{X}$ be a complete metric space with the metric $d(\cdot, \cdot)$. For $T : \mathcal{X} \to \mathcal{X}$ let

$$L(T) := \sup_{x \neq y \in \mathcal{X}} \frac{d(Tx, Ty)}{d(x, y)} \in [0, \infty].$$

We assume here that $a \cdot \infty = \infty \cdot a = \infty$ for any $a \in \mathbb{R}_+$ and $0 \cdot \infty = \infty \cdot 0 = 0$. Note that $L(T) = 0$ if and only if T is a constant operator. For any $T, Q : \mathcal{X} \to \mathcal{X}$ $L(TQ) \leq L(T)L(Q)$. T is called Lipschitz continuous if $L(T) < \infty$. T is called a contraction if $L(T) < 1$. Assume that T is a contraction. Then it is well known that T has a unique fixed point ξ. Furthermore the sequence $T^i, i \in \mathbb{N}$ converges pointwise to a constant operator $Q : \mathcal{X} \to \{\xi\}$. That is $T^i x \to \xi$ for any $x \in \mathcal{X}$.

Lemma 3.1 *Let $T_i, i \in \mathbb{N}$ be a sequence of operators on a complete metric space $\mathcal{X}$. Let $Q_i := T_1 T_2 ... T_i, i \in \mathbb{N}$ be a sequence of operators. Assume that the following two conditions hold:*

$$\lim_{i \to \infty} L(Q_i) = 0. \tag{3.1}$$

$$\lim_{i \to \infty} \sup_{j \in \mathbb{N}} L(Q_i) d(T_{i+1} ... T_{i+j} x, x) = 0, \quad \text{for some } x \in \mathcal{X}. \tag{3.2}$$

Then $Q_i, i \in \mathbb{N}$ converges pointwise to a constant operator $Q : \mathcal{X} \to \{\xi\}$ for some $\xi \in \mathcal{X}$.

Proof. Since

$$d(Q_{i+j}x, Q_i x) = d(Q_i(T_{i+1}...T_{i+j}x), Q_i x) \le L(Q_i)d(T_{i+1}...T_{i+j}x, x) \le \sup_{k\in\mathbb{N}} L(Q_i)d(T_{i+1}...T_{i+k}x, x)$$

the condition (3.2) implies that $Q_i x, i \in \mathbb{N}$ is a Cauchy sequence. Hence $\lim_{i\to\infty} Q_i x = \xi$. Clearly

$$d(Q_i y, \xi) \le d(Q_i y, Q_i x) + d(Q_i x, \xi) \le L(Q_i)d(x, y) + d(Q_i x, \xi). \quad (3.3)$$

The condition (3.1) implies the lemma. □

Recall that a metric spaces $\mathcal{X}$ has a finite diameter if $\sup_{x,y\in\mathcal{X}} d(x, y) < \infty$. Clearly any compact metric space has a finite diameter. Note that if $\mathcal{X}$ is has a finite diameter then (3.1) implies (3.2).

Corollary 3.2 *Let $T_i, i \in \mathbb{N}$ be a sequence of operators on a complete metric space $\mathcal{X}$ of finite diameter. Let $Q_i := T_1T_2...T_i, i \in \mathbb{N}$ be a sequence of operators. Assume that the condition (3.1) holds. Then $Q_i, i \in \mathbb{N}$ converges pointwise to a constant operator $Q : \mathcal{X} \to \{\xi\}$ for some $\xi \in \mathcal{X}$. In particular, if $T_i, i \in \mathbb{N}$ is a sequence of uniform contractions, i.e. $L(T_i) \le a < 1$ for all $\in \mathbb{N}$, on a complete metric space $\mathcal{X}$ of finite diameter then (3.1) holds.*

$A = (a_{ij})_1^n \in \mathrm{M}_n(\mathbb{R})$ is called a nonnegative matrix if $a_{ij} \ge 0, i, j = 1, ..., n$. A is called *row allowable* (*column allowable*) if A is nonnegative and $A\mathbb{R}_+^n \subset \mathbb{R}_+^n$ ($A^{\mathrm{T}}\mathbb{R}_+^n \subset \mathbb{R}_+^n$) , i.e. each row (column) of A contains a positive element. A is called primitive if A is nonnegative and there is $m \in \mathbb{N}$ such that $A^m \in \mathrm{M}_n(\mathbb{R}_+)$. From here and to the end of this section we assume that A is row allowable unless stated otherwise. Then A acts on $\mathbb{PR}_+^n$, i.e. $\hat{A} : \mathbb{PR}_+^n \to \mathbb{PR}_+^n$. It is known that $\hat{A}$ is Lipschitz continuous and $L(\hat{A}) \le 1$ [7]. It was shown by Birkhoff [1] that for $A \in \mathrm{M}_n(\mathbb{R}_+)$ $\hat{A}$ is a contraction. It is known [12] that

$$L(\hat{A}) = \frac{1 - \sqrt{\psi(A)}}{1 + \sqrt{\psi(A)}}, \text{ where } \psi(A) := \min_{i,j,k,l\in[1,n]} \frac{a_{ik}a_{jl}}{a_{il}a_{jk}}, \ A = (a_{ij})_1^n \in \mathrm{M}_n(\mathbb{R}_+). \quad (3.4)$$

For a row allowable nonpositive A $\psi(A) = 0 \iff L(\hat{A}) = 1$. (For a nonnegative non row allowable A we let $\psi(A) = -1 \iff L(A) = \infty$.) Note that $L(\hat{A}) = 0$ if and only if A is a positive rank one matrix. Thus $L(\hat{A}) = 0 \iff L(\widehat{A^{\mathrm{T}}}) = 0$. Furthermore if $A_k, k \in \mathbb{N}$ is a sequence of row allowable matrices then the equivalence of the two conditions stated in (1.3) holds.

Let $A_k = (a_{ij,k})_{i,j}^n, B_k = (b_{ij,k})_{i,j=1}^n \in \mathrm{M}_n(\mathbb{R}_+)$ for $k = N, N+1, ...$ and some $N \in \mathbb{N}$. We say that $A_k, B_k, k \in \mathbb{N}$ are *asymptotically equal*, and denote it by $\{A_k\} \sim \{B_k\}$, if

$$\lim_{k\to\infty} \frac{a_{ij,k}}{b_{ij,k}} = 1, \quad \text{for } i, j = 1, ..., n.$$

The following result is known, e.g. [7].

Lemma 3.3 *Let $A_k, k \in \mathbb{N}$ be a sequence of nonnegative row allowable matrices. Then $\lim_{k\to\infty} L(A_k) = 0$ if and only if there exists a sequence of positive rank one matrices $B_k \in \mathrm{M}_n(\mathbb{R}_+), k \in \mathbb{N}$ such that $\{A_k\} \sim \{B_k\}$.*

Theorem 3.4 *Let $A_k \in \mathrm{M}_n(\mathbb{R}), k \in \mathbb{N}$ be a sequence of nonnegative row (column) allowable matrices. Then $\lim_{k\to\infty} L(\widehat{A_1...A_k}) = 0$ ($\lim_{k\to\infty} L(\widehat{A_k^{\mathrm{T}}...A_1^{\mathrm{T}}}) = 0$) if and only if there exists $\mathbf{u} \in \mathbb{R}_+^n, \mathbf{v}_k \in \mathbb{R}_+^n, k$ for $k > N$ such that $\{A_1...A_k\} \sim \{\mathbf{u}\mathbf{v}_k^{\mathrm{T}}\}$ ($\{A_k^{\mathrm{T}}...A_1^{\mathrm{T}}\} \sim \{\mathbf{v}_k\mathbf{u}^{\mathrm{T}}\}$).*

Proof. Lemma 3.3 implies that if $\{A_1...A_k\} \sim \{\mathbf{u}\mathbf{v}_k^{\mathrm{T}}\}$, where $\mathbf{u}, \mathbf{v}_k \in \mathbb{R}_+^n$, then $\lim_{k\to\infty} L(A_1...A_k) = 0$. Assume that $A_k, k \in \mathbb{N}$ are row-allowable and $\lim_{k\to\infty} L(A_1...A_k) = 0$. Hence there exists $k \in \mathbb{N}$ such that $Q_k = (q_{ij,k})_{i,j=1}^n = A_1...A_k \in \mathrm{M}_n(\mathbb{R}_+)$ for $k > N$. Then [7, Thm 1] implies that $Q_k, k \in \mathbb{N}$ tends to row proportionality. That is there exists $U = (u_{ij}) \in \mathrm{M}_n(\mathbb{R}_+)$ such that

$$\lim_{k\to\infty} \frac{q_{il,k}}{q_{jl,k}} = u_{ij}, \quad i, j \text{ and } i, j = 1, ..., n.$$

Clearly $u_{ii} = 1$ and $u_{ij} = \frac{1}{u_{ji}}$. As $\frac{q_{il,k}}{q_{jl,k}} = \frac{q_{il,k}}{q_{ml,k}} \frac{q_{ml,k}}{q_{jl,k}}$ it follows that $u_{ij} = u_{im}u_{mj}$. Hence $u_{ij} = u_{i1}u_{1j} = \frac{u_{i1}}{u_{j1}}$. Let

$$\mathbf{u} = (u_1, u_2, ..., u_n)^{\mathrm{T}} := (u_{11}, u_{21}, ..., u_{n1})^{\mathrm{T}}, \quad \mathbf{v}_k = (q_{11,k}, q_{12,k}, ..., q_{1n,k})^{\mathrm{T}}$$

and the theorem follows. □

Corollary 3.5 *Let $A_k \in \mathrm{M}_n(\mathbb{R}), k \in \mathbb{N}$ be a sequence of nonnegative row allowable matrices. Assume that $\lim_{k\to\infty} L(\widehat{A_1...A_k}) = 0$. Then $\widehat{A_1...A_k} : \mathbb{PR}_+^n \to \mathbb{PR}_+^n$ converges to a constant operator $Q : \mathbb{PR}_+^n \to \{\hat{\mathbf{u}}\}$ for some $\mathbf{u} \in \mathbb{R}_+^n$.*

Since $\mathbb{PR}_+^n$ is not compact under the hyperbolic metric it follows that Corollary 3.5 is a stronger version of Lemma 3.1. We now give an example which shows that the condition (1.3) for $A_k \in \mathrm{M}_n(\mathbb{R}_+)$ does not imply (1.2). Let $A_k \in \mathrm{M}_n(\mathbb{R}_+)$ be a periodic sequence, i.e. $A_{k+m} = A_k$ for all $k \in \mathbb{N}$ and some $m > 1$. Since $L(\widehat{A_k...A_1}) \le L(\hat{A}_k)...L(\hat{A}_1)$ we deduce that (1.3) holds. Assume the normalization $\rho(A_m...A_1) = 1$. Then $(A_m...A_1)^k \to \mathbf{u}\mathbf{v}^{\mathrm{T}}$, where $A_m...A_1\mathbf{u} = \mathbf{u}, \mathbf{v}^{\mathrm{T}} A_m...A_1 = \mathbf{v}^{\mathrm{T}}, \mathbf{v}^{\mathrm{T}}\mathbf{u} = 1$ for some $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n_+$. Then for $p \in [1, m-1] \cap \mathbb{Z}$ $\lim_{k\to\infty} A_{km+p}...A_{mk+1}A_{mk}...A_1 = A_p...A_1\mathbf{u}\mathbf{v}^{\mathrm{T}}$. Clearly, we can choose $A_1, ..., A_m$ such that (1.2) does not hold.

A special version of the following weak generalization of Theorem 3.4 will be needed to prove Theorem 1.2 in the complex case. Recall that $\mathbb{PC}^n$ is a compact complex manifold. Let $d(\cdot,\cdot)$ be the Fubini-Study metric on $\mathbb{PC}^n$ [6]. Then $\mathbb{PC}^n$ has a finite diameter.

Theorem 3.6 *Let $\mathcal{X} \subset \mathbb{PC}^n$ be a compact set with respect to the Fubini-Study metric d on $\mathcal{X}$ and assume that $\mathcal{X}$ has a nonempty interior. Let $B_k \in \mathrm{M}_n(\mathbb{C})$, $k \in \mathbb{N}$ be a sequence of matrices such that $\widehat{\ker B_k} \cap \mathcal{X} = \emptyset$ and $T_k := \hat{B}_k : \mathcal{X} \to \mathcal{X}$ for each $k \in \mathbb{N}$. Let $Q_k := T_1...T_k$, $k \in \mathbb{N}$. Assume that $\lim_{k\to\infty} L(Q_k) = 0$. Then Q_k converges converges pointwise to a constant operator $Q : \mathcal{X} \to \{\hat{\mathbf{w}}\}$ for some $\hat{\mathbf{w}} \in \mathcal{X}$. Furthermore the limit of any convergent subsequence $\lim_{l\to\infty} \widehat{B_1...B_{k_l}} = \hat{C} \in \mathbb{P}\mathrm{M}_n(\mathbb{C})$ is of the form $\widehat{\mathbf{w}\mathbf{z}^{\mathrm{T}}}$, where $\mathbf{z}$ depends on a subsequence.*

Proof. Corollary 3.2 yields that Q_k, $k \in \mathbb{N}$ converges to a constant operator Q such that $Q\mathcal{X} = \{\hat{\mathbf{w}}\}$. Assume that $\lim_{l\to\infty} \widehat{B_1...B_{k_l}} = \hat{C} \in \mathbb{P}\mathrm{M}_n(\mathbb{C})$. Since $\mathcal{X}$ has an interior, there exists an interior point $\hat{\mathbf{x}} \in \mathcal{X}$ such that $\mathbf{x} \notin \ker C$. Hence $\hat{\mathbf{w}} = \lim_{l\to\infty} Q_{k_l}(\hat{\mathbf{x}}) = \hat{C}\hat{\mathbf{x}} = \widehat{C\mathbf{x}}$. Since this result holds for any $\mathbf{y}$ in the small neighborhood of $\mathbf{x}$ it follows that C is a rank one matrix of the form $\mathbf{w}\mathbf{z}^{\mathrm{T}}$. □

Corollary 5.2 gives a family of examples for which Theorem 3.6 applies.

4 Proof of Theorem 1.1 and Theorem 1.2 for $\mathbb{R}$

To prove Theorems 1.1 and 1.2 we use the following well known fact:

Proposition 4.1 *Let $\mathcal{X}$ be a compact metric space. Then a sequence $x_k \in \mathcal{X}, k \in \mathbb{N}$ converges to ξ if and only if from any convergent subsequence $x_{l_i}, i \in \mathbb{N}$ there exists a subsequence $x_{p_j}, j \in \mathbb{N}$ which converges to ξ.*

Proof of Theorem 1.1. From the definition of $\psi(A^{\mathrm{T}})$ in (3.4) it follows that $\lim_{k\to\infty} \psi(A_k^{\mathrm{T}}) = \psi(A^{\mathrm{T}}) \in (0,1)$. Hence $L(A_1^{\mathrm{T}}...A_k^{\mathrm{T}}) \le L(A_1^{\mathrm{T}})...L(A_k^{\mathrm{T}}) \to 0$. Theorem 3.4 yields the existence of $\mathbf{w}, \mathbf{x}_k \in \mathbb{R}^n_+$ such that $\{A_1^{\mathrm{T}}...A_k^{\mathrm{T}}\} \sim \{\mathbf{w}\mathbf{x}_k^{\mathrm{T}}\}$. Hence $\{A_k...A_1\} \sim \{\mathbf{x}_k\mathbf{w}^{\mathrm{T}}\}$.

Let $C_k := A_k A_{k-1} ... A_2 A_1, k \in \mathbb{N}$. Assume that $\hat{C}_{k_l} \to \hat{C}$. Since $\mathbb{PR}^n$ is compact from each subsequence $\hat{\mathbf{x}}_{k_l}$ we can find as subsequence $\hat{\mathbf{x}}_{l_i}$ such that $\hat{\mathbf{x}}_{l_i} \to \hat{\mathbf{y}}$ where $\mathbf{y} \in \mathbb{R}^n$ is a probability vector. Since $\{A_k ... A_1\} \sim \{\mathbf{x}_k \mathbf{w}^{\mathrm{T}}\}$ it follows that $\hat{C}_{l_i} \to \widehat{\mathbf{y}\mathbf{w}^{\mathrm{T}}} \Rightarrow \hat{C} = \hat{\mathbf{y}}\hat{\mathbf{w}}^{\mathrm{T}}$.

We first deduce the theorem in the case A is a rank one matrix $A = \mathbf{u}\mathbf{v}^{\mathrm{T}}$. Assume that $\hat{C}_{k_l} \to \hat{\mathbf{y}}\hat{\mathbf{w}}^{\mathrm{T}}$. From the sequence k_l pick up a subsequence p_q such that $\hat{C}_{p_q - 1} \to \hat{\mathbf{z}}\hat{\mathbf{w}}^{\mathrm{T}}$ for some probability vector $\mathbf{z}$. Under the above assumptions

$$\lim_{p_q \to \infty} \hat{C}_{p_q} = \lim_{p_q \to \infty} \widehat{A_{p_q} C_{p_q - 1}} = \widehat{\mathbf{u}\mathbf{v}^{\mathrm{T}}} \widehat{\mathbf{z}\mathbf{w}^{\mathrm{T}}} = \widehat{\mathbf{u}\mathbf{w}^{\mathrm{T}}}.$$

Therefore $\lim_{k \to \infty} \hat{C}_k = \widehat{\mathbf{u}\mathbf{w}^{\mathrm{T}}}$ and the theorem follows.

We now consider the general case. Without loss of generality we assume that the spectral radius of A is equal to 1. Then $A^m \to \mathbf{u}\mathbf{v}^{\mathrm{T}}$, where $\mathbf{u}^{\mathrm{T}}\mathbf{v} = 1$. Choose $\epsilon_m, m \in \mathbb{N}$ a sequence of positive decreasing numbers tending to zero with the following property:

$$X_1, ..., X_m \in \mathrm{M}_n(\mathbb{R}) \text{ and } ||X_i - A|| < \epsilon_m, i = 1, ..., m \Rightarrow ||X_1 X_2 ... X_m - A^m|| < \frac{1}{m}.$$

Let N_m the following increasing sequence: $||A_k - A|| < \epsilon_m$ for each $k > N_m$. Hence $||A_{j+m} ... A_{j+1} - A^m|| < \frac{1}{m}$ for any $j > N_m$.

Let $C_k := A_k A_{k-1} ... A_2 A_1, k \in \mathbb{N}$. Assume that $\hat{C}_{k_l} \to \widehat{\mathbf{y}\mathbf{w}^{\mathrm{T}}}$. First choose a subsequence $\{q_j\}$ of $\{k_l\}$ such that $q_{j+1} - q_j > N_{j+1} + j + 1$, where $q_0 = 0$. Let $r_j = q_j - j$ for $j \in \mathbb{N}$. Note that $r_{j+1} > q_j + N_{j+1}$. Hence

$$||A_{q_j} ... A_{r_j + 1} - A^j|| < \frac{1}{j}, \quad \text{for all } j \in \mathbb{N}.$$

From the sequence $r_j, j \in \mathbb{N}$ choose a subsequence r_{j_m} such that $\hat{C}_{r_{j_m}} \to \widehat{\mathbf{z}\mathbf{w}^{\mathrm{T}}}$ for a probability vector $\mathbf{z} \in \mathbb{R}^n$. Note that since $r_{j_m} + j_m = q_{j_m}$ it follows that $\hat{C}_{q_{j_m}} \to \widehat{\mathbf{y}\mathbf{w}^{\mathrm{T}}}$. On the other hand $\hat{C}_{q_{j_m}} = \widehat{A_{q_{j_m}} ... A_{r_{j_m}+1}} \hat{C}_{r_{j_m}}$. Our assumptions yield that the second factor converges to $\widehat{\mathbf{z}\mathbf{w}^{\mathrm{T}}}$. Our construction yields that the first factor converges to $\widehat{\mathbf{v}\mathbf{u}^{\mathrm{T}}}$. Hence $\hat{C}_{k_l} \to \widehat{\mathbf{u}\mathbf{w}^{\mathrm{T}}}$ and the theorem follows in this case too. □

Let $A \in \mathrm{M}_n(\mathbb{R})$ be a primitive matrix. Then A is row and column allowable. Furthermore $\rho(A) > 1$ and there exists $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n_+, \mathbf{v}^{\mathrm{T}}\mathbf{u} = 1$ such that $A\mathbf{u} = \rho(A)\mathbf{u}, \mathbf{v}^{\mathrm{T}} A = \rho(A)\mathbf{v}^{\mathrm{T}}$. Moreover $\lim_{m \to \infty} \rho(A)^{-m} A^m = \mathbf{u}\mathbf{v}^{\mathrm{T}}$. The arguments of the proof of Theorem 1.1 yield:

Corollary 4.2 *Let $A_k, k \in \mathbb{N}$ be a sequence of column allowable matrices such that* $\lim_{k\to\infty} A_k = A$*, where A is a primitive matrix. Then (1.2) and (1.4) hold.*

Proof of Theorem 1.2 in the real case. We assume that $A_k \in \mathrm{M}_n(\mathbb{R}), k \in \mathbb{N}$. Hence $lim_{k\to\infty} A_k = A \in \mathrm{M}_n(\mathbb{R})$. Since the nonreal eigenvalues of A come in pairs $z, \bar{z}$, it follows that the unique eigenvalue of A on the circle $\{z : |z| = \rho(A)\}$ is equal to $\pm\rho(A)$. By multiplying each A_k and A by $\pm\rho(A)^{-1}$ we may assume that $\rho(A) = 1$ and 1 is an eigenvalue of A. 1 is a simple eigenvalue of the characteristic polynomial of A and all other eigenvalues of A lie inside the unit disk $|z| < 1$. By considering TA_kT^{-1} instead of A_k and TAT^{-1} instead of A it is enough to prove the theorem in the case

$$A\mathbf{e} = \mathbf{e},\ A^{\mathrm{T}}\mathbf{v} = \mathbf{v}, \quad \mathbf{e} = (1, ..., 1)^{\mathrm{T}},\ \mathbf{v} = (v_1, ..., v_n)^{\mathrm{T}} \in \mathbb{R}^n_+,\ v_1 + ... + v_n = 1.$$

Indeed, since 1 is a simple root the characteristic polynomial of A, there exists $Q \in \mathrm{GL}_n(\mathbb{R})$ such that $B := QAQ^{-1} = (1) \oplus B'$ for some $B' \in \mathrm{M}_{n-1}(\mathbb{R})$. Hence $B\mathbf{e}_1 = B\mathbf{e}_1^{\mathrm{T}} = \mathbf{e}_1 = (1, 0, ..., 0)^{\mathrm{T}}$ We claim that for $n \geq 2$ there exists $S \in \mathrm{GL}_n(\mathbb{R})$ such that

$$S\mathbf{e}_1 = \mathbf{e}, \quad S^{\mathrm{T}}\mathbf{v} = \mathbf{e}_1, \quad \text{for any } \mathbf{v} \in \mathbb{R}^n_+,\ \mathbf{e}^{\mathrm{T}}\mathbf{v} = 1.$$

The first equation yields that the first column of S is $\mathbf{e}$. The second equation yields that the last $n-1$ columns of S orthogonal to $\mathbf{v}$. Pick any $n-1$ linearly independent vectors in $\mathbf{s}_2, ..., \mathbf{s}_n \in \mathbb{R}^n$ which are orthogonal to $\mathbf{v}$. Then $S := (\mathbf{e}_1, \mathbf{s}_2, ..., \mathbf{s}_n) \in \mathrm{GL}_n(\mathbb{R})$ satisfies the above condition. Now let $T = SQ$.

Our assumptions yield

$$\lim_{m\to\infty} A^m = \mathbf{e}\mathbf{v}^{\mathrm{T}}.$$

As in the proof of Theorem 1.1, let us consider first the case $A = \mathbf{e}\mathbf{v}^{\mathrm{T}}$. As $\lim_{k\to\infty} A_k = A$ and A is a positive matrix it follows that $A_k \in \mathrm{M}_n(\mathbb{R}_+)$ for $k \geq M$. Theorem 1.1 yields that $\widehat{A_k...A_M}$ converges to $\widehat{\mathbf{e}\mathbf{w}_0}$. Hence $\lim_{k\to\infty} \widehat{A_k...A_1} = \widehat{\mathbf{e}\mathbf{w}^{\mathrm{T}}}$, where $\mathbf{w}^{\mathrm{T}} = \mathbf{w}_0^{\mathrm{T}} A_{M-1}...A_1$. This proves the theorem in this case.

Assume that $A \neq \mathbf{e}\mathbf{v}^{\mathrm{T}}$. As $lim_{m\to\infty} A^m = \mathbf{e}\mathbf{v}^{\mathrm{T}}$ it follows that there exists $m \in \mathbb{N}$ such that $A^m \in \mathrm{M}_n(\mathbb{R}_+)$. Hence $A_{k+m-1}...A_k \in \mathrm{M}_n(\mathbb{R}_+)$ for $k \geq N$. Theorem 1.1 yields that $\lim_{k\to\infty} \widehat{A_{km+N}...A_{N+1}} \to \widehat{\mathbf{e}\mathbf{w}_0^{\mathrm{T}}}$. Hence

$$\lim_{k\to\infty} \widehat{A_{km+j+N}...A_{N+1}} = \widehat{A^j}\,\widehat{\mathbf{e}\mathbf{w}_0^{\mathrm{T}}} = \widehat{\mathbf{e}\mathbf{w}_0^{\mathrm{T}}}$$

and the theorem follows in this case too. $\square$

5 Proof of Theorem 1.2 in the complex case.

Since $\mathrm{M}_n(\mathbb{C}) \sim \mathbb{C}^{n^2}$ it follows that $\mathbb{P}\mathrm{M}_n(\mathbb{C}) \sim \mathbb{P}\mathbb{C}^{n^2}$. Let d_1 be the Fubini-Study metric on $\mathbb{P}\mathrm{M}_n(\mathbb{C})$. Let $\hat{A} \in \mathbb{P}\mathrm{M}_n(\mathbb{C})$. Then $\hat{A} : \mathbb{P}\mathbb{C}^n \backslash \widehat{\ker A} \to \mathbb{P}\mathbb{C}^n$ is a holomorphic map.

Lemma 5.1 *Let $E \in \mathrm{M}_n(\mathbb{C})$ be rank one matrix with $\rho(E) > 0$, i.e. $E = \mathbf{v}\mathbf{u}^{\mathrm{T}}, \mathbf{u}^{\mathrm{T}}\mathbf{v} \neq 0$. Let $O_r := \{\hat{\mathbf{x}} \in \mathbb{P}\mathbb{C}^n : \ d(\hat{\mathbf{x}}, \hat{\mathbf{v}}) \leq r\}$ such that $O_r \cap \widehat{\ker E} = \emptyset$. Then $\hat{E} : O_r \to \{\hat{\mathbf{v}}\}$. Assume that $E_k \in \mathrm{M}_n(\mathbb{C})\backslash\{0\}, k \in \mathbb{N}$ converges to E. Then there exists N such that $\hat{E}_k : O_r \to O_r$ is a sequence of uniform contractions for $k > N$, i.e. $d(\hat{E}_k\hat{\mathbf{x}}, \hat{E}_k\hat{\mathbf{y}}) \leq \kappa d(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ for all $\hat{\mathbf{x}}, \hat{\mathbf{y}} \in O_r$ some $\kappa \in (0, 1)$ and $k > N$. Moreover there exists $\epsilon > 0$, depending on E, r and $\kappa \in (0, 1)$, such that for each $\hat{B} \in \mathbb{P}\mathrm{M}_n(\mathbb{C})$ satisfying $d_1(\hat{B}, \hat{E}) \leq \epsilon$ one has $\hat{B} : O_r \to O_r$ and $L(\hat{B}) \leq \kappa$.*

Proof. Clearly $\widehat{E\mathbf{x}} = \hat{\mathbf{v}}$ if $\mathbf{u}^{\mathrm{T}}\mathbf{x} \neq 0$. Hence $\hat{E} : O_r \to \{\hat{\mathbf{v}}\}$. Since $\widehat{\ker E_k}$ converges $\widehat{\ker E}$ it follows that $\hat{E}_k|O_r$ converges uniformly to $\hat{E}|O_r$. In particular $\hat{E}_k : O_r \to O_r$ for $k > M$.

Let $B \in \mathrm{M}_n(\mathbb{C})\backslash\{0\}$. Then for each $\mathbf{x} \in \mathbb{P}\mathbb{C}^n \backslash \widehat{\ker B}$ we can define the local distortion of $\hat{B}$ at $\hat{\mathbf{x}}$:

$$\delta(\hat{B}, \hat{\mathbf{x}}) =: \lim_{m\to\infty} \sup_{\hat{\mathbf{y}}\neq\hat{\mathbf{z}}, d(\hat{\mathbf{y}},\hat{\mathbf{x}})\leq\frac{1}{m}, d(\hat{\mathbf{z}},\hat{\mathbf{x}})\leq\frac{1}{m}} \frac{d(\hat{B}\hat{\mathbf{y}}, \hat{B}\hat{\mathbf{z}})}{d(\hat{\mathbf{y}}, \hat{\mathbf{z}})}.$$

For any $\mathcal{Y} \subset \mathbb{P}\mathbb{C}^n \backslash \widehat{\ker B}$ let

$$\delta(\hat{B}, \mathcal{Y}) := \sup_{\hat{\mathbf{x}}\in\mathcal{Y}} \delta(\hat{B}, \mathbf{x}).$$

Recall that a set $\mathcal{Y}$ is called convex if any two points $\mathbf{x}, \mathbf{y} \in \mathcal{Y}$ can be connected by a geodesic that completely lies in $\mathcal{Y}$. It is a standard fact that if $\mathcal{Y} \subset \mathbb{P}\mathbb{C}^n \backslash \widehat{\ker B}$ is a convex set then

$$d(\hat{B}\hat{x}, \hat{B}\hat{y}) \leq \delta(\hat{B}, \mathcal{Y}) d(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \quad \text{for all } \hat{\mathbf{x}}, \hat{\mathbf{y}} \in \mathcal{Y}.$$

Clearly $\delta(\hat{E}, \hat{\mathbf{x}}) = 0$ for all $\hat{\mathbf{x}} \in \mathbb{P}\mathbb{C}^n \backslash \widehat{\ker E}$. As $E_k \to E$ it follows that $\lim_{k\to\infty} \delta(\hat{E}_k, \hat{\mathbf{x}}) = 0$ for all $\hat{\mathbf{x}} \in \mathbb{P}\mathbb{C}^n \backslash \widehat{\ker E}$. Use this fact and the fact that O_r can be covered by a finite number of convex balls $\{\hat{\mathbf{y}} : \ d(\hat{\mathbf{y}}, \hat{\mathbf{x}}) < r(\hat{\mathbf{x}})\}, \hat{\mathbf{x}} \in O_r$ to deduce the the first part of the lemma.

We now deduce the second part of the lemma. Since O_r and $\widehat{\ker E}$ closed and disjoint it follows that $d(O_r, \widehat{\ker E}) = 2a > 0$. Hence there exists ϵ_1 such that $d(O_r, \widehat{\ker B}) \geq a$ if $d_1(\hat{B}, \hat{E}) \leq \epsilon_1$. It is not difficult to show that

$$\lim_{t \searrow 0} \max_{\hat{B}, d_1(\hat{B}, \hat{E}) \leq t} \delta(\hat{B}, O_r) = \delta(\hat{E}, O_r) = 0.$$

Hence for ϵ small enough and $d_1(\hat{B}, \hat{E}) \leq \epsilon$ one has $d(\hat{B}O_r, \hat{E}O_r) = d(\hat{B}O_r, \hat{u}) < r$ and $L(\hat{B}) = \delta(\hat{B}, O_r) < \kappa$. □

Corollary 5.2 *Let $E \in \mathrm{M}_n(\mathbb{C})$ be a rank one nonnilpotent matrix and let $r > 0, \kappa \in (0,1)$ be given as in Lemma 5.1. Let $B_k \in \mathrm{M}_n(\mathbb{C}) \backslash \{0\}$ and assume that $d_1(\hat{B}_k, \hat{E}) \leq \epsilon$ for each $k \in \mathbb{N}$. Then for $\mathcal{X} = O_r$ the assumptions of Theorem 3.6 hold.*

In what follows we use the concepts of the exterior products $\wedge_k \mathbb{F}^n \subset \mathbb{F}^{\binom{n}{k}}$ and the operators $\wedge_k A \in \mathrm{M}_{\binom{n}{k}}(\mathbb{F})$ induced by $A \in \mathrm{M}_n(\mathbb{F})$. In matrix theory $\wedge_k A$ is called $k-th$ compound matrix, and its entries are given as the $k \times k$ minors of A. For any $\mathbf{x}_1, ..., \mathbf{x}_k \in \mathbb{F}^n$ the coordinates of $\mathbf{x}_1 \wedge ... \wedge \mathbf{x}_k \in \wedge \mathbb{F}^{\binom{n}{k}}$ are $\binom{n}{k}$ minors of the $n \times k$ matrix $(\mathbf{x}_1 ... \mathbf{x}_k)$ arranged in the lexicographical order. Note any nonzero vector $\mathbf{x}_1 \wedge ... \wedge \mathbf{x}_k$ represents a unique subspace $\mathbf{X} = \mathrm{span}(\mathbf{x}_1, ..., \mathbf{x}_k)$ of dimension k, which is an element of the Grassmannian $\mathrm{Gr(k, n, \mathbb{F})}$. Then $\mathbf{y}_1 \wedge ... \wedge \mathbf{y}_k$ represents $\mathbf{X}$ if and only if $\mathrm{span}(\mathbf{x}_1, ..., \mathbf{x}_k) = \mathrm{span}(\mathbf{y}_1, ..., \mathbf{y}_k)$. See for example [5] for the properties of the compound matrices and [3] for a concise survey of mulitilinear algebra used in this paper.

In particular we use the following facts. Let $A, B \in \mathrm{M}_n(\mathbb{C})$. Then
(a) $\wedge_k AB = \wedge_k A \wedge_k B$.
(b) $A\mathbf{x}_1 \wedge ... \wedge A\mathbf{x}_k = \wedge_k A(\mathbf{x}_1 \wedge ... \wedge \mathbf{x}_k)$. If $\mathbf{x}_1, ..., \mathbf{x}_k$ spans a k-dimensional invariant subspace of A then $\mathbf{x}_1 \wedge ... \wedge \mathbf{x}_k$ is an eigenvector of $\wedge_k A$. In particular if $\mathbf{x}_1, ..., \mathbf{x}_k$ are k-linearly independent eigenvectors of A corresponding to the eigenvalues $\lambda_1, ..., \lambda_k$ then $\mathbf{x}_1 \wedge ... \wedge \mathbf{x}_k$ is an eigenvector of $\wedge_k A$ corresponding to the eigenvalue $\lambda_1 ... \lambda_k$.
(c) Let $\lambda_1, ..., \lambda_n$ be the eigenvalues of A counting with their multiplicities. Then $\lambda_{i_1} ... \lambda_{i_k}$ for all $1 \leq i_1 < ... < i_k \leq n$ are all $\binom{n}{k}$ eigenvalues of $\wedge_k A$.

Proof of Theorem 1.2 in the complex case.

By our assumptions the spectral circle $\{z : \ |z| = \rho(A)\}$ contains exactly one eigenvalue λ of algebraic multiplicity 1. By considering $\rho(A)^{-1}A$ we may assume that 1 is a simple algebraic eigenvalue of A, while other eigenvalues of A are in the open unit disk. Hence $\lim_{k \to \infty} A^m = \mathbf{u}\mathbf{v}^{\mathrm{T}}, \mathbf{u}^{\mathrm{T}}\mathbf{v} = 1$. Let $E := \mathbf{v}\mathbf{u}^{\mathrm{T}}$. Lemma 5.1

yields that there exists $\epsilon > 0$ so that for each $B \in \mathrm{M}_n(\mathbb{C})$ satisfying $d_1(\hat{B}, \hat{E}) \le \epsilon$ one has $\hat{B} : O_r \to O_r$ and $L(\hat{B}, O_r) \le \frac{1}{2}$. From the arguments of the proof of Theorem 1.1 it follows that there exists $m \in \mathbb{N}, N \in \mathbb{Z}_+$ such that $d_1(\widehat{A_{k+1}^{\mathrm{T}}...A_{k+m}^{\mathrm{T}}}, \hat{E}) \le \epsilon$ for any $k \ge N$.

Note that for $k > N$ we have $C_k = A_k...A_1 = C_{N+1,k}Q_N$, where $C_{p,k} := A_kA_{k-1}...A_{p+1}A_p, p \le k \in \mathbb{N}$ and $Q_0 = I$ if $N = 0$ and $Q_N := A_N...A_1$ if $N \ge 1$. Since $A_j \in \mathrm{GL}_n(\mathbb{C})$ for $j \in \mathbb{N}$ to prove the theorem it is enough to consider the case $N = 0$. That is we assume that the sequence $A_k^{\mathrm{T}}...A_{k+m-1}^{\mathrm{T}} : O_r \to O_r, k \in \mathbb{N}$ is a sequence of uniform contractions on O_r. Corollary 3.2 implies that $\lim_{k\to\infty} \hat{C}_{j,mk+j-1}^{\mathrm{T}}\hat{\mathbf{x}} = \hat{\mathbf{w}}_j$ for any $\hat{\mathbf{x}} \in O_r$ and some $\hat{\mathbf{w}}_j \in O_r$ for $j = 1, ..., m$.

Assume that $\hat{C}_{k_l} \to \hat{C} \in \mathbb{P}\mathrm{M}_n(\mathbb{C})$, where $C \in \mathrm{M}_n(\mathbb{C})\backslash\{0\}$. We claim that

$$\lim_{l\to\infty} \widehat{A_{k_l}...A_1} = \hat{C} = \widehat{\mathbf{z}\mathbf{y}^{\mathrm{T}}}, \ \texttt{for some}\ \mathbf{y}, \mathbf{z} \in \mathbb{C}^n\backslash\{0\}. \tag{5.1}$$

Choose a subsequence $\{p_q\}_{q\in\mathbb{N}}$ of $\{k_l\}_{l\in\mathbb{N}}$ such that each $p_q - (j-1)$ is divisible by m for some $j \in [1, m] \cap \mathbb{N}$. Then Theorem 3.6 yields that $\lim_{q\to\infty} \widehat{C_{j,p_q}^{\mathrm{T}}} = \widehat{\mathbf{w}_j\mathbf{z}^{\mathrm{T}}}$. Hence $C = \mathbf{z}\mathbf{y}^{\mathrm{T}}$ where $\mathbf{y} = A_1^{\mathrm{T}}...A_{j-1}^{\mathrm{T}}\mathbf{w}_j$, where $A_0 = I$.

To prove the theorem it is enough to show that $\mathbf{z} \in \mathrm{span}(\mathbf{u})$ and $\mathbf{y} \in \mathrm{span}(\mathbf{w})$ for some fixed $\mathbf{w} \in \mathbb{C}^n\backslash\{0\}$. This is done by converting the complex matrices to the real matrices of double dimension, taking the second compounds of the corresponding matrices and using the results of Theorem 1.2 for the real case.

Recall that any linear transformation of $\mathbb{C}^n$ to itself represented by a matrix $L \in \mathrm{M}_n(\mathbb{C}),\ L = P + \sqrt{-1}Q,\ P, Q \in \mathrm{M}_n(\mathbb{R})$ can be presented by $\tilde{L} := \begin{pmatrix} P & -Q \\ Q & P \end{pmatrix}$. This is done by representing any $\mathbf{z} \in \mathbb{C}^n,\ \mathbf{z} = \mathbf{x} + \sqrt{-1}\mathbf{y},\ \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ by $\tilde{\mathbf{z}} := (\mathbf{x}^{\mathrm{T}}, \mathbf{y}^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{2n}$. Then $\widetilde{L\mathbf{z}} = \tilde{L}\tilde{\mathbf{z}}$ and $\widetilde{L_1L_2} = \tilde{L}_1\tilde{L}_2$ for any $L_1, L_2 \in \mathrm{M}_n(\mathbb{C})$. Note that one dimensional subspace $\mathrm{span}(\mathbf{z}) \in \mathbb{C}^n,\ \mathbf{z} \ne 0$ corresponds to the two dimensional subspace $\mathrm{span}(\tilde{\mathbf{z}}, \widetilde{\sqrt{-1}\mathbf{z}}) \in \mathbb{R}^{2n}$. Assume that $\lambda_1, ..., \lambda_n$ are the eigenvalues of L counted with their multiplicities. It is straightforward to show $\lambda_1, \overline{\lambda}_1, ..., \lambda_n, \overline{\lambda}_n$ are the eigenvalues of $\tilde{L}$ counted with their multiplicities. (For a diagonable L the proof reduces to the case where $L \in \mathrm{M}_1(\mathbb{C})$.) Moreover if L is rank one nonnilpotent then $\tilde{L}$ is rank two diagonable.

The assumptions of the theorem yield that $\tilde{A}_k \in \mathrm{GL}_{2n}(\mathbb{R}), k \in \mathbb{N}$ and $\lim_{k\to\infty} \tilde{A}_k = \tilde{A}$. Hence $\wedge_2\tilde{A}_j \in \mathrm{GL}_{\binom{2n}{2}}(\mathbb{R})$ and $\lim_{k\to\infty} \wedge_2\tilde{A}_2 = \wedge_2\tilde{A}$. Since A was rank one non-nilpotent matrix $\tilde{A}$ is a rank two diagonable matrix. Hence $\wedge_2\tilde{A}$ is a rank one matrix with the eigenvector $\tilde{\mathbf{u}} \wedge \widetilde{\sqrt{-1}\mathbf{u}}$ corresponding to the eigenvalue $|\lambda|^2 > 0$. Thus we can apply real version Theorem 1.2 for the sequence $\wedge_2 \tilde{A}_k, k \in \mathbb{N}$. Hence

$$\lim_{k\to\infty} \widehat{\wedge_2 \tilde{A}_k ... \wedge_2} \tilde{A}_1 = \hat{F},\ F = (\tilde{\mathbf{u}} \wedge \widetilde{\sqrt{-1}}\mathbf{u})\mathbf{s}^{\mathrm{T}},\ \text{ for some } \mathbf{s} \in \mathbb{R}^{\binom{2n}{2}} \backslash \{0\}.$$

Compare that with (5.1) to deduce that $\widehat{\wedge_2 \widetilde{\mathbf{z}\mathbf{y}}^{\mathrm{T}}} = \hat{F}$. Equivalently $\wedge_2 \widetilde{\mathbf{z}\mathbf{y}}^{\mathrm{T}} = aF$ for some $a \neq 0$. This shows that first that $\mathbf{z} \in \mathrm{span}(\mathbf{u})$. Second that $\tilde{\mathbf{y}} \wedge \widetilde{\sqrt{-1}}\mathbf{y} = \mathbf{s}$. Since $\mathbf{s}$ is fixed the one dimensional subspace $\mathrm{span}(\mathbf{y})$ does not depend on the convergent subsequence $C_{k_l}, l \in \mathbb{N}$. Thus we can choose $\mathbf{w}$ to be equal to $\mathbf{y}$ for one convergent subsequence $C_{k_l}, l \in \mathbb{N}$. □

6 Finer results

The aim of this section is to consider the convergence of $A_k...A_1\mathbf{x}_0$ under the assumptions of Theorems 1.1 and 1.2 when $\mathbf{w}^{\mathrm{T}}\mathbf{x}_0 = 0$. In this case we need to pass to the exterior products. In this section we assume that the vector and operator norms on $\mathbb{F}^n$ and $\mathrm{M}_n(\mathbb{F})$ for $\mathbb{F} = \mathbb{R}, \mathbb{C}$ are the l_2 norms $|| \cdot ||_2$.

To extend the results of Theorem 1.1 one needs to recall the notions of strictly totally positive matrices and (*discrete*) Tchebyshev systems. See for example [5] or [10] for the notion of strictly totally positive matrices and [10] for the classical notion of Tchebyshev systems. We call $\mathbf{x}_1, ..., \mathbf{x}_p \in \mathbb{R}^n$ a p-Tchebyshev system if $\mathbf{x}_1 \in \mathbb{R}_+, \mathbf{x}_1 \wedge \mathbf{x}_2 \in \mathbb{R}_+^{\binom{n}{2}}, ..., \mathbf{x}_1 \wedge ... \wedge \mathbf{x}_p \in \mathbb{R}_+^{\binom{n}{p}}$. A vector $\mathbf{x} \in \mathbb{R}^n$ is said to have exactly k-changes of signs, denoted by $S(\mathbf{x}) = k$, if by replacing any zero coordinate of $\mathbf{x}$ by a positive or negative number one obtains a vector $\mathbf{y}$ whose coordinates have exactly k changes of signs. It is straightforward to show that if $S(\mathbf{x}) = k \leq n-1$, the there exists a k-Tchebyshev system $\mathbf{x}_1, ..., \mathbf{x}_k$ such that $\mathbf{x}_k = \pm\mathbf{x}$.

Recall that $A \in \mathrm{M}_n(\mathbb{R})$ is called strictly totally positive of order $p \in [1, n] \cap \mathbb{Z}$ (STP_p) if $\wedge_k A \in \mathrm{M}_{\binom{n}{k}}(\mathbb{R}_+)$ for $k = 1, ..., p$. (Here $\wedge_1 A := A$.) That is A and all its $k \leq p$ compounds are positive. The spectrum of A spec A is of the form $\{\lambda_1, ..., \lambda_p\} \cup \mathrm{spec}_{\,p+1} A$. Here $\lambda_1 > ... > \lambda_p > 0$ are p positive real numbers and $\mathrm{spec}_{\,p+1} A \subset \{z \in \mathbb{C} :\ |z| < \lambda_p\}$ if $p < n$. ($\mathrm{spec}_{\,n+1} A = \emptyset$.) Each λ_i is a simple root of det (zI − A) for $i = 1, ..., p$. Furthermore one can choose the signs of the eigenvectors of A and A^{T} corresponding to $\lambda_1, ..., \lambda_k$ such that they form Tchebyshev systems:

$$A\mathbf{u}_i = \lambda_i \mathbf{u}_i, ||\mathbf{u}_i|| = 1, S(\mathbf{u}_i) = i-1, i = 1, ..., p, \mathbf{u}_1 \in \mathbb{R}_+^n, ..., \mathbf{u}_1 \wedge ... \wedge \mathbf{u}_p \in \mathbb{R}_+^{\binom{n}{p}},$$

$$A^{\mathrm{T}}\mathbf{v}_i = \lambda_i \mathbf{v}_i,\ S(\mathbf{v}_i) = i-1,\ i = 1,...,p, \quad \mathbf{v}_1 \in \mathbb{R}_+^n, ..., \mathbf{v}_1 \wedge ... \wedge \mathbf{v}_p \in \mathbb{R}_+^{\binom{n}{p}},$$
$$\mathbf{v}_i^{\mathrm{T}}\mathbf{u}_j = \delta_{ij}, \quad i,j = 1,...,p. \tag{6.1}$$

Theorem 6.1 *Let $A_k \in \mathrm{M}_n(\mathbb{R}_+), k \in \mathbb{N}$ be a sequence of STP_p matrices which converge to a STP_p matrix $A \in \mathrm{M}_n(\mathbb{R}_+)$ for some $p \in [2, n]$ satisfying (6.1). Then there exists a p-Tchebyshev system $\mathbf{w}_1, ..., \mathbf{w}_p$ such that the following conditions hold. Let $C_k = A_k ... A_1$ for $k \in \mathbb{N}$. Then*

$$\lim_{k\to\infty} \frac{\lambda_{i+1}(C_k)}{\lambda_i(C_k)} = 0,\ i = 1,...,p-1, \tag{6.2}$$

$$\lim_{k\to\infty} \frac{\wedge_i C_k}{\prod_{j=1}^{i} \lambda_j(C_k)} = \mathbf{u}_1 \wedge ... \wedge \mathbf{u}_i (\mathbf{w}_1 \wedge ... \wedge \mathbf{w}_i)^{\mathrm{T}},\ i = 1,...,p, \tag{6.3}$$

$$C_k = \sum_{i=1}^{p} \lambda_i(C_k)\mathbf{u}_{i,k}\mathbf{w}_{i,k}^{\mathrm{T}} + o(|\lambda_p(C_k)|),\ \mathbf{w}_{i,k}^{\mathrm{T}}\mathbf{u}_{j,k} = \delta_{ij}, \tag{6.4}$$

$$C_k \mathbf{u}_{i,k} = \lambda_i(C_k)\mathbf{u}_{i,k},\ ||\mathbf{u}_{i,k}|| = 1,\ C_k^{\mathrm{T}}\mathbf{w}_{i,k} = \lambda_i(C_k)\mathbf{w}_{i,k}, \tag{6.5}$$

$$\mathbf{u}_{1,k} \wedge ... \wedge \mathbf{u}_{i,k},\ \mathbf{w}_{1,k} \wedge ... \wedge \mathbf{w}_{i,k} \in \mathbb{R}_+^{\binom{n}{i}},$$
$$\lim_{k\to\infty} \mathbf{u}_{i,k} = \mathbf{u}_i,\ \lim_{k\to\infty} \mathbf{w}_{i,k} = \mathbf{w}_i,\ \mathbf{w}_i^{\mathrm{T}}\mathbf{u}_j = \delta_{ij},\ i,j = 1,...,p. \tag{6.6}$$

Proof. Assume first the assumptions of Theorem 1.1. Let $\mathbf{u}_{1,k}, \mathbf{w}_{1,k}$ be as above. Assume furthermore let $||\mathbf{u}_{1,k}|| = 1$. From the proof of Theorem 1.1 it follows that $\mathbf{u}_{1,k} \to \mathbf{u}_1 = \mathbf{u}$. Let E be defined by (1.2). Then $\rho(E) = \mathbf{w}^{\mathrm{T}}\mathbf{u}$. Hence $\frac{\lambda_1(C_k)}{||C_k||} \to \frac{\mathbf{w}^{\mathrm{T}}\mathbf{u}}{||\mathbf{u}||\,||\mathbf{w}||}$. Hence (6.3) holds for $p = 1$. The proof of Theorem 1.1 yields that one has the equality (6.4) for $p = 1$. Here $\mathbf{w}_1 = (\mathbf{w}^{\mathrm{T}}\mathbf{u})^{-1}\mathbf{w}$.

We now show the theorem for the case $p = 2$. Let $\mathrm{M}_{\binom{n}{2}}(\mathbb{R}_+) \ni B_k := \wedge_2 A_k \to B := \wedge_2 A \in \mathrm{M}_{\binom{n}{2}}(\mathbb{R}_+)$. As $\frac{\lambda_1(C_k)}{||C_k||} \to ||\mathbf{w}_1||^{-1}$ (1.2) yields (6.2) for $p = 2$. Let $D_k := B_k ... B_1, k \in \mathbb{N}$. Clearly $\lambda_1(D_k) = \lambda_1(C_k)\lambda_2(C_k)$ and the corresponding Perron eigenvectors of D_k, D_k^{T} are $\mathbf{u}_{1,k} \wedge \mathbf{u}_{2,k}, \mathbf{w}_{1,k} \wedge \mathbf{w}_{2,k} \in \mathbb{R}_+^{\binom{n}{2}}$. Then Theorem 1.1 applied to $B_k, k \in \mathbb{N}$ yields that

$$\mathrm{span}(\mathbf{u}_{1,k}, \mathbf{u}_{2,k}) \to U_2 = \mathrm{span}(\mathbf{u}_1, \mathbf{u}_2) \in \mathrm{Gr}(2, \mathrm{n}, \mathbb{R}),$$
$$\mathrm{span}(\mathbf{w}_{1,k}, \mathbf{w}_{2.k}) \to W_2 \in \mathrm{Gr}(2, \mathrm{n}, \mathbb{R}).$$

As $\mathbf{w}_{1,k}^{\mathrm{T}}\mathbf{u}_{2,k} = 0$, $||\mathbf{u}_{2,k}|| = 1$ and $\mathbf{w}_{1,k} \to \mathbf{w}_1$ it follows that $\mathrm{span}(\mathbf{u}_{2,k}) \to \mathrm{span}(\mathbf{u}_2)$. As $\mathbf{u}_{1,k} \wedge \mathbf{u}_{2,k} \in \mathbb{R}_+^{\binom{n}{2}}$ it follows that $\mathbf{u}_{2,k} \to \mathbf{u}_2$. Clearly $\mathbf{w}_1 \in W_2$. As $\mathbf{w}_{2,k}^{\mathrm{T}}\mathbf{u}_{1,k} = 0$, $\mathbf{w}_{2,k}^{\mathrm{T}}\mathbf{u}_{2,k} = 1$ it follows that $\mathbf{w}_{2,k} \to \mathbf{w}_2$, which is the unique vector in W_2

satisfying the conditions $\mathbf{w}_2^{\mathrm{T}}\mathbf{u}_1 = 0,\ \mathbf{w}_2^{\mathrm{T}}\mathbf{u}_2 = 1$. So $\mathbf{w}_{1,k} \wedge \mathbf{w}_{2,k} \to \mathbf{w}_1 \wedge \mathbf{w}_2 \in \mathbb{R}_+^{\binom{n}{2}}$, which is the positive eigenvector of the following rank one matrix

$$E_2 := \frac{\mathbf{u}_1 \wedge \mathbf{u}_2 (\mathbf{w}_1 \wedge \mathbf{w}_2)^{\mathrm{T}}}{||\mathbf{u}_1 \wedge \mathbf{u}_2||\ ||\mathbf{w}_1 \wedge \mathbf{w}_2||} = \lim_{k\to\infty} \frac{D_k^{\mathrm{T}}}{||D_k^{\mathrm{T}}||}.$$

The above equality is equivalent to (6.3) for $i = 2$.

Recall that all the eigenvalues of D_k are of the form $\lambda\mu,\ \lambda,\mu \in \mathrm{spec}\ (C_k)$, where either $\lambda \neq \mu$ or $\lambda = \mu$ is a multiple eigenvalue of D_k. Thus if $|\lambda| \geq |\mu|$ then $\lambda_2(D_k) > |\mu|$ unless $\lambda = \lambda_1(D_k),\ \mu = \lambda_2(C_k)$. Combine all these facts to obtain (6.4) for $p = 2$.

Assume now that $p > 2$. By considering the compound matrices $\wedge_i A_k, k \in \mathbb{N}$ for $i = 3, ..., p$ we deduce the rest of theorem as in the case $p = 2$. ☐

Assume the assumptions of Theorem 6.1. Let $\mathbf{z} \in \mathbb{R}^n$ and $S(\mathbf{z}) = p - 1$. Since $\pm\mathbf{z}$ can be completed to a p-Tchebyshev $\mathbf{z}_1, ..., \mathbf{z}_p$ it follows that it is impossible that $\mathbf{w}_i^{\mathrm{T}}\mathbf{z} = 0$ for $i = 1, ..., p$. Thus one can estimate the behavior of $\widehat{C_k\mathbf{z}}$ as $k \to \infty$.

Theorem 6.2 *Let $A_k \in \mathrm{GL}_n(\mathbb{C}), k \in \mathbb{N}$. Assume that for $k > N$ the following conditions satisfied: For $p \in [1, n] \cap \mathbb{Z}$ there exists $\alpha \in (0, 1)$ and:*
(a) biorthonormal sets $\mathbf{x}_{1,k}, ..., \mathbf{x}_{p,k}, \mathbf{y}_{1,k}, ..., \mathbf{y}_{p,k} \in \mathbb{C}^n$ such that

$$||\mathbf{x}_{i,k}|| = 1,\ \mathbf{y}_{i,k}^{\mathrm{T}}\mathbf{x}_{j,k} = \delta_{ij},\ i,j = 1, ..., p,\ k > N,$$
$$\lim_{k\to\infty} \mathbf{x}_{i,k} = \mathbf{u}_i,\ ||\mathbf{u}_i|| = 1,\ \lim_{k\to\infty} \mathbf{y}_{i,k} = \mathbf{v}_i,\ \mathbf{v}_i^{\mathrm{T}}\mathbf{u}_j = \delta_{ij},\ i,j = 1, ..., p.$$

(b) $\lambda_{1,k}, ..., \lambda_{p,k} \in \mathrm{spec}\ (A_k)$ are simple roots of the characteristic polynomial of A_k such that

$$A_k\mathbf{x}_{i,k} = \lambda_{i,k}\mathbf{x}_{i,k},\ A_k^{\mathrm{T}}\mathbf{y}_{i,k} = \lambda_{i,k}\mathbf{y}_{i,k},\ |\lambda_{i,k}| \geq \alpha|\lambda_{i+1,k}|,\ i = 1, ..., p, \quad \text{for any } k > N,$$

where $\lambda_{p+1,k}$ is any eigenvalue of A_k different from $\lambda_{1,k}, ..., \lambda_{p,k}$. Furthermore, there exists an operator norm $|||\cdot||| : \mathrm{M}_n(\mathbb{C}) \to [0, \infty)$ such that

$$|||A_k - \sum_{i=1}^{p} \lambda_{i,k}\mathbf{x}_{i,k}\mathbf{y}_{i,k}^{\mathrm{T}}||| \leq \alpha|\lambda_{p,k}|, \quad k > N. \tag{6.7}$$

Let $C_k := A_k...A_1, k \in \mathbb{N}$. Then there exists $N_1 > N$ that for $k > N_1$ the following conditions hold. C_k has p simple eigenvalues $\lambda_1(C_k), ..., \lambda_p(C_k)$ such that $|\lambda_1(C_k)| > ... > |\lambda_p(C_k)|$. It is possible to choose the corresponding eigenvectors of C_k, C_k^{T} as $\mathbf{u}_{1,k}, ..., \mathbf{u}_{p,k}, \mathbf{w}_{1,k}, ..., \mathbf{w}_{p,k}$ such that equalities (6.2) - (6.6) hold.

Proof. We first consider the case $p = 1$. By considering the matrices $\lambda_{1,k}^{-1} A_k$ it is enough to prove the above theorem in the case $\lambda_{1,k} = 1$ for $k > N$. Let $R_k := A_k - \mathbf{x}_{1,k}\mathbf{y}_{1,k}^{\mathrm{T}}$ for $k > N$. The spectral decomposition of A yields and (6.7) yields

$$R_k\mathbf{x}_{1,k} = R_k^{\mathrm{T}}\mathbf{y}_{1,k} = 0, \quad |||R_k||| \le \alpha, \quad k > N. \tag{6.8}$$

In order to use the arguments of the proof of Theorem 1.1 it is enough to show that for each $m > 1$ there exists $K(m)$ such that if $j > K(m)$

$$|||A_{j+m}...A_{j+1} - \mathbf{u}_1\mathbf{v}_1^{\mathrm{T}}||| < |\alpha|^m + \frac{1}{m}. \tag{6.9}$$

Consider the product

$$A_{j+m}...A_{j+1} = (\mathbf{x}_{1,j+m}\mathbf{y}_{1,j+m}^{\mathrm{T}} + R_{j+m})...(\mathbf{x}_{1,j+1}\mathbf{y}_{1,j+1}^{\mathrm{T}} + R_{j+1}). \tag{6.10}$$

Expand this product to 2^m terms. The first term in this product is

$$\mathbf{x}_{1,j+m}\mathbf{y}_{1,j+m}^{\mathrm{T}}...\mathbf{x}_{1,j+1}\mathbf{y}_{1,j+1}^{\mathrm{T}} = (\prod_{i=j+1}^{j+m-1} \mathbf{y}_{1,i+1}^{\mathrm{T}}\mathbf{x}_{1,i})\mathbf{x}_{1,j+m}\mathbf{y}_{1,j+1}^{\mathrm{T}}.$$

Hence it converges to $\mathbf{u}_1\mathbf{v}_1^{\mathrm{T}}$ as $j \to \infty$. Consider the last term in (6.10). Since $|||\cdot|||$ is an operator norm

$$|||R_{j+m}...R_{j+1}||| \le |||R_{j+m}|||...|||R_{j+1}||| \le \alpha^m.$$

It is left to show that that all other $2^m - 2$ terms in (6.10) tend to zero. Each of this term contains either a factor $\mathbf{x}_{j+i+1}\mathbf{y}_{j+i+1}^{\mathrm{T}}R_{j+i}$ or $R_{j+i+1}\mathbf{x}_{j+i}\mathbf{y}_{j+i}^{\mathrm{T}}$. Use (6.8) to deduce

$$\mathbf{x}_{j+i+1}\mathbf{y}_{j+i+1}^{\mathrm{T}}R_{j+i} = \mathbf{x}_{j+i+1}(\mathbf{y}_{j+i+1} - \mathbf{y}_{j+i})^{\mathrm{T}}R_{j+i},$$
$$R_{j+i+1}\mathbf{x}_{j+i}\mathbf{y}_{j+i}^{\mathrm{T}} = R_{j+i+1}(\mathbf{x}_{j+i} - \mathbf{x}_{j+i+1})\mathbf{y}_{j+i}^{\mathrm{T}}.$$

If a term contains more then one of such factors choose the above modification at one factor exactly. Now estimate the norm of this term by taking the products of the norms of m factors. It now follows that each of this terms tends to zero. Hence (6.9) follows. Now we can repeat the arguments of the proof of Theorem 1.2 to prove the theorem for $p = 1$.

To prove the theorem for $p > 1$ we consider the wedge products $\wedge_i A_k, k \in \mathbb{N}$ for $i \in [2, p]$. The spectral analysis of $\wedge_i A_k$ implies that $\wedge_i A_k, k \in \mathbb{N}$ satisfy the above conditions for $p = 1$. Use the arguments of the proof of Theorem 6.1 to deduce the theorem in this case. □

Assume that $R \in \mathrm{M}_n(\mathbb{C})$ has a spectral radius $\rho(R) \in [0,1)$. It is well known that for any $\alpha \in (\rho(R), 1)$ there exists an operator norm $|||\cdot||| : \mathrm{M}_n(\mathbb{C}) \to [0,\infty)$ such that $|||R||| \le \alpha$.

Corollary 6.3 *Let* $A_k \in \mathrm{GL}_n(C), k \in \mathbb{N}$. *Assume that* $\lim_{k\to\infty} A_k = A \in \mathrm{M}_n(\mathbb{C})$. *Suppose furthermore that* $\lambda_1, ..., \lambda_p$ *are* p *simple roots of* $\det\ (\mathrm{zI} - \mathrm{A})$, *where* $\rho(A) = |\lambda_1| > ... > |\lambda_p| > 0$. *Assume furthermore that any other eigenvalue* $\lambda \in \mathrm{spec}\ A\backslash\{\lambda_1, ..., \lambda_p\}$ *satisfies* $|\lambda| < |\lambda_p|$. *Then* $A_k, k \in \mathbb{N}$ *satisfy the assumptions of Theorem 6.2, where* $A\mathbf{u}_i = \lambda_i \mathbf{u}_i,\ i = 1, ..., p$.

References

[1] G. Birkhoff, Extensions of Jentzsch's theorem, *Trans. Amer. Math. Soc.* 85 (1957), 219-227.

[2] J. Borcea and B. Shapiro, On families of functions satisfying finite recurrence relations, *preprint* January 14, 2004.

[3] R.A. Brualdi, S. Friedland and A. Pothen, The sparse basis problem and multilinear algebra, *SIAM J. Matrix Anal. Appl.* 16 (1995), 1-20.

[4] S. Friedland, Invariant measures of groups of homeomorphisms and Auslander's conjecture, *J. Ergod. Th. & Dynam. Sys.* 15 (1995), 1075-1089.

[5] F.R. Gantmacher, *The Theory of Matrices*, Chelsea Publ. Company, vol. II, 1964.

[6] P. Griffiths and J. Harris, *Principles of Algebraic Geometry*, Wiley Interscience, 1978.

[7] J. Hajnal, On product of non-negative matrices, *Math. Proc. Camb. Phil. Soc.* 79 (1976), 521-530.

[8] D.J. Hartfiel, *Nonhomogeneous Matrix Products*, World Scientific Publishing Co., 2002.

[9] D. Hilbert, Neue Begründung der Bolya-Lobatschefskyschen Geometrie, *Math. Ann.* 57 (1903), 137-150.

[10] S. Karlin, *Total Positivity*, Stanford Univ. Press, 1968.

[11] V. I. Oseledec, A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems, *Trans. Moscow Math. Soc.* 19 (1968), 197-221.

[12] E. Seneta, *Nonnegative Matrices and Markov Chains*, 2nd ed., Springer, 1981.